MOLECULAR CHARACTERISATION OF

Burkholderia pseudomallei

by

CLARE DAVIES B.Sc. (Hons)

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

Department of Biological Sciences
Faculty of Science

In collaboration with
D.E.R.A. PORTON DOWN

March 2001

MOLECULAR CHARARCTERISATION OF *Burkholderia pseudomallei*
CLARE DAVIES Bsc. (Hons.)

ABSTRACT

A programme of research was carried out to attempt the molecular characterisation of the human and animal pathogen, *Burkholderia pseudomallei*, the causative agent of melioidosis and the newly described avirulent species, *B.thailandensis* for comparative purposes. Melioidosis is still little understood, and so the clinical approach to the prevention and control of melioidosis must ultimately rest upon the basic understanding of the causative organism, particularly the pathogenic properties of *B.pseudomallei*.

A range of *B.pseudomallei* and *B.thailandensis* isolates were cultured and the extracellular products were isolated and concentrated and an initial study conducted to identify potential target molecules for cloning. Those isolates tested were shown to have somewhat differing ECP profiles when analysed with SDS-PAGE and antigenic profiles when subject to immunoblotting using convalescent human serum although isolates within and between species shared a number of common bands. The ECPs were also tested for a range of activities and it was established that both species had proteolytic and phospholipase activities neither had a haemolytic activity and only isolates of *B.pseudomallei* had a hexosaminidase activity a putative pathogenicity determinant.

Genomic DNA of *B.pseudomallei* was used to construct genomic libraries in a range of *E.coli* host vector systems. A λGT11 genomic library was screened with antisera for the presence of *B.pseudomallei* antigens and a number of natural and synthetic substrates for the presence of haemolytic and proteolytic components. Screening yielded one stable immunopositive clone with a novel positive reaction in the form of a "halo" of reaction around the plaque. The 5 kbp cloned fragment was subcloned into a plasmid vector, and the resulting recombinant molecule, pBPGT2 was DNA sequenced and found to contain a putative pilin gene. Attempts were made to determine the size of the recombinant antigen and to further express the pilin gene product all of which were unsuccessful. A southern blot procedure confirmed the fidelity of the cloning procedure proving that the fragment was from the host organism, *B.pseudomallei*. A further southern blot procedure tested for the presence of the pilin sequence in a range of *B.pseudomallei* and *B.thailandensis* isolates proving the presence of the gene in only isolates of *B.pseudomallei*.

PCR primers were designed to amplify the DNA encoding the active site of the ADP-ribosylating toxin (ETA) of *Pseudomonas aeruginosa* and a PCR reaction was carried out on a number of *B.pseudomallei* and *B.thailandensis* isolates. The reaction yielded a 500 bp product in only *B.pseudomallei* isolates and DNA sequencing of the product revealed no obvious homology to ETA of *P.aeruginosa* but was used as a probe to isolate a larger fragment of DNA which was found to encode a number of interesting putative genes. These included one with homology to a porin similar to that of the pathogen *Neisseria gonorrhoea*, with a role in virulence.

During attempts to digest the genomic DNA of *B.pseudomallei* isolate 4845 with the restriction enzyme *Sau*3A two 12 kbp bands of DNA were resistant to the endonuclease activity. Attempts were made to clone these bands into a range of plasmid vectors with two clones containing deleted products. DNA sequencing proved inadequate with only a small amount of sequence information obtained. However, towards the final stages of the research project sequence information from the *B.pseudomallei* genome sequencing project facilitated the recognition of a 38 kbp fragment containing the sequence information from one of the clones, which encodes an alkaline protease and a putative haemagglutinin and is postulated to be a Pathogenicity Island encoding secreted virulence factors.

The sequencing project also facilitated the isolation of two putative hexosaminidase genes postulated to be responsible for the activities observed when testing the *B.pseudomallei* isolates concentrated ECPs. Future studies for the putative genes identified and other components of *B.pseudomallei* are discussed.

LIST OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF FIGURES:

LIST OF TABLES:

ACKNOWLEDGMENT

I would like to acknowledge the following people for without their help this work would not have been possible:

Dr. Martyn Gilpin for giving me the opportunity to carry out this research project and for his supervision, guidance, expert assistance and great patience.

Dr. Hilton Grayson, Dr. Michele Kiernan and Mrs Lynne Cooper for their invaluable assistance, advice and good humour.

Dr. David Dance for his invaluable help and supervision, and for providing bacterial strains and genomic DNA.

Drs. Richard Titball, Kerry Mack and Petra Oysten of Porton Down for providing genomic DNA, suicide vectors, carrying out DNA sequencing and for their general kind help and advice.

Drs. Ty Pitt and Hazel Aucken of CPHL, Colindale for providing bacterial strains, antisera and advice.

The postgraduate research students. Magdy, who was not only by my side from the very beginning but without whose help and friendship I would never have got this far. Sarah and Laura, for their relentless friendship, humour, support and above all therapeutic shopping trips.

The staff of Biological Sciences for all their expert help with a special mention to Tara Arrowsmith, Jo Carter, Cynthia Palmer, Marion Pill, Nick Crocker, Annette Wrathnell and Janet Roper.

To Howie whose love and patience knows no end and whose support has got me through the past year.

Last but most certainly not least is my parents and it is to them I dedicate this thesis for without their love, support (and computer facilities), especially in these last few months, I would never have completed this thesis.

Authors Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other university award.

This study has been financed with the aid of a studentship from the University of Plymouth and in collaboration with DERA, Porton Down, Salisbury, Wiltshire.

A programme of advanced study was also undertaken involving a final year honours module in Bacterial Pathogenicity and Microbial Biotechnology. Training was also given in the use of DNA computer software packages, for example DNasis, Amplify and in computer-aided nucleotide and amino acid sequence analysis

Meetings and Conferences Attended;

International Congress on Melioidosis, November 22-25, 1998. Bangkok, Thailand.
Society of General Microbiology, general meeting, April 12-16, 1999. Edinburgh.

UK Melioidosis Meetings;

The UK Melioidosis Network is composed of a group of clinicians and scientists from throughout the UK who meet every 6 months to discuss matters relating to melioidosis and report their latest research work on *B.pseudomallei*.

22 November, 1997 London School of Hygiene and Tropical Medicine.
25 May, 1998 Oxford.
17 December, 1998 Colindale Public Health Laboratory, London.*
7 May, 1999. Derriford Hospital, Plymouth.
26 January, 2000 Imperial College, London.*
28 June, 2000 CBD Porton Down, Salisbury.

* Indicates when work was presented by the author.

Signed..C.A.Dawson...

Date..13/3/01.....

Abbreviations

A	adenine
ADP	adenine diphosphate
mA	milliampere
Ara	arabinose
ATP	adenosine triphosphate
BHIB	brain heart infusion broth
bp	base pairs
C	cytosine
°C	degrees centigrade
$CaCl_2$	calcium chloride
cm	centimetre
CsCl	caesium chloride
DMF	dimethylformamide
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothreitol
ECP	extracellular products
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbant assay
fg	femtogram
G	guanine
g	gramme
GTP	guanosine triphosphate
x g	acceleration due to gravity
h	hour
HCl	hydochloric acid
IAA	isoamyl alcohol
IPTG	isopropylthio-β-D-galactosidase
kbp	kilobase pairs
kDa	kilodaltons
kg	kilogram
LB	Luria-Bertani medium
M	molar
mA	milliamp
Mbp	megabase pairs
mg	milligram
µg	microgram
µm	micrometre
ml	mililitre
µl	microlitre
min	minute
mm	milimetre
mM	milimolar
mRNA	messenger ribonucleic acid
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanogram
OD	optical density

OD_{600}	optical density at 600 nm wavelength
^{32}P	phosphorus-32
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
%	percent
pg	picogram
PMSF	phenylmethyl sulfonyl fluoride
RAPD	randomly amplified polymorphic DNA
RNA	ribonucleic acid
mRNA	messenger RNA
rRNA	ribosomal RNA
RNase	ribonuclease
*Sau*3A	*Sau*3AI
SDS	sodium dodecyl sulphate
sec	second
T	thymine
TBE	tris borate EDTA buffer
TE	tris EDTA buffer
TEMED	*N,N,N',N'*-tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
U	uracil
UTP	uridine triphosphate
UV	ultra violet
V	volt
v/v	volume for volume
w/v	weight for volume
Xgal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

CHAPTER 1

INTRODUCTION

1.1 HISTORY OF MELIOIDOSIS AND THE CAUSATIVE AGENT, *Burkholderia pseudomallei.*

1.1.1 The First Descriptions of Melioidosis.

Melioidosis was first reported in April 1911 by a British pathologist, Captain A.Whitmore working in Rangoon General Hospital, Burma (Whitmore and Krishnaswamy, 1912). While carrying out a post-mortem on a 40-year-old "morphia" addict, Whitmore discovered a glanders-like infectious disease and cultures grown yielded an organism the same size and shape as "*Bacillus mallei*", the name used at the time for the causative agent of glanders. Hence the organism was originally named "*Bacillus pseudomallei*" due to this resemblance. Over the next 6 years in Rangoon, Knapp and Krishnaswamy described further cases of the disease based on Whitmore's initial observations (Dance, 1991). In fact, the disease proved to be surprisingly common, being identified in approximately 5% of post-mortem reports in Rangoon General Hospital (Krishnaswamy, 1917).

The disease was again recognised later by Stanton and Fletcher (1921) working in Kuala Lumpur, who were responsible for naming the disease melioidosis, which was derived from the Greek term for a variety of conditions resembling glanders. Stanton and Fletcher (1932) predicted that the disease would prove to be far more prevalent than appreciated at that time; a claim that has been substantiated in the past twenty years or so, as melioidosis has gained greater recognition, particularly in the tropics. In the same report, the authors developed the concept that the wild rat was the natural reservoir of infection and that humans contracted the

disease through ingestion of food or water contaminated with rat excreta. Several years later, French workers in Indo-China made a number of observations of the disease and its causative agent, which has led to current understanding of the ecology of *B.pseudomallei* (Dance, 2000c). The French workers noted that infections in humans and animals often followed exposure to mud and water and that the disease was very rarely observed in wild rodents (Dance, 2000c). Chambon (1955) isolated the organism from mud and water of ponds and rice paddies in southern Vietnam and thus it became apparent that the organism was actually an environmental saprophyte, and that disease in man and animals ensued as a consequence of exposure to the organism in soil and surface water (Dance, 2000c).

The first confirmed indigenous case of melioidosis in Thailand was reported in 1955 (Chittivej *et al.*, 1955), despite previous descriptions of an imported case in 1928 and cases of melioidosis in two prisoners of war in 1947 (Dance, 1991). In Australia, melioidosis was first identified in sheep in Queensland in 1949 (Dance, 1990) and the first human case was described by Rimington (1962) in a 32-year-old diabetic man. Cases of melioidosis were then described in many areas of the tropics and by 1957 more than 300 human cases of melioidosis had been documented worldwide (Smith *et al.*, 1987).

1.1.2 Military Experience of Melioidosis.

A high incidence of melioidosis has been reported in army personnel who have participated in military campaigns taken place throughout history in Southeast Asia (Smith *et al.*, 1987), which has consequently led to an increased recognition of the disease worldwide. During the Civil War in Vietnam, there was a notable increase in the number of infections occurring in troops (Rubin *et al.*, 1963). However, melioidosis first gained recognition in the Western world during the French-Indochina War, dating from 1948 to 1954, where at least 100 cases were reported in French forces (Dance, 1990). Some years later, melioidosis again provoked

interest when 343 American military personnel on active duty in Vietnam from 1965 to 1971 were treated for overt disease (Chodimella *et al.*, 1997). Clayton *et al.* (1973) postulated that some of the American military personnel who had been exposed to *B.pseudomallei* during the conflict may go on to develop the disease at a later date, due to its long latency period. Indeed there have now been a number of reports of Vietnam veterans developing melioidosis more than 20 years after the conflict (Chodimella *et al.*, 1997).

More recently in a report by Embi *et al.* (1992) the incidence of antibody to exotoxin was over 40% and 30% for military personnel in the age group 26 to 32 years serving in Sarawak and Sabah, Malaysia, respectively. Also in a recent study in Singapore between 1987 and 1994, 23 cases of melioidosis were diagnosed in persons serving in the Singapore Armed Forces (Lim, 1997). Melioidosis therefore poses a significant health threat to military personnel serving in endemic areas, particularly if they sustain injury during this period.

1.1.3 Recent History of Melioidosis.

Little further progress was made with the disease until the mid 1970's (Dance, 2000b). In 1975 at the meeting of the Infectious Disease Group of Thailand, Punyagupta *et al.* (1976) reported on 10 culture proven cases of melioidosis. The first national workshop on melioidosis was then held 10 years later, compiling more than 700 reported cases that led to research activities on various aspects of the disease (Leelarasamee, 2000). Melioidosis now qualifies as an emerging infection, since it has been recognised with increased frequency over the past two decades both within established endemic areas and elsewhere (Dance, 2000b). Increasing recognition of the disease has partly been due to an increased awareness on the part of the clinicians and microbiologists, especially as diagnostic facilities have improved greatly, allowing the identification of more cases of melioidosis. However, Dance (2000b) suggested that there could actually have been a genuine increase in the disease in recent times, perhaps

due to changing environmental factors and farming practices.

Enthusiasm and interest for this disease reached the point that an International Symposium on "Melioidosis: Prevailing Problem and Future Direction" was organised in 1994 (Leelarasamee, 2000), and more recently an International Congress on Melioidosis was held in Bangkok in November, 1998 where researchers from all over the world discussed significant progress in different aspects of the disease.

1.1.4 Taxonomy of The Genus, *Burkholderia* and *Burkholderia pseudomallei.*

Under current nomenclature, the agent of melioidosis is known as *Burkholderia pseudomallei.* However, the same organism has been known as *Bacillus pseudomallei, Pfeifferella whitmori, Bacillus whitmore, Flavobacterium pseudomallei, Actinobacillus pseudomallei, Loefflerella* and *Malleomyces pseudomallei* (Tan *et al.*, 1997). Furthermore, based on assessment of nutritional and biochemical characteristics, Haynes (1957) reassigned the organism to the genus *Pseudomonas*. More recently, investigations based on 16S ribosomal RNA (rRNA) sequences, DNA-DNA homology values, cellular lipid and fatty acid composition, and phenotypic characteristics, led Yabuuchi *et al.* (1992) to propose a new genus, *Burkholderia*, for the RNA homology group II of the genus *Pseudomonas*, which included *P.pseudomallei.* Seven species in this group were transferred to the new genus *Burkholderia* at the time, those being *B.cepacia, B.mallei, B.pseudomallei, B.caryophylli, B.gladioli, B.pickettii* and *B.solanacearum*. Further work on internal transcribed spacer (ITS) sequence homology by Tyler *et al.* (1995) supported the transfer of these former *Pseudomonads* to a new genus, as well as work by Li and Hayward (1994) comparing whole cell protein profiles. The genus was named *Burkholderia*, after W.H.Burkholder, the American bacteriologist who first discovered the etiologic agent of rotten onion (Burkholder, 1950), now known as *Burkholderia cepacia.*

Since the initial transfer, other species have been added and conversely removed from *Burkholderia*. *B.solanacearum* and *B.pickettii* were found to belong to a separate RNA subgroup establishing another genus, *Ralstonia* (Yabuuchi *et al.*, 1995). Other species that have now been transferred to the genus *Burkholderia* include *B.plantarii*, *B.glumae*, *B.vandii* (Urakami *et al.*, 1994), *B.pyrrocinia* (Viallard *et al.*, 1998) and *B.vietamensis* (Gillis *et al.*, 1995). Recently, new species have been identified and deemed members of the genus *Burkholderia*. One, *B.caribensis* was isolated on the island of Martinique in the French West Indies by Achouak *et al.* (1999) and another, *B.graminis* identified by Viallard *et al.* (1998). However, of this genus at present, only *B.mallei*, *B.pseudomallei*, *B.cepacia* and *B.gladioli* are pathogenic for humans or animals, whereas the other members are mostly found in, or on, plants (Otterbein *et al.*, 1998).

1.1.5 Bacteriology.

B. pseudomallei is an oxidase positive, gram-negative rod, which is motile by means of a polar tuft of 2 to 4 flagella and able to use a wide range of substrates as carbon sources (Redfearn *et al.*, 1966; Dance, 1990; Sanford, 1995; DeShazer *et al.*, 1997). It is obligately aerobic and non spore-forming (Leelarasamee and Bovornkitti, 1989), and when stained with methylene blue, Wayson or Wright stain, marked irregularities with a bipolar "safety pin" pattern are observed (Sanford, 1995). Growth of the organism is accompanied by a distinctive sweet, earthy smell (Dance, 1990) and it is readily identified by its growth and characteristic morphology on Ashdown's selective medium (Chaowagul, 1996). The bacterium appears as cream to yellow-coloured colonies, 2 to 3 mm in diameter with a reported metallic sheen over areas of confluent growth (Smith *et al.*, 1987; Dance *et al.*, 1989b). On Blood agar, colonies are opaque white and after 48 hours a narrow zone of beta haemolysis becomes visible around each colony (Jesudason *et al.*, 1997). Colonies can vary from smooth to rough in appearance and often become wrinkled after a few days of incubation (Leelarasamee and Bovornkitti, 1989). The

respective virulence of these colonial types has been studied by Veljanov and Najdenski (1993), who observed that bacteria from smooth colonies survived and multiplied more actively in macrophages than the alternative, rough colonial bacteria. Other characteristic properties of *B.pseudomallei* include the reduction of nitrates, the ability to accumulate poly-β-hydroxybutyrate, and arginine dihydrolase and gelatinase activity (Dance, 1990). Plasmids have not been identified in the bacterium (Dance, 1990), but 32 *B.pseudomallei* bacteriophages have been isolated (Manzenyuk *et al.*, 1994).

The guanine + cytosine (G+C) content of the DNA of *B.pseudomallei* is 69% (Redfearn *et al.*, 1966). Recent work by Songsivilai and Dharakul (2000) has shown that two large replicons constitute the genome of *B.pseudomallei*, which appear to be chromosomes as each one contains rRNA sequences. The combined genome size of *B.pseudomallei* estimated from pulsed field gel electrophoresis (PFGE) separation of undigested chromosomal DNA was calculated to be approximately 6.54 megabase pairs (Mbp), which makes it one of the largest bacterial genomes (Songsivilai and Dharakul, 2000). Songsivilai and Dharakul (2000) suggest the average gene size as being about 1031 bp and that the genome is gene rich, with about 89% of the potential coding capacity being used as coding sequences

1.1.6 *Burkholderia mallei.*

When first describing *B.pseudomallei* in 1911, Whitmore thought it bore a striking resemblance to *B.mallei*, the causative agent of glanders. Glanders is primarily an infectious disease of equine animals, for example the horse, mule or donkey (Sanford, 1995), but it is also infectious for humans and is characterised by a high mortality rate. Melioidosis does indeed bear a remarkable resemblance to glanders both clinically and pathologically, but it is epidemiologically dissimilar (Sanford, 1995). Once found worldwide, glanders has now been eliminated in Western Europe and Northern America through the use of the mallein test, but

recently Arun *et al.* (1999) confirmed that equine glanders is still endemic in Turkey. Glanders was recognised as early as 400BC and it has been postulated that *B.pseudomallei* may have unwittingly been the etiological agent of many past cases of "glanders" in man and animals before its own recognition (Rubin *et al.*, 1963).

The organism, *B.mallei* is a non-motile, gram-negative bacillus (Sanford, 1995) and so differs from *B.pseudomallei* with respect to motility. The two organisms have very similar antibiotic susceptibility profiles (Kenny *et al.*, 1999) and although biochemical testing distinguishes between *B.mallei* and *B.pseudomallei*, it creates a high risk of infection through working with live cultures. For this reason, a polymerase chain reaction (PCR) procedure has been developed to discriminate between *B.mallei* and *B.pseudomallei*, detecting a nucleotide difference in the 23S rRNA sequences at position 2143 (Bauernfeind *et al.*, 1998).

1.1.7 *Burkholderia thailandensis.*

Wuthiekanun *et al.* (1996) investigated an observed discrepancy between the distribution of the disease and the distribution of the organism in the environment, which compared the biochemical and antigenic features, and antibiotic susceptibilities of clinical and environmental isolates of *B.pseudomallei*. Results showed that although clinical and environmental isolates of *B.pseudomallei* were morphologically similar and antigenically indistinguishable, all clinical isolates of *B.pseudomallei* failed to utilise the carbohydrate, L-arabinose (Ara^-), whereas some soil isolates from areas of endemic infection could utilise this carbohydrate (Ara^+). The Ara^+ biotype was also shown by Wuthiekanun *et al.* (1996) to differ from the Ara^- isolates by its positive reaction for adonitol, 5-keto-gluconate and xylose and by its failure to assimilate dulcitol, erythritol and trehalose. To determine whether the varieties of *B.pseudomallei* were genomically related or were sufficiently different to warrant further investigation of species identity, the variation in rRNA gene loci of isolates with both

phenotypes was examined by Trakulosomboon *et al.* (1997). At the species level, these sequences are sufficiently conserved that differences within them often reflect intra-species variation or types (Trakulosomboon *et al.*, 1997). Results showed that approximately half of the isolates from soil specimens were of ribotypes similar to those from patients with melioidosis, but were markedly different from other environmental isolates, which is inconsistent with the single species concept (Trakulosomboon *et al.*, 1997). Investigations by Dharakul *et al.* (1999) on the 16S rRNA-encoding genes showed that there were differences between the Ara^+ and Ara^- isolates, although the difference observed was smaller than amongst the other species of the genus *Burkholderia*. Further supportive evidence for the Ara^+ isolates being a different species was shown by Brett *et al.* (1998), using 16S rDNA-based phylogenetic analysis, and the new species name suggested for this organism was *Burkholderia thailandensis* (Brett *et al.*, 1998).

Perhaps the most significant difference associated with the Ara^+ isolates was discovered in an investigation by Brett *et al.* (1997), which showed that these isolates had a markedly decreased virulence in the Syrian golden hamster model; in fact demonstrating more than a 10^5-fold decrease in virulence relative to the clinical Ara^- isolates. At present there has only been one reported case of "melioidosis" having been caused by an Ara^+ isolate, but it was suggested that this was due to an unusually heavy inoculation at the time of the accident responsible for the infection (Lertpatanasuwan *et al.*, 1999).

Further work has now been carried out investigating other differences between the two species. Work by Winstanley *et al.* (1998) on conservation amongst flagellin genes, showed high levels of conservation amongst the Ara^- clinical isolates (over 99%) which contrasted with less than a 90% similarity on comparison with an Ara^+ isolate. A conserved deletion of 5 amino acids in the flagellin gene sequence within 41 Ara^+ isolates has been demonstrated by Wajanarogana *et*

al. (1999), thus providing one form of differentiation between Ara^+ and Ara^- isolates. Another recent finding is that Ara^- and Ara^+ isolates have different macrorestriction patterns on PFGE (Chaiyaroj *et al.*, 1999), and investigations by Winstanley and Hart (2000) have shown that type III secretion genes were absent from most of the Ara^+ isolates tested but present in all Ara^- isolates. Further differences will inevitably be highlighted between the two species, especially with respect to virulence determinants and the genes that encode them, and so the newly discovered species *B.thailandensis* will prove an invaluable tool for investigating the pathogenicity of *B.pseudomallei*.

1.1.8 *Burkholderia uboniae.*

Recently, Yabuuchi *et al.* (2000) have identified another new species, which was originally identified as *B.thailandensis*, and although able to assimilate L-arabinose, revealed a DNA-DNA reassociation rate of 36.7% under stringent conditions with the type strain of *B.thailandensis*. The authors noted that this strain differed from the *B.thailandensis* type strain by its physiological, biochemical and nutritional characteristics without any significant difference in cellular fatty acid and lipid composition, and has named the new species *Burkholderia uboniae*.

1.2 ECOLOGY OF *B.pseudomallei.*

1.2.1 Geographical Distribution of Melioidosis Cases and of *B.pseudomallei.*

The majority of cases of melioidosis reported in the world literature have arisen in Southeast Asia and Northern Australia, with most cases in general tending to occur between the latitudes 20°N and 20°S (Leelarasamme and Bovornkitti, 1989; Dance 1991). Fig. 1.1 shows the worldwide distribution of *B.pseudomallei* and *B.pseudomallei*-like isolates (Dance, 1991).

Fig.1.1 The worldwide distribution of *B.pseudomallei* and *B.pseudomallei*-like isolates (Dance, 1991). Shaded areas represent the main endemic areas and hatched areas and asterisks represent sporadic isolates.

Considering that there are still limited microbiological facilities in some areas of the tropics, coupled with a lack of awareness on the part of some clinicians, melioidosis may actually be under-diagnosed in many rural areas of the tropics and so the overall picture has yet to be fully elucidated. Melioidosis, however, is rarely encountered in Europe but as worldwide travel becomes easier and more frequent, particularly with regard to the transportation of animals, it is likely that the disease could spread from endemic areas to those areas not presently believed to be at threat from melioidosis. Infected animals may excrete *B.pseudomallei* in sputum, pus, urine and faeces, which may then contaminate their new environment, providing it is suitable (Dance, 2000b). There have been many cases where travel to endemic areas has caused melioidosis, which becomes apparent on return to the Western world (Dance, 1990). Of the 15 human cases diagnosed in the UK between 1988 and 1998, five originated from Bangladesh and one each from India and Pakistan (Dance *et al.*, 1999). A similar problem has also been shown to occur in Australia, where imported melioidosis cases are seen in southern hospitals, normally from tourists returning from the Northern Territory of Australia as well as Southeast Asia. There have also been documented cases of melioidosis occurring as a result of animal transportation from endemic areas. This occurred in France during the mid 1970s, after an imported panda donated by Mao-Tse-Tung probably caused an outbreak of melioidosis (Mollaret, 1988). There have also been cases of melioidosis in imported primates in Britain originating from Indonesia and the Philippines (Trakulsomboon *et al.*, 1994). In temperate Southwest Australia between 1966 and 1991, melioidosis occurred in a farmer and his livestock, which was investigated by Currie *et al.* (1994). In this investigation molecular typing supported the theory of clonal introduction of *B.pseudomallei* into the non-endemic region, local dissemination and persistence for over 25 years (Currie *et al.*, 1994).

Geographical distribution of the causative organism itself has not been studied in any great detail worldwide. Studies carried out by Vuddhakul *et al.* (1999) investigating the distribution

of *B.pseudomallei* in the soil collected from four regions in Thailand (Table 1.1), showed that the organism was far more prevalent in the Northeast region, which correlated with a higher incidence of melioidosis.

Table 1.1 Isolation rate of *Burkholderia pseudomallei* from soil samples in Thailand (Vuddhakul *et al.*, 1999).

Region	Number of soil samples	Number of positive samples	Number of collection sites	Number of positive sites
North	720	32 (4.4%)	180	25 (13.8%)
Central	196	12 (6.1%)	49	12 (24.5%)
Northeast	1428	291 (20.4%)	357	179 (50.1%)
South	1241	73 (5.9%)	310	57 (18.4%)
Overall	3585	408 (11.4%)	896	273 (30.5%)

1.2.2 Soil Ecology.

In nature, members of the genus *Burkholderia* multiply in the rhizosphere, defined as the region of soil modified as a result of the uptake and deposition of substances growing in the plant root, where the bacteria grow in close association with root nodules and assist in the fixation of nitrogen and promote uptake of minerals, particularly phosphate by the plant (Pitt *et al.*, 2000). A specific plant host has yet to be described for *B.pseudomallei* but it is considered plausible that the species may play a similar role to *B.cepacia* in the rice paddy field (Pitt *et al.*, 2000). The distribution and behaviour of *B.pseudomallei* within the soil environment is a significant factor in the epidemiology of the disease, and in determining pathways of melioidosis transmission (Brook *et al.*, 1997). A number of factors that may influence the distribution of *B.pseudomallei* in soil have been investigated, including climatic factors such as rainfall, temperature and sunlight. However, Dance (2000c) stressed recently that any environmental surveys using artificial media pre-1997 may not discriminate between

B.pseudomallei (Ara^-) and *B.thailandensis* (Ara^+), and so the isolation of *B.pseudomallei* in some of these cases may not indicate endemicity for pathogenic *B.pseudomallei*, since only Ara^- isolates cause infection.

1.2.2.1 Temperature.

Considering that melioidosis is predominantly a disease of tropical and subtropical climates, it is probable that environmental temperature contributes to the distribution of the organism and disease (Dance, 2000c). *B.pseudomallei*-positive areas investigated by Tong *et al.* (1996) were usually places where the atmospheric temperatures in January, the coldest month in the South, were above 12°C. The study also showed that the optimum temperature for growth of the organism *in vitro* was between 24°C and 32°C (Tong *et al.*, 1996). Indeed, most workers find the organism unable to grow at temperatures below 21°C (Dance, 2000c). Further investigations have shown the organism to be very susceptible to heat, as well as the cold, as exposure for 10 minutes at 50°C, and at 4°C for 2 to 3 weeks were equally lethal, which may explain the restriction of the free-living organism to tropical regions (Smith *et al.*, 1987). However, if climatic conditions continue to change, with the threat of global warming, it may eventually mean that the organism will not be as environmentally restricted as it currently is. Persistence of *B.pseudomallei* within the environment has been studied and the organism has been shown to remain viable and to retain its virulence even after 27 days of being kept desiccated, mixed with garden soil at 27°C (Smith *et al.*, 1987). Wuthiekanun *et al.* (1995b) monitored the successful survival of *B.pseudomallei* in sterile triple-distilled water at tropical room temperature, without any additional nutrients for more than three years, showing that the bacterium has a remarkable capacity to survive in a nutrient-free medium.

1.2.2.2 Rainfall and Moisture Content of the Soil.

Most cases of melioidosis present during the rainy season months in endemic areas and in the

case of Thailand, this is presumably as a result of direct infection during rice farming in wet paddy fields (White, 1994). However, the possibility of other seasonal factors, for example, dietary changes, stress, or concomitant viral infection that may precipitate relapse of latent melioidosis cannot be excluded (Dance, 2000c). It is thought that during periods of high rainfall, *B.pseudomallei* is physically leached from its normal environmental habitat of soil and surface water, temporarily creating a more widespread distribution (Leakey *et al.*, 1998). Tong *et al.* (1996) showed that soil with water contents of below 10% were highly detrimental to *B.pseudomallei*, but increasing water content up to 40% increased survival. Studies by Wuthiekanun *et al.* (1995a) on recovering *B.pseudomallei* from soil has similarly shown it to be dependent on moisture content; as no strain was isolated from the soil surface in the dry season. The association between isolation rate and soil depth may partly be related to the killing of *B.pseudomallei* by sunlight, which has been investigated by Tong *et al.* (1996) who found that the organism was more easily killed by UV light *in vitro* than the "permanent soil bacteria" (Tong *et al.*, 1996).

1.2.2.3 Other Factors Affecting Soil Distribution of *B.pseudomallei.*

Studies carried out by Tong *et al.* (1996) on the correlation between a variety of soil conditions and the survival of *B.pseudomallei* showed that pH values below 4 and above 8 were detrimental to the organism. The pH of soil in the rice field is usually between 5 and 6.8 and so is well within the range that *B.pseudomallei* can tolerate (Tong *et al.*, 1996). Extensive investigations into soil type have not been carried out, but in a study in Australia by Thomas *et al.* (1981), it was found that the isolation rate of the organism was higher in clay soils as opposed to sandy; possibly attributed to the greater water retention of clay. It has also been shown that cultivated areas and wet rice fields provided higher isolation rates of *B.pseudomallei* than forested areas (Straus *et al.*, 1969), and considering that *B.pseudomallei* is able to reduce nitrate in order to grow in an anaerobic environment, the use of nitrate

fertilisers may actually contribute to its proliferation in agricultural land (Dance, 2000c). Further investigations are obviously needed, but it is clear that temperature, water content and soil pH are major ecological conditions governing the environmental distribution of *B.pseudomallei.*

1.3 EPIDEMIOLOGY OF MELIOIDOSIS.

1.3.1 Incidence of Melioidosis.

Melioidosis is especially prevalent in Northeast Thailand where approximately 20% of community-acquired septicaemia can be attributed to this bacterial pathogen (Chaowagul *et al.*, 1989). In Thailand, it has been estimated that between 2000-5000 cases of melioidosis occur each year (Dharakul *et al.*, 1996). The disease is also prevalent in Northern Australia and in a period between October 1989, and February 1997, 170 cases of melioidosis were diagnosed in the Northern Territory of Australia (Edmond *et al.*, 1998). Findings over a 9-year period carried out in the top end of the Northern Territory of Australia by Currie *et al.* (2000b) showed that the average annual incidence of melioidosis was 16.5/100,000 population, and the overall mortality was 21%. This compares to an annual incidence of 4.4/100,000 population in Ubon Ratchathani, in Northeast Thailand (Suputtamongkol *et al.*, 1994). The higher incidence rate reported in Australia has been suggested by Currie *et al.* (2000b) to be a reflection of a possible underassessment in the study in Thailand, where less severely ill patients are treated in different hospitals and not referred to the larger central hospital and so missed. An interesting observation in the study by Currie *et al.* (2000b) was that although Aboriginal Australians account for just 24% of the population of the region, they represented 50% of the melioidosis cases reported. Surprisingly, in Singapore between 1989 and 1996, the incidence of melioidosis was relatively high, 1.7/100,000 population, despite the highly urbanised environment with 88% of the population living in high-rise flats (Heng *et al.*, 1998). In a

recent study carried out by Vuddhakul *et al.* (1999) it showed that of in-patients observed in Northeast Thailand, 137.9/100,000 admitted had melioidosis compared with 13.4/100,000 in Central Thailand, 18/100,000 in the North and 14.4/100,000 in the South, which reflects the authors findings that the organism has a much higher isolation rate in the Northeast region (Table 1.1).

The male to female ratio of melioidosis cases in Thailand is approximately 3:2 respectively, a fact that has previously been attributed to work-related exposure, although many women in rural Thailand participate actively in rice farming (Dance, 1990). However, Currie *et al.* (2000b) found in Australia that 74% of the cases observed were in males. Melioidosis may present at any age, although the disease is very rarely observed in neonates (Halder *et al.*, 1998). In Northeast Thailand the mean age of presentation is in the fifth decade (White, 1994).

1.3.2 Inoculating Events and Mode of Transmission.

B.pseudomallei is a free-living bacterium found widely in soil and surface water of rice paddies, fields newly planted with oil palm, monsoon drains, gardens and playgrounds in endemic areas (Leelarasamee and Bovornkitti, 1989). In humans, it is thought that melioidosis is mainly acquired by contact with infectious soil and water through a pre-existing skin abrasion or ulcer, but infection is possible through ingestion, or inhalation of infectious dust particles (Leelarasamee and Bovornkitti, 1989). Melioidosis is found more frequently amongst the rice farming community in Thailand as rice farmers often plough and plant seedlings without any protective clothing, therefore minor skin trauma is very common (Wuthiekanun *et al.*, 1995a). However, in 15% of cases there is no obvious focus of primary infection (Chaowagul *et al.*, 1989). In fact, relatively few patients can identify a specific incident before they become ill, probably because rice farmers, for example, are continually exposed to infection and because of the variable incubation period of the disease, which may remain latent

for weeks, months or years (Dance, 2000a). In a study by Currie *et al.* (2000b) only 25% of cases had likely inoculating events, those being situations of exposure to soil or muddy water specifically recalled.

Recently, an outbreak of melioidosis in Northwest Australia was investigated by Inglis *et al.* (2000) where the source of contamination was tracked-down to a potable water supply and the aerator was identified as the probable source. This is concerning as many remote communities in Australia have similar water supplies and so should be considered in future as possible sources of infection. Infection due to inhalation of contaminated dust or soil particles became apparent during the Vietnam conflict, where a disproportionately high number of helicopter crewman developed melioidosis in comparison to other soldiers, and it was proposed that the rotors of the helicopter disturbed infectious dust particles facilitating pulmonary infection (Sanford, 1990).

Human-to-human transmission is rare and only one sexually transmitted case has been described in a man who had chronic prostatitis and his wife subsequently had a raised haemagglutination titre (McCormick *et al.*, 1975). A case described by Kunakorn *et al.* (1991a) involved a man with melioidosis who was believed to have passed it on to his sister who had cared for him during his illness, and the isolates were later shown to be indistinguishable by PFGE (Dance, personal communication). In a recent report, a goat with melioidosis from the Darwin rural region aborted twins and died shortly afterwards and *B.pseudomallei* was isolated from the goat's uterus and from the spleen of one of the aborted twins, confirming transplacental infection in animals (Choy *et al.*, 2000).

1.3.3 Host Risk Factors.

Cellular immunity is postulated to play important roles in immunity to melioidosis that may

influence the severity and clinical outcome of the disease (Dharakul *et al.*, 1998). The majority of adult cases have an underlying predisposition to infection; for example, diabetes mellitus and chronic renal disease are the two most common predisposing conditions (White, 1994). Other pre-existing circumstances such as steroid or cytotoxic drug therapy, leukaemia, tuberculosis, alcoholism, malignancy, connective tissue diseases and cirrhosis can also predispose individuals (Dance, 1990; Chaowagul, 1996; Dance, 2000c). Recently, Currie *et al.* (2000b) reported that in a 9-year study, 36% of cases observed were in diabetics, 37% were heavy alcohol consumers, 24% had chronic lung disease and 11% had chronic renal failure. A study by Ahmed *et al.* (1999) revealed that *B.pseudomallei* has low attachment ability to pharyngeal epithelial cells, which may partly explain why melioidosis mainly affects malnourished and immunodeficient hosts. Perhaps surprisingly, Thummakul *et al.* (1999) reported that melioidosis was not an important opportunistic infection in acquired immune deficiency syndrome (AIDS) patients in endemic regions.

Genetic factors, including human leukocyte antigen (HLA) polymorphism, play important parts in the cellular immunity against infectious diseases (Dharakul *et al.*, 1998). Work by Dharakul *et al.* (1998) has showed evidence of the association of certain HLA Class II alleles in particular groups of melioidosis patients, especially those with severe clinical disease and relapse cases. As there is a very strong association of melioidosis with diabetes, investigations into the interaction of *B.pseudomallei* and insulin was carried out by Woods *et al.* (1993), who found that insulin apparently inhibited the growth of *B.pseudomallei*. However, Simpson and Wuthiekanun (2000) showed that *B.pseudomallei* growth was in fact inhibited by m-cresol, a preservative found in the insulin preparations, whereas growth of the bacterium was not inhibited by pure insulin. Thus the findings of Woods *et al.* (1993) may have been due to the preservative rather than the insulin. Currie (1995) also stated that most diabetic patients with melioidosis, in their experience, did not actually have the insulin-dependent type I diabetes and

that many diabetics are probably hyperinsulinemic due to control over diet and drug therapies.

Woods *et al.* (1993) provided the first evidence for the presence of a specific insulin receptor on *B.pseudomallei*, through radiolabelled insulin binding studies. Studies by Kanai *et al.* (1996) showed that *B.pseudomallei* and *B.cepacia* stained positively with fluorescently labelled insulin, whereas *Pseudomonas aeruginosa* did not, and reported findings that insulin exposure reduced phospholipase C and acid phosphatase activities in *B.pseudomallei*. The findings by Kanai *et al.* (1996) thus support those of Woods *et al.* (1993), concluding that *B.pseudomallei* may possess insulin receptors which could be associated with a signal transfer system involving phospholipase and protein tyrosine phosphatase.

1.3.4 Molecular Epidemiology of *B.pseudomallei*.

Epidemiological studies of *B.pseudomallei* have been hampered in the past by the lack of discriminatory typing methods (Lew and Desmarchelier, 1993). Ribotyping is now a common method as an epidemiological tool in the study of a number of bacterial pathogens, whereby patterns of restriction fragment length polymorphisms (RFLPs) in rRNA genes from different isolates are compared. Lew and Desmarchelier (1993) carried out *Bam*HI restriction enzyme digests of 100 isolates of *B.pseudomallei* from a number of sources, and on the basis of the patterns observed, the strains were classified into 22 different groups. However, this was before the discovery of *B.thailandensis* and so some of the ribotype groups represented this species rather than *B.pseudomallei*. Interestingly, Lew and Desmarchelier (1993) found that some of the ribotypes were restricted to a particular host species, geographical area or period of isolation, as all of the earliest strains were represented as a single clone, whereas strains from 1940 onwards belong to increasingly divergent ribotype patterns. This typing method was later utilised by Trakulsomboon *et al.* (1997) to investigate differences between clinical and environmental strains, showing that the ribotype patterns fell into two groups which are now

clearly identified as *B.pseudomallei* (Group I) and *B.thailandensis* isolates (Group II). Pitt *et al.* (2000) have also carried out ribotyping on 350 isolates and discovered 44 ribotypes, whereby types 1 and 3 are the most common, accounting for half of the collection and prevalent worldwide.

Norton *et al.* (1998) applied random amplification of polymorphic DNA (RAPD) and multilocus enzyme electrophoresis (MEE) to investigate whether disease presentation was clonally related in 18 cases of melioidosis observed in four geographically separate communities in North Queensland, Australia. The authors found that isolates segregated into two groups that actually correlated with disease presentation rather than geographical location, with the first group including patients mainly with respiratory disease and the second, from patients with abscesses in viscera other than the lung or brain and those with bone or joint disease. Norton *et al.* (1998) suggested that different strains of *B.pseudomallei* might carry genetic regions capable of influencing tissue tropism of the infecting strain, although a study using a larger group of clinical isolates would be required to develop this clonal hypothesis. Koonpaew *et al.* (2000), however, found no significant association between clinical manifestations of melioidosis and the level of relatedness of *B.pseudomallei* isolates, whose DNA had been digested with a restriction enzyme, *Nco*I and then subjected to PFGE.

1.4 MELIOIDOSIS.

1.4.1 Clinical Manifestations.

When humans and animals encounter *B.pseudomallei* in the environment one of the following outcomes are possible; firstly there may be no effect, secondly, asymptomatic seroconversion may occur and thirdly, a clinical infection becomes apparent (Dance, 2000a). Melioidosis has actually been termed "the great imitator" of every infectious disease, as virtually every organ

can be affected (Leelarasamee and Bovornkitti, 1989). Attempts have been made to classify "subtypes" of the disease, but none of these are entirely satisfactory (Dance, 1990). Some descriptions tend to classify the illness manifesting as an acute, subacute, or chronic process (Sanford, 1995). The acute form is the most severe, with a rapid onset (24 to 48 hours) of non-specific symptoms (Leakey *et al.*, 1998) that may result in death within 72 hours (Smith *et al.*, 1987). Acute disease may involve one or a combination of several different types of disease: septicaemia, pneumonia, emphysema, localised abscess or wound infection, urinary tract infection, septic arthritis and visceral abscess formation (Walsh and Wuthiekanun, 1996). The lung is the most common organ involved, either as a primary focus of infection, or secondary to metastatic spread from bloodstream infection (White, 1994), and the severity of infection can range from a mild bronchitis to a disseminated bronchopneumonia (Dance, 1990). The subacute and chronic types take the form of a prolonged febrile illness associated with the formation of multiple abscesses with death ensuing in a few weeks or months (Smith *et al.*, 1987).

Around 60% of melioidosis patients are bacteraemic, with most of these presenting with community-acquired sepsis syndrome and evidence of multifocal metastatic infection, with the remaining 40% having localised infections usually abscesses or granulomata (Dance, 2000a). Abscesses in the lungs, liver and spleen are common and 5% of patients also have superficial skin abscesses (Chaowagul, 1996). There have also been cases of abscesses occurring in the brain (Lee and Chua, 1986; Lath *et al.*, 1998; Padiglione *et al.*, 1998), but it is thought that the blood-brain barrier is breached only in overwhelming widespread infection in a debilitated host. Many patients require over one month's hospitalisation and the overall mortality of melioidosis is approximately 40% (Chaowagul, 1996). Subclinical infections, defined by the detection of haemagglutinating antibody in people residing in areas of endemicity, are very common (Anuntagool *et al.*, 1998). However with the recent recognition of *B.thailandensis*, a

proportion of these may represent asymptomatic infection with the newly described species, rather than *B.pseudomallei*.

Other manifestations have been observed, such as neurological melioidosis, which has been described in a number of cases (Woods *et al.*, 1992; Howe *et al.*, 1997; Maguire *et al.*, 1998). Clinical features of this initially manifest as a headache and in some cases stiffness in the neck is observed, but eventually the brainstem, cerebellum and spinal cord are all affected. Currie *et al.* (2000a) found in a 9-year study that 5% of melioidosis cases were neurological and that the mortality rate was 25% in those cases.

1.4.2 Acute Suppurative Parotitis.

Acute suppurative parotitis accounts for nearly 7% of the disease in Ubon Ratchatani, and almost exclusively affects children, manifesting normally as a unilateral parotitis, presenting with painful facial swelling and fever (Dance, 1990). Parotitis constitutes 38% of melioidosis cases in children in Northeast Thailand (Dance *et al.*, 1989a). The underlying reasons for this are unclear, since children are usually previously healthy, but presumably reflects oral contamination with soil or water during play (Dance, 1990). Dance *et al.* (1989a) postulated that superinfection with *B.pseudomallei* may follow mumps parotitis, as a small number of children suffering with this had previously had mumps, but further investigation is needed to prove this theory. Reports by Edmond *et al.* (1998) show that cases of parotid abscesses were not shown in children in the Northern Territory of Australia between 1989 and 1997, and suggest that a comparison of behaviour patterns and perhaps dietary habits of the children in Australia and Thailand is needed to explain such a finding. Acute suppurative parotitis may represent primary infection, but there is a good prognosis provided there is early incision, drainage and appropriate antibiotic therapy (Chaowagul, 1996).

1.4.3 Latency and Relapse of Melioidosis.

Failure of antimicrobial therapy resulting in acute deterioration and death, the appearance of new abscesses whilst patients are on treatment, or relapse of the infection after completing a course of antimicrobial treatment are all common (Chaowagul, 1996). Relapse of melioidosis occurs especially in hosts who are immunocompromised or who have an associated condition such as diabetes mellitus, cirrhosis of the liver, chronic pyelonephritis and alcoholism (Leelarasamee and Bovornkitti, 1989), or in patients with large abscesses and a multi-organ involvement (Thummakul *et al.*, 1999). Moreover, improvement in acute treatment has resulted in a greater number of successfully treated patients who survive and leave hospital, but eventually relapse (Chaowagul *et al.*, 1993). In a study carried out by Chaowagul *et al.* (1993) the overall relapse rate was found to be around 15.3% per year of follow up; but this figure may be inaccurate as large numbers were lost to follow up. The same study showed that the median time to relapse was 21 weeks and the longest period between recovery and relapse was five and a half years. RAPD typing and RFLP analysis has confirmed that the majority of relapse cases of melioidosis are in fact a re-emergence of the initial infecting strain, and that relapsed melioidosis is almost always associated with failure to complete maintenance therapy (Desmarchelier *et al.*, 1993; Haase *et al.*, 1995; Vadivelu *et al.*, 1998). However, Vadivelu *et al.* (1998) discovered that in one case, two distinct strains were isolated from a patient with a single episode of melioidosis, suggesting that some patients may be infected with more than one strain of *B.pseudomallei*. The antimicrobial susceptibility patterns of the relapse isolate are usually unchanged, although a simultaneous resistance to chloramphenicol, doxycycline, and trimethoprim-sulfamethoxizole is occasionally recognised (Chaowagul, 1996). Thus, Chaowagul (1996) suggests that all relapse isolates should be reanalysed for their antimicrobial susceptibility patterns.

B.pseudomallei may remain dormant in asymptomatic individuals only to recrudesce, if

conditions are favourable, and in many cases years after initial exposure, as an acute exacerbation (Ashdown and Koehler, 1990). The longest period of latency, described by Chodimella *et al.* (1997), is in a Vietnam veteran who developed melioidosis after an inapparent infection 29 years earlier in an endemic area.

1.4.4 Disease in Animals.

The disease occurs naturally in rats, rabbits, birds, pandas, kangaroos, guinea pigs, cats, dogs, sheep, goats, swine, horses, squirrels, crocodiles, koalas and seals, and occasional isolates have been reported from cows (Smith *et al.*, 1987; Veljanov and Najdenski, 1993; Sanford, 1995; Markova *et al.*, 1998). Melioidosis has also been identified in a camel in the United Arab Emirates (Werney *et al.*, 1997). Vedros *et al.* (1988) reported cases of melioidosis in captive cetaceans kept at Ocean Park in Hong Kong, where the highest mortality occurs during the rainy season. During this period considerable amounts of soil contaminated with *B.pseudomallei* are washed down from the mountains into bays that used as a source of water for the Ocean Park cetacean pools, hence causing infection (Vedros *et al.*, 1988). Arthropod-borne infection has never been known to occur naturally, although experimental transmission has been accomplished in guinea pigs, from the bite of the mosquito and of the rat flea (Sanford, 1995).

Infected animals may play a role in the dissemination of the organism to new environments (Dance, 1990), but apparently do not represent a major reservoir for human disease (Sanford, 1995). Treatment of animals suffering with melioidosis can become an expensive, prolonged and often unsuccessful venture, and in the case of farming it is dealt with by prevention and control rather than treatment (Choy *et al.*, 2000).

1.5 DIAGNOSIS AND TREATMENT OF MELIOIDOSIS.

1.5.1 Diagnosis of Melioidosis.

In acute cases of melioidosis rapid diagnosis is urgent so that effective chemotherapy can be instituted as early as possible, as most patients will die within 2 to 3 days if they are not properly diagnosed and treated (Pongsunk *et al.*, 1999). Melioidosis is confirmed by growth of the organism from blood or body fluids, such as purulent exudates, sputum, spinal fluid, urine or faeces (Smith *et al.*, 1987; Chaowagul, 1996). This approach is time-consuming and consequently it is often too late to administer effective therapy (Rattanathongkom *et al.*, 1997). Problems are also encountered with specimens such as sputum and throat swabs, as they are likely to be heavily contaminated (Walsh *et al.*, 1994), a factor additionally complicated by the relative slow growth of *B.pseudomallei* in comparison with other flora. Walsh and Wuthiekanun (1996) suggested that on admission to hospital a patient suspected of suffering with melioidosis should have the following performed: 3 blood cultures (15ml of blood), tracheal or sputum aspirate, throat swab and urine sample as well as a sample of any other fluid, pus or wound exudates that may be present.

Identification of *B.pseudomallei* can be achieved by using the commercial API 20NE biochemical kit (Dance *et al.*, 1989b). Dance *et al.* (1989b) found the API 20NE to be extremely reliable, identifying over 97% of strains on the first testing and over 99% on the second. One disadvantage is that the cost of the kit is quite high per test, and may not be economically acceptable in many areas where melioidosis is endemic (Dance, 1989b). Inglis *et al.* (1998) found that the API 20NE test would sometimes mistakenly identify *B.pseudomallei* as *Chromobacterium violacieum* and found it more economically viable to use the Microbat 24E system.

Serological diagnosis however is fraught by the high background seroconversion in the

population and therefore it lacks specificity. However, there have been significant improvements in recent years as enzyme-linked immunosorbant assay (ELISA) tests based on specific antigens have been developed (Chaowagul, 1996). Examples include an ELISA developed by Petkanjanapong *et al.* (1992) with endotoxin preparations of *B.pseudomallei* as the antigen, which gave satisfactory results for sensitivity and specificity (95.7% and 94.2% respectively), but is still at an experimental level. Anuntagool *et al.* (1996) have developed a simple antigen detection test based on the use of a specific monoclonal antibody (mAb) in a sandwich ELISA, and an ELISA using a fluorescein isothiocyanate (FITC)-anti-FITC amplification system has been developed by Desakorn *et al.* (1994) to detect *B.pseudomallei* antigen in urine. However, these methods require either expensive equipment or reagents, making them difficult to set up in poor areas where infection is endemic (Anuntagool *et al.*, 1996).

Pongsunk *et al.* (1999) have developed a monoclonal antibody specific to the 30 kilodalton (kDa) protein of *B.pseudomallei* that can be used to identify the bacterium in blood culture, shortening the time for results by at least two days with the added advantage of not identifying *B.thailandensis*. Recently, a latex agglutination test based on an exopolysaccharide-specific antibody has been developed by Steinmetz *et al.* (1999) to which *B.pseudomallei*-like organisms did not react and which can be used to identify *B.pseudomallei* from different geographical locations.

Alternative methods of detection being developed include a PCR devised by Rattanathongkom *et al.* (1997), indirect immunofluorescence microscopy (Naigowit *et al.*, 1993) and a gold blot to detect immunoglobulin M (IgM) specific antibodies (Kunakorn *et al.*, 1991b). However, microbiological isolation and identification of *B.pseudomallei* is still used as conclusive evidence for the diagnosis of melioidosis.

1.5.2 Antibiotic Susceptibility of *B.pseudomallei.*

B.pseudomallei has an unusual antibiotic susceptibility pattern, as it is susceptible to co-amoxyclav, ceftazidime, the carbapenems, chloramphemicol, doxycycline and co-trimoxazole, but intrinsically resistant to many other antibiotics, including most early β-lactams, for example ampicillin, the macrolides, aminoglycosides, and most cephalosporins (Walsh and Wuthiekanun, 1996; Dance, 2000a). A multidrug efflux system has been identified in *B.pseudomallei* by Moore *et al.* (1999), which is active for both aminoglycoside and macrolide antibiotics and hence accounts for the organism's resistance to these antibiotics.

Recently, investigations have been carried out by Haussler *et al.* (1999) on slow growing "small colony variants" (SCVs), which are isolated after exposure to ceftazidime, ciprofloxacin and gentamicin. The SCVs isolated show a significant increase in the minimal inhibitory concentrations of various unrelated classes of antimicrobial agent and this phenotype was stable throughout numerous passages on antimicrobial-free media, although interestingly, revertants to the parental phenotype did occur, and these revertants led to rapid overgrowth in liquid media (Haussler *et al.*, 1999). Thus, an *in vivo* selection of SCVs under antibiotic treatment would be difficult to detect in clinical specimens and so further investigations are needed to assess whether SCV revertants contribute to the frequently observed relapses of melioidosis.

1.5.3 Treatment of Melioidosis.

Prior to the use of antimicrobials, the case fatality rate from apparent *B.pseudomallei* infections was 95%, and French experience in Indochina indicated that with chloramphenicol therapy the mortality was still 20% (Woods *et al.*, 1999). Traditional treatment of melioidosis with a combination of chloramphenicol, doxycycline and trimethoprim-sulphamethoxazole is still of value in the maintenance treatment of the infection, but was replaced by ceftazidime for

acute treatment of severe melioidosis as it was shown to halve the mortality of severe melioidosis in a trial carried out by White *et al.* (1989). Since then, Thamprajamchit *et al.* (1998) have found that cefoperazone/sulbactam plus co-trimoxazole was of potential use as an alternative treatment for severe melioidosis, giving similar results to ceftazidime. Furthermore, Simpson *et al.* (1999) showed imipenem to be a safe and effective treatment for severe melioidosis and suggested its use as an alternative to ceftazidime, if necessary. Recently, in one case of melioidosis reported by Minassian *et al.* (1999) imipenem had to be used as the patient did not respond to ceftazidime. Investigations by Chaowagul *et al.* (1997) using oral fluoroquinolones yielded poor results and it was suggested that they could only be used as a third-line drug for those patients who are either intolerant of the other available antibiotics, or who have organisms that are resistant to them. More recently, Vorachit *et al.* (2000) have suggested, based on initial experiments that the following agents are successful in the treatment of acute melioidosis and should be used in clinical trials: imipenem, meropenem, and imipenem plus azithromycin.

Parenteral antimicrobial therapy should be given for at least 10 to 14 days and continued until there is clear evidence of symptomatic improvement, with maintenance treatment given for at least 18 weeks (Chaowagul, 1996). The cost of effective treatment is therefore considerable, and would substantially exceed annual per-capita health expenditure in the majority of the tropical countries (White, 1994).

1.6 PATHOGENESIS.

1.6.1 Immunology of *Burkholderia pseudomallei.*

The features of melioidosis suggest that there may be unique microbial virulence factors that not only promote tissue invasion and destruction, but also permit evasion from the normal humoral and cell-mediated immunity (Egan and Gordon, 1996). *B.pseudomallei* was shown by Pruksachartvuthi *et al.* (1990) to remain undigested and able to multiply within human phagocytes. Jones *et al.* (1996) have similarly shown that *B.pseudomallei* is capable of invading non-professional phagocytes and has the ability to survive intracellularly in professional phagocytic cells, and postulate that resistance to defensins may contribute towards intracellular survival. The ability of *B.pseudomallei* to survive and multiply in both professional and non-professional phagocytes may provide an explanation for both the occurrence of latent infections, and the relapses of infection that result from reactivation of a persistent endogenous source of infection (Jones *et al.*, 1996). Harley *et al.* (1994) showed that uptake of *B.pseudomallei* appeared to be by conventional phagocytosis, but the organism would then escape from vacuoles, postulated to be as a result of damage to the vacuolar membrane. Jones *et al.* (1996) reached a similar conclusion, but further work is required to fully elucidate the mechanisms of intracellular survival of *B.pseudomallei.*

1.6.2 Immunopathogenesis of melioidosis.

The immunopathogenesis of melioidosis, particularly with respect to the development of acute versus chronic infection, is poorly understood (Ulett *et al.*, 1998). However, regarding the mechanisms responsible for the acute and chronic clinical presentations of *B.pseudomallei* infection, it is possible that the balance between T helper 1 (Th1) / T helper 2 (Th2) responses determines the clinical course of infection in humans (Ulett *et al.*, 1998).

Cytokines are critical immunoregulatory determinants of the pathway of disease pathogenesis

and progression (Ulett *et al.*, 2000). Investigations carried out by Friedland *et al.* (1992) showed that elevated interleukin 6 (IL-6) concentration was the best predictor of mortality in *B.pseudomallei* sepsis and that 50% of patients with detectable IL-8 levels died, but found that plasma tumour necrosis factor (TNF) bioactivity did not relate to outcome. In a recent study by Ulett *et al.* (2000) it was demonstrated that high levels of TNF-α, IL-1β and IL-6 contribute towards disease pathogenesis and progression in *B.pseudomallei* infection, hence showing that proinflammatory cytokine responses are important in the immunopathogenesis of melioidosis (Ulett *et al.*, 2000).

1.6.3 Endotoxin (Lipopolysaccharide) and Serum Sensitivity.

Endotoxin is a lipopolysaccharide (LPS) that is an integral component of the outer membrane of gram-negative bacteria (Salyers and Whitt, 1994). Matsuura *et al.* (1996) suggested that because *B.pseudomallei* is a gram-negative bacterium, it is likely that endotoxin plays a role as a virulence factor, similar to that of *Enterobacteriaceae*. Results by Anuntagool *et al.* (1998) showed that LPS from clinical Ara^- isolates and Ara^- soil isolates were indistinguishable and were identical to the LPS of Ara^+ isolates with regards to sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) profiles and immunoreactivities with melioidosis sera. Due to the antigenic similarity between pathogenic Ara^- isolates and non-pathogenic Ara^+ isolates, Anuntagool *et al.* (1998) suggested that this provides an explanation to the observed discrepancy between seroprevalence and disease prevalence in melioidosis. This then poses the question of whether LPS has a role in virulence. Studies carried out by Matsuura *et al.* (1996) showed that the LPS of *B.pseudomallei* clinical isolates appeared to be less toxic than those from other gram-negative pathogens. The endotoxic activities of the LPS of *B.pseudomallei* (BP-LPS), such as pyrogenic activity in rabbits and lethal toxicity in Galactosamine (GalN)-sensitised mice, were shown to be some 10 to 100 times weaker than those of the reference, *Salmonella abortus equi* (SEA-LPS) in an investigation by Matsuura *et*

al. (1996). Similarly, in macrophage activation experiments, BP-LPS exhibited activities 10 times weaker than SEA-LPS (Matsuura *et al.*, 1996).

Anuntagool *et al.* (2000) recently proposed the possibility that *B.pseudomallei* may release higher amounts of LPS into the surrounding tissues during growth in patients. Investigations carried out by Anuntagool *et al.* (2000) showed that the shedding pattern and quantity of LPS found in the culture of fluid during growth of Ara^- and Ara^+ isolates were similar, but the Ara^- isolates contained a previously described 200 kDa antigen (Sirisinha *et al.*, 1998) in large quantities during the early phase of growth.

B.pseudomallei has been found to produce two S-type LPSs, differing in the chemical structures of their O-polysaccharide (O-PS) components and termed, O-PS I, and O-PS II (Perry *et al.*, 1995). Many gram-negative bacteria are susceptible to the serum bactericidal system, but some pathogens possess mechanisms to evade killing by this system (Woods *et al.*, 1999). Initial work by Egan and Gordon (1996) revealed that *B.pseudomallei* rapidly activated, and then consumed complement and that the organism was resistant to killing by polymorphonuclear leukocytes. Recently, DeShazer and Woods (1999) using Tn5-OT182 mutagenesis revealed that serum-sensitive mutants were deficient in the type II O-antigenic polysaccharide moiety of LPS and discovered a cluster of 15 genes required for the biosynthesis of type II O-PS and serum resistance. Thus concluding that the type II O-PS moiety of LPS is essential for *B.pseudomallei* serum resistance and virulence (DeShazer and Woods, 1999).

1.6.4 The Capsular Polysaccharide of *B.pseudomallei.*

Some bacterial pathogens are able to combat the bactericidal activity of complement and phagocytes by encapsulation with structurally diverse polysaccharides and so only late into the

infection, when anticapsular antibodies are produced by the host, can the protective encapsulation activity be overcome (Masoud *et al.*, 1997). Little is known about the capsular polysaccharide (CPS) of *B.pseudomallei*. In a study carried out by Vorachit *et al.* (1995) it was shown that cells of *B.pseudomallei* tend to produce exopolysaccharide materials that constitute a highly hydrated glycocalyx. When cells were grown in liquid media or in animal tissues, the production of the fibrous glycocalyces facilitated the formation of microcolonies or "sister" cells that adhered readily and non-specifically to available surfaces, showing that in infected tissue the organism actually grows in glycocalyx-enclosed biofilms (Vorachit *et al.*, 1995). Structural analysis of the CPS by Masoud *et al.* (1997) revealed it to be an unbranched polymer of repeating tetrasaccharide units composed of D-galactose and 3-deoxy-D-manno-octulosonic acid (KDO), containing an O-acetyl substituent. The authors also found that sera from patients with melioidosis contained high titres of antibodies to the CPS, but sera from the controls did not, which suggests that the CPS functions as an immunogen and thus has potential use in serological diagnosis as well as vaccine development (Masoud *et al.*, 1997).

1.6.5 Exotoxin of *B.pseudomallei.*

Nigg *et al.* (1955) first described the presence of a thermolabile, lethal toxin in culture supernatants of *B.pseudomallei*. Later, Heckly and Nigg (1958) described the presence of two thermolabile toxins, one that produced necrotic lesions and the other, which was non-necrotising, but still lethal for mice and hamsters. A thermolabile toxin produced as a 36 kDa polypeptide chain by *B.pseudomallei* is postulated to inhibit intracellular protein synthesis by the same mechanism as diphtheria toxin (DT) of *Corynebacterium diphtheriae* and the exotoxin A (ETA) of *P.aeruginosa* (Ismail *et al.*, 1991). The mode of action of these exotoxins is the catalysis of adenosine-5-diphosphate ribose from oxidised nicotinamide adenine dinucleotide (NAD) to elongation factor 2 (EF2), and such toxins are termed ADP-ribosylating toxins (Mohamed *et al.*, 1989). Investigations by Mohamed *et al.* (1989) showed that a lag

period was not required before toxic activity was expressed in *B.pseudomallei*, and so it would seem that the toxin was not of the membrane-damaging class of bacterial protein toxins, but that it acted on an intracellular mechanism, such as protein synthesis, supporting the theory that the toxin could indeed be an ADP-ribosylating toxin. This theory was further supported in the same study, whereby the toxin produced by *B.pseudomallei* was found to inhibit both DNA and protein synthesis in macrophages (Mohamed *et al.*, 1989).

The extent to which the toxic product of *B.pseudomallei* is involved in human and animal infections remains unclear (Ismail *et al.*, 1987), but the discovery that antibodies to the toxin are produced in individuals during infection implies that the toxin is elaborated *in vivo* during the course of the disease (Smith *et al.*, 1991). The uncertainty concerning the role of the exotoxin in the pathogenesis of melioidosis stems in part from the lack of a simple, precise and highly sensitive test (Ismail *et al.*, 1987). Ismail *et al.* (1987) developed a competitive ELISA procedure for exotoxin measurement over 16 ng/ml, and found that exotoxin production was higher at 37°C than at room temperature, using Brain Heart Infusion Broth (BHIB), supplemented with glycerol at 2%. Interestingly, supplementing with glycerol has also been shown to increase production of ETA in *P.aeruginosa* (Liu, 1973).

An exotoxin lethal to cells in culture has been identified in culture filtrates of *B.pseudomallei* by Haase *et al.* (1997). The cytolethal toxin (CLT) was able to pass through a 10 kDa cut-off ultrafilter, and may represent the second thermolabile toxin previously reported by Heckly and Nigg (1958). In the investigation carried out by Haase *et al.* (1997), during the process of ultrafiltration, the upper chamber retained a cell-elongating activity (CEA), which was concentrated to one-tenth of the initial volume and when incubated with McCoy cells DNA synthesis was shown to be inhibited by 50% and protein synthesis by 25% (Haase *et al.*, 1997). The CLT was shown to be produced at different amounts by different strains, which appeared

to be correlated with their respective virulence. In the light of recent investigations, this may now actually correspond to some strains having been the avirulent species, *B.thailandensis*. However, in the study all soil isolates tested did produce the CLT at low amounts, including therefore any *B.thailandensis* isolates. Further investigations are required to elucidate the role of the CLT and the CEA in the pathogenic processes of melioidosis and if they are different from other known toxins.

1.6.6 Motility of *B.pseudomallei*

Flagellin proteins from several different strains of *B.pseudomallei* have been isolated and purified by Brett *et al.* (1994), and analysis on SDS-PAGE revealed that flagellin monomer protein bands were around 43,400 Da in size. A number of genes involved in the synthesis and operation of the flagellar apparatus in *B.pseudomallei* have now been identified by DeShazer and Woods (1999) using Tn5-OT182 mutagenesis. The flagellin structural gene, *fli*C, has been cloned and characterised by DeShazer *et al.* (1997) and investigations carried out by DeShazer and Woods (1999) using *fli*C mutants showed that no significance difference in virulence was observed in mutants of the flagellar structural gene. They concluded that the flagellum is probably not a virulence determinant, but is postulated to serve as a protective immunogen against *B.pseudomallei* infection (DeShazer *et al.*, 1997).

1.6.7 Secretion Products of *B.pseudomallei*

Ashdown and Koehler (1990) tested the extracellular products from a number of clinical isolates and found that most strains of *B.pseudomallei* produced a lecithinase, lipase and protease, however none tested positive for elastase. The authors also discovered that most strains produced a haemolysin that was weakly cytolytic and another, which although it occurred infrequently, was found to be far more cytolytic. Haussler *et al.* (1998) also identified an extracellular product from *B.pseudomallei* that was both cytotoxic and haemolytic, from

culture supernatants of *B.pseudomallei* and this has been identified as a rhamnolipid. The rhamnolipid is believed to be the extracellular product responsible for producing the heat-stable haemolytic activity first identified by Liu (1957) (Haussler *et al.*, 1998). However, further studies are needed to fully elucidate the role of the acidic rhamnolipid as a virulence determinant in the pathogenesis of melioidosis.

Sexton and Jones, (1994) confirmed the presence of a 36 kDa antigen associated with proteolytic activity in *B.pseudomallei* culture supernatants, and later Percheron *et al.* (1995) characterised a 42 kDa protease produced by most *B.pseudomallei* strains. However, it has been shown by Gauthier *et al.* (2000) that the expression of a high proteolytic activity by strains of *B.pseudomallei* was neither sufficient nor necessary for virulence when bacteria were injected via the intraperitoneal route.

DeShazer and Woods (1999) used transposon mutagenesis with Tn5-OT182 to obtain mutants unable to secrete protease, lipase or lecthinase and a number of these mutants were mapped to a genetic locus spanning 12 kb. This region encoded proteins with a high degree of homology to the general secretory, or type II secretion pathway proteins, and others to an unlinked gene encoding a protein with homology to type IV prepilin peptidases. Thus indicating that protease, lipase and lecthinase are secreted by a type II secretion pathway (DeShazer and Woods, 1999). The relative virulence of secretion mutants was then studied but allegedly these mutants were not seriously attenuated in their ability to cause fulminating illness (Brett and Woods, 2000).

1.6.8 Adherence of *B.pseudomallei.*

Woods *et al.* (1999) claim that *B.pseudomallei* possesses at least two types of pili and that they have cloned genes with strong homology to *pil*B, *pil*C and *pil*D of *P. aeruginosa* for type IV

pili and a 15 kDa pilin structural protein, *Fim*C, a chaperone protein, and an outer membrane usher protein *Fim*D, all for type I pili. Vorachit *et al.* (1995) observed that strains which produced pili had a much reduced capacity for glycocalyx formation, tending not to form microcolonies in fluid media but adhering to available surfaces as dispersed individual cells rather than as coherent biofilms.

1.6.9 Other Potential Virulence Factors of *B.pseudomallei.*

A phosphatidylcholine-hydrolysing phospholipase C (PC-PLC) gene has been cloned and characterised by Korbsrisate *et al.* (1999). The PC-PLC was demonstrated to be non-haemolytic and the authors suggested that it might have a role in the intracellular survival of the organism and cell-to-cell spread, although further studies are needed to prove this theory.

A siderophore was identified by Yang *et al.* (1991), which was found to belong to the hydroxymate class of siderophores, and was termed malleobactin. Further investigations by Yang *et al.* (1993) demonstrated that the siderophore (malleobactin) was capable of scavenging iron from both lactoferrin and transferrin *in vitro*. Additionally, malleobactin was shown to enhance the growth of *B.pseudomallei* and thought to contribute towards virulence by allowing the organism to multiply in host tissue or survive in the bloodstream (Yang *et al.*, 1991).

Recently, Brett and Woods (2000) claim to have isolated a mutant devoid of acid phosphatase activity and that the translated product of the gene involved was highly homologous to an acid phosphatase expressed by *Francisella tularensis*.

1.6.10 Vaccine Development.

There is currently no available licensed vaccine for use against *B.pseudomallei*. However,

B.pseudomallei produces several surface-associated molecules that have been proposed for use as vaccine components (DeShazer *et al.*, 1998), which include LPS and flagella. Woods *et al.* (1999) believe, in light of their studies with *B.pseudomallei*, that the strongest and most promising vaccine candidates are the flagellin protein and the O-polysaccharide (PS) moieties derived from endotoxin. In a study by Brett and Woods (1996) it was shown that when the O-polysaccharide moiety was covalently linked to flagellin protein it elicited a high titre IgG response, which was then capable of protecting diabetic rats from challenge with *B.pseudomallei*. Similar findings have been shown by Charuchaimontri *et al.* (1999), where a high level of anti-LPS II was a significant factor protective against fatal melioidosis and postulate it to be a potentially useful component of a vaccine. Steinmetz *et al.* (2000) have attempted passive immunisation using a specific monoclonal antibody with specificity for the tetrasaccharide repeating unit of the capsular polysaccharide of *B.pseudomallei* and their results indicate that it may be a promising candidate for use as a protective antigen in active immunisation experiments. Experimental vaccines have been used in animals with some degree of success. Vedros *et al.* (1988) prepared an experimental vaccine consisting of a protein-polysaccharide mixture given to captive cetaceans and claim to have observed a reduced mortality over a 5-year period from 45% to less than 1%.

1.7 MOLECULAR BIOLOGY: APPLICATIONS IN THE STUDY OF BACTERIAL PATHOGENS.

Molecular biology has become an essential analytical tool in all fields of biology and has led to current understanding of the molecular basis of life. The field of molecular biology is extensive and encompasses not only the study of DNA, but also the biochemical basis of gene expression and its regulation, as well as the characterisation and manipulation of the protein products encoded. Thus, molecular biology involves not only recombinant DNA

techniques but biochemical, immunological and protein analysis techniques. It is gene cloning though, perhaps more than any other factor that has made the biggest impact on biological research and so its origins and development with respect to the study of bacterial pathogens is discussed next.

1.7.1 The Origins and Development of Gene Cloning.

Modern recombinant DNA technology has mainly developed over the past thirty years or so from the early experiments carried out. Molecular biology arose from Oswald Avery's initial observations on bacterial transformation in the 1940's (Shostak, 1998), which in turn led to other attempts to investigate the way in which bacteria can acquire new genetic information. Thus, three main methods of gene transfer in bacteria were discovered:

1) Transformation, whereby bacteria can take up and express exogenous DNA (Mandel and Higa, 1970)

2) Conjugation, where DNA is transferred directly from one bacterium to another (Williams and Skurray, 1980).

3) Bacteriophage-mediated transduction, where genetic material is transferred from one bacterium to another via a bacteriophage particle (Singer and Berg, 1991).

Whilst characterisation of these processes was in progress, attempts were made at introducing foreign DNA into prokaryotic and eukaryotic host cells. These attempts were fairly unsuccessful, which was due to a number of reasons. Firstly, there was the problem that bacteria possess host restriction and modification systems and so any DNA introduced into the host bacterial cells was degraded. Secondly, and most importantly, even assuming that the foreign DNA could be taken up by recipient cells, it could not be maintained because it could not replicate in the new host, and so would become lost during successive rounds of multiplication of the host cells. In addition to the problems of replication, if subsequent detection of DNA uptake was dependent on gene expression, failure could be due to lack of

accurate transcription and translation or failure of any post-translational modification.

Research thus focused on overcoming such fundamental problems that eventually led to a number of solutions that form the basis of modern gene cloning technology. For example, when genes specifying host controlled restriction and modification were identified, this facilitated the production of restrictionless bacteria for cloning purposes, and there now exists a number of strains of *Escherichia coli* with various restriction/modification phenotypes. Also, in order to ensure the replication of foreign DNA, it needed to be linked to an origin of replication, or replicon, and in bacteria there is only one per genome. However, due to the size of the bacterial genome, it would not be possible to manipulate it in an intact form. Therefore, smaller replicons were sought, and as a consequence of the huge amount of microbial genetic studies performed with bacterial species such as *E.coli*, the discovery of more suitable replicons in the form of small plasmids and bacteriophage DNA was achieved. These are now known in gene cloning as vectors, and it is derivatives of the wild types discovered that are now widely used, as well as man-made cosmid vectors. These make suitable cloning vectors because their maintenance is not dependent on integration into the host's genome and it is possible to isolate their DNA in an intact form. So, not only do these vectors provide the replicon function, but they also permit the easy bulk preparation of foreign DNA sequence, free from host cell DNA (Old and Primrose, 1989). Molecules in which foreign DNA has been inserted into a vector molecule are known as recombinant molecules, and the process has been termed molecular cloning or gene cloning because it facilitates the propagation of a line of genetically identical organisms, all of which contain the same inserted foreign DNA fragment. Hence amplifying not only the composite molecule but any gene product whose synthesis it directs.

The basic procedure of *in vitro* gene cloning in prokaryotes was also facilitated by the

simultaneous development of the following techniques. 1. Cutting and joining DNA molecules as a result of the discovery of restriction endonucleases and DNA ligases (Gefter *et al.*, 1967; Zimmerman *et al.*, 1967; Smith and Wilcox, 1970 and Mertz and Davis, 1972). 2. Monitoring the cutting and joining of molecules through the development of agarose gel electrophoresis (Sharp *et al.*, 1973; Helling *et al.*, 1974). 3. Introducing recombinant DNA into the host cell through a number of experiments on transforming *E.coli*. Cohen *et al.*, (1972) showed that $CaCl_2$ treated *E.coli* were effective recipients for plasmid DNA. All these techniques were developed in the late 1960's and early 1970's and facilitated the first cloning experiments carried out by Jackson *et al.* (1972) and Lobban and Kaiser (1973).

1.7.2 Plasmids as Cloning Vectors.

The development of plasmids to serve as vectors for recombinant DNA cloning was essential and it was Cohen and Chang who developed the first plasmid for use as a cloning vector, pSC101, in the 1970's (Cohen *et al.*, 1973 and Cohen and Chang, 1977). Plasmids are usually covalently-closed circles (CCC) of double-stranded DNA and have been detected in a wide range of bacterial genera that are stably inherited in an extrachromosomal state.

The desirable properties of plasmid cloning vectors are as follows:

1) A low molecular weight,

2) An ability to confer readily selectable phenotypic traits on host bacteria,

3) Single sites for a number of restriction endonucleases, preferably in genes with a readily scorable phenotype (Old and Primrose, 1989).

The ideal plasmid cloning vector should be within the size range of 2 to 10 kb, as transformation efficiency decreases steeply with increasing plasmid size and as a consequence it is desirable to limit the amount of foreign DNA to be cloned accordingly.

Additionally, high molecular weight plasmids are more likely to be physically damaged when handled in gene cloning experiments. Whereas the low molecular weight plasmid is often present in multiple copies, which not only facilitates it own isolation but means that, via a gene dosage effect, it can direct the synthesis of high levels of cloned gene products. In some cases though, certain genes cannot be cloned on high copy number plasmids because their presence seriously disrupts the normal physiology of the host and so the strategy here may be to use a low copy number plasmid instead.

1.7.3 Bacteriophage λ as a Cloning Vector.

Bacteriophage λ vectors have been used extensively in recombinant DNA cloning (Murray, 1991). The DNA of bacteriophage λ is a linear duplex molecule of about 48.5 kbp and at the ends are short, single stranded 5' projections of 12 nucleotides, which are complementary in sequence (known as the *cos* site), whereby the DNA adopts a circular structure once injected into a host cell (Old and Primrose, 1989). Genes located in the central region of the DNA sequence are concerned with recombination and lysogenisation and therefore non-essential for lytic growth of the phage, and so can be deleted or replaced with little adverse effect on the propagation of the phage. This region is known as the "stuffer" fragment. Viable phage particles can only be produced if the total DNA is between 37-52 kbp in length this is due to a physical requirement for that amount of DNA by the packaging protein for production of the phage particle. Bacteriophage λ has been the focus of extensive study over the years and so a myriad of derivatives have been developed which can be categorised as either insertional vectors, whereby there exists a single target site where foreign DNA may be inserted or alternatively, replacement vectors where a pair of sites define a fragment that can be removed and replaced by foreign DNA.

1.7.4 Cosmids as Cloning Vectors.

Cosmids are plasmids that contain the cohesive end sites (*cos*) of bacteriophage λ, where the *cos* region is essential for the packaging of λ DNA into its capsid. So, providing that the *cos* regions are located in parallel in the DNA and are 37 to 52 kb apart, these vectors can be incorporated into and propagated as complete λ particles (Collins and Hohn, 1978). Cosmids are therefore useful cloning vectors as they facilitate the study of large genetic regions. They also contain a plasmid origin of replication, at least one selectable marker, and useful restriction endonuclease sites. Packaged DNA is transfected into λ-sensitive *E.coli* by means of the normal λ infection process and then once introduced into the *E.coli* cell, it circularises via the *cos* sites and replicates as a plasmid. An advantage of using cosmids is that because only DNAs that have the *cos* sites the correct distance apart are packaged into phage heads, only recombinant cosmids will be transduced.

1.7.5 Benefits of Gene Cloning.

It is accepted that gene cloning is carried out routinely in the study of bacterial pathogens but it is important to address the fundamental reasons for carrying out such a procedure. Why clone DNA? Basically, DNA is cloned to obtain a specific fragment of DNA in large enough amounts to analyse. In the case of bacterial pathogens, as it is virulence factors that are directly involved in the mechanisms of pathogenicity, their genes represent targets for gene cloning procedures to analyse their potential contribution to pathogenicity. The subsequent identification of such genes is central to understanding how bacterial pathogens circumvent the immune system and cause disease. Gene cloning has made possible numerous types of detailed analyses of DNA, which will be discussed along with the main benefits of cloning genes with respect to the study of bacterial pathogens. The techniques applied to characterise cloned virulence genes and their products is too vast and complex to be fully covered in this review, but the reader can gain suitable knowledge from a number of excellent publications, for

example, Harb and AbuKwaik (1999) and Pelicic *et al.* (1998).

1.7.5.1 Moving Genes from Organism-to-Organism.

Perhaps the most important advantage of cloning genetic material from bacterial pathogens comes from the ability to move genes encoding putative virulence factors into a host cell in which their structure, function and regulation may then be readily studied (Singer and Berg, 1991). This enables gene(s) of interest to be separated, subsequently propagated and possibly expressed away from the confusing and often interfering molecular mechanisms of the organism from which the gene was isolated. The hosts most widely used for prokaryotic gene cloning are based upon *E.coli* strain K12, which has been aptly named as the "workhorse" of molecular biologists and has contributed majorly to modern knowledge of elementary molecular biological mechanisms and the genetic basis of life (Kunhert *et al.*, 2000). This is one example and there exists in practice a range of host/vector systems, but their ultimate function is still the same.

1.7.5.2 Investigations into the Structure of a Gene.

Gene cloning has made analysis of nucleotide sequence possible through methods developed, for example, by Sanger *et al.* (1977). Sequence information is a prerequisite for planning any substantial manipulation of the DNA, and the deduced nucleotide sequence may then be processed in a number of ways. It is possible to search for all known restriction endonuclease target sites to produce a comprehensive physical map that facilitates procedures such as subcloning. Features such as tandem repeats, inverted repeats and potential promoters may be searched for. Analyses are assisted by computer software that allows the nucleotide sequence to be translated into six possible polypeptides, using all three reading frames on both strands so that "open" reading frames (ORFs), which are sequences that do not contain premature terminating codons may be identified that could indicate a polypeptide coding region. It is now

possible to ascertain or confirm a possible function to a sequence by comparing it with other sequences already deposited in DNA databases from other organisms. This can be done at the nucleotide level as well as the amino-acid level, however, due to the degeneracy of the code, similarities may be found between two polypeptide sequences, which may not be so apparent if carried out at the DNA level. Therefore, if sequence similarity is identified this can be an indication of similar function. The internet is now an essential tool in this process as it allows rapid homology searches such as Web-based BLAST searches (Altschul *et al.*, 1997).

1.7.5.3 Investigating the Control of a Gene.

A common aspect of the many strategies adopted by pathogenic bacteria is the ability to vary the expression of the molecules and structures involved in virulence by either the induction or repression of gene expression in response to environmental stimuli. Hence the expression of a number of virulence genes in a closely co-ordinated manner is one of the main factors in determining the success of a pathogen, and so it must recognise signals from the host and express key virulence determinants accordingly to ensure rapid growth and/or survival in a host environment. Virulence gene expression can be followed indirectly by the use of reporter gene fusion technologies in which the regulatory sequences of virulence genes direct the expression of a readily assayable product such as β-galactosidase, alkaline phosphatase, chloramphenicol acetyl transferase, or the green fluorescent protein (Harb and AbuKwaik, 1999). β-galactosidase (*lacZ*) gene fusions have been used to detect the expression of genes in response to laboratory conditions that are thought to mimic the intracellular environment that a pathogen encounters, such as low pH or changes in temperature, upon entry to the host. Leimeister-Wachter *et al.* (1992) studied the virulence genes of *Listeria monocytogenes* and through gene fusions, showed the virulence genes to be thermoregulated, suggesting that the virulence of the organism is triggered on entry to the host.

1.7.6 Other Techniques Used in the Study of Bacterial Pathogenesis.

1.7.6.1 Creating Mutants.

Mutants are an essential pre-requisite for any genetic study and no more so than in the study of gene structure and function relationships (Old and Primrose, 1989). With the development of gene cloning it is possible to precisely modify DNA sequences *in vitro*, prior to studies carried out either *in vivo* or *in vitro* on the genetic function of the modified DNA. A variety of strategies for mutant construction have been developed and the method of choice for a given DNA sequence depends not only on the questions being asked, but on the information already available on the DNA sequence. Once a gene has been cloned and sequenced it is possible to determine whether that particular bacterial gene is involved in pathogenicity by creating a mutant through allelic exchange mutagenesis (Pelicic *et al.*, 1998). However, there are other techniques that can be performed without the prior knowledge of the nucleotide sequence of the gene, such as transposon mutagenesis. These two methods will be discussed in turn: allelic exchange mutagenesis and transposon mutagenesis.

1.7.6.1.1 Allelic Exchange Mutagenesis.

It is possible to use allelic exchange mutagenesis, whereby the gene under scrutiny is disrupted by exploiting the homologous recombination properties of the cell: the functional allele is thereby replaced with an inactivated copy and the virulence of the mutant is compared with that of the parental strain (Pelicic *et al.*, 1998). Mutations can be generated by introducing a point mutation via site-directed mutagenesis as well as null mutations caused by insertions of large fragments of DNA, typically carrying a drug resistance marker. Site-directed mutagenesis is a highly specific predetermined change and the simplest method employed is the single primer method (Gillam *et al.*, 1980; Zoller and Smith, 1983). This method involves priming *in vitro* DNA synthesis with a chemically synthesised oligonucleotide, around 7 to 20 nucleotides in length that carries a base mismatch with the complementary sequence. The synthetic

oligonucleotide primes DNA synthesis and is itself incorporated into the resulting heteroduplex molecule. After transformation of the host *E.coli*, this heteroduplex gives rise to homoduplexes whose sequences are either that of the wild type DNA or that of the mutant, which can be identified by screening by nucleic acid hybridisation with a ^{32}P-labelled oligonucleotide as the probe.

1.7.6.1.2 Transposon Mutagenesis.

Another approach is to make mutations at random in the genome of a particular pathogen using a transposon, and to screen the mutants generated by this process to ascertain whether there has been a loss of virulence, hence indicating that the mutant is lacking gene(s) required for virulence. A transposon is a genetic entity that promotes its own movement or transposition. They are otherwise known as mobile genetic elements and can insert as a discrete segment of DNA at random into plasmids or the bacterial chromosome (Karcher, 1995). The subsequent insertion of the transposon disrupts the nucleotide sequence and thus a mutant is generated. The transposon contains a genetic marker, often an antibiotic resistance gene so that selection of cells that contain transposons is easily conducted. Transposon mutagenesis has been the most commonly used tool for the identification of virulence genes (Valdivia and Falkow, 1998). Early experiments using this method were applied to *Salmonella typhimurium*, whereby genes essential for intracellular survival were identified (Miller *et al.*, 1989). Here, mutants were screened in a macrophage infection model and this led to the identification of a 2-component regulatory system (*Pho*P/*Pho*Q) required for intracellular survival (Groisman *et al.*, 1989; Miller *et al.*, 1989).

1.7.6.2 Polymerase Chain Reaction.

The polymerase chain reaction (PCR) has proved to be an invaluable tool in molecular biology. Fundamentally the method involves the *in vitro* enzymatic amplification of a specific

DNA fragment. Various techniques utilising PCR have been developed for identifying differentially transcribed genes, such as RNA subtractive hybridisation (RSH) and differential display PCR (DD-PCR) (Harb and AbuKwaik, 1999). RSH relies on the hybridisation of cDNAs from bacterial populations exposed to different conditions, the cDNAs found under both conditions are subtracted, and unique cDNAs are studied further (Straus and Ausubel, 1990; Timblin *et al.*, 1990). RSH was used by Plum and Clark-Curtiss (1994) to identify a macrophage-induced gene (*mig*) that is transcribed between one and five days postinfection, but its role in virulence has yet to be fully established.

1.7.6.3 *In Vivo* Expression Technology.

Recently, there has been the development of *in vivo* expression technology (IVET), which has the advantage over classical gene fusion technologies of positively selecting bacterial genes that are induced within their host (Mahan *et al.*, 1993). IVET requires a strain with a mutation in a biosynthetic gene that greatly attenuates its growth *in vivo*. A library of DNA fragments are fused to an operon formed of the same promoter-less gene and a reporter gene is constructed and introduced into the auxotrophic mutant (Pelicic *et al.*, 1998). Clones that show expression of the reporter activity are discarded, and the remaining recombinants are used to infect the host, so that only clones with the DNA fragments that can complement the auxotrophic mutant growth, which correspond to promoters of *in vivo* up-regulated genes, are recovered. Several variations of the IVET method have been used successfully to detect genes that are expressed preferentially during infection (Camilli *et al.*, 1994; Mahan *et al.*, 1995). Some of the genes identified by IVET are involved in general biosynthetic processes or transcriptional regulation, and whilst mutations in some of the *in vivo* induced (*ivi*) genes led to a decrease in virulence, the role of many of these genes remains unclear (Valdivia and Falkow, 1998). Thus, IVET methodology is still being developed and refined.

1.7.6.4 Signature Tagged Mutagenesis.

A new approach has been developed that is less labour-intensive than transposon mutagenesis, which is signature tagged mutagenesis (STM), developed by David Holden and co-workers working on *S.typhimurium* (Hensel *et al.*, 1995). STM combines mutational analysis with the ability to follow the fate of a large number of different mutants within a single animal. It allows the identification of genes that are required for the infective process, but excludes all genes required for survival on laboratory media (Pelicic *et al.*, 1998). Each transposon is engineered to contain a unique, short DNA tag that allows mutants to be distinguished from each other, therefore it is possible to determine which tag from a bacterial input pool used to infect a host is not retrieved in the pools recovered from various organs, hence possible to identify avirulent mutants, unable to persist in animals (Pelicic *et al.*, 1998). The application of STM to the study of *S.typhimurium* resulted in the identification of 19 new virulence genes, most of which were located on a new Pathogenicity Island (SP12) on the *Salmonella* chromosome (Hensel *et al.*, 1995; Shea *et al.*, 1996).

1.8 VIRULENCE FACTORS OF BACTERIAL PATHOGENS.

Virulence factors can be loosely classified into two categories: those factors that cause damage to the host, and those factors that promote colonisation and invasion of the host (Salyers and Whitt, 1994). A myriad of virulence factors exist in bacterial pathogens and it is beyond the scope of this study to review all the mechanisms of microbial pathogenesis, but there are comprehensive reviews available such as Salyers and Whitt (1994) and Patrick and Larkin (1995).

1.8.1 Virulence Factors that Damage the Host.

At a chemical level there are two types of bacterial toxins, firstly, lipopolysaccharides (LPS), which are associated with the cell walls of gram-negative bacteria (endotoxin), and secondly,

proteins which may be released into the extracellular environment of pathogenic bacteria (exotoxin).

1.8.1.1 Exotoxins.

There are different types of exotoxins, grouped together on the basis of their modes of action. One type of exotoxin are the membrane-disrupting, of which there are two types. Firstly, there are those that form channels, or pores in the host cell membrane, which leads to lysis of the host cell, for example, the α-toxin of *Staphylococcus aureus* and streptolysin O produced by *Streptococcus pyogenes*. The second are enzymes known as phospholipases that destabilise the membrane by removing the charged head group from the lipid portion of phospholipids (Salyers and Whitt, 1994). An example is the phospholipase C or α-toxin of *Clostridium perfringens*, which causes damage to the membrane of leukocytes, fibroblasts and muscle cells, and causes platelets to aggregate and lyse (Patrick and Larkin, 1995). Lysis of host bacterial cells can be an advantage to the invading pathogen, as it is a way of releasing nutrients not otherwise available to the bacterium and facilitates the spread of the bacterium through host tissues (Patrick and Larkin, 1995).

A different type of bacterial exotoxin are those that have an intracellular target and have a bifunctional AB structure; where the A component contains the enzymatically active domain, and the B component mediates the binding to the membrane and translocation to the cytosol (Balfanz *et al.*, 1996). Examples are transferases that basically function to inhibit protein synthesis of target cells. This involves the transfer of the ADP-ribosyl moiety of NAD to GTP-binding proteins, examples of these are cholera toxin (*Vibrio cholerae*), diphtheria toxin (DT) of *C.diphtheriae*, the Exotoxin A (ETA) and Exoenzyme S (EXOS) of *P.aeruginosa*, the C2 toxin of *Clostridium botulinum* and the iota toxin of *C.perfringens* (Kruger and Barbieri, 1995). *P.aeruginosa*, for example produces two ADP-ribosyltransferases that differ in their

biochemical properties as ETA ADP-ribosylates elongation factor 2 (Iglewski *et al.*, 1977), while EXOS ADP-ribosylates several different eukaryotic proteins including vimentin and low molecular weight GTP-binding proteins, *in vitro* (Coburn *et al.*, 1989). The overall effect of these toxins is to kill the cell by stopping protein synthesis.

A final type of exotoxin are those involved with the modulation of physiological receptors of cells, which have an affect on intracellular reaction cascades (Balfanz *et al.*, 1996). An example is the heat stable group of toxins (ST) of enterotoxigenic *E.coli* that stimulate cyclic-GMP synthesis in the intestinal brush border, producing an overall effect of fluid loss and lack of fluid adsorption (Patrick and Larkin, 1995).

1.8.1.2 Endotoxins.

The biological, immunological and serologic properties of endotoxin, or LPS, is determined by its structure, and so a full elucidation of such structures is therefore an essential element in establishing a role in virulence at the molecular level. The lipid portion of LPS, which is the toxic portion, is embedded in the outer membrane and exerts its effect on lysis of the bacterial cell, as released LPS is thought to play a major role in the development of the septic-shock that accompanies gram-negative caused infectious diseases (Glauser *et al.*, 1991). Functionally, LPS provides a barrier to heavy metals, lipid-disrupting agents, and larger molecules such as lytic enzymes. It also contains a highly variable carbohydrate (O-antigen) that functions to reduce complement binding and presents the host with multiple antigenic structures in various strains, thus rendering the organisms more resistant to serum-mediated host defences (Proctor *et al.*, 1995).

1.8.1.3 Enzymes that Degrade the Host Extracellular Matrix.

Some pathogenic bacteria produce enzymes that degrade the extracellular matrix in host

tissues, which functions to hold tissues together to give them their physical structure, and is composed of a network of fibrous proteins such as collagen, elastin and fibronectin embedded in a hydrated polysaccharide gel (Patrick and Larkin, 1995). Elastin accounts for nearly 30% of the protein in lung tissue and is responsible for the elasticity of other tissues such as blood vessels and skin. *P.aeruginosa* has elastolytic activity that appears to be due to the activity of two enzymes: LasA and LasB (Salyers and Whitt, 1994). The elastase activity is thought to be involved in the damage that the organism causes to the lungs, particularly common in people with cystic fibrosis. Collagenases are also produced by some pathogens, for example *Clostridium histolyticum* expresses at least two collagenases that cause severe tissue damage and are lethal if injected intravenously into animals where they cause pulmonary damage (Patrick and Larkin, 1995).

1.8.2 Virulence Factors that Promote Colonisation.

1.8.2.1 Pili (Fimbriae).

Adherence to host tissues is required for successful colonisation and subsequent infection by most bacterial pathogens and so to achieve this many pathogenic bacteria possess filamentous surface structures termed pili or fimbriae, of which two major types will be described, Type I and Type IV. It is the tip of the pilus that mediates adherence of bacteria, by attaching to a molecule on the host cell surface, commonly carbohydrate residues of glycoproteins or glycolipids (Salyers and Whitt, 1994). Type I pili are found on the majority of *E.coli*, and *Salmonella* species as well as *Klebsiella pneumonia* and are thin, 7 nm wide and approximately 1μm long surface polymers that bind to D-mannose residues on eukaryotic cells (Clegg and Gerlach, 1987). Type I pilin structures have been implicated in the role of *E.coli* colonisation of the urinary tract and the large bowel during infection (Finlay and Falkow, 1989).

Type IV pili are found in *Pseudomonas* (Pasloske *et al.*, 1985), *Neisseria* (Meyer *et al.*, 1984), *Moraxella* (Marrs *et al.*, 1985), *Bacteriodes* (Elleman *et al.*, 1986) and *Vibrio* (Taylor *et al.*, 1987). Type IV pili have been grouped together on the basis of amino acid sequence similarities among the major pilin component, where homology is highest at the amino-terminus (Fernandez and Berenguer, 2000). The second conserved characteristic is the occurrence of N-methylated amino acids (phenylalanine or methionine) as the first amino acid of the mature pilin structural subunit (Strom and Lowry, 1993). In *P.aeruginosa* it was shown that piliated organisms were 10 times more virulent than non-piliated strains in a burned mouse model (Strom and Lowry, 1993). Type IV pili are thought to have other possible functions in pathogenesis, such as twitching motility, which allows bacteria to spread in infected tissue, as seen in *P.aeruginosa.*

Bacteria also possess proteins on the surface of the cell that aid adherence, but are clearly not assembled as pili and are often referred to as afimbrial adhesins and are thought to be proteins that mediate the tighter binding of bacteria to host cells that follows initial binding via pili (Salyers and Whitt, 1994).

1.8.2.2 Flagella.

Motility in non-filamentous bacteria is generally provided by flagella, which are helical surface structures that rotate from membrane-bound "motors", which push bacteria like boat propellers by anticlockwise rotation (Fernandez and Berenguer, 2000). Potential benefits include increased efficiency of nutrient acquisition, avoidance of toxic substances, the ability to translocate to preferred hosts and access optimal colonisation sites within them, and dispersal in the environment during the course of transmission (Ottemann and Miller, 1997). In the case of *Campylobacter jejuni* and *Helicobacter pylori* the flagellum is essential for pathogenesis, as these organisms have adapted to survival in the gastrointestinal mucus, where colonisation is

imperative in the avoidance of removal by the flow of intestinal chyme, and for negotiating mucus that covers the epithelial surface of the gut (Ottemann and Miller, 1997).

1.8.2.3 Capsule.

A capsule is a loose network of polymers, mostly composed of polysaccharides, that covers the surface of a bacterium (Salyers and Whitt, 1994). The capsule is an important structure, aiding the colonisation of pathogenic bacteria by preventing phagocytosis and protecting the bacteria against complement activation as part of the host's inflammatory response. For example, the capsule of *Streptococcus pneumoniae* has long been recognised as its major virulence factor (AlonsoDeVelasco *et al.*, 1995).

1.8.2.4 Iron Acquisition.

A prerequisite for bacterial colonisation of any environment is the biochemical machinery necessary to obtain adequate nutrition (Patrick and Larkin, 1995). For example, iron is essential for bacterial growth, and in a host such as the human body, the concentration of free iron is low because lactoferrin, transferrin, ferritin, and haemin bind most of the available iron (Salyers and Whitt, 1994). Thus, bacteria have developed mechanisms to acquire iron by producing low-molecular weight structures with a high affinity for Fe^{3+}, known as siderophores. A notable example is that the majority of enterobacteria produce an enterochelin-phenocatechol siderophore, termed enterobactin (Siegfried and Kmetova, 1997).

1.8.2.5 sIgA Proteases.

Mucus layers in the human body contain secretory IgA, whereby its Fc tail attaches to mucin and traps bacteria in the mucus layer. However, some pathogenic bacteria attempting to colonise mucosal surfaces possess a sIgA proteases that break down secretory IgA allowing these bacteria to avoid being trapped in the mucin layer (Salyers and Whitt, 1994).

1.9 AIMS OF THE RESEARCH PROJECT.

In this research project, the techniques of molecular biology are used to study the bacterial pathogen, *B.pseudomallei* and for comparative purposes, the newly discovered avirulent species, *B.thailandensis* the overall aims of which are:

1. To use genomic DNA from *B.pseudomallei* to construct various genomic libraries in a number of *E.coli* host/vector systems.
2. To screen constructed genomic libraries for recombinant bacterial and viral clones producing antigens, enzymes or other components of *B.pseudomallei*.
3. To isolate and characterise any cloned *B.pseudomallei* antigens of interest and to determine their gene sequences and presence in other isolates of *B.pseudomallei* and *B.thailandensis* to elucidate a role in the pathogenicity of *B.pseudomallei*.
4. To use sequences encoding extracellular virulence factors available from other bacterial pathogens to construct PCR primers to test for the presence of homologous sequences in *B.pseudomallei* and *B.thailandensis*.
5. To directly compare the ECPs of *B.pseudomallei* and *B.thailandensis* to elucidate differences that may reflect virulence factors and targets of any future cloning strategies.

In attempting these aims it is hoped that the molecular characterisation of *B.pseudomallei* will elucidate virulence factors and hence the pathogenic processes involved with melioidosis, which has been previously little understood in the hope that a vaccine may be produced. Initially experiments proceeded using donated *B.pseudomallei* and *B.thailandensis* DNA kindly supplied by collaborators. Towards the end of the project, a Containment Level 3 facility, necessary for the safe handling of category 3 pathogens, such as *B.pseudomallei* became available in the department. Subsequently, some work with *B.pseudomallei* cultures was possible. Consequently, the order in which experimental chapters appear in the thesis do not reflect the chronological order in which experimental work was carried out.

CHAPTER 2

MATERIALS AND METHODS

2.1 REAGENTS AND MEDIA.

Details of formulation and preparation of laboratory reagents and growth media (including antibiotics) are given in Appendix I.

2.2 BACTERIOLOGY.

2.2.1 Bacterial Strains.

Isolates of *B.pseudomallei* and *B.thailandensis* used in this study are listed in Table 2.1 and bacterial species closely related to *B.pseudomallei*, which were used as controls and for comparative purposes in certain investigations, are listed in Table 2.2.

Strains of *E.coli* K12 used in gene cloning experiments, together with their genotypes are given in Table 2.3.

Table 2.1 Isolates of *B.pseudomallei* and *B.thailandensis* Used in Experiments in this Research Project (Provided by Dr. D.A.B. Dance PHLS, Derriford Hospital, Plymouth, Dr. T. Pitt, CPHL, Colindale, London and Dr. R. Titball, Porton Down, Salisbury).

STRAIN	SOURCE	SITE	GEOGRAPHICAL ORIGIN	DATE	ARA UTILISATION
19	Environment	Soil	Singapore	1991	-
22	Environment	Soil	Burkinafaso	1973	-
25	Environment	Soil	Madagascar	1977	-
33	Environment	Manure	France	1976	-
46	Human	Blood	NE Thailand	1988	-
53	Human	Urine	NE Thailand	1987	-
97	Environment	Soil	Australia	-	-
98	Environment	Soil	Australia	-	-
102	Environment	Soil	Australia	-	-
112	Human	Multiple	NE Thailand	1992	-
212	Environment	Soil	NE Thailand	1990	-
216	Environment	Soil	NE Thailand	1990	-
217	Environment	Soil (wet)	NE Thailand	1990	-
392	Human	Pus	NE Thailand	1989	-
426	Environment	Soil	Vietnam	-	-
448	Environment	Soil	Vietnam	-	-
E254	Environment	Soil	Thailand	-	+
E255	Environment	Soil	Thailand	-	+
E256	Environment	Soil	Thailand	-	+
E260	Environment	Soil	Thailand	-	+
441398	Human	Blood	Abergavenny (ex-Malaysia)	1999	-
E8	Environment	Soil	Thailand	-	-
E25	Environment	Soil	Thailand	-	-
E27	Environment	Soil	Thailand	-	+
E82	Environment	Soil	Thailand	-	+
204	Human	Blood	Thailand	-	-
576	Human	Blood	Thailand	-	-
Hainan 1	Human	Pus	Hainan Island	1996	*
Hainan 2	Human	Pus	Hainan Island	1996	*
Hainan 4	Human	Blood	Hainan Island	1996	*
Hainan 55	Human	Blood	Hainan Island	1996	-
Hainan 706	Human	Pus	Hainan Island	1996	-
Hainan 106	Human	Blood	Hainan Island	1996	-
Hainan 93	Human	Blood	Hainan Island	1996	-
Jie 187	Human	Blood	China	-	-
Zhan 1	Human	Pus	China	-	*
4845	Monkey	Blood	Unknown	1936	-
MAL6	Unknown	-	-	-	-

* These isolates are claimed by Prof. Song Yang, Hainan Peoples Hospital, China, to be Ara^{+}s which were all isolated from clinical cases (Dance, personal communication), but their arabinose utilisation has not yet been confirmed by either Dr. D.A.B. Dance or Dr. T. Pitt through the API 20NE biochemical testing kit.

Table 2.2 Other Strains of the Genus *Burkholderia* Used in Experiments in this Research Project (Provided by Dr T Pitt, CPHL, Colindale London).

SPECIES	GENOMOVAR or STRAIN	SOURCE
B.cepacia	I	Non-epidemic, Cystic Fibrosis (sputum)
B.cepacia	II	Non-epidemic, Cystic Fibrosis (sputum)
B.cepacia	IIIa epi	Epidemic from New-Zealand, Cystic Fibrosis (sputum)
B.cepacia	IIIa non-epi	Non-epidemic, Cystic Fibrosis (sputum)
B.cepacia	IIIb	Non-epidemic, Cystic Fibrosis (sputum)
B.cepacia	IIIc	Asaccharolytic epidemic, Cystic Fibrosis (sputum)
B.cepacia	IV	Non-epidemic, Cystic Fibrosis (sputum)
B.cocovenans	LMG 11626	
B.plantarii	LMG 10908	
B.vietnamensis	LMG 6998	
B.vietnamensis	LMG 6999	
B.vietnamensis	LMG 10926	
B.vandii	LMG 10620	
B.gladioli	NCTC 12378	

Table 2.3 Strains of *E.coli* K12 Used in Gene Cloning Experiments in this Study. (All strains were obtained from the Department of Biological Sciences, University of Plymouth culture collection).

STRAIN	GENOTYPE	USE
JM109	*rec*A1 *sup*E44 *end*A1 *hsd*R17 *gyr*A96 *rel*A1 *thi* Δ *(lac-pro*AB*)*	A recombinant-deficient *lac*⁻ strain for use with vectors carrying the β-galactosidase system.
DH1	*sup*E44 *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1	A recombinant deficient strain used for routine plating and growth of plasmids and cosmids.
DH5α	F- *end*A1 *hsd*R17 [rK- mK+] *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 ϕ80d*lac*ZΔM15 Δ[*lac*ZYA-*arg*F] U169	A recombinant deficient strain used as a host in plasmid gene cloning.
LE392	F'*hsd*S20 [$r_K^- m_K^-$] *lac*Y1 *gal*K2 *met*B1 *trp*R55 *sup*E44 *sup*F58 λ^-	A strain used for propagation of λGEM11.
C600	*sup*E44 *hsd*R *thi*-1 *thr*-1 *leu*B6 *lac*Y1 *ton*A21	A suppresser strain used in transfection experiments.
Y1088	Δ (*lac*U 169) *sup*E *sup*F *hsd*R (r- m+) *met*B *trp*R *ton*A21 *pro*C::Tn 5 (pMC 9)	Used for the construction of gene libraries in λGT11.
Y1089	*ara*D139 Δ*lac*U169 *pro*A+ Δ*lon rps*L *hfl*A150 [chr::Tn10(tet^r)] *pMC9*	Used for gene cloning in λGT11, deficient in the lon protease, which may allow increased stability of foreign proteins.
Y1090	*sup*F *hsd*R *ara*D139 Δ*lon* Δ*lac*U169 *rps*L *trp*C22::Tn10(tet^r) pMC9	Used for screening expression libraries and propagation of λGT11.

2.2.2 Laboratory Culture and Storage of Bacteria.

2.2.2.1 Strains of *E.coli.*

Strains of *E.coli* K12 were routinely cultured using Luria-Bertani (LB) medium. Those strains that bore plasmids were cultured using LB supplemented with the appropriate antibiotic(s). Bacterial strains were streaked onto LB agar, and subsequently single colonies for broth culture were used to inoculate a given volume of LB media. Bacterial cultures were incubated aerobically at 37°C unless otherwise stated. Stock cultures were maintained in 50% glycerol (which served as a cryoprotectant), in 1ml volumes stored in polypropylene microcentrifuge tubes (MCC tubes) kept at -20°C. Some strains were flash frozen in liquid nitrogen and stored at -70°C. Petri dishes containing agar cultures of *E.coli* in regular use were stored, inverted at 4°C and subcultured every 4 to 6 weeks.

2.2.2.2 Strains of *B.pseudomallei*, *B.thailandensis* and other *Burkholderia* species.

B.pseudomallei is classified by the Advisory Committee on Dangerous Pathogens (ACDP) as a category 3 pathogen and thus requires special containment facilities. All strains of *B.pseudomallei* and *B.thailandensis* were cultured in a Containment Level 3 (CL 3) laboratory unit either by Dr. D.A.B. Dance at the PHLS, Derriford Hospital, Plymouth or by Dr. M.L. Gilpin in 401A, Davy Building, University of Plymouth. Isolates of *B.pseudomallei*, *B.thailandensis* and other *Burkholderia* species used were initially streaked onto LB media whereby single colonies were used to inoculate 10 ml of LB broth in sterile polypropylene Universal bottles. Cultures were incubated statically at 37°C for 48 h.

2.2.3 Characterisation of Bacterial Cultures.

Cultures of *B.pseudomallei* supplied by Dr. D.A.B. Dance from PHLS, Derriford hospital, Plymouth, and cultures of other *Burkholderia* species supplied by Dr. T. Pitt, CPHL,

Colindale, London had been previously identified by means of the API 20NE biochemical testing kit.

2.3 BACTERIOPHAGE λ.

2.3.1 Titration of Bacteriophage λ Suspensions.

This procedure enabled the titre (the number of plaque forming units per ml) of a bacteriophage λ suspension to be determined. A 10 ml culture of the appropriate *E.coli* host strain was incubated overnight at 37°C in NZCYM broth supplemented with 0.2% (w/v) maltose. Bacteria grown in the presence of maltose adsorb bacteriophage λ more efficiently as it induces the maltose operon, which contains the gene (*lamB*) encoding the bacteriophage λ receptor. Hence, incubation with maltose functions to increase the number of bacteriophage receptors (Sambrook *et al.*, 1989). A set of serial, 10-fold dilutions were set up (of the order 10^{-1} to 10^{-7}) of the phage suspension using SM buffer as the diluent. A 0.1 ml volume of phage dilution was placed into a sterile, thick-walled Universal tube to which 0.2 ml of the appropriate *E.coli* host strain was added. The contents were mixed gently and incubated at 37°C for 15-20 min. to allow the phage to absorb to the bacteria, after which a 5 ml volume of molten soft NZCYM agar was added. The tube was mixed gently, so as not to introduce any air bubbles, and overlayed onto NZCYM agar plates. Plates were left for a few minutes to set and then incubated, inverted at 37°C for 18-24 h.

2.3.2 Preparation of Bacteriophage λ Stocks.

2.3.2.1 Bacteriophage λ Stock from a Single Plaque.

A small-scale stock of phage could be made from just one plaque and this stock was stable stored at 4°C for long periods of time. A sterile Pasteur pipette with a rubber bulb attached was

stabbed through a chosen plaque, and by applying mild suction the plaque and underlay was drawn up into the pipette. Well-separated plaques were preferentially picked, as bacteriophage λ is able to diffuse considerable distances through the top layer of soft agar (Sambrook *et al.*, 1989). The agar plug was placed into 1 ml of SM buffer, held in a sterile MCC tube and one drop of chloroform was added to inactivate any remaining bacteria. The tube was left for at least 1-2 h. to enable phage to diffuse out of the agar plug.

2.3.2.2 Bacteriophage λ Stock from Confluent Lysis Plates.

Larger stocks of phage were prepared from whole plate lysates. Dilutions of the appropriate phage were plated out as in 2.3.1, and those plates exhibiting near confluent lysis were chosen. Approximately 20 plates were used and after incubation each plate was flooded with a 3 ml volume of SM buffer and placed at 4°C for 1-2 h. to soften the agar overlay. Afterwards, the soft agar overlay was macerated using a sterile spreader to ensure good recovery of the phage. The preparation was transferred into MCC tubes and centrifuged at 13,000 x g for 10 min. to pellet the soft agar, and the resulting supernatant was collected into fresh MCC tubes to which 3 drops of chloroform from a Pasteur pipette was added.

2.3.2.3 Bacteriophage λ Stock From Liquid Cultures.

This method yielded a high titre stock of phage that could be used for DNA extraction and subsequent use in gene cloning experiments. This procedure involved large volumes of cultures, initially containing a low concentration of phage and a known number of host bacterial cells, which were incubated so that uninfected bacterial cells would continue to divide for several hours. The ratio of phage to host bacterial cells, or multiplicity of exposure (MOE) was calculated in each experiment. A MOE of 1 was a 1:1 ratio of phage to bacterial cells. MOE's ranging from 0.01 to 100 were used in experiments and calculated by firstly enumerating the number of bacterial cells in an overnight culture using a haemocytometer, and

using the phage titre which was previously obtained by performing an assay as detailed in 2.3.1. Flasks (2 litres) containing 500 ml of Phage Propagation Media were incubated with a known number of phage particles and bacterial cells for 16-18 h. at 37°C with shaking. During this time, successive rounds of infection led to the production of increasing quantities of phage until complete lysis of the culture occurred. This was easily visualised as the culture would be clear in colour with strings of cell debris at the bottom of the flask.

A 0.25 ml volume of chloroform was added to flasks showing evidence of complete lysis and left shaking at 37°C for 10 min. Afterwards, the culture was centrifuged at 5000 x g for 10 min. to remove cell debris. The volume of supernatant was measured before transferring to a sterile flask and NaCl was added to 0.5M and polyethylene glycol (PEG) 6000 to 10% (w/v). The contents were shaken to dissolve the PEG 6000 and allowed to incubate overnight, on ice to precipitate the phage. Precipitated phage was pelletted by centrifugation at 15,000 x g for 15 min. at 4°C before resuspending in 35 ml of SM buffer. Once resuspended, the contents were transferred to a 40 ml Beckman Quickseal tube and heat-sealed. This was ultracentrifuged for 12 h. at 27,000 x g at 20°C. The bacteriophage particles appeared as a glassy pellet at the bottom of the tube. The supernatant was carefully discarded and the pellet resuspended in a 500µl volume of SM buffer.

2.3.3 Extraction of Bacteriophage λ DNA.

A high titre stock of phage (as produced in 2.3.2.3) was aliquoted in 100 µl volumes into sterile MCC tubes. Next, a 10 µl volume of 2 M Tris-HCl and 0.2 M Na_2EDTA (pH 8.5) solution was added and mixed thoroughly. An equal volume (110 µl) of formamide was added and the preparation was incubated at room temperature for 1-2 h. Afterwards, an equal volume (220 µl) of sterile analar water was added and the phage DNA was precipitated with ethanol (2.4.1). The phage DNA was pelletted by centrifugation at 13,000 x g for 10 min., the

supernatant was removed and the DNA pellet washed with 70% ethanol, allowed to dry and resuspended in TE buffer. The DNA was further purified by extraction with phenol (2.4.3.2).

2.3.3.1 Wizard Lambda Prep-DNA Purification System (Promega).

Alternatively, phage DNA was isolated from lysates (prepared as in 2.3.2.2) using a commercial kit as follows. To a 10 ml volume of phage lysate, 40 µl of resuspended Nuclease Mixture was added and the tube incubated for 15 min. at 37°C. A 4 ml volume of Phage Precipitant Solution was added, the contents mixed gently and left on ice for 30 min. Next, the contents were centrifuged at 10,000 x g for 10 min. The resulting supernatant was discarded and the pellet resuspended in 500 µl of Phage Buffer Solution and transferred to a sterile MCC tube. Any insoluble particles were removed by centrifuging for 10 sec. at 10,000 x g and the resulting supernatant transferred to a fresh MCC tube, where 1 ml of thoroughly resuspended Purification Resin was added and mixed by inverting the tube. The phage DNA was recovered using Wizard minicolumns attached to a vacuum manifold. The columns were prepared by attaching a syringe barrel to the Luer-lok extension of each column. The resin/lysate mixture was loaded into the syringe barrel and a vacuum applied to draw this mixture through the column. The column was washed using 2 ml of 80% isopropanol (propan-2-ol), drawn through the column and the resin was dried by drawing a vacuum for an additional 30 sec. after the isopropanol had been pulled through the column. The column was removed, placed into a 1.5 ml MCC tube and centrifuged at 10,000 x g for 2 min. to remove any remaining isopropanol. The column was then placed into a fresh MCC tube and a 100 µl volume of sterile analar water, or TE buffer pre-heated to 80°C was added and the column immediately centrifuged for 20 sec. at 10,000 x g to elute the DNA. A second elution was carried out using a 50 µl volume of water or TE buffer, to ensure maximal recovery.

2.4 DNA ISOLATION AND PURIFICATION METHODS.

2.4.1 Concentrating DNA solutions.

DNA solutions were concentrated by either ethanol or isopropanol precipitation. Firstly, the DNA solution was measured into a sterile MCC tube and a 0.1 volume of 3 M sodium acetate was added and mixed gently. To this either a 0.6 volume of isopropanol or 2.5 volumes of ice-cold analytical grade ethanol was added. To facilitate precipitation, DNA precipitated with ethanol was incubated at -80°C for 15 min. and DNA precipitated with isopropanol was incubated at room temperature for up to 2 h. Resulting precipitated DNA was pelletted by centrifugation in a microcentrifuge at 13,000 x g for 10 min. Once pelletted the supernatant was removed and the DNA pellet was washed with 70% (v/v) ethanol and centrifuged again at 13,000 x g for 5 min. The DNA pellet was dried and resuspended in TE buffer or sterile analar water and stored at 4°C.

2.4.2 Plasmid DNA Isolation.

2.4.2.1 Miniprep Method.

A variation on the alkaline lysis method of Birnboim and Doly (1979) was used to produce plasmid DNA of a sufficient quality to enable restriction enzyme digestion and transformation. A 1.5 ml volume of an overnight bacterial culture was transferred into a MCC tube and centrifuged at 13,000 x g for 30 sec. to pellet the bacterial cells. The supernatant was discarded and the cell pellet was resuspended in 100 µl of B&D Solution I and incubated at room temperature for 5 min. Next, 150 µl of B&D Solution II was added and the contents mixed gently by inverting the tube several times. At this stage the solution appeared clear and viscous and was incubated on ice for 5 min., after which time 150 µl of ice-cold B&D Solution III was added, mixed gently and the preparation incubated on ice for a further 5 min. A white flocculent precipitate formed which was pelletted by centrifugation at 13,000 x g for 5 min. to yield a clear supernatant of approximately 400 µl, which was transferred to a fresh MCC tube.

To this 1 ml of ethanol (-20°C) was added to precipitate the plasmid DNA, and after mixing the tube was incubated on ice for 10 min. followed by centrifugation at 13,000 x g for 10 min. The pellet of plasmid DNA was washed with 70% ethanol, centrifuged at 13,000 x g for 2 min. and the pellet was dried before resuspending in either TE buffer or sterile analar water. Often RNA was present in the preparation and therefore 1 µl of RNase (1 µg/µl) would be added and the contents incubated at 37°C for 1-2 h. The RNase and any residual proteins were then removed by extraction with phenol (2.4.3.2) and the plasmid DNA finally precipitated with ethanol and resuspended as before.

2.4.2.2 Maxiprep Method.

The maxiprep method was used if processing larger cultures from 500 ml up to 2 litres and is again a variation of Birnboim and Doly (1979). Bacterial cells were harvested by centrifugation at 5000 x g for 15 min. at room temperature and resuspended in 16 ml of B&D Solution I, which was supplemented with 0.8 ml of lysozyme at 20 mg/ml, to weaken the cell walls of the bacteria. Next, 16 ml of B&D II was added and mixed gently by swirling the tube, the solution appeared clear and viscous at this stage. The solution was left on ice for 10 min., after which 12 ml of ice-cold B&D III was added and the contents mixed gently to form a white flocculent precipitate and left on ice for a further 10 min. The precipitate was pelletted by centrifugation at 5000 x g for 15 min. at 4°C and the resulting supernatant was removed and allowed to warm to room temperature before adding a 0.6 volume of room temperature isopropanol. The contents were left for 2-3 h. at room temperature to precipitate the plasmid DNA. Afterwards, the plasmid DNA was pelletted by centrifugation at 10,000 x g for 15 min. and resuspended in 2.5 ml of TE buffer ready for Caesium Chloride-Ethidium Bromide (CsCl-EtBr) density centrifugation to purify the plasmid DNA.

2.4.2.3 DNA Wizard Plus Maxiprep DNA Purification System (Promega)

This commercial kit was used to isolate plasmids from bacterial cultures that were 100-500 ml in size as an alternative method of isolating plasmid DNA of a quality for restriction enzyme analysis and ligation. The culture was pelletted by centrifugation at 5000 x g for 10 min. at 20°C and resuspended in 15 ml of Cell Resuspension Solution. Next, 15 ml of Cell Lysis Solution was added and mixed gently by inverting the centrifuge tubes. The solution became clear and viscous once cell lysis was completed (approximately 20 min.). To this 15 ml of Neutralisation Solution was added and mixed immediately by gently inverting the centrifuge tubes. The solution was then centrifuged at 14,000 x g for 15 min. The cleared supernatant was filtered through miracloth, the volume measured and transferred to a clean centrifuge bottle. A 0.5 volume of room-temperature isopropanol was added and mixed in order to precipitate the plasmid DNA. The solution was centrifuged at 14,000 x g for 15 min. to pellet the plasmid DNA. The supernatant was discarded and the DNA pellet resuspended in 2 ml of TE buffer.

A 10 ml volume of DNA Purification Resin was added to the solution and mixed. One Wizard maxicolumn was used for each extraction and inserted into a vacuum manifold port. The DNA/resin mix was transferred into the maxicolumn and to this 25 ml of Column Wash Solution was added and a vacuum was applied. A 5 ml volume of 80% ethanol was added to the column and a vacuum applied once again until the ethanol had passed through. The maxicolumn was removed and inserted into a 50 ml Falcon tube and centrifuged in a Denley benchtop centrifuge at 1,300 x g for 5 min. after which both the Falcon tube and contents were discarded. The maxicolumn was replaced back onto the vacuum source and the resin allowed to dry to completion by drawing a vacuum for 5 min. before placing into a provided tube to which 1.5 ml of pre-heated (65-70°C) TE buffer was added, and after waiting for one minute the plasmid DNA eluted by centrifuging at 1300 x g for 5 min. To eliminate any resin fines

that may have been present, the DNA solution was passed through a 0.2 μm filter with the aid of a syringe. The resulting plasmid DNA was stored at 4°C.

2.4.2.4 Qiagen Plasmid Isolation (Minikit).

Qiagen kits were often used as an alternative to the CsCl/EtBr density centrifugation method (2.4.3.1) as a quick method of isolating plasmid DNA of a sufficient quality for DNA sequencing. A 3 ml volume of an overnight culture was centrifuged at 13,000 x g, for 1 min. to pellet the bacterial cells, the supernatant was discarded and the cells were resuspended in 0.3 ml of Buffer P1. To this 0.3 ml of Buffer P2 was added, mixed gently and incubated at room temperature for 5 min. Afterwards, 0.3 ml of ice-cold Buffer P3 was added and mixed immediately, before incubating on ice for a further 5 min. The preparation was centrifuged at 13,000 x g for 10 min. and the supernatant was discarded. A Qiagen-tip 20 column was allowed to equilibrate by applying 1 ml of Buffer QBT and allowing the column to empty by gravity flow. Afterwards, the column was washed with 1 ml of Buffer QC, which was repeated for a total of 4 washes. The plasmid DNA was eluted from the column with 0.8 ml of Buffer QF into a sterile MCC tube and precipitated with a 0.7 volume of isopropanol. The plasmid DNA was pelletted by centrifugation at 13,000 x g for 10 min., the supernatant was removed and after washing the pellet with 1 ml of 70% ethanol, the plasmid DNA was dried and resuspended in 20 μl of TE buffer or sterile analar water.

2.4.3 Plasmid DNA Purification Methods.

2.4.3.1 Caesium chloride/Ethidium bromide density ultracentrifugation.

This process allowed the isolation of higher and purer concentrations of plasmid DNA, free of contaminating molecules such as chromosomal DNA, high molecular weight RNA and protein, which was important for successful ligation reactions and for DNA sequencing.

To resuspended plasmid DNA in a volume of 2.5 ml of TE buffer, 2.6 g of CsCl was added, and the contents warmed slightly to aid solution of the salt. Next, 0.25 ml of ethidium bromide solution (10 mg/ml) was added and the tube was left in the dark for around 2 h. to allow proteins to complex with the ethidium bromide. The tube was spun briefly to pellet the unwanted protein material and the supernatant was loaded into a 3.5 ml Beckman Quickseal tube and heat-sealed before ultracentrifuging at 80,000 x g for 24 h. at 20°C. After this time, two distinct bands of DNA were visible using UV light. The top band, which contained less material, represented any remaining chromosomal DNA and nicked plasmid DNA, whereas the lower band represented the closed circular plasmid DNA. Any RNA in the preparation was pelletted at the bottom of the tube. A hypodermic needle was inserted into the top of the tube to allow air to enter. A fresh needle attached to a syringe was inserted into the side of the tube, bevelled side up and positioned just below the lower plasmid band. The contents of the band were drawn into the syringe carefully so as not to remove any of the upper chromosomal DNA.

The volume of resuspended plasmid DNA recovered from the CsCl/EtBr density centrifugation procedure was measured and to this an equal volume of isopropanol, previously saturated with water and CsCl was added and mixed gently. Two distinct layers formed, of which the upper, containing the isopropanol and ethidium bromide was removed carefully with a pipette and discarded. This process was repeated until the lower aqueous layer containing the plasmid DNA was clear in colour, indicating that the ethidium bromide had been removed. The plasmid DNA and CsCl were precipitated with ethanol after adding a 0.1 volume of 3 M sodium acetate, and the resulting pellet was washed with 1 ml of 70% ethanol, which functioned to dissolve the CsCl. The tube was then centrifuged at 13,000 x g for 5 min. to pellet plasmid DNA and any undissolved CsCl. The supernatant was discarded and another 1 ml of 70% ethanol was added to fully dissolve all of the remaining salt. The tube was

centrifuged again at 13,000 x g for 5 min. to pellet the plasmid DNA, which was resuspended in TE buffer or analar water.

2.4.3.2 Phenol Extraction of DNA.

Extraction of DNA with liquefied phenol facilitates the removal of proteins from DNA solutions, which is often necessary to improve the quality of DNA for techniques such as restriction analysis or ligations. An equal volume of phenol solution was added to a DNA solution of known volume in a MCC tube, mixed thoroughly until an emulsion formed. The mixture was centrifuged at 10,000 x g for 2 min. in a microcentrifuge to separate into 2 layers. The upper aqueous layer of DNA solution was removed carefully with a pipette so as not to disturb the band of proteinaceous material that had formed at the interface, and was transferred to a new MCC tube. To this an equal volume of a phenol solution previously mixed with chloroform plus isoamylalcohol (IAA) in the ratio: 24:24:1 was added, mixed to form an emulsion and separated by centrifugation at 10,000 x g for 2 min. Again, the top aqueous layer containing the DNA was removed and transferred to a new MCC tube. This step was repeated until no band of proteinaceous material was visible at the interface. Finally, an equal volume of a water-saturated solution of chloroform plus IAA was added to the DNA solution, mixed and centrifuged at 10,000 x g for 2 min. The upper aqueous layer was removed and transferred to a new MCC tube. This was repeated twice to ensure that no phenol remained in the DNA solution, which might interfere with subsequent processes such as restriction digestion and ligation reactions.

2.4.4 Chromosomal DNA Isolation Procedures.

2.4.4.1 RapidPrep^{TM} Genomic DNA Isolation Kit (Pharmacia Biotech)

This commercial kit was used to isolate chromosomal DNA from *Burkholderia* species that did not require Containment level 3 facilities and any *E.coli* strains used as negative controls

in Southern blots. The kit was also used to further purify genomic DNA isolated from *B.pseudomallei* and *B.thailandensis* in 2.4.4.2.

A volume of culture that contained 2 x10^9 bacterial cells was centrifuged at 13,000 x g for one minute in a 1.5 ml MCC tube and the supernatant was discarded. A 50 µl volume of Modified Extraction Buffer was added to the sedimented cells and vortexed until a homogeneous suspension formed, and the preparation was incubated at 55°C for 30 min. Afterwards, 800 µl of Application Buffer was added and mixed by inverting the tube several times, followed by incubation at room temperature for 5 min.

Microspin columns were prepared by inverting several times to resuspend the resin. The bottom closure was snapped off, the top cap removed and the column placed into a MCC tube and centrifuged at 735 x g for 1 min. The MCC tube was discarded and the bottom closure replaced on the column. A 400 µl volume of the bacterial preparation was added to the column and inverted several times before leaving for 1 min. to allow the resin to settle. The top cap and bottom closure were again removed and the microcolumn centrifuged in a clean MCC tube at 735 x g for 2 min. The support tube was emptied and the remaining 400 µl of the preparation was added to the same pre-spun column, again resuspending the resin before centrifuging at 735 x g for 2 min. A 400 µl volume of Wash Buffer was added to the top of the resin bed and the microspin column centrifuged at 735 x g for 2 min. The support tube was emptied and this process repeated once for a total of 2 washes. The support tube was replaced with a clean MCC tube and 200 µl of Elution Buffer was added to the column and centrifuged at 735 x g for 2 min., which was repeated for a total of two elutions. The genomic DNA was precipitated by adding 320 µl of isopropanol and incubating at room temperature for 10 min. The genomic DNA was pelletted by centrifugation for 10 min. at 13,000 x g, and resuspended in 10 µl of TE buffer.

2.4.4.2 Chromosomal DNA Isolation Procedures for *B.pseudomallei* and *B.thailandensis* Cultures.

These procedures were carried out either by Dr. D.A.B. Dance or Dr. M.L. Gilpin.

2.4.4.2.1 Miniprep Method (10 ml Cultures).

This method was used for small cultures of *B.pseudomallei* and *B.thailandensis* to produce DNA of a quality for PCR reactions. To a 10 ml culture of *B.pseudomallei* 5 ml of Solution A was added, mixed and left at room temperature for 5 min. Next, 1 ml of Solution B was added and the preparation was left at room temperature for a further 5 min. After which time a 4 ml volume of Solution C was added and mixed, followed by 2 ml of Solution D. The preparation was placed at 80°C with occasional mixing for 1-2 h. From this solution, 0.2 ml aliquots were removed and spread onto Nutrient agar plates to check for viability. Plates were incubated at 37°C for 72 h., while the remaining lysate was stored at 4°C until viability checks were completed to a satisfactory conclusion. This ensured lysates were free of the pathogen before being released from the CL 3 laboratory for use elsewhere. These preparations were sometimes subjected to further purification (2.4.4.1).

2.4.4.2.2 Maxiprep Method (200 ml Cultures)

To a 200 ml culture of *B.pseudomallei* or *B.thailandensis,* 100 ml of Solution A was added, mixed and left for 5 min. at room temperature. A 20 ml volume of solution B was added, the contents mixed and left at room temperature for a further 5 min. After which time 80 ml of Solution C was added, followed by 40 ml of Solution D, the tube contents mixed and left at 80°C with occasional mixing for 1-2 h. Next, 0.2 ml volumes were removed from the preparation and spread onto Nutrient agar plates to check for viability. Plates were for incubated for 48 h. at 37°C and the lysate was stored at 4°C until viability checks were complete (as in 2.4.4.2.1).

2.4.5 Agarose Gel Electrophoresis.

2.4.5.1 Determination of DNA Concentration.

The amount of DNA in a given solution was determined using ethidium bromide fluorescent quantification, which utilises the ultraviolet fluorescence emitted by ethidium bromide molecules intercalated into the DNA (Sambrook, *et al.*, 1989). This was achieved by subjecting the DNA sample to electrophoresis in agarose gels containing ethidium bromide at a concentration of 0.5 μg/ml. Whereby DNA samples to be checked were run alongside DNA standards of known concentration. Bacteriophage λ DNA was used, whereby 4 μl of the λ standard DNA was equivalent to 0.5 μg of DNA. A variety of amounts of the standard were used on an agarose gel and the unknown was compared against these on the basis of the amount of fluorescence emitted when exposed to UV light (310nm) on a UV transilluminator (UV Products inc., California USA).

2.4.5.2 Preparation of an Agarose Gel.

In order to analyse, identify and purify DNA fragments, electrophoresis through an agarose gel was the standard method used. DNA's from 200 bp up to approximately 50 kbp in length can be separated in agarose gels of various concentrations (Sambrook *et al.*, 1989). Gels of either 0.8% (w/v) or 1% (w/v) were used in most cases. This was achieved by adding the desired percentage of agarose to either 30 ml or 60 ml of Tris-Borate EDTA (TBE) buffer in a conical flask. TBE buffer was prepared at 10 X strength and diluted with distilled water accordingly. The agarose and TBE preparation was heated in a microwave oven for approximately 3 to 4 min. until all of the agarose had melted. Molten agarose was placed in a waterbath at 45°C until equilibrated. Once cooled, ethidium bromide was added, 5 μl (at a concentration of 10 mg/ml) per 100 ml of agarose gel solution to give an overall concentration of 0.5 μg/ml. The molten agarose was poured into a taped UV-transparent gel mould into which a comb was inserted to create subsequently, wells in which to load the DNA samples.

2.4.5.3 Agarose Gel Electrophoresis.

Once an agarose gel had been made and allowed to set, DNA samples were prepared by mixing them with a Ficoll DNA loading buffer. The taped ends of the tray mould containing the agarose gel were removed and the tray was placed into a Pharmacia GNA 100 electrophoresis tank, filled with 350 ml of TBE buffer (X1) so that it was submerged by the buffer, before loading the DNA samples into the wells. The gel was run at a constant 120 V after which the DNA bands were visualised using a UV transilluminator and photographed using a Polaroid CU 5 camera system with a type 665 film, or by using a UVI pro documentation system.

2.4.6 Recovery of DNA from Agarose Gels.

2.4.6.1 Electroelution into Dialysis Tubing.

Dialysis tubing was prepared by boiling in 100 ml of TBE buffer (X1) for 3 min. in a microwave and allowing it to cool. Once a DNA sample had been run on an agarose gel, the DNA band of interest was located on the transilluminator and by using a clean, sterile razor blade, a slice of agarose was excised containing the desired DNA band. This was placed into dialysis tubing with a knot at one end. The tubing was filled with 1X TBE buffer and another knot was placed in the opposite end to seal the contents, care was taken not to introduce any air bubbles into the tubing. The tubing was immersed in 1X TBE buffer in an electrophoresis tank and an electric current applied to electroelute the DNA out of the gel and onto the inner wall of the bag. A constant 120 V was applied for 30 min. to ensure that all the DNA was eluted out of the agarose gel. The polarity of the current was reversed for 1 min. to release the DNA from the inner wall of the tubing back into the TBE buffer and the DNA solution was removed and placed into a MCC tube. The tubing was washed out with a small amount of TBE buffer to enhance DNA retrieval and the DNA recovered was purified using the BIO-RAD, Prep-A-Gene DNA Purification System (2.4.6.2).

2.4.6.2 DNA Isolation from Agarose Slices Using the BIO-RAD Prep-A -Gene Kit.

This method was sometimes used as a quicker alternative to isolating DNA from agarose gels as described in 2.4.6.1. The Prep-A-Gene kit was also adapted to purify DNA fragments as an alternative to phenol extraction.

The desired DNA band was excised from an agarose gel using a sterile razor blade as in 2.4.6.1 and placed into a MCC tube. The amount of DNA in the band was estimated by comparing the fluorescence to that of the standards prior to excision. The gel slice was centrifuged for several seconds so that it was situated on bottom of the tube and the volume of the gel slice was estimated. The amount of Prep-A-Gene Matrix to be used was calculated, given that 1 µl of Matrix binds 0.2 µg of DNA. Hence, based on the estimated volume of the gel slice and the calculated amount of Matrix needed to bind all of the DNA, 3 volumes of this combined amount of DNA purification kit Binding Buffer was added to the gel slice and the tube agitated gently so that it fully dissolved. After which, the predetermined amount of Matrix was added and the tube mixed gently at room temperature for 10 min. The Matrix was pelletted by centrifuging 13,000 x g for 30 sec. and the supernatant was discarded. The pellet was resuspended in an amount of purification kit Binding Buffer equivalent to 25 X the Matrix volume. The tube was centrifuged again at 13,000 x g for 30 sec., the supernatant disposed of and the pellet washed with a 25 X Matrix volume of Wash Buffer for a total of 2 washes, whereupon after the second wash, all traces of liquid were carefully removed by centrifuging twice. The DNA was eluted by resuspending the pellet in at least 1 pellet volume of Elution Buffer and incubated between 37-50°C for 5 min., which was repeated for a total of 2 elutions.

2.5 DNA MANIPULATION METHODS.

2.5.1 Transformation of Calcium Chloride Treated *E.coli* with Plasmid DNA.

Plasmid DNA was introduced into *E.coli* by the process of transformation, whereupon treatment of bacterial cells with divalent cations makes the bacteria temporarily permeable to small DNA molecules, such as plasmids. A 1.5 ml volume of appropriate bacteria, grown to an OD of 0.6 (A_{590}) was centrifuged at 5,000 x g for 1 minute in a sterile MCC tube. The supernatant was discarded and the bacterial cell pellet resuspended in 200 µl of ice-cold $CaCl_2$ solution (75 mM). Once resuspended, the bacteria were left at 4°C for 40 min. and re-pelletted by centrifugation at 5,000 x g for 1 min. The supernatant was again discarded and the bacterial cell pellet resuspended in 100 µl of ice-cold $CaCl_2$ (75 mM). Plasmid DNA was added to the bacteria, mixed gently and the preparation left on ice for 10 min., before placing at room temperature for a further 10 min. and finally returning to ice for 30 min. Unless otherwise stated, 50 ng of plasmid DNA was added in a volume no greater than 10 µl to be transformed. Next, the bacteria were heat-shocked by placing at 42°C for 2 min. and a 0.9 ml volume of LB broth was added, mixed gently and the transformation mix was placed at 37°C for 90 min. After this time, samples were spread onto LB plates supplemented with the appropriate antibiotic required for selection of successful transformants. Plates were incubated at 37°C for 24-36 h. although longer periods of incubation were used if more than one antibiotic was being selected for, or alternatively, if lower incubation temperatures were being used in certain gene cloning experiments.

2.5.2 Blue/White Screening to Detect Recombinant Plasmids in *E.coli* K12 Hosts that Contain the *lac* Genes.

To check for recombinants, isolation plates would often be supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) and isopropylthio-β-D-galactosidase (IPTG) if the vector contained the *lacZ* gene, which codes for β-galactosidase. Recombinants disrupt the

lacZ gene, as foreign DNA is ligated into it, and so a non-active form of β-galactosidase is produced. Recombinants appear white as opposed to non-recombinants, which appeared blue-green in colour. When β-galactosidase is assayed with X-gal, which is normally colourless, it cleaves off the galactose moiety from X-gal, forming a blue colour. However, if there is a disruption of the *lacZ* gene so that an inactive form of β-galactosidase is formed, the reaction cannot happen and so the galactose moiety remains on X-gal, hence the recombinant colonies are white. To screen recombinants in this way, LB agar containing the appropriate antibiotic was prepared and the plates dried at 45°C for 10-15 min. The plates were overlayed with 5 ml of LB agar supplemented with the same antibiotic, as well as X-gal solution (at 40 μg/ml) and IPTG solution (at 120 μg/ml). The overlay was allowed to set before the transformation mixture was spread plated onto the plates, which were incubated for 24 h. at 37°C, unless otherwise stated, and examined for presence of blue (*lac*$^+$, non-recombinant), and white (*lac*$^-$, recombinant) colonies.

2.5.3 Preparation of Transformation Competent *E.coli* K12 Cells.

2.5.3.1 Preparation of Frozen Competent *E.coli* Cells.

A procedure for making competent cells was carried out, which had been developed by Viesturs Simanis, Imperial College, London, previously developed from a procedure by Hanahan (1985). The bacterial strain used was DH5α, of which a 400 ml culture was incubated at 37°C with shaking until an OD of 0.6 (A_{600}). The culture was chilled on ice for 15 min. before centrifugation at 1,000 x g, 15 min. at 4°C to pellet the bacterial cells. The supernatant was discarded and the bacteria were resuspended in one-third of the initial culture volume of RF1 (Appendix I) and kept on ice for 15 min. before centrifuging at 1,000 x g for 15 min. at 4°C. The bacteria were resuspended in 1/12.5 of the initial culture volume of RF2 (Appendix I). Competent cells were aliquoted into 200 μl volumes in MCC tubes and flash frozen in liquid nitrogen for storage at -70°C until needed.

2.5.3.2 Transformation of Frozen Competent *E.coli* Cells.

A 200 µl volume of frozen competent DH5α bacteria was removed from -70°C and allowed to thaw on ice. The bacteria were mixed gently by flicking the tube and to this the plasmid DNA was added. A maximum of 50 ng of DNA was added in a volume no greater than 10 µl per 100 µl of competent cells. After gently mixing, the transformation mixture was left on ice for 20-40 min. Next, the bacteria were heat shocked at 42°C for 2 min. before placing on ice for a remaining 2 min. An 800 µl volume of cold (4°C) LB media was added to the bacteria and the tube incubated at 37°C for 90 min. before plating out onto fresh LB media containing appropriate antibiotics.

2.5.4 Restriction Endonuclease Digestion.

Restriction endonucleases are enzymes that cleave both strands of double stranded DNA at specific nucleotide sequences. The digestion reaction was carried out using an excess of restriction enzyme in the supplier's recommended buffer system. The amount of restriction enzyme required is calculated in units, whereby 1 unit of enzyme cuts 1 µg of λ DNA in 1 h. at 37°C in a volume of 50 µl. Typically a reaction contained the following solutions:

DNA solution (typically 1 µg of DNA)	1 µl
Restriction enzyme buffer (x10)	1 µl
Analytical grade H_2O	7 µl
Restriction enzyme	1 µl

The components of the reaction were added in the above order into a MCC tube, the contents mixed thoroughly, centrifuged briefly and incubated at 37°C for 3-18 h. If larger quantities of DNA were digested the reaction could be scaled up, ensuring that the restriction enzyme buffer was present at a 1/10 overall concentration. Double enzyme digests were carried out using a suitable buffer, deemed by the supplier's guide, to provide suitable conditions for both

enzymes and the quantity of each restriction enzyme added was halved. Products of restriction enzyme digests were visualised using agarose gel electrophoresis and compared with the migration of DNA molecular size markers.

2.5.5 Alkaline phosphatase treatment of plasmid DNA.

Calf-intestinal alkaline phosphatase (CIAP) was used in this procedure, which functioned to remove the 5'-phosphate groups from linear DNAs. This process prevented the recircularization of plasmid DNA in gene cloning experiments. Cleaved plasmid DNA was ethanol precipitated and resuspended in 44 µl of analar water in a sterile MCC tube, to which 5 µl of CIAP buffer (X10) and 1 µl of CIAP was added. The contents of the tube were incubated at 37°C for 30 min. After the incubation time, a further 1 µl of CIAP was added to the tube and incubated for 30 min. at 37°C. The CIAP was inactivated by heating the contents of the tube to 75°C for 10 min. A phenol extraction was carried out to remove the alkaline phosphatase.

2.5.6 Ligation of Foreign DNA into Plasmid Vectors.

Plasmid DNA that has been cleaved with a restriction enzyme can be joined *in vitro* to foreign DNA in a ligation reaction. Successful ligation depended on the compatibility of the termini of the foreign DNA fragment and the restriction sites in the plasmid vector. Additionally, the relative concentration of plasmid and foreign DNA in the reaction was calculated based upon their relative nucleotide sizes. Unless otherwise stated, the ratio of plasmid DNA to foreign DNA was 2:1 respectively. The reaction was performed using T4 DNA ligase in the supplier's recommended buffer system. Ligations, unless otherwise stated were incubated at 15°C for 18 h. in an overall volume of 50 µl. Agarose gel electrophoresis comparing ligated DNA with unligated DNA solutions was used to confirm successful ligations.

In a reaction with blunt-ended termini, where both plasmid and vector in the ligation reaction have blunt ends with no overhanging bases, the efficiency of ligation tends to be reduced and so higher concentrations of DNA and ligase were used.

2.5.7 SURECLONE™ Ligation Kit (Pharmacia Biotech).

This commercial kit was used as an alternative method to clone PCR products quickly for DNA sequencing. The first step of the method involved blunt-ending the PCR product, which was carried out in a reaction as follows:

PCR Product	1 µg
Klenow fragment	1 µl
Buffer (10X)	2 µl
Polynucleotide kinase	1 µl
Analar H_2O	15 µl added to give a total volume of 20 µl.

The components were added into a sterile MCC tube, mixed gently and centrifuged briefly and incubated at 37°C for 30 min. After which the product was purified using the BIO-RAD Prep-A-Gene Kit (2.4.6.2) and resuspended in 10 µl of sterile analar water. The following ligation reaction was then set up:

DNA (PCR product)	2 µl
Vector (pUC18)	1 µl
2 X Ligation Buffer	10 µl
DTT	1 µl
T4 Ligase	1 µl
Sterile Analar Water	5 µl

This gave an overall total of 20 µl and the preparation was incubated at 15°C for 90 min. before using in a transformation procedure (2.5.1) using blue/white screening (2.5.2).

2.6 PROTEIN ANALYSIS METHODS.

2.6.1 SDS-PAGE.

Proteins were electrophoretically separated for analysis by SDS-PAGE using a discontinuous buffer system. The electrophoretic equipment used in this study was either the Mini-Protean II system (BIO-RAD, Herts., UK) or the Biometra system. The SDS-PAGE gel consisted of two gels; a main, separating gel and an upper stacking gel formed between two glass slides. Polymerisation of gels was carried out using the ammonium persulphate (APS)/TEMED system. Details of gel formulation and reagents used can be found in Appendix I.

2.6.1.1 Formation of SDS-PAGE Gels.

The two glass slides that formed the sandwich support for the formation of the SDS-PAGE gel were cleaned with 70% IMS, dried and assembled. The main gel was made at the desired concentration with the APS and TEMED added lastly. This mixture was pipetted into the space between the plates using a glass pipette and filled to about two-thirds the height of the largest glass plate. Once filled to the desired level, a small volume of water-saturated isobutanol was added, forming a layer on top of the main gel to exclude air, hence promoting gel polymerisation. A period of 30 min. was allowed for polymerisation of the main gel to occur. After this time, the water-saturated isobutanol was poured off and the top of the gel was washed with sterile analar water before the second, stacking gel was added. The stacking gel was made to the desired concentration, as with the main gel by adding APS and TEMED lastly and was pipetted onto the set, main gel. A comb was placed into the stacking gel prior to polymerisation so that the protein samples could be added to the wells formed by the comb. A further 30 min. was allowed for polymerisation of the stacking gel.

2.6.1.2 Electrophoresis of SDS-PAGE Gels.

Once set, the whole gel was transferred to the electrophoresis tank. Electrophoresis buffer was poured into the tank and the comb was removed. Any air bubbles at the bottom of the gel were removed carefully with a syringe. Protein samples were mixed with SDS-PAGE loading buffer and boiled for 5 min. before loading onto the gel. A maximum of 20 µl of sample could be loaded into each well. Coloured molecular weight markers (Sigma) were also boiled and loaded onto the gel. The gel was then run at around 120 V for approximately 1 h., or until the dye front had run off the bottom of the gel.

2.6.1.3 Staining of SDS-PAGE Gels.

Gels were stained with coomassie blue to visualise protein bands, or used in a Western blotting procedure for immunological detection of proteins. The gel was first removed from the electrophoresis tank and carefully separated from the glass slides and placed in coomassie blue staining solution for around 40 min. with gentle agitation before destaining. The gel was left in destaining solution with gentle agitation, whereby the destaining solution was changed several times over a period of 3-18 h.

2.6.2 Western Blotting.

Proteins on an SDS-PAGE gel could be transferred onto nitrocellulose for the purpose of immunological detection, which was carried out using the BIO-RAD Transblot system. A piece of nitrocellulose was cut to the size of the gel, along with 3 pieces of blotting paper and left in transfer buffer for 10 min. along with two foam pads. These were then arranged so that the foam pads were on the outside, a piece of blotting paper was placed onto one of the foam pads and the gel was carefully placed onto this followed by the nitrocellulose and a couple of pieces of blotting paper and the second foam pad. Any air bubbles on the nitrocellulose were rolled out using a clean glass rod. The gel was then electroblotted overnight at 30 mA at 4°C.

Once electroblotting was complete, the nitrocellulose membrane was washed in PBS for 5 min. to remove any methanol left from the transfer buffer, and incubated with blocking solution for 1 h. After this time, the nitrocellulose was removed and incubated with the desired concentration of primary antibody for 2 h. followed by washing in 50 ml of PBS for a total of 3 washes. Afterwards, the nitrocellulose was left incubating for 2 h. with the secondary antibody conjugate at the appropriate concentration to label bound antibody, which was, unless otherwise stated, peroxidase-conjugated swine anti-rabbit immunoglobulins, before finally washing with PBS 3 times. Labelled antibody was visualised by development using the 3'3' diaminobenzidine tetrahydrochlorate (DAB) system with nickel enhancement (Harlow and Lane, 1988), where a positive reaction was indicated by a purple/brown colour.

2.7 SOUTHERN BLOTTING.

Southern blotting allows the detection of homologous DNA sequences (Southern, 1975) that have been transferred from an agarose gel to a membrane support, through capillary action and, using a nucleic acid probe that hybridises to the transferred DNA, and depending on the type of label, hybridisation bands of interest can then be detected. Sodium saline citrate (SSC) buffer is used, in which nucleic acids are highly soluble, and it is drawn up through the gel taking with it single stranded DNA which becomes trapped in the membrane.

2.7.1 Digoxigenin Non-radioactive Labelling of DNA.

Southern blotting was carried out using a non-radioactive nucleic acid detection system, as opposed to the alternative, radioactively labelled DNA detection system. The system used was the digoxigenin-anti-digoxigenin system, which uses DIG (digoxigenin) (Boehringer Mannheim), a cardenolide steroid isolated from Digitalis plants (Martin *et al.*, 1990). A nucleotide triphosphate analogue containing the digoxigenin moiety (DIG-UTP) is incorporated into DNA through random primer labelling in a PCR reaction. The DIG-labelled

probe is then detected using an enzyme-linked immunoassay using an antibody to digoxigenin (anti-DIG), to which alkaline phosphatase has been conjugated. A chromogenic substrate for alkaline phosphatase was used to detect the DIG-labelled probe, the 5-bromo-4-chloro-3-indolyl phosphate (BCIP) with nitroblue tetraazolium chloride (NBT) system. Unless otherwise stated a 5 µl volume of DIG-UTP was incorporated into the standard PCR reaction designed for the probe.

2.7.2 Southern Blotting Procedure.

The DNA samples to be used were firstly digested with the appropriate restriction enzymes (2.5.4) and separated by agarose gel electrophoresis. Specific markers were run on the gel, which were composed of a size standard of λ DNA cut with *Hind*III (Gibco/BRL). Once agarose gel electrophoresis of the DNA samples was complete and a photograph was taken, the wells were removed and the gel was placed into 0.2 M HCl for 10 min. as a depurination step. The gel was removed and placed in a separate container along with 100 ml of Denaturation Solution and left to soak for 20 min. Finally, the gel was placed into a 200 ml of Neutralisation Solution. A piece of nylon membrane was cut to the size of the agarose gel along with 2-3 pieces of filter paper. Numerous paper towels were also cut to this size for use in the blotting procedure. The blotting unit was prepared by firstly washing with water. The unit was then left to soak in sterile water before priming the wick by adding 25-50 ml of 10X SSC solution over the surface, and 450 ml of 10X SSC was added to the blotting base. The gel was placed, facedown, onto the centre of the wick and cling film was used to cover the remaining surface of the wick not covered by the gel to prevent a "short-circuit" during transfer. The membrane and filter paper were soaked in water for 2-3 min. followed by 10X SSC for 5 min. Next, the membrane was placed onto the gel, followed by the filter paper and a stack of paper towels were placed on top of the filter paper high enough so that the lid would sit above the blotting base. This was left overnight to facilitate the transfer of DNA onto the

nylon membrane. After this time, the towels and filter paper were removed and the membrane was washed by immersion in 5X SSC for 1-5 min. to remove any residual agarose and left to air dry for around 30 min. before binding the nucleotides to the membrane by UV crosslinking. This was achieved by placing the membrane on a transilluminator and subjecting it to UV light of 254 nm for two and a half min. The membrane was removed and prehybridised with 30 ml of Prehybridisation Solution at 68°C for 2 h. During this time, the probe was denatured by boiling for 10 min. in a waterbath before placing directly onto ice. The Prehybidisation Solution was replaced with Hybridisation Solution containing around 100 ng of the freshly denatured probe, and was left to hybridise overnight. Next, the membrane was washed twice with 50 ml of Wash Buffer I for 5 min. at room temperature and a further 2 times with 50 ml of Wash buffer II for 15 min. at 68°C. The detection procedure was started by washing the membrane for 1 min. in Detection Buffer I, followed by 50 ml of Detection Buffer II for 30 min. and finally the membrane was washed again in 50 ml of Detection Buffer I for 2 min. The antibody conjugate (Anti-Digoxigenin-AP) was diluted to 150 mU/ml, a 1 in 5000 dilution. Thus, 5 µl of the conjugate was added to 25 ml of Detection Buffer I and the membrane was incubated in this before adding the Coloured Detection Solution. The reaction was stopped by adding Detection Buffer III.

2.8 POLYMERASE CHAIN REACTION (PCR).

PCR is the *in vitro* amplification of specific sequences of nucleic acid. It allows an exponential increase in the amount of any given DNA template by the simultaneous primer extension of the two complementary strands of DNA. This is achieved by the use of synthetic oligonucleotides, termed primers that are complementary to sequences within the target DNA and flank the piece of DNA to be amplified. DNA polymerase then carries out the synthesis of a complementary strand of DNA.

2.8.1 Primer Design.

One primer is designed to be complementary to the 3' end of the antisense strand of DNA (forward primer) and the other is designed to be complementary to the 3' end of the sense strand of DNA (reverse primer). Generally, PCR generates shorter DNA sequences (200-400 nucleotides in length) more reliably than longer sequences, and this was taken into account when designing primers to the target DNA. Ideally, primers were designed so that they had a 50% G+C content, or a G+C content similar to the fragment to be amplified, and so that they would not contain extensive secondary structure nor be complementary to each other. Design of PCR primers was carried out using the DNasis computer program (Hitachi) and the resulting choices of primers were checked by the computer program Amplify to identify the best suited forward and reverse primers. Primers were synthesised by Sigma-Genosys Ltd. (Cambridgeshire, UK).

2.8.2 Storage of Primers.

Primers were received lyophilised, and resuspended in 500 µl of 1X TE buffer, boiled in a waterbath for 10 min. and placed directly onto ice to ensure that the primers were fully denatured. The concentrations of the primers were calculated, and the primer solutions were further diluted with sterile analar water to give stock concentrations of around 20 pmol/µl. The primer stocks were stored in sterile, thin-walled PCR tubes in 50 µl aliquots at -80°C.

2.8.3 The PCR Reaction.

The basic steps of the PCR were as follows, although the time and temperature of the annealing step sometimes differed, depending on the melting point of the primers and often it was necessary to carry out a number of PCRs to ascertain the ideal reaction conditions:

1.	Denaturation of templated DNA	95°C, 1 min.
2.	Annealing primers to templated DNA	45-65°C, 1-2 min.
3.	Extension of primers by DNA polymerase	72°C, 2 min.

Steps 1-3 were repeated 30 times, unless otherwise stated, followed by 10 min. at 72°C to allow filling-in of the fragments. The reaction conditions were programmed into the thermocycler, which was, unless otherwise stated, a Perkin Elmer TC1.

The overall volume of the PCR reaction was 50 µl and the reaction components generally consisted of the following, although exact quantities differed slightly:

Reaction Buffer (x10)	5 µl
Forward primer	10-100 pmol
Reverse primer	10-100 pmol
dNTPs	2 mM
Taq DNA polymerase	2 units
Template DNA	2-50 ng
Analar H_2O	Added to bring the final volume of the reaction to 50 µl.
Mineral oil	50 µl

Components were added into 0.5 ml thin-walled PCR tubes and throughout the procedure all components and tubes were kept on ice. Mineral oil was the final component to be added to form a layer above the reaction mixture to minimise any evaporation during the heating and cooling cycles. The tubes were then placed into the thermocycler, previously programmed to carry out the desired steps of the PCR cycle automatically.

The purity of DNA was crucial, as any DNA contaminants would be amplified if the primers were able to anneal to a part of the sequence. Therefore, to avoid nuclease contamination, gloves were worn throughout the procedure, which was carried out in a specialised PCR cabinet where all equipment was first sterilised under UV light prior to use. In addition to these precautions, micropipettors were used with tips that contained a filter to prevent contamination, and all solutions and equipment used for PCR were kept separate from other laboratory uses. In most cases, once the PCR reaction was complete, a one-tenth volume of the reaction mixture was run on a 1% agarose gel to examine the products produced and to estimate their relative sizes.

2.9 DNA SEQUENCING.

DNA sequencing in this project was carried out either by Dr. R. Titball at DERA, Porton Down, Salisbury or commercially by MWG Biotech Ltd.

2.9.1 Sequencing Methods.

The sequencing method utilised in both cases are variations on the chain terminator, or dideoxy procedure, which is now favoured for large-scale DNA sequence determination and considered to be more simple, rapid and accurate. The chain terminator DNA sequence procedure was developed by Sanger *et al.* (1977) and basically involves the synthesis of a DNA strand by a DNA polymerase *in vitro* using a single-stranded template, which capitalises on two properties of DNA polymerase:

1. Its ability to synthesise faithfully a complementary copy of a single stranded DNA template.
2. Its ability to use 2'3'-dideoxynucleoside triphosphates (ddNTP) as substrates (Old and Primrose, 1989).

Dideoxynucleotides lack the 3' hydroxyl group necessary for the formation of the next phosphodiester bond during chain elongation, and so is therefore terminated whenever a dideoxynucleotide is incorporated into the strand. Initiation of DNA synthesis requires a primer and usually this is a chemically synthesised oligonucleotide. DNA synthesis is then carried out in the presence of the four dideoxynucleoside triphosphates (ddNTPs) and deoxynucleotides (dNTPs) in four separate incubation mixes containing a low concentration of one of the ddNTP analogues so that a ddNTP is incorporated into the growing chain rather than a dNTP. This means that in each reaction there is a population of partially synthesised DNA molecules, each having a common 5' end, but each varying in length to a base specific 3' end. The Klenow fragment of DNA polymerase I lacks the 5'→3' exonuclease activity of the intact enzyme, and was therefore used in the original protocol to synthesise a complementary copy of the single-stranded target sequence, and keeps the 5' end of the primer intact. Sequencing methods now make use of thermostable DNA polymerases, which, unlike the previous DNA polymerases allows the template DNA to be used many times in a single reaction. The four different reaction mixtures are separated by size on a polyacrylamide gel and the sequence is determined by correlating the order of the bands on the gel with the dideoxynucleotide used to generate each band. Labels, which can be radiolabels or dyes, are used for the detection of the DNA synthesised in a sequencing reaction and can be incorporated into the primer, the deoxynucleotides or dideoxynucleotides.

2.9.2 Automated Sequencing.

The development of instruments for automated sequencing has dramatically increased the throughput of individual laboratories. Unlike manual sequencing, which generally uses a radioactive label and visualisation of the banding pattern is by autoradiography the sequencing carried out in this project was by an automated sequencer. Automated sequencers use a scanning laser to detect DNA fragments labelled with fluorescent dyes. The health risks and

inconvenience of radioisotope disposal are sufficient incentive for many laboratories to automate much or all of their DNA sequence analysis. The fluorescent dyes can be linked to the primer (dye primer), the deoxynucleotides (internal labelling) or the dideoxynucleotides (dye terminator). The products are loaded onto a sequencing gel and separated according to size. A laser beam continuously scans across the lower portion of the gel, which functions to excite the fluorescent dyes attached to the fragments, and they emit light at a specific wavelength for each dye. The light is collected and separated according to wavelength by a spectrograph, where data collection software collects the light intensities and stores them as electrical signals for eventual processing.

2.9.3 Template Preparation for DNA Sequencing by MWG-BIOTECH Ltd.

To achieve optimum results, DNA cloned into a plasmid to be sent for sequencing, was cultivated in an *end*A$^-$ *E.coli* strain and in the case of this research project, this was always carried out in DH5α. The DNA purification methods used to prepare all DNA samples was either the CsCl-EtBr density centrifugation or by using a Qiagen plasmid kit. The following amounts were sent according to the size of the insert to be sequenced and the service required:

Single sequencing reaction		10 μg
Complete double strand: Insert length	<5 kbp	30 μg
	<10 kbp	50 μg
	>10 kbp	100 μg

All samples were sent as a freeze-dried pellet of the appropriate quantity of DNA.

CHAPTER 3

COMPARISON OF THE EXTRACELLULAR PRODUCTS OF *B.pseudomallei* AND *B.thailandensis.*

3.1 INTRODUCTION.

Many bacterial pathogens produce extracellular products that are essential for pathogenicity. This may also be true of *B.pseudomallei,* as its virulence appears to be partly reliant on the production of a number of extracellular products during the course of infection. For example, of those already identified, there exists a thermolabile exotoxin (Nigg *et al.*, 1955), a cytolethal toxin (Haase *et al.*, 1997) a lecthinase, lipase, protease (Ashdown and Koehler, 1990), a haemolytic product (Liu, 1957) and a phospholipase (Korbsrisate *et al.*, 1999). However, there is little evidence in current literature to show that an extensive study on all of the extracellular products (ECPs) produced by *B.pseudomallei* has been carried out. Therefore, it was decided to investigate the ECPs produced by a range of isolates of the bacterium. In addition to this, in light of the recent identification of the species *B.thailandensis*, there has not yet been a direct comparison of the ECPs produced by the two different species. Considering that the two species are antigenically indistinguishable and share the same LPS profile, but differ with respect to their virulence, there may exist certain factor(s) produced by *B.pseudomallei* responsible for this, of which some may be extracellularly produced. Thus, any differences observed between their respective ECP profiles could be representative of potential virulence factors in *B.pseudomallei.* There is also the possibility of variation between isolates of *B.pseudomallei* obtained from infected humans and isolates obtained from the environment, which again have not been compared with respect to the ECPs produced. Any differences

observed between such isolates may indicate virulence factors that are elicited or essential during the course of infection.

In addition to ECP SDS-PAGE profiles, there is little published work concerning direct testing of *B.pseudomallei* isolates, in particular the ECPs, for a range of activities, such as haemolysis. Therefore, microtitre plate assays to detect a range of activities were performed on *B.pseudomallei* and *B.thailandensis* isolates. It was hoped that this would highlight activities of *B.pseudomallei* that had yet to be discovered and represent potential targets for any future gene cloning work with the pathogen.

3.2 Isolation and Concentration of Extracellular Products.

3.2.1 Bacterial Strains.

The isolates of *B.pseudomallei* and *B.thailandensis* used in this study are shown in Chapter 2 (Table 2.1). Other *Burkholderia* species were also used for comparative purposes in this study (Table 2.2).

3.2.2 Isolation and Concentration of ECPs.

Ten ml volumes of cultures were centrifuged at 4,400 x g in an IEC centrifuge at 4°C for 20 min. The resulting supernatant was aliquoted into MCC tubes and centrifuged at 15, 000 x g at 4°C for 10 min. The supernatants were carefully removed and subjected to concentration in AMICON MINICON mini-plus units to X 250. For most isolates, 15 µl of concentrated supernatant was added to 15 µl of SDS-PAGE loading buffer. Exceptions, due to less bacterial growth, were as follows; *B.cepacia* genomovars IIIb, II and III non-epi, and *B.vietnamensis* LMG 6998, 30 µl of concentrated supernatant was added to 30 µl of loading buffer and for

B.cepacia genomovars IIIc, IIIa epi and IV, and *B.vietnamensis* LMG 6999, 40 µl of concentrated supernatant was added to 40 µl of loading buffer. Samples were boiled in the CL3 laboratory for 10 min. and stored at -70°C until use. Viability checks were carried out before use outside the CL3 laboratory.

3.3 SDS-PAGE ANALYSIS.

3.3.1 SDS-PAGE Analysis to Test the Concentration of the Isolated ECPs.

A representative group of isolates were chosen to test on SDS-PAGE (2.6.1) to ascertain the respective concentrations of the ECPs isolated. The isolates tested were *B.pseudomallei* 19, *B.thailandensis* E82, *B.cepacia* II and *B.vietnamensis* LMG 6999. A range of volumes were used as follows: 0.25 µl, 0.5 µl, 1 µl, 2 µl and 5 µl of concentrated supernatant mixed with more loading buffer to bring the final volume up to 20 µl. Samples were placed in a boiling waterbath for 5 min. and subjected to electrophoresis in 13% polyacrylamide gels and the resulting protein bands were visualised with coomassie blue staining. The results showed that there was a very low concentration of ECPs isolated and even when using the largest volume of concentrated supernatant (5µl), protein bands were just visible (results not shown).

3.3.2 SDS-PAGE Analysis of Concentrated ECPs Isolated from *B.pseudomallei* and *B.thailandensis* Isolates.

Profiles of the ECPs from all the isolates of *B.pseudomallei* and *B.thailandensis* were obtained using SDS-PAGE analysis (2.6.1). The volume of concentrated supernatant used was determined by the test performed (3.3.1), but as the concentration of the proteins was found to be very low, a maximum 15 µl of concentrated supernatant was mixed with 5 µl of loading buffer and boiled for 10 min. Samples were subjected to electrophoresis in 13%

polyacrylamide gels in conjunction with coloured molecular weight markers (Sigma). Protein bands were visualised by staining with coomassie blue.

The results of all the SDS-PAGE gels can be seen in Appendix III (Gels 1- 6), as some of the concentrations of ECPs were so low that protein profiles could not be distinguished. In addition to this, some samples were difficult to load, probably due to high carbohydrate content. Therefore, a representative group of *B.pseudomallei* and *B.thailandensis* ECP profiles can be seen in Fig.3.1. Relative molecular weights of the protein bands were calculated by measuring the running distance of the molecular weight markers from the top of the gel and plotting the logged values against their respective logged molecular weights, which was carried out for each individual gel. Protein bands from the samples were measured from the same point as the markers and the logged value used to calculate the relative molecular weight graphically. Results of this can be seen in Table 3.1.

An example of a SDS-PAGE gel showing whole-cell profiles of *B.pseudomallei* isolates, carried out by Mr M. Mahfouz, has been included in Fig.3.2 for comparative purposes and to determine whether any of the ECP samples had been inadvertently contaminated with whole-cell proteins. It seemed unlikely that an overspill of whole-cell proteins had occurred, as the pattern of protein bands observed from whole-cells were not evident in the ECP protein profiles.

3.4 WESTERN BLOTTING OF ECPs.

The ECPs of a representative number of isolates from both *B.pseudomallei* (53, H706, E8, 448, 576 and 217) and *B.thailandensis* (E255 and E260) were analysed by immunoblotting.

The *B.pseudomallei* isolates chosen represented both clinical and environmental isolates. The ECPs were separated by SDS-PAGE (2.6.1) before being electrotransferred onto a nitrocellulose membrane. Immunoblotting was carried out using convalescent human serum as the primary antibody (2.6.2). This was kindly donated by Dr. T. Pitt, CPHL, Colindale. Little background information was available concerning the human sera, including its relative concentration and so it was used experimentally at a concentration of 1:200. Goat anti-human immunoglobulins (Sigma) with peroxide conjugate were used as the secondary at a concentration of 1:400. Resulting bound antibody was visualised using DAB/$NiCl_2$ colour development. The results of the immunoblotting can be seen in Fig.3.3. The ECPs are somewhat overshadowed by the immune response to what is considered to be polysaccharide or LPS, which is represented by the darker coloured "ladder" of bands in each lane, which were equally as strong in both species. The exceptions to this were *B.pseudomallei* strains 576 and 217 where there was no evidence of the ladder in the lanes and so major protein bands were easily observed. However, with the use of a light box, major ECPs could be identified in the other isolates and their relative sizes estimated as with the SDS-PAGE gels (3.3.2). This was carried out firstly for *B.pseudomallei* 576 as these bands were the clearest and the other isolates were compared to this.

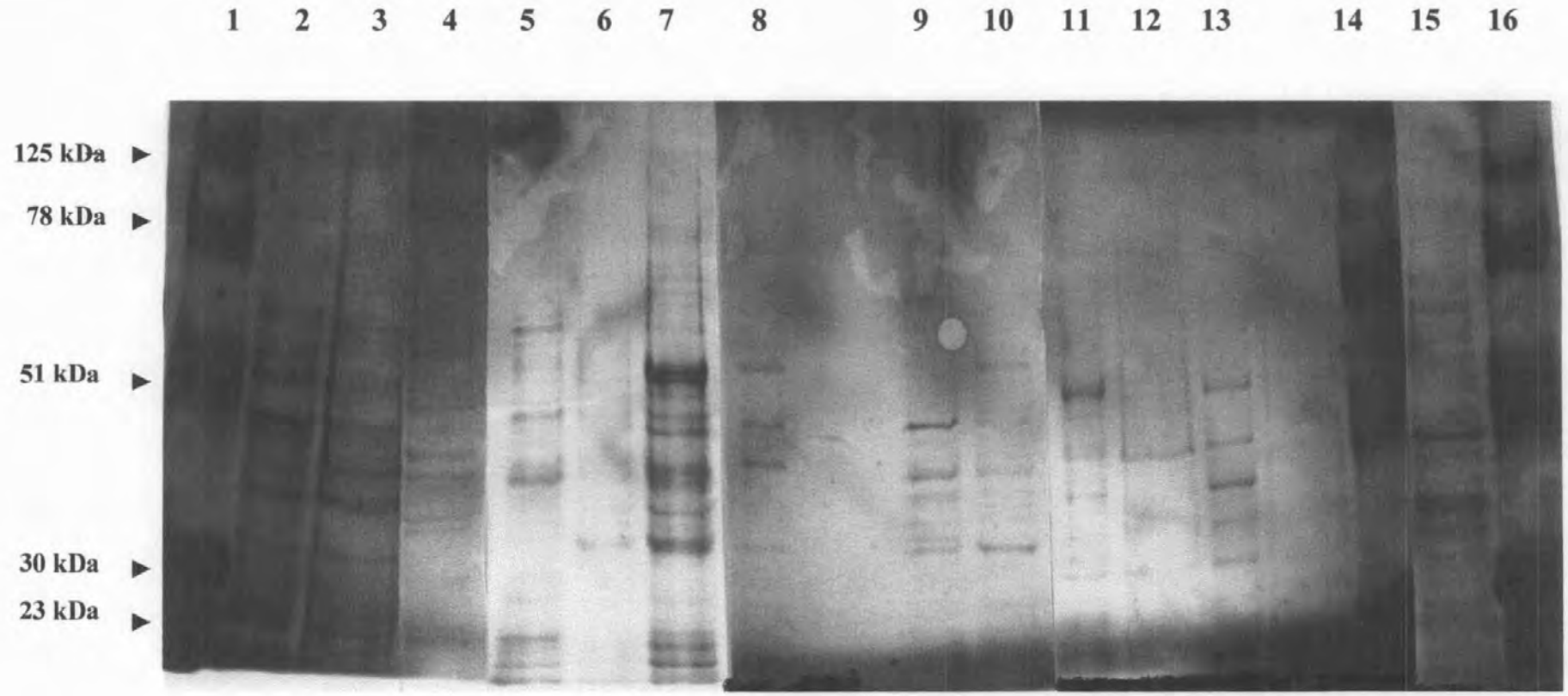


Fig. 3.1Concentrated ECPs from a Representative Group of *B.pseudomallei* and *B.thailandensis* Isolates. LANE 1: Coloured Molecular Weight Markers (Sigma) (from top to bottom): 205 kDa, 125 kDa, 78 kDa, 51 kDa, 30 kDa, 23 kDa, 16.5 kDa. LANE 2: 448 (Gel 3). LANE 3: 576 (Gel 3). LANE 4: 217 (Gel 3). LANE 5: 25 (Gel 1). LANE 6: 46 (Gel 1). LANE 7: 53 (Gel 1). LANE 8: E27 (Gel 4). LANE 9: JIE187 (Gel 4). LANE 10: H706 (Gel 4). LANE 11: E260 (Gel 5). LANE 12: E256 (Gel 5). LANE 13: E255 (Gel 5). LANE 14: Coloured MolecularWeightMarkers (Sigma). LANE 15: E8 (Gel 2). LANE 16: Coloured Molecular Weight Markers (Sigma).

Table 3.1 Relative Molecular Weights Calculated for Protein Bands of *B.pseudomallei* and *B.thailandensis* Isolates Visualised by SDS-PAGE Analysis .

SPECIES/ISOLATE	LOG DISTANCE (mm)	LOG MWt (kDa)	RELATIVE MWt (kDa)
***B.pseudomallei* 448** (Gel 3)	1.04 1.18 1.29 1.48 1.53 1.56 1.57 1.61 1.63 1.66 1.69 1.7 1.74 1.78 1.79 1.8 1.81 1.83	2.18 2.15 2.02 1.84 1.78 1.76 1.75 1.7 1.68 1.65 1.6 1.59 1.56 1.48 1.47 1.46 1.44 1.41	151.4 141.3 104.7 69.2 60.3 57.5 56.2 50.2 47.9 44.7 39.8 38.9 36.3 30.2 29.5 28.8 27.5 25.7
***B.pseudomallei* 576** (Gel 3)	1.39 1.43 1.46 1.48 1.51 1.53 1.56 1.57 1.63 1.66 1.69 1.7 1.74 1.77 1.79 1.8 1.81 1.83	1.92 1.89 1.86 1.84 1.8 1.78 1.76 1.75 1.68 1.65 1.6 1.59 1.56 1.5 1.47 1.46 1.44 1.41	83.2 77.6 72.4 69.2 63.8 60.3 57.5 56.2 47.9 44.7 39.8 38.9 36.3 31.6 29.5 28.8 27.5 25.7
***B.pseudomallei* E8** (Gel 3)	1.28 1.31 1.34 1.43 1.48 1.51 1.52 1.56 1.59 1.62 1.64 1.66 1.67 1.68 1.69 1.71 1.72 1.74 1.75 1.78 1.79 1.81 1.82	2.2 1.99 1.97 1.89 1.84 1.805 1.795 1.76 1.72 1.68 1.66 1.65 1.63 1.62 1.6 1.58 1.57 1.54 1.53 1.48 1.47 1.4451.42	158.5 98.9 93.3 77.6 69.2 63.8 62.4 57.5 52.5 47.9 45.7 44.7 42.7 41.7 39.8 38.0 37.2 34.7 33.9 30.2 29.5 27.9 26.3
***B.pseudomallei* 25** (Gel 1)	1.34 1.39 1.42 1.43 1.47 1.5 1.51 1.525 1.57 1.58 1.59 1.63 1.65 1.66 1.67 1.69 1.7 1.78 1.79 1.8 1.81	1.94 1.87 1.85 1.84 1.86 1.78 1.77 1.75 1.705 1.695 1.68 1.65 1.62 1.6 1.59 1.56 1.55 1.42 1.41 1.38 1.36	87.1 74.1 70.8 69.2 63.2 60.3 58.9 56.2 50.7 49.5 47.9 44.7 41.7 39.8 36.3 35.5 26.3 25.7 24.0 22.9
***B.pseudomallei* 46** (Gel 1)	1.13 1.23 1.34 1.39 1.43 1.48 1.5 1.51 1.54 1.55 1.57 1.59 1.62 1.64 1.66 1.67 1.69 1.7 1.71 1.75 1.78 1.79 1.8 1.81	2.1 2.02 1.94 1.87 1.84 1.795 1.78 1.77 1.74 1.73 1.705 1.68 1.65 1.63 1.6 1.59 1.56 1.55 1.54 1.48 1.42 1.41 1.38 1.36	125.9 104.7 87.1 74.1 69.2 62.4 60.3 58.9 55.0 53.7 50.7 47.9 44.7 42.7 39.8 38.9 36.3 35.5 34.7 30.2 26.3 25.7 24.0 22.9
***B.pseudomallei* 53** (Gel 1)	0.7 0.93 1.08 1.13 1.23 1.27 1.32 1.34 1.37 1.39 1.43 1.47 1.48 1.5 1.51 1.525 1.54 1.57 1.59 1.61 1.63 1.65 1.67 1.69 1.7 1.71 1.72 1.74 1.75 1.76 1.78 1.79 1.8 1.81	2.25 2.195 2.12 2.1 2.02 1.98 1.945 1.94 1.905 1.87 1.84 1.805 1.795 1.78 1.77 1.75 1.74 1.705 1.68 1.66 1.65 1.62 1.59 1.56 1.55 1.54 1.52 1.49 1.48 1.46 1.42 1.41 1.38 1.36	177.8 156.7 131.8 125.9 104.7 95.5 88.1 87.1 80.4 74.1 69.2 63.8 62.4 60.3 58.9 56.2 55.0 50.7 47.9 45.7 44.7 41.7 38.9 36.3 35.5 34.7 33.1 30.9 30.2 28.8 26.3 25.7 24.0 22.9
***B.thailandensis* E27** (Gel 4)	1.5 1.57 1.62 1.71	1.87 1.8 1.74 1.62	74.1 63.1 55.0 41.7
***B.pseudomallei* JIE187** (Gel 4)	1.4 1.45 1.56 1.57 1.6 1.61 1.63 1.65 1.66 1.69 1.695 1.7 1.74	1.96 1.905 1.805 1.8 1.77 1.76 1.74 1.7 1.695 1.65 1.64 1.635 1.58	91.2 80.4 63.8 63.1 58.9 57.5 55.0 50.7 49.5 44.7 43.7 43.2 38.0
***B.pseudomallei* H706** (Gel 4)	1.4 1.45 1.5 1.56 1.57 1.6 1.61 1.63 1.65 1.66 1.69 1.695 1.7 1.76 1.8	1.96 1.905 1.87 1.805 1.8 1.77 1.76 1.74 1.7051.695 1.65 1.64 1.635 1.55 1.48	91.2 80.4 74.1 63.8 63.1 58.9 57.5 55.0 50.7 49.5 44.7 43.7 43.2 35.5 30.2
***B.thailandensis* E260** (Gel 5)	1.15 1.23 1.32 1.42 1.49 1.52 1.56 1.57 1.61 1.63 1.65 1.66 1.68 1.69 1.72 1.74 1.76 1.765	2.03 2.005 1.945 1.87 1.805 1.78 1.74 1.72 1.68 1.65 1.62 1.61 1.58 1.56 1.52 1.48 1.43 1.42	107.2 101.2 88.1 74.1 63.4 60.3 55.0 52.5 47.9 44.7 41.7 40.7 38.0 36.3 33.1 30.2 26.9 26.3
***B.thailandensis* E256** (Gel 5)	1.57 1.65 1.69	1.72 1.62 1.56	52.5 41.7 36.3
***B.thailandensis* E255** (Gel 5)	1.23 1.32 1.42 1.48 1.49 1.57 1.61 1.65 1.69 1.72	2.005 1.945 1.87 1.82 1.805 1.72 1.68 1.62 1.56 1.52	101.2 88.1 74.1 66.1 63.4 52.5 47.9 41.7 36.3 33.1
***B.pseudomallei* 217** (Gel 2)	1.5 1.57 1.63 1.65 1.69 1.71 1.75 1.785 1.79	1.78 1.7 1.65 1.62 1.56 1.54 1.48 1.42 1.41	60.3 50.2 44.7 41.7 36.3 34.7 30.2 26.3 25.7

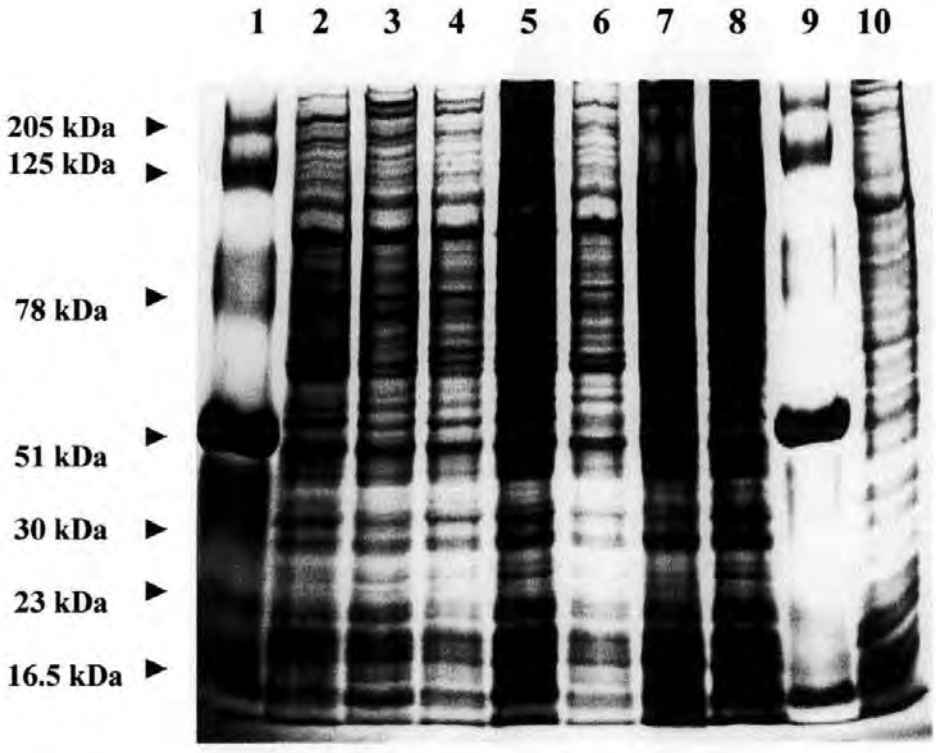


Fig.3.2 SDS-PAGE Analysis of Whole-Cell Proteins from Isolates of *B.pseudomallei* and *B.thailandensis*. LANE 1: Coloured Molecular Weight Markers (Sigma). LANE 2: E8. LANE 3: 19. LANE 4: 22. LANE 5: 25. LANE 6: E25. LANE 7: 35. LANE 8: 46. LANE 9: Coloured Molecular Weight Markers (Sigma). LANE 10: E260.

The relative molecular weights calculated from the individual gels were analysed for protein bands common to the two different species. The general pattern appeared to be that although isolates of both species shared a number of common bands, there were differences in the SDS-PAGE ECP profiles between species, with the *B.thailandensis* isolates generally having fewer protein bands than *B.pseudomallei* isolates. *B.thailandensis* isolates E260, E256 and E255 were run on the same gel and obvious similarities can be seen. Similar protein bands were then looked for in the *B.thailandensis* isolate E27 and all the molecular weight values calculated for each *B.thailandensis* isolate were compared to E260, which had the clearest, and most resolved protein bands present. The results of which can be seen in Table 3.2.

Table 3.2 Protein Bands Common to *B.thailandensis* Isolates E260, E256, E255 and E27.

Protein Bands of E260 (kDa)	**E256**	**E255**	**E27**
107.2			
101.2		*	
88.1		*	
74.1		*	*
63.4		*	(63.1)
60.3			
55.0			*
52.5	*	*	
47.9		*	
44.7			
41.7	*	*	*
40.7			
38.0			
36.3	*	*	
33.1		*	
30.2			
26.9			
26.3			

*Asterisks denote identical molecular weight values.

From Table 3.2, the isolates from this species clearly share a number of common protein bands, and those bands common to two or more of the *B.thailandensis* isolates tested are as follows: 101.2, 88.1, 74.1, 63.4 (63.1), 55.0, 52.5, 47.9, 41.7, 36.3 and 33.1 kDa.

The *B.pseudomallei* isolates similarly share a number of common protein bands. Although isolates were run on different gels all molecular weight values calculated for the *B.pseudomallei* isolates were compared to *B.pseudomallei* 53 (Table 3.3), which contained the clearest and most resolved protein profile.

Table 3.3 Protein Bands Common to *B.pseudomallei* Isolates 53, JIE187, H706, 576, 46, 448, 25, E8 and 217.

	Clinical				Soil			
Protein Bands of 53 (kDa)	**JIE187**	**H706**	**576**	**46**	**448**	**25**	**E8**	**217**
177.8								
156.7								
131.8								
125.9				*				
104.7				*	*			
95.5								
88.1								
87.1				*		*		
80.4	*	*						
74.1		*		*		*		
69.2			*	*	*	*	*	
63.8	*	*	*			(63.2)	*	
62.4				*			*	
60.3			*	*	*	*		*
58.9	*	*		*		*		
56.2			*		*	*		
55.0	*	*		*				
50.7	*	*		*	(50.2)	*		(50.2)
47.9			*	*	*	*	*	
45.7							*	
44.7	*	*	*	*	*	*	*	*
41.7						*	*	*
38.9	(38.0)		*	*	*		(38.0)	
36.3			*	*	*	*		*
35.5		*		*		*		
34.7				*			*	*
33.1							(33.9)	
30.9								
30.2		*		*	*		*	*
28.8			*		*			
26.3				*		*	*	*
25.7			*	*	*	*		*
24.0				*		*		
22.9				*		*		

*Asterisks denote identical molecular weight values.

Those bands common to two or more isolates of *B.pseudomallei* include the following: 125.9, 104.7, 87.1, 80.4, 74.1, 69.2, 63.8 (63.2), 62.4 (63.1), 60.3, 58.9, 56.2, 55.0, 56.7, 47.9, 45.7, 44.7, 41.7, 38.9 (38.0), 36.3, 35.5, 34.7, 30.2, 28.8, 26.5, 25.7, 24.0 and 22.9 kDa. Those values in brackets are of a similar molecular weight and common to a number of isolates. In addition to these, a number of other bands are found in more than one of the *B.pseudomallei* isolates but not found in 53, which are 39.8 and 57.5 kDa. The following common protein bands are found in both species: 88.1, 74.1, 55.0, 47.9, 41.7, 36.3 and 33.1 kDa. Virtually all of the protein bands that are common to the *B.thailandensis* isolates tested are also found in *B.pseudomallei* isolates. However, there are many other differing bands both between species and within species.

Dividing the *B.pseudomallei* isolates into clinical and environmental, or soil isolates facilitated the discovery of other similarities, as the protein bands of 125.9 and 80.4 kDa were found only in clinical isolates and so may represent virulence factors elicited on infection.

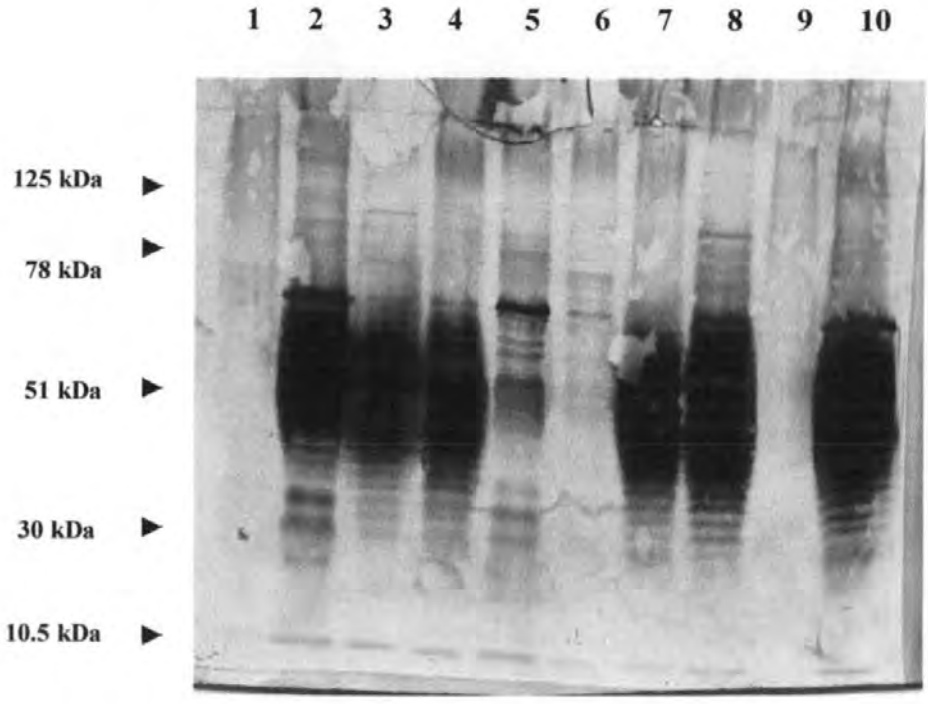


Fig 3.3 Western Blot of ECPs Isolated from *B.pseudomallei* and *B.thailandensis*. LANE 1: Coloured Molecular Weight Markers (Sigma). LANE 2: 53. LANE 3: H706. LANE 4: 448. LANE 5: 576. LANE 6: 217. LANE 7: E8. LANE 8: E255. LANE 9: Coloured Molecular Weight Markers (Sigma). LANE 10: E26.

Table 3.4 Relative Molecular Weights of Extracellular Protein Bands from a number of *B.pseudomallei* and *B.thailandensis* Isolates Visualised by Western Blotting.

Species/Isolate	Log Distance (mm)	Log MWt. (kDa)	Relative MWt. (kDa)
***B.pseudomallei* 53**	0.6 0.7 0.78 0.81 0.85 0.93 0.95 0.98 1.02 1.08 1.11 1.19 1.31 1.62 1.65 1.69 1.74	2.22 2.2 2.18 2.17 2.15 2.12 2.1 2.06 2.08 2.03 2.005 1.96 1.87 1.52 1.47 1.38 1.2	166.0 158.5 151.4 147.9 141.3 131.8 125.9 120.2 114.8 107.2 101.2 91.2 74.1 33.1 29.5 24.0 15.9
***B.pseudomallei* H706**	1.02 1.19 1.23 1.27 1.31 1.38 1.41 1.45 1.53 1.5	2.06 1.96 1.93 1.895 1.87 1.86 1.8 1.76 1.705 1.68 1.65	114.8 91.2 85.1 78.5 74.1 72.4 63.1 57.5 50.7 44.7 45.7
***B.pseudomallei* 448**	1.08 1.29 1.31 1.35 1.38 1.45	2.03 1.88 1.87 1.82 1.8 1.705	107.2 75.9 74.1 66.1 63.1 50.7
***B.pseudomallei* 576**	0.95 0.98 1.02 1.08 1.11 1.15 1.16 1.19 1.23 1.31 1.32 1.35 1.38 1.41 1.45 1.50 1.53 1.55 1.60 1.62 1.65 1.69 1.71	2.1 2.08 2.06 2.03 2.005 1.99 1.98 1.96 1.93 1.87 1.86 1.82 1.80 1.76 1.70 1.68 1.65 1.62 1.56 1.52 1.47 1.38 1.305	125.9 120.2 114.8 107.2 101.2 97.7 95.5 91.2 85.1 74.1 72.4 66.1 63.1 57.5 50.7 45.7 44.7 41.7 36.3 33.1 29.5 24.0 20.2
***B.pseudomallei* 217**	1.02 1.11 1.15 1.19 1.23 1.27 1.31 1.41 1.45 1.5	2.06 2.0051.99 1.96 1.93 1.895 1.87 1.76 1.705 1.68	114.8 101.2 97.7 91.2 85.1 78.5 74.1 57.5 50.7 45.7
***B.thailandensis* E255**	0.6 0.81 0.93 1.02 1.06 1.1 1.15 1.16 1.19 1.22 1.31 1.35 1.4 1.45 1.51 1.56 1.65	2.22 2.18 2.12 2.06 2.04 2.02 1.99 1.98 1.96 1.94 1.87 1.82 1.77 1.72 1.67 1.605 1.47	166.0 151.4 131.8 114.8 109.6 104.7 97.7 95.5 91.2 87.1 74.1 66.1 58.9 52.5 46.8 40.3 29.5
***B.thailandensis* E260**	0.6 0.93 1.02 1.1 1.15 1.31 1.65	2.22 2.12 2.06 2.02 1.99 1.87 1.47	166.0 131.8 114.8 104.7 97.7 74.1 29.5

The molecular weight values calculated for the ECPs in the Western blot were then analysed for bands of similar molecular weight in the different isolates. As *B.pseudomallei* isolate 576 had no 'LPS' like reaction, the bands were clearer and so the other isolates were compared against this (Table 3.5).

Table 3.5 Isolates of *B.pseudomallei* and *B.thailandensis* that Share Common Protein Bands with *B.pseudomallei* Isolate 576 as Visualised by Western Blotting.

	B.pseudomallei			***B.thailandensis***		
Protein Bands of 576 (kDa)	53 (Clinical)	H706 (Clinical)	448 (Soil)	217 (Soil)	E255	E260
125.9	*					
120.2	*					
114.8	*	*		*	*	*
107.2	*		*			
101.2	*			*		
97.7				*	*	*
95.5					*	
91.2	*	*		*	*	
85.1		*		*		
74.1	*	*	*	*	*	*
72.4		*				
66.1			*		*	
63.1		*	*			
57.5		*		*		
50.7		*	*	*		
45.7		*		*		
44.7		*				
41.7						
36.3						
33.1	*					
29.5	*				*	*
24.0	*					
20.2						

*Asterisks denote identical molecular weight values.

Table 3.5 reveals that there are a number of extracellular protein products common to both species. Those include proteins at a molecular weight of 114.8, 101.2, 97.7, 95.5, 91.2, 85.1, 74.1, 66.1, 57.5, 50.7, 45.7 and 29.5 kDa.

In addition to these, bands of 166.0, 151.4 and 131.8 kDa were visualised in *B.pseudomallei* isolate 53 and in *B.thailandensis* E255, and so too could be considered common to both species. Interestingly, all isolates tested contained a protein band of 74.1 kDa. Those protein products found only in *B.thailandensis* isolates were bands of 109.6, 104.7, 87.1, 58.9, 52.5, 46.8 and 40.3 kDa. However, only five *B.pseudomallei* isolates were tested and so other isolates of *B.pseudomallei* not tested may contain these bands. Of the *B.pseudomallei* isolates, these were further divided into clinical and soil isolates, which revealed a number of bands found in clinical isolates, not found in those from the soil. These included bands at 125.9, 120.2, 72.4, 44.7, 41.7, 36.3, 33.1, 24.0 and 20.2 kDa. It is possible that these may represent virulence factors elicited during the course of an infection.

3.5 Microtitre Plate Assays.

In each assay 1µl of the concentrated ECP isolated was mixed with 2.5 µl of the appropriate substrate in a microtitre plate well. Stock solutions of substrates (1 mg/ml) were made in dimethyl formamide (DMF) and tests were carried out according to the manufacturers (Sigma) instructions. In each case, 21.5 µl of PBS was added to the ECP-substrate mixture and left at 37°C for 1310 h. and. This was carried out, unless otherwise stated, with a panel of 4 representative isolates which were as follows: *B.pseudomallei* 204, *B.pseudomallei* 576, *B.pseudomallei* 448 and *B.thailandensis* 82. The microtitre tests carried out and results can be found in Table 3.6.

Table 3.6 Microtitre Plate Assays to Test ECPs from Isolates of *B.pseudomallei* and *B.thailandensis* for Various Activities.

SUBSTRATE	ACTIVITY	ISOLATES TESTED	RESULTS *B.pseudomallei*	RESULTS *B.thailandensis*
Sheep red blood cells	Haemolysis,	30 Ara^- 6Ara^+	-	-
p-nitrophenylphosphoryl-choline	Phospholipase C	30 Ara^- 6Ara^+	+	+
N-trinitrophenylphosphatidyl ethanolamine (chromogenic)	Phospholipase A2	PANEL	-	-
5-bromo-4-chloro-3-indolyl phosphate (chromogenic)	Phosphatase	PANEL	- (+ 576 only)	-
AZO-COLL (chromogenic)	General Protease	PANEL	+	+
BANA (chromogenic)	Trypsin	PANEL	-	-
BAPNA (chromogenic)	Chymotrypsin	PANEL	-	-
AZO-CASEINE (chromogenic)	Protease	PANEL	+	+
AZO-ALBUMIN	Protease	PANEL	+	+
4-phenylazobenyloxycarbonyl-PRO-LEU-GLY-PRO-ARG (chromogenic)	Peptidase	PANEL	-	-
N-CB2-GLYGLY-PRO-LEU-GLY-PRO (chromogenic)	Elastase	PANEL	-	-
N-CB2-PRO-ALA-GLY-PRO-4-methoxy-β-naphthylamide (chromogenic)	Collagenase	PANEL	-	-
MU-acetyl-neuramide (fluorogenic)	Neuraminidase	PANEL	-	-
MU-diacetyl-chitobioside	Hexosaminidase	PANEL	+	-
MU-N-acetyl-galactosaminidine (fluorogenic)	Hexosaminidase	PANEL	+	-
MUTMAC (fluorogenic)	Chymotrypsin	PANEL	-	-
MU-N-acetyl-glucosaminidine (fluorogenic)	Hexosaminidase	PANEL	+	-
MU-mannopyranoside (fluorogenic)	Mannosidase	PANEL	-	-
MU-guanidinobenzoate (fluorogenic)	Trypsin	PANEL	-	-
MU-fucoside (fluorgenic)	Fucosidase	PANEL	-	-
N-t-Boc-LEU-GLY-ARG 7-amido 4 methyl coumarin (fluorogenic)	C5 peptidase	PANEL	-	-
MU-oleate (fluorogenic)	Lipase	PANEL	+	+

The results from the microtitre plate assays revealed that in the ECPs of *B.pseudomallei* and *B.thailandensis* there is phospholipase C activity, protease activity and lipase activity, which have all been previously described for *B.pseudomallei* but not confirmed in any isolates of *B.thailandensis*. No haemolytic activity, trypsin or chymotrypsin activity was found in either species. However, in only the *B.pseudomallei* isolates there was a hexosaminidase activity, which has previously been undescribed for the pathogen. Activity was found when screening with all of the following substrates: MU-diacetyl-chitobioside, MU-N-acetyl-galactosaminidine and MU-N-acetyl-glucosaminidine. It may be that *B.pseudomallei* possesses three separate enzymes or a single hexosaminidase enzyme responsible for the activities observed.

3.6 DISCUSSION.

The SDS-PAGE ECP profiles revealed that there were differences between the true *B.pseudomallei* isolates and isolates of the proposed species, *B.thailandensis*. However, there were also differences between isolates of the same species. In the case of the *B.thailandensis* isolates it may be possible that some of these may unknowingly be members of the very newly recognised species, *B.uboniae* (Yabuuchi *et al*, 2000). The results of the immunoblotting highlighted some interesting findings. Firstly, in that the strains 576 and 217 had no ladder of what was considered to be LPS reacting with the convalescent human serum. Isolate 576 reportedly has an atypical LPS pattern, shared with 5% of *B.pseudomallei* isolates (Dr. T. Pitt, personal communication). Therefore it is possible that the sera contained antibodies raised to the typical LPS, which is found in 95% of *B.pseudomallei* isolates and indeed *B.thailandensis* isolates. It was also observed that both species had the same ladder of LPS reaction, which is in keeping with the findings that LPS is immunologically indistinct from the two species (Anuntagool *et al.*, 1998).

The results of the SDS-PAGE analysis and the Western blot revealed that of the isolates tested, the two species shared a number of common protein bands, although obviously it is possible that the two species may share protein bands of exact size but that they may not actually be related proteins. However, it may be relevant as many asymptomatic infections are believed to be due to *B.thailandensis*, which reacts similarly to *B.pseudomallei* in the older tests first developed for melioidosis, which have now been superseded by more discriminating tests (Pongsunk *et al.*, 1999; Steinmetz *et al.*, 1999). This is most likely due to LPS, but a false positive as a result of similar ECPs possessed by the two species cannot be ruled out. Additionally, there were differences observed between the clinical and environmental isolates of *B.pseudomallei* tested. It may be the case therefore that these differences represent virulence factors elicited on infection. Unfortunately, the reaction to LPS was so strong in some of the isolates that they could not be analysed between particular molecular weights, which may affect conclusions regarding the presence of certain bands in particular isolates, or between the species. This reaction however has shown the overwhelming response in infections to LPS and indicates the sheer mass of this material that must surround bacterial cells, which may be the reason that some samples were found to float out of the wells during SDS-PAGE.

Comprehensive literature regarding such investigations is not available and so only tentative explanations regarding the likely character of these proteins can be attempted by analysing published literature. A phosphatidylcholine-hydrolysing phospholipase C (PC-PLC) has been characterised by Korbsrisate *et al.* (1999), which was revealed to be produced as a 73 kDa mature protein. It was not established in which isolates of *B.pseudomallei* this was produced nor if the avirulant species, *B.thailandensis* also produced the PC-PLC. Results from the Western blot indicated that all isolates tested, including *B.thailandensis*, produced a protein at around 74.1 kDa. In addition to this, from the microtitre plate assays it was shown that all isolates of both species produced phopholipase C activity and so the 74.1 kDa band may

represent the PC-PLC, allowing for any error involved with manually calculating an approximate value for the molecular weight. In *Listeria monocytogenes* a similar phospholipase has been shown to be involved in intracellular survival and cell-to-cell spread (Korbsrisate *et al.*, 1999), a factor perhaps common to both species of *Burkholderia.*

Perhaps most importantly regarding potential extracellular virulence factors, it has been reported that *B.pseudomallei* produces a 36 kDa exotoxin (Nigg *et al.*, 1955; Ismail *et al.*, 1991). Again, no extensive investigations have yet been carried out to establish which isolates produce the toxin or whether the newly defined species, *B.thailandensis* also produces this toxin. On analysing the protein bands on SDS-PAGE gels, many isolates produce a band around 36 kDa, including *B.thailandensis* isolates. However, on analysis of the Western blot it was only 576, a clinical *B.pseudomallei* isolate that contained a protein band of 36.3 kDa. Obviously it has not been proven that this band represents the toxin but it was the only band produced of such size and moreover, in a clinical isolate. As there was such a strong reaction to LPS in the clinical isolate 53 it was not possible to establish whether this too possessed a similarly sized band.

Among other extracellular products that have been reported is a protease produced by 94% of clinical isolates tested by Gauthier *et al.* (2000) at a molecular weight of 42 kDa (Percheron *et al.*, 1995). However, the report by Gauthier *et al.* (2000) did not state whether the protease was produced in environmental isolates or in isolates of *B.thailandensis.* Korbsrisate *et al.* (1999) similarly described the presence of a protease in isolates of *B.pseudomallei* but again did not mention if the protease was found in *B.thailandensis* isolates. SDS-PAGE profiles indicated that all of the *B.thailandensis* isolates tested and *B.pseudomallei* 53 produced a band of around 42kDa, which may represent the protease described by Percheron *et al.* (1995). However, in the Western blot only *B.pseudomallei* 576 produced a band near to this value of 42 kDa (41.7

kDa). The protease may be common to both species, as tests by Gauthier *et al.* (2000) showed the protease was not necessary for virulence. Indeed, the microtitre tests in 3.5 revealed that proteolytic activity was present in both *B.pseudomallei* and *B.thailandensis* isolates and so it is possible that both species may possess the same protease. Obviously further testing of *B.thailandensis* isolates for the protease is required to prove this. However the gene(s) encoding the protease have yet to be cloned and characterised in *B.pseudomallei* and so the protease may still represent a potential candidate for future gene cloning studies.

The microtitre plate assays (3.5) revealed no haemolytic activity in any of the *B.pseudomallei* or *B.thailandensis* isolates. Ashdown and Koehler (1990) found that of 100 clinical isolates tested, 93 produced a haemolysin but they only found the activity around very heavy growth on saline-washed sheep erythrocyte brain heart infusion agar and failed to observe this around individual colonies or in broth culture filtrate. In current literature very little has been investigated concerning culture supernatants and extracellular products but it must be stressed that different groups have all used different media to culture *B.pseudomallei* in such experiments. Both Heckly and Nigg (1958) and Colling *et al.* (1958) when assaying for a toxic product used media containing 4% glycerin beef extract to enhance expression. Further studies are therefore needed assaying isolates on a range of media to test for enhanced expression of certain activities.

Additionally, the microtitre plate assays revealed the presence of a potential hexosaminidase only in the *B.pseudomallei* isolates, which has not been described for this pathogen and therefore represents an ideal target for future cloning work with the pathogen. The microtitre tests also showed that activities already described for *B.pseudomallei* such as the phospholipase C, protease and lipase activities were also present in isolates of *B.thailandensis* and may not, therefore, have an essential role in virulence.

CHAPTER 4

CONSTRUCTION AND SCREENING OF *B.pseudomallei* GENOMIC LIBRARIES IN *E.coli* K12 HOST/VECTOR SYSTEMS.

4.1 INTRODUCTION.

Creating a genomic library enables the study of the complete genome of a bacterium, through fragmenting its DNA and randomly cloning the resulting fragments into a vector such as a plasmid, cosmid or a bacteriophage. The resulting pool of recombinant molecules is then introduced into a suitable bacterial host (usually *E.coli*) and the mixture of recombinant bacterial colonies, or phage plaques obtained represent a genomic library or gene bank. This process is also referred to as shotgun cloning and in theory, given enough clones it is possible to find every gene possessed by the organism under investigation.

In the case of this research, that is applying the techniques of molecular biology to the study of *B.pseudomallei*, the work progressed as in the manner of other studies with such pathogens, with the process of creating genomic libraries. This allowed random fragments of *B.pseudomallei* genomic DNA to be transferred into *E.coli* to facilitate the isolation and subsequent characterisation of genes encoding proteins of interest from *B.pseudomallei*. Thus enabling the identification of gene sequences and their products, with the potential aim of elucidating a role for them in the pathogenesis of melioidosis.

In order to create a genomic library the initial step is to fragment the genome of the bacterium to produce appropriately sized, random DNA fragments for subsequent ligation into the cloning vector of choice. The most suitable method for achieving this is by using restriction enzymes. It is possible to create a genomic library with a single restriction enzyme, such as *Eco*RI, with a six-base recognition site and so fragments will be on average several kilobases long. The disadvantage with the use of this is that there is the chance that a gene of interest may well have a recognition site within it and so will not be cloned intact. Or conversely, it could be surrounded by a region of DNA containing no *Eco*RI sites and so the resulting fragment will be so large that it will be cloned with a reduced efficiency in certain vectors, notably plasmids. However by using an enzyme with a four-base recognition site, such as *Sau*3A, the randomness of the library can be improved. This can be achieved by partial digestion of the genomic DNA with the enzyme, because allowing complete digestion of the DNA will yield fragments too small to be of use, but a partial digest will yield fragments of differing sizes. Thus, if a gene of interest is cut in one fragment the chance is that there will be several others where it is intact. The disadvantage to using such a frequently cutting enzyme is that multiple inserts could be cloned due to the re-ligation of fragments, which although may not prevent detection of a specific DNA sequence, may cause considerable confusion when characterising the cloned fragment and its relationship to adjacent regions of the bacterial DNA. Alternatively, to generate a completely random library it is possible to subject the genomic DNA to mechanical shearing such as sonication. However, this will generate fragments of which there is no knowledge of their termini, which may well prove incompatible to those on the cloning vector.

The choice of vector for use in a genomic library is based on the construction of a conveniently sized library so that only a relatively small number of individual recombinant bacterial or phage clones need to be screened to ensure a high probability of obtaining the gene sequence

of interest. Additionally, individual vectors have a specific capacity for foreign DNA inserts and as the genome itself is of a finite size it is often useful to have some idea of the number of independent recombinant clones that are required in the library for screening purposes. This is so that the number of clones used is representative of the genome so to be reasonably confident of finding the gene of interest. This can be calculated by the following formula devised by Clarke and Carbon (1976):

$$N = \frac{\log (1 - P)}{\log (1 - x/y)}$$

where:

N = the number of clones needed; P = the probability of a specific gene being present; x = the size of the foreign DNA insert (bp) and y = the size of the genome (haploid) in bp.

Thus, for example, for a genomic library of *E.coli* (4×10^6 bp) if the average insert size was 4 kbp the number of clones needed for a 99% probability (P = 0.99) is about 4600, but if inserts are larger, a smaller number of corresponding clones are needed. However, this figure assumes a random distribution of clones with no sequences being under or over-represented and so was used only as a token guide to the minimum number of bacterial colonies or phage plaques to be screened. Thus, in practice far more recombinants were screened than the equation predicted necessary.

The genomic library constructed can be screened immediately (an unamplified genomic library) by using selective media or overlays or by transferring the colonies or plaques to a membrane such as nylon or nitrocellulose that can be tested for the presence of the gene of interest. Positive clones can then be picked, purified, checked and confirmed. Alternatively, clones can be harvested and stored as a mixture for future plating out and screening; this

constitutes an amplified genomic library. There are three main ways of screening genomic libraries which make use of two detectable and unique properties of the clone of interest, one being the nucleotide sequence of the foreign DNA and the other any product of expression of the cloned gene. The former is through the use of gene probes and the latter is through immunological or enzymatic activity, respectively.

In order to generate a gene probe, some information regarding the likely sequence of the gene in question is required. If the corresponding gene from a closely related organism has already been cloned, it is possible to use the sequence information available from various databases, particularly if the organisms are very closely related, because the likelihood is that the sequence homology will be high. Alternatively, if the protein product of the gene has been purified, then determination of a short region of the amino acid sequence at the amino-terminus (N-terminus) allows prediction of the DNA sequence of the corresponding portion of the gene, although there will be some mis-match owing to the degeneracy of the universal code (Grange *et al.*, 1991). Probes can then be generated through a PCR reaction or a custom-made oligonucleotide and either radiolabelled or labelled with a non-radioactive chromogenic system, such as DIG.

An alternative to the use of gene probes is the use of indirect methods of screening that are based on the selection of a given phenotype. In the case of this study there were no available probes for the direct screening of recombinants. Thus, indirect methods of screening were utilised whereby the target gene product is in fact an enzyme whose activity can be detected, or a protein that can be detected immunologically. With both of these procedures the gene of interest must be expressed, or at least that part containing the epitope to be detected, and so this too has an important influence over the choice of vector. A gene may not be transcribed in a foreign host, or the gene product may be degraded, especially if only one part of the protein

is produced. This is why expression vectors such as pUC18 and λGT11 are used but the resulting product may still be subject to proteolytic degradation. This chapter describes the construction of gene libraries in plasmid, cosmid and bacteriophage cloning vectors and their subsequent screening with both enzymatic and immunological detection methods for the expression of haemolytic or proteolytic activity, or the expression of *B.pseudomallei* antigens recognised by polyclonal rabbit antiserum raised against whole cells and formalised cells of the bacterium.

4.2 SOURCE OF *B.pseudomallei* DNA AND ANTISERUM.

The *B.pseudomallei* DNA used in constructing all of the genomic libraries prepared in this study was kindly donated by Dr. R.Titball, DERA, Porton Down. The DNA had been prepared by a cell lysis procedure, followed by purification with repeated phenol extractions. The DNA used was obtained from isolates 204 and 576 (Table 2.1). These isolates were used, as they were representative examples of each of the two LPS profiles found in this bacterium.

Antiserum used in screening gene libraries was kindly donated by Dr. T Pitt, CPHL, Colindale, London, which was generated in rabbits against two isolates of *B.pseudomallei*, 204 heat-killed whole cells and 576 formalised cells.

4.3 ATTEMPTS TO CONSTRUCT *B.pseudomallei* GENOMIC LIBRARIES IN PLASMID AND COSMID VECTORS.

As *B.pseudomallei* probably has a number of genes encoding toxic products, a range of plasmids were used to attempt the construction of genomic libraries, those being the high copy

number expression plasmid, pUC18, the low copy number expression plasmid pGD103 and a medium copy number plasmid, pBR328. Genomic libraries using pUC18 were attempted with 576 *B.pseudomallei* DNA digested with *Eco*RV, *Hpa*I and *Sca*I. With the cloning vector, pGD103 a genomic library was attempted using 204 *B.pseudomallei* DNA digested with *Sau*3A and with pBR328 attempts were made with both 204 and 576 *B.pseudomallei* DNA with the following enzymes: *Bam*HI, *Eco*RI and *Eco*RV. Ligation reactions were often unsuccessful and the number of recombinants obtained when successful genomic libraries were transformed was not of a number high enough to represent the entire genome of *B.pseudomallei*. Hence none of these gene banks was used for screening.

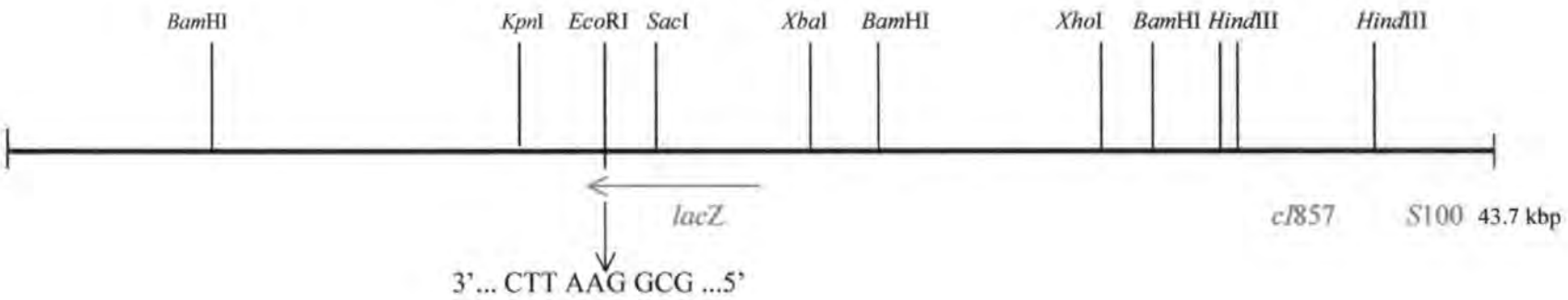
A cosmid genomic library was attempted using pHC79 ligated to *Sau*3A digested *B.pseudomallei* 204 genomic DNA. However, this too was unsuccessful.

4.4 CONSTRUCTION OF A *B.pseudomallei* GENOMIC LIBRARY IN A BACTERIOPHAGE LAMBDA VECTOR.

4.4.1 Bacteriophage λGT11.

The bacteriophage cloning vector, λGT11 (Fig.4.1) is a λ-derived insertional vector that is capable of cloning foreign DNA of up to 7.2 kbp into a unique *Eco*RI site within a *lac*Z gene. As the insertion site for foreign DNA is within the β-galactosidase gene, under appropriate conditions, the foreign DNA is expressed as part of a fusion protein with β-galactosidase. However, as it makes use of an inducible promoter, expression of a fusion protein will only occur if an inducing agent is added, such as IPTG. Without the addition of an inducing agent host cells produce large quantities of *lac* operon repressor (encoded by the *lac*I gene) which functions to prevent P-*lac*Z-directed expression of a fusion protein during the initial hours of

plaque formation, which is advantageous if the foreign protein is toxic to the host cells, as insufficient amounts of the antigen will be produced for detection. IPTG functions to inactivate the *lac*I repressor protein hence *lac*Z-directed expression occurs. Additionally, a *lon*$^-$ strain is used as a host to increase the stability of foreign proteins that may be expressed because *E.coli* produce *lon* protease as one of a number of proteases responsible for lowering the stability of foreign or abnormal proteins that are not normally produced in the cell. Foreign protein may also be expressed without being fused to β-galactosidase as a result of transcription and translation of the foreign DNA sequence by means of promoter regions within the DNA insert (*lac* independent) or alternatively the expression may be as a result of the *lac*Z promoter (*lac* dependent). It is also possible to use antibodies as probes to screen the genomic library constructed in λGT11 for recombinants producing proteins that may be of interest. The phage vector produces a temperature-sensitive repressor (*c*I857), which is activated at 42°C and carries an amber mutation (*S*100) that renders it lysis-defective in hosts lacking the amber repressor (*sup*F) (Huynh *et al.*, 1985).



λGT11 43.7 kbp (Huynh *et al.*, 1985)

Figure 4.1 Bacteriophage cloning vector λGT11.

Restriction enzyme cleavage sites are shown, including the unique *Eco*RI site for the insertion of foreign DNA into the *lacZ* gene for the expression of foreign DNA as fusion proteins. Also shown is the *c*I857 temperature-sensitive repressor region and the *S*100 amber mutation.

4.4.2 Construction of a λGT11/*B.pseudomallei*/*Eco*RI Genomic Library.

Approximately 2 μg of *B.pseudomallei* 204 genomic DNA was digested with the restriction enzyme, *Eco*RI (2.5.4) and afterwards a sample of the digest was checked on a 0.8% agarose gel (2.4.5) to ensure complete digestion had occurred. *Eco*RI-cleaved λGT11 was obtained in the form of alkaline phosphatase phosphorylated vector arms (Stratagene, La Jolla, CA., USA). A ligation reaction (2.5.6) was set up using the vector arms obtained and the *B.pseudomallei* 204 *Eco*RI-cleaved genomic DNA and incubated at 15°C for 18 h. and stored at 4°C. A sample of the ligation reaction was checked on an agarose gel to confirm successful ligation. Ligated DNA was packaged using a commercial *In Vitro* packaging kit (Boheringer Manneheim). A ten-fold dilution series of the packaged bacteriophage particles was made in SM buffer and samples were plated with the *E.coli* host strain, Y1088 in an overlay on NZCYM plates. The resulting phage, which represented a *B.pseudomallei*/λGT11 genomic library were recovered from the soft agar overlay (as in 2.3.2.2), treated with a drop of chloroform and titrated (2.3.1) using NZCYM plates containing X-gal and IPTG to ascertain the number of recombinant phage. The number of non-recombinants was calculated, appearing as blue plaques, and the recombinant frequency was calculated as being approximately 99%. The genomic library was stored at 4°C until needed.

4.5 IMMUNOLOGICAL SCREENING OF A λGT11/*B.pseudomallei*/*Eco*RI GENE LIBRARY.

The *B.pseudomallei* gene library constructed in the λGT11 expression vector (4.4.2) was screened immunologically as recombinant phage plaques on a lawn of *E.coli* K12 host bacteria. Proteins released by the lysis of the cells within the plaques were immobilised on nitrocellulose membranes placed over the lawn. Protein bound to the nitrocellulose was

screened with the two different types of antisera donated (4.2). Antisera was preadsorbed overnight with Acetone powders (made by Dr. A. Evenden) to pre-adsorb any anti-*E.coli* antibodies that may have been present in the antiserum. The genomic library was plated out so that it gave rise to the formation of approximately 1000 plaques per plate. The actual number of individual plaques needed to constitute a λGT11/*B.pseudomallei*/*Eco*RI genomic library was estimated by using the formula described in 4.1, using a probability value of 0.99 and taking an average insert size of 4 kbp and the size of the *B.pseudomallei* genome as 6.54 x 10^6 bp (Songsivilai and Dharakul, 2000), which gave a number of phage particles to be screened as 7527. However, the actual number of phage particles screened was far higher and of the order of 20,000 plaques. The λGT11 genomic library was plated out onto NZCYM plates, supplemented with 0.2% lactose, with an overnight culture of *E.coli* Y1090. Once the plates were set they were incubated, inverted at 37°C for 18 h. After incubation, plates were overlayed with gridded, 82mm nitrocellulose filters (BA 85/23 Schleicher and Schuell, Dassel, Germany). The nitrocellulose filters were numbered and marked with a ball-point pen in the centre with a cross and a line was marked at the centre of the top of the grid vertically down to indicate the top of the nitrocellulose and smaller lines were marked at the centre of the other three outer edges. The nitrocellulose was then transferred, wearing surgical gloves and using forceps so as not to contaminate the plates, gridded side down, onto the plates. This was done carefully and the filter was folded slightly so that the middle of the filter touched the plate first so as not to introduce any air bubbles. The bottom of the petri dish was marked in the same fashion as the filter so that the orientation of the filter and plate could be easily aligned after screening. The filters were left to absorb for 30 min. at 37°C before carefully removing with forceps and washed for 5 min. in a 50 ml volume of PBS with gentle agitation before removing the PBS and replacing it with 200 ml of PBS containing 1% (w/v) casein as a membrane-blocking reagent. Blocking was carried with gentle agitation for 1 h. after which the blocking reagent was replaced, and primary antibody was added at a concentration of 1:200

and incubated for 90 min. at room temperature. The filters were then washed with 200 ml of PBS for 5 min. at room temperature for a total of three washes. Next, 200 ml of blocking solution was added and to this the secondary antibody, swine anti-rabbit immunoglobulins (P217, Dako Ltd., High Wycombe, UK) at a concentration of 1:500. Filters were left incubating with the secondary antibody for 90 min. followed by three washes with 200 ml of PBS for 5 min. at room temperature. After washing, antibody binding was visualised using DAB/$NiCl_2$ colour development. The λGT11/*B.pseudomallei*/*Eco*RI genomic library was screened in this manner on four occasions, twice where the primary antibody was rabbit antisera raised against *B.pseudomallei* 204 heat-killed whole cells and twice against rabbit antisera raised against *B.pseudomallei* 576 formalised cells. Therefore, of 80 filters each containing approximately 1000 plaques, two were found to contain one immunopositive plaque each, of which an example can be seen in Fig. 4.2, both of which were the result of screening with antisera to 576 formalised cells. Both immunopositives were identified as a "halo" of colour that developed around the plaque.

4.6 PURIFICATION OF AN IMMUNOPOSITIVE RECOMBINANT λGT11 BACTERIOPHAGE.

The immunopositive plaques, designated λBPGT1 and λBPGT2 were identified on their respective plates using the grids on the filters and picked with a sterile Pasteur pipette (as in 2.3.2.1) along with several negative plaques. A wire loopful of the resulting phage suspensions were streaked out onto NZCYM plates and overlayed with soft agar continuing Y1090 host cells, supplemented with 0.2% lactose and incubated at 37°C for 18 h. Afterwards, areas with well separated plaques were overlayed with nitrocellulose, including several negative controls and subsequently probed with antiserum as before (4.5). The recombinant phage, λBPGT1 remained immunopositive and was retained for further analysis, whereas λBPGT2 did not.

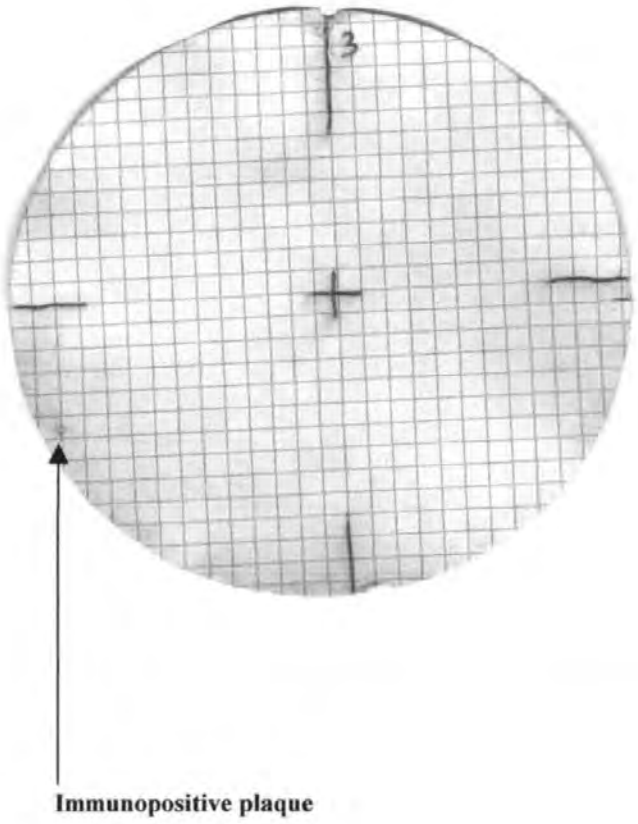


Fig.4.2. Nitrocellulose Filter Containing an Immunopositive Plaque-Lift from a λGT11/*B.pseudomallei*/*Eco*RI Genomic Library. The immunopositive plaque is arrowed and appears as a "halo" of colour around the plaque. The filter shown was probed with rabbit anti-*B.pseudomallei* 576 formalised cell serum as the primary antibody using the method in 4.5.

4.7 SCREENING OF A λGT11/*B.pseudomallei/Eco*RI GENOMIC LIBRARY.

In addition to the immunological screening of the λGT11 library, the library was also screened for enzymatic activity in a number of ways. As previously calculated (4.5), around 7527 plaques were required to screen the entire λGT11 genomic library and around 10,000 recombinant plaques were screened for each substrate used. Recombinant phage plates were set up as in 4.5.

4.7.1 Screening of a λGT11/*B.pseudomallei/Eco*RI Genomic Library for Haemolytic Activity.

While ECPs of *B.pseudomallei* showed no detectable haemolytic activity (3.5), blood agar plates of the bacterium did show some haemolysis and so recombinants were screened with blood for this reason. Plates of recombinant λGT11 plaques and a control were screened for haemolytic activity using an erythrocyte/agarose overlay. This was essentially a modification of Kehoe *et al.* (1983) where erythrocyte/agarose overlays were used to detect genes coding for the α-haemolysin from *S.aureus* in λ47.1. The genomic library was screened using sheep erythrocytes, carried out as follows. A 5 ml volume of overlay which consisted of 10% sheep erythrocytes suspended in a PBS/agarose (45°C) was carefully poured onto the recombinant λGT11 plaques and allowed to set. The overlayed plates were then incubated either at room temperature for 24h., or at 37°C for 4 h. and then placed room temperature. Recombinants were analysed for any exhibiting zones of haemolytic activity around the plaques. However, no zones of haemolytic activity were observed in any of the overlayed plates.

4.7.2 Screening of a λGT11/*B.pseudomallei*/*Eco*RI Genomic Library for Proteolytic Activity.

The recombinant λGT11 phage plaques were also screened for the presence of proteolytic activity. Plates were screened along with vector phage controls and each plate was overlayed with molten agarose/PBS containing a protease substrate, allowed to set and incubated for 24h. prior to examination.

4.7.2.1 Natural Protease Substrates.

Recombinant λGT11 clones were screened with two natural protease substrates, casein and gelatin. Both substrates were incorporated into 5 ml of agarose/PBS overlay solution at a concentration of 1% (w/v), overlayed onto the phage-containing plates and incubated at either room temperature or 37°C. Plates containing casein as a substrate were examined by flooding the surface with 3% (v/v) acetic acid solution, whilst gelatin-containing plates were flooded with 5% (w/v) tannic acid solution. Both acetic acid and tannic acid function to precipitate the protein substrate so that any clones producing a proteolytic activity may be identified by a zone of proteolysis. No proteolytic activity was detected in the overlayed plates.

4.7.2.2 Synthetic Protease Substrates.

Recombinant λGT11 clones were also screened with synthetic protease substrates, which included chromogenic azocollagen (azocoll), Nα-benzoyl-DL-arginine-p-nitroamide (BAPNA) and Nα-benzoyl-DL-arginine β-napthylamide (BANA) and fluorogenic 4-methylumbelliferyl p-guanidinobenzoate (MUGB). The library and controls were overlayed with 5 ml of agarose/PBS solution containing 0.5% (w/v) azocoll and 0.1% (w/v) BAPNA, BANA and MUGB. All followed incubation either at room temperature or 37°C. Plates containing the azocoll overlay were examined for the release of red/purple chromagen in areas of proteolysis and those containing BAPNA and BANA were examined for an orange-red colour. Plates

containing the fluorogenic substrate (MUGB) were examined under UV light for a blue-white fluorescence. However, no proteolytic activity was observed.

4.7.3 Other Substrates.

Recombinant λGT11 plaques were also screened with MU-N-acetyl-glucosaminidine a fluorogenic substrate to screen clones for any producing hexosaminidase activity. Microtitre plate testing revealed that all isolates of *B.pseudomallei* tested had a hexosaminidase activity (3.5). However, no clones with hexosaminidase activity were observed.

4.8 DISCUSSION.

Attempts to create plasmid genomic libraries were unsuccessful; mainly due to problems in ligating *B.pseudomallei* DNA with plasmid vector DNA, hence not producing sufficient clones when transformed to warrant screening. That is, enough to represent the entire genome of *B.pseudomallei*. The *B.pseudomallei* DNA used in the attempts to create plasmid genomic libraries was donated, and the case would seem that this DNA was not of a quality for successful ligation. It was suspected that the *B.pseudomallei* genomic DNA was contaminated with a fair amount of complex carbohydrates not removed during the cell lysis procedure. Moreover, as the DNA was purified with phenol extractions, which functions to dissociate DNA from protein, the contaminating carbohydrate would still be present. Additionally, only small quantities of *B.pseudomallei* DNA were received and so it was not feasible to further purify the DNA, for example thorough a CsCl/EtBr density ultracentrifugation procedure.

Only one of the *B.pseudomallei* genomic libraries created was successful, that being the λGT11/*B.pseudomallei*/*Eco*RI genomic library. However, the screening procedures then used

with this genomic library yielded little in the way of results, and only two immunopositive clones were identified using rabbit antiserum to *B.pseudomallei* 576 formalised cells and only one of these retained its immunopositivity. In retrospect, it was perhaps not surprising that antiserum to whole cells yielded no positives when considering the previous chapter's results when ECPs isolated from *B.pseudomallei* isolates were immunoblotted using convalescent human sera (3.4); as the main immune response was to LPS. Therefore, the antiserum to whole cells and probably that to formalised cells as well, was likely to be mainly against LPS. Additionally, when generating antiserum to both whole cells and formalised cells it was likely that many of the ECPs were lost in a washing procedure and so there would therefore be little in the way of results.

In addition to this, the general lack of success of screening was probably due to a number of other factors that influence the outcome of any gene screening procedure. Firstly, those resulting from the construction of the genomic library; secondly, inadequacies in the screening procedure itself and thirdly, and probably most fundamentally with respect to the present study, is the expression of the gene sequences contained in the genomic library. It was unlikely, with respect to the λGT11 genomic library, to be a result of a deficiency in the construction of the gene bank. Each stage of its construction was carefully monitored and the proportion of recombinants was calculated at being 99%, hence the library being composed of a high proportion of vector-only ligations was not the case. In addition, each screening procedure involved considerably more recombinants than the number predicted by Clarke's formula.

The screening procedures employed were based on published reports that the bacterium had haemolytic activity and although no activity was observed with the ECPs in this study, blood agar plates of the bacterium did yield some haemolysis. Also, that *B.pseudomallei* produces

proteolytic activity, although the exact nature of the protease has yet to be elucidated and so a range of proteolytic substrates were used for screening. Microtitre plate assays of ECPs from *B.pseudomallei* isolates also revealed a hexosaminidase activity (3.5) but screening the λGT11 genomic library with one of the substrates used to detect hexosaminidase activity yielded nothing. Two temperature ranges were also employed in screening as some components may only be expressed at specific temperatures, for example on entering the human body, at 37°C, but neither revealed any results.

The construction of the λGT11/*B.pseudomallei*/*Eco*RI gene library was a fairly straightforward process, however the successful expression of the sequences cloned is a somewhat more difficult procedure. Although steps were taken to encourage expression by using the expression vector λGT11 and inducing the fusion of β-galactosidase with any foreign proteins in frame by the addition of an inducing agent, the lack of expression of *B.pseudomallei* genes in the *E.coli* K12 host/vector system was probably one of the biggest problems, in addition to the antisera, which led to the lack of success experienced in screening the library. There was limited published information on cloning *B.pseudomallei* DNA into *E.coli* K12 host/vector systems and so it was not known how successful the expression of *B.pseudomallei* genes in these systems would be, although it was assumed that considering such a large number of genes from a wide range of bacterial taxa have been successfully cloned and expressed in these systems that it would be possible to do the same with *B.pseudomallei*. However, although *B.pseudomallei* is a gram-negative bacterium it is still possible that incompatibilities between the machinery of gene expression exist, despite the fact that steps were taken by using a cloning vector where *lac*-gene fusions could occur.

The immunopositive clones were identified as a "halo" of colour that formed around the plaques. This was unusual in that most immunopositives screened in this way are identified as

a complete colour development over the entire plaque. It was not known what exactly this represented, but it could be as a result of a breakdown of the protein responsible for the immunopositivity, or perhaps representative of export of the protein or it may represent the relative solubility of the protein. As this was the only positive result from the construction and screening of the genomic libraries and the fact it was continually and strongly immunopositive compared to negative controls, and of the interesting nature of the colour development, it was decided to pursue this clone.

CHAPTER 5

MANIPULATION AND STRUCTURAL ANALYSIS OF THE IMMUNOPOSTIVE CLONE λBPGT1.

5.1 INTRODUCTION.

Once a clone exhibiting a desired genetic property is selected from a genomic library, analysis can begin on the recombinant. In the case of the immunopositive clone, λBPGT1 obtained (Chapter 4), attempts to determine the size and level of expression of the immunoreactive antigen(s) was investigated without any knowledge of the function of the genes responsible. This may be achieved most simply on an SDS-PAGE gel, by comparing the protein profiles of the immunopositive against a negative control in order to identify the synthesis of any novel proteins in cells carrying the recombinant molecules. However, normal *E.coli* proteins may mask the detection of recombinant gene products, in which case Western blotting may be applied. Other methods exist which allow the detection of novel proteins without the often confusing concurrent expression of the host genome, of which there are three main methods, two of which are *in vivo*:

1. Minicells for the expression of plasmid-borne genes (Frazer and Curtiss, 1975),

2. Maxicells also applicable to genes cloned in λ vectors (Sancar *et al.*,1979; Newman *et al.*, 1979), and

3. *In vitro* transcription and translation systems.

Bacterial cell-free systems for coupled transcription and translation of *in vitro* genes contained on bacterial plasmids or bacteriophage genomes are available commercially and rely on the

presence of all the biochemical components of protein synthesis extracted from certain cell types, but without a functional genome. These systems have advantages over the *in-vivo* methods with respect to their sensitivity, as they allow higher expression levels of proteins that are normally expressed at low levels *in vivo* due to the action of host-encoded repressors. An additional advantage is that results can be visualised after just a few hours. The system used in this project to attempt to detect proteins expressed in the immunopositive clone not present in *E.coli* controls was the *E.coli* S30 Extract System for Circular DNA that incorporates the Transcend^{TM} Non-Radioactive Translation Detection System (Promega). This particular system allows the non-radioactive detection of proteins synthesised *in vitro*, as biotinylated lysine residues are incorporated into nascent proteins during translation, thus eliminating the need for labelling with [^{35}S] methionine. The biotinylated lysine is added to the translation reaction as a precharged ε-labelled biotinylated lysine-tRNA complex rather than a free amino acid. After subjecting the reaction products to SDS-PAGE and electroblotting, the biotinylated proteins can be visualised by binding Streptavidin-Alkaline Phosphatase (AP) or Steptavidin-Horse Radish Peroxidase (HRP), followed by either colorimetric or chemiluminescent detection. Typically, 0.5 to 5 ng of protein can be detected within 3 to 4 hours after gel electrophoresis and so the sensitivity of this non-radioactive detection system is equivalent to that achieved with [^{35}S] methionine incorporation and autoradiographic detection, 6 to 12 hours after gel electrophoresis. The system thus confers the advantage of not only eliminating the need for radioisotope handling, storage and disposal but also that the biotin tag is stable for 12 months unlike ^{35}S-labelled proteins, whose label decays over time.

Once the gene product of interest has been detected, the characterisation process can begin. To aid characterisation of a DNA sequence it is often necessary to subclone the fragment into a more appropriate vector, for example from a λ vector to a plasmid, so that restriction mapping can be facilitated. Restriction mapping itself is usually one of the first steps in the

characterisation of a cloned DNA fragment and is a prerequisite for planning a sequencing strategy and selecting appropriate fragments for further subcloning. Restriction maps themselves can be determined either directly using partial digestion techniques, or by indirect approaches such as those based on the analysis of fragments created by single and double digests with a set of restriction enzymes (Brown, 1993). The latter, indirect method eventually arrives at a map compatible with the measured fragment sizes by a process of exclusion, which is relatively straightforward, providing the number of sites for each enzyme is small. To start the map, the DNA is digested with a series of single restriction enzymes and the fragments produced are analysed by agarose gel electrophoresis, and their sizes estimated against a ladder of known molecular weights. Once achieved, commonly the DNA is then digested with combinations of restriction enzymes whose single profiles have been characterised and the results are used to create a restriction map.

5.2 ATTEMPTS TO DETECT IMMUNOREACTIVE ANTIGENS IN λBPGT1.

The initial stage in the characterisation of the immunopositive clone involved the detection of the antigen(s) responsible for the immunogenicity of the clone, λBPGT1. To aid the detection of the unknown antigen(s), negative controls of Y1088 cells infected with λGT11 phage were used.

5.2.1 Preparation of a Concentrated Lysate of λBPGT1.

Phage lysate stocks of λBPGT1 and two negative controls were made from plates supplemented with 0.2% (w/v) lactose that exhibited confluent lysis (2.3.2.2). The resulting stock of phage was titrated (2.3.1) and stored at 4°C until use.

5.2.2 SDS-PAGE and Western Blotting of λBPGT1.

The concentrated phage lysates (5.2.1) were subjected to SDS-PAGE analysis (2.6.1) and a Western blotting procedure (2.6.2) using antisera to *B.pseudomallei* 576 formalised cells. No difference was observed in the protein profiles between the immunopositive and the negative controls in either technique.

5.3 SUB-CLONING OF λBPGT1 INTO A PLASMID VECTOR.

It was decided to isolate the DNA insert contained within λBPGT1 to ascertain its size and to subclone the fragment into a plasmid vector to aid manipulation and further analysis of the clone.

5.3.1 Isolation of λBPGT 1 DNA.

A high titre stock of the recombinant phage, λBPGT1 was obtained from liqiud cultures (2.3.2.3). The recombinant phage DNA was extracted with formamide (2.3.3) and a sample checked on an agarose gel.

5.3.2 Determination of the Insert Size for λBPGT1.

There exists one *Eco*RI cleavage site in λGT11, which was used to create the λGT11/*B.pseudomallei*/*Eco*RI genomic library (4.4.2) hence digesting the DNA of recombinants with *Eco*RI releases the foreign fragment. Approximately 0.5 μg of λBPGT1 DNA was used in a restriction enzyme digest (2.5.4) with *Eco*RI and the products were visualised on agarose gel electrophoresis (Fig 5.1), which shows clearly an insert of approximately 5 kbp.

5.3.3 Subcloning the 5 kbp Fragment into the Plasmid Vector, pUC18.

5.3.3.1 The Plasmid Vector, pUC18.

The plasmid vector, pUC18 (Appendix II) is a small vector of 2.7 kbp, carrying the selectable phenotype of resistance to ampicillin. The vector also contains the *lacZ*α region of *E.coli*, which renders recombinants *lac*$^-$, facilitating their selection via *lac* gene complementation (blue-white screening) in *E.coli* hosts deleted for *lacZ*α. Additionally, due to the strong inducible *lac* promoter, coupled with a potentially high copy number of up to 700 plasmids per cell, cloning into pUC18 facilitates the increased production of any foreign proteins.

5.3.3.2 The Subcloning Procedure.

Approximately 1 μg of λBPGT1 DNA and 2 μg of pUC18 DNA were digested with *Eco*RI. Products of digestion were checked on an agarose gel and the digested λBPGT1 DNA was purified by extraction with phenol (2.4.3.2). Linearised pUC18 DNA was treated with alkaline phosphatase (2.5.5) before purification by phenol extraction. A standard ligation reaction was set up (2.5.6) and incubated at 15°C for 18 h. Ligation products were checked on an agarose gel before carrying out a transformation using DH5α competent cells (2.5.3.2). The transformation mix was plated out onto LB media supplemented with ampicillin and an overlay of X-gal and IPTG facilitating blue/white screening (2.5.2). Plates were incubated at 37°C for 18 h. A number of resulting white colonies were cultured in 10 ml of LB broth supplemented with ampicillin and incubated at 37°C for 18 h. with shaking. Small-scale plasmid extractions (2.4.2.1) were performed on six subclones (pBPGT1-6). Plasmid DNA isolated was digested with *Eco*RI and subjected to agarose gel electrophoresis to check for the presence of an appropriately sized DNA insert (Fig.5.2). All recombinant plasmids pBPGT1-6 contained an insert of approximately 5kbp.

5.4 COLONY BLOTS.

Subclones pBPGT2, 3, 4, and 6 grew successfully on LB media supplemented with ampicillin and 0.2% lactose and were used in a colony blot to determine whether they were immunopositive. Nitrocellulose filters were layered over well-spaced recombinant colonies and a comparative control of DH5α recently transformed with pUC18. Filters were left on the plates for 2 min. before carefully removing with forceps and placing onto filter paper saturated with 0.1% SDS for 15 min. to lyse the bacterial cells. Afterwards, the filters were suspended in a tank filled with chloroform for 20 min. before placing into Blocking Solution for 18 h. Screening and developing after this time was carried out using the same procedure used to screen the λGT11/*B.pseudomallei*/*Eco*RI genomic library (4.5), with the exception that the primary antibody was used at a concentration of 1:100. The subclones did develop earlier and slightly darker than the control, especially pBPGT2 which was chosen for further analysis but the recombinants did not develop the "halo" effect as observed with the immunopositive phage clone, λBPGT1. However, after 20 min. of developing time, there was little difference between the recombinants and the control.

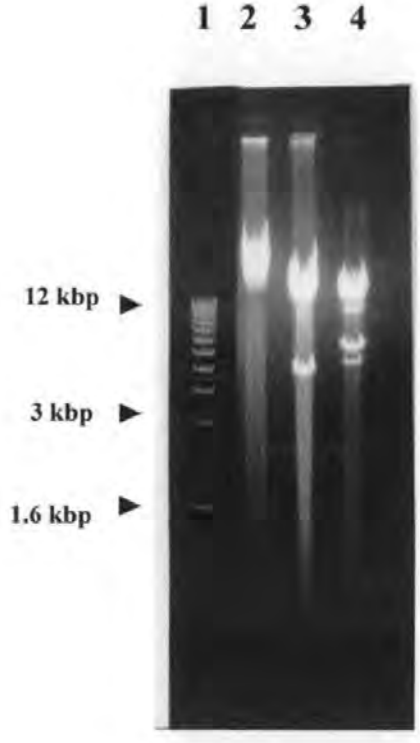


Fig.5.1 Agarose Gel Electrophoresis of *Eco*RI and *Bam*HI Cleaved DNA from a λGT11/*B.pseudomallei*/*Eco*RI Immunopositive Clone, λBPGT1. LANE 1: 1 kpb ladders (from top to bottom) 12 kbp, 11 kbp, 10 kbp, 9 kbp, 8 kbp, 7 kbp, 6 kbp, 5 kbp, 4 kbp, 3 kbp, 2 kbp and 1.6 kbp. LANE 2: λBPGT1/*Eco*RI DNA. LANE 3: λBPGT1/*Bam*HI DNA.

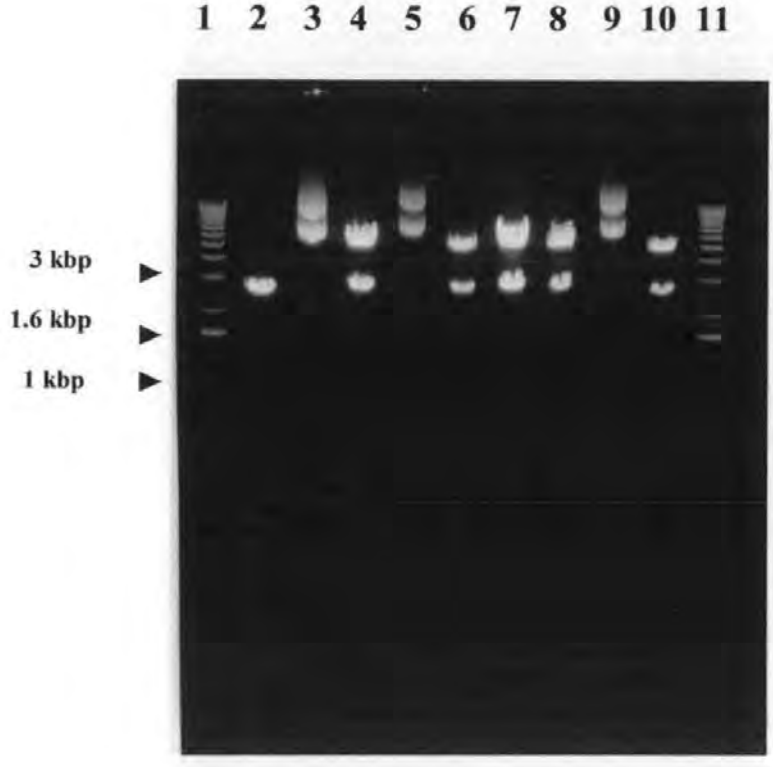


Fig.5.2 Agarose Gel Electrophoresis of Recombinant Plasmid Subclones (pBPGT1-6) Digested with *Eco*RI. LANE 1: 1 kbp ladders. LANE 2: pUC18/*Eco*RI. LANE 3: pBPGT2. LANE 4: pBPGT2/*Eco*RI. LANE 5: pBPGT3. LANE 6: pBPGT3/*Eco*RI. LANE 7: pBPGT4/*Eco*RI. LANE 8: pBPGT5/*Eco*RI. LANE 9: pBPGT6. LANE 10: pBPGT6/*Eco*RI. LANE 11: 1 kbp ladders.

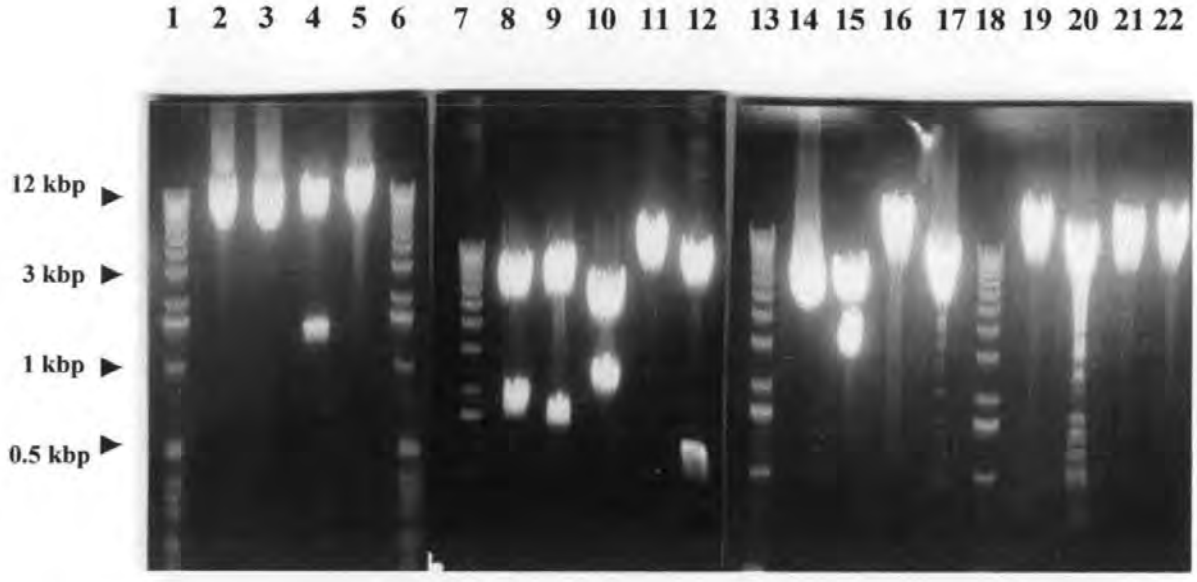
5.5 RESTRICTION ENZYME MAPPING OF pBPGT2.

5.5.1 Single Restriction Enzyme Digests.

Single restriction enzyme digests were set up with approximately 0.5 μg of pBPGT2 DNA to determine if there were any sites for these enzymes on the cloned insert and their relative value for use in double enzyme digests. Restriction enzymes chosen were those known to cleave the vector once or twice or not at all. The results of which were analysed on agarose gel electrophoresis (Fig.5.3) and the resulting fragment sizes were assessed by comparison to 1 kbp ladders. Cleavage with *Eco*RV and *Xho*I, for example, produced two fragments and these enzymes are known not to cleave the vector and so it can be inferred from this that there are two sites on the cloned insert for each of these enzymes.

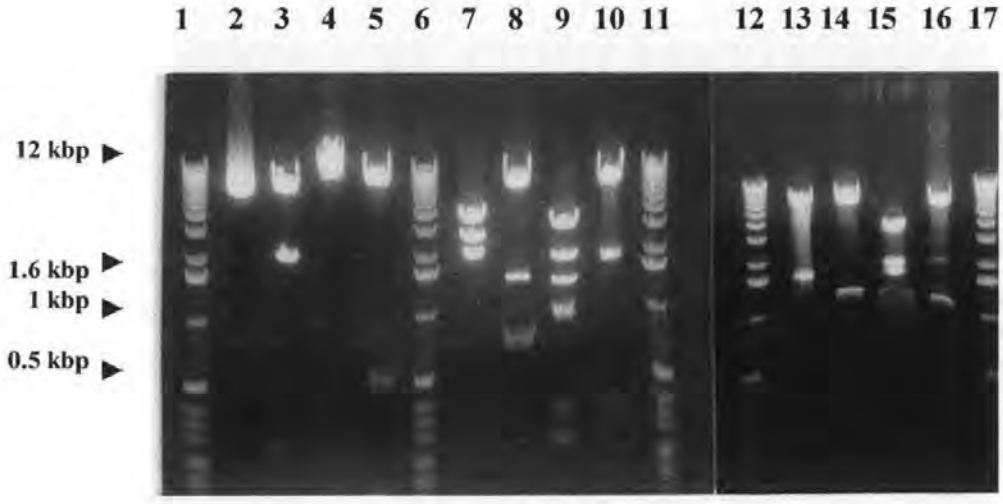
5.5.2 Double Enzyme Digests.

Using information that the single enzyme digests provided double enzyme digests were performed, whereby two restriction enzymes were used in combination. These were mainly carried out using both a "primary" enzyme, known to cleave the insert in pBPGT2, and a "secondary" enzyme whose cleavage site was undetermined and far enough from the insert to produce fragments easily visible after agarose gel electrophoresis. The double enzyme digests were set up (2.5.4) and as with single enzyme digests the results were analysed by agarose gel electrophoresis (Fig.5.4). Using the results from single and double enzyme digests (Fig.5.3 and 5.4), i.e. the number of bands produced and their relative sizes and relating this information to sites that exist on the vector and those that can be deduced on the insert, a map of restriction sites on the foreign 5 kbp fragment was produced (Fig.5.5).



Lane	Enzyme	No. of Fragments	No. of Cleavage Sites on VECTOR	No. of Cleavage Sites on INSERT	Size of Fragments (kbp)	Total
1	Markers	-	-	-	-	-
2	Uncut pBPGT1	-	-	-	-	-
3	*Eco*RI	2	1	0	2.7, 5.0	7.7
4	*Bam*HI	1	1	0	7.7	7.7
5	*Sal*I	-	1	-	-	-
6	Markers	-	-	-	-	-
7	*Hind*III	1	1	0	7.7	7.7
8	*Hinc*II	-	1	-	-	-
9	*Eco*RV	2	0	2	7.5, 0.2	7.7
10	*Sca*I	1	1	0	7.7	7.7
11	Markers	-	-	-	-	-
12	*Xho*I	2	0	2	6.0, 1.7	7.7
13	*Sma*I	2	1	1	6.3, 1.4	7.7
14	*Pvu*II	2	2	0	5.1, 2.3	7.4
15	*Xba*I	1	1	0	7.7	7.7
16	*Pst*I	2	1	1	7.0, 0.7	7.7
17	Markers	-	-	-	-	-
18	*Hpa*I	0	0	0	-	-
19	*Bcl*I	0	0	0	-	-
20	*Bgl*II	2	0	2	6.2, 1.5	7.7
21	*Nco*I	1	0	1	7.7	7.7
22	Markers	-	-	-	-	-

Fig.5.3. Agarose Gel Electrophoresis of Single Enzyme Digests of the Recombinant Plasmid, pBPGT2.



Lane	Restriction Enzymes	Sizes of Fragments Produced (kbp)	Total
1	Markers	-	-
2	Uncut pBPGT1	-	-
3	*Xba*I/*Nco*I	5.7, 2.0	7.7
4	*Xba*I/*Bsc*I	7.7	7.7
5	*Xba*I/*Eco*RV	7.0, 0.5, 0.2	7.7
6	*Xba*I/*Xho*I	3.5, 2.2, 2.0	7.7
7	Markers	-	-
8	*Xba*I/*Bgl*II	6.4, 1.5, 0.6	8.5
9	*Ava*I	3.2, 1.6, 1.4, 0.9, 0.3, 0.2	7.7
10	*Sma*I/*Nco*I	7.0, 1.9, 0.4, 0.3	9.6
11	Markers	-	-
12	Markers	-	-
13	*Sma*I/*Bsc*I	-	-
14	*Sma*I/*Eco*RV	6.0, 1.2, 0.6	7.8
15	*Sma*I/*Xho*I	3.4, 2.0, 1.8, 0.5	7.7
16	*Sma*I/*Bgl*II	-	-
17	Markers	-	-

Fig.5.4. Agarose Gel Electrophoresis of Double Restriction Enzyme Digests of the Recombinant Plasmid, pBPGT2.

5.6 FURTHER SUBCLONING OF THE RECOMBINANT PLASMID pBPGT2.

Once a restriction map had been obtained for pBPGT2 (Fig.5.5) it was decided to further subclone the 5 kbp insert into smaller clones for attempts to aid expression of the immunoreactive antigen(s) and to facilitate sequencing.

5.6.1 Cloning Strategy.

The cloning strategy is shown in Fig.5.5 whereby the 5 kbp fragment was digested with predetermined restriction enzymes, sites for which were previously mapped, to create specific DNA fragments that were to be cloned into a suitable vector thus creating the five desired subclones pCD1-5, as indicated in Table 5.1.

Table 5.1 Strategy Used for Further Subcloning of the 5 kbp Fragment into Plasmid Vectors.

Restriction Enzymes	Expected Size of the fragment (kbp)	Plasmid Vector	Recombinant
*Eco*RI/*Pst*I	1.0	pGD103	pCD1
*Pst*I/*Sma*I	0.7	pUC18	pCD2
*Sma*I/*Xho*I	0.9	pUC18	pCD3
*Xho*I	1.8	pUC18	pCD4
*Xho*I/*Eco*RI	0.6	pUC18	pCD5

5.6.2 Isolation of DNA Fragments for Further Subcloning.

Approximately 3 μg of pBPGT2 was digested with each combination of restriction enzymes as indicted in Table 5.1 and the products were analysed on agarose gel electrophoresis. The desired band was identified by size in each case, excised from the agarose gel and the DNA electroeluted (2.4.6.1) and purified with the BIO-RAD Prep-A-Gene kit (2.4.6.2).

5.6.3 Subcloning Procedure.

The DNA fragments isolated for cloning into pUC18 (Table 5.1) were ligated using the Sureclone Kit (2.5.7). The *Eco*RI/*Pst*I fragment was cloned into the plasmid vector pGD103 (Appendix II), which carries the selectable phenotype of kanamycin resistance. Approximately 0.5 μg of pGD103 DNA was digested with the restriction enzymes *Eco*RI and *Pst*I, and after checking the digest on an agarose gel the linearised plasmid was treated with alkaline phosphatase (2.5.5), purified by extraction with phenol (2.4.3.2) and used in a ligation reaction with the *Eco*RI/*Pst*I fragment at a ratio of 5:1 respectively. All ligations were used in a transformation procedure with DH5α competent cells (2.5.3.2) and the transformation mix was spread onto LB media containing the appropraite antibiotics and incubated at 37°C for 18 h. Recombinants were identified using blue/white screening and recombinant white colonies were cultured in 10 ml LB broths supplemented with the appropriate antibiotics. Small-scale plasmid extractions were carried out on cultures (2.4.2.1). All plasmid DNA was digested with the restriction enzymes *Eco*RI and *Pst*I to ensure that they contained a correctly sized insert. Subclones containing the desired size insert were identified, results of which can be seen in Fig.5.6.

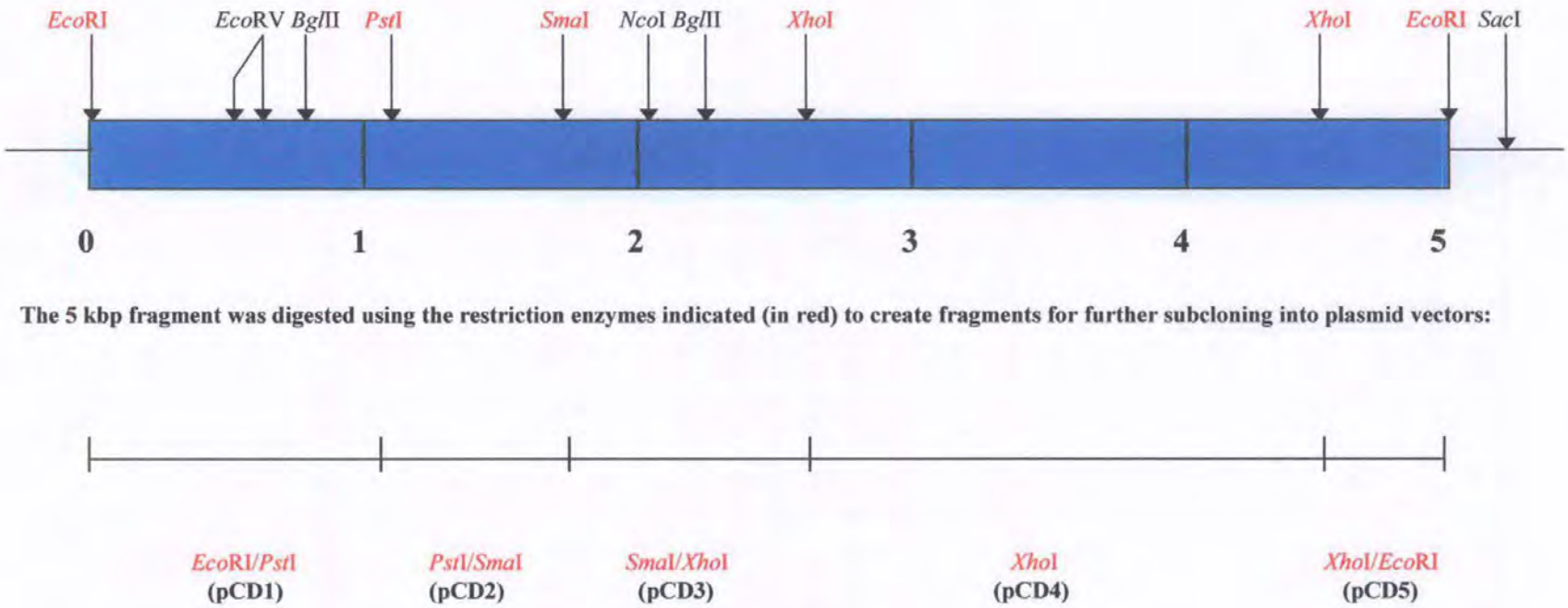


Fig.5.5 Restriction Map of the 5 kbp foreign *B.pseudomallei* DNA fragment and Further Subcloning Strategy. The restriction map shows all tested restriction sites within the fragment and shows those restriction enzymes used for further subcloning of the 5 kbp fragment into other plasmid vectors (pCD1-5).

Fig.5.6 Agarose Gel Electrophoresis of Subclones (pCD1-5) Digested with the Restriction Enzymes *Eco*RI/*Pst*I to Release a Foreign Insert. LANE 1: 1 kpb ladders. LANE 2: λ Standard (0.125 μg) LANE 3: pCD1. LANE 4: pCD1/*Eco*RI/*Pst*I. LANE 5: pCD3. LANE 6: pCD3/*Eco*RI/*Pst*I. LANE 7: pCD4. LANE 8: pCD4/*Eco*RI/*Pst*I. LANE 9: 1 kbp ladders. LANE 10: λ Standard (0.125 μg) LANE 11: pCD2. LANE 12: pCD2/*Eco*RI/*Pst*I. LANE 13: 1 kbp ladders LANE 14: pCD5/*Eco*RI/*Pst*I.

5.7 ATTEMPTS TO USE AN *In Vitro* TRANSCRIPTION AND TRANSLATION KIT TO IDENTIFY NOVEL PROTEINS PRODUCED BY THE SUBCLONE, pBPGT2.

No protein product(s) could be identified from the immunopositive clone λBPGT1. For this reason, the *E.coli* S30 Extract System for Circular DNA that incorporates the Transcend™ Non-Radioactive Translation Detection System (Promega) was used to try to identify any novel proteins produced by the subclone, pBPGT2 and further subclones of the fragment, pCD1-5. In this way the gene(s) responsible for the immunopositivity observed may have become separated away from other, interfering sequences contained on the fragment.

5.7.1 Standard Protocol.

For each DNA sample used the following basic reaction was set up:

DNA template	≤ 4 μg
Amino acid mixture (complete)	1 μl
S30 Premix without amino acids	20 μl
S30 extract, Circular	15 μl
Nuclease-free water to a final volume of	49 μl

The components were added into a MCC tube, vortexed gently and centrifuged for a few sec. to bring the reaction mixture to the bottom. A 1 μl volume of Transcend™ tRNA was added to the translation reaction and immediately incubated at 37°C for 2 h. The reaction was terminated by placing the tubes into an ice bath for 5 min. A provided positive control was set up in each experiment which was the plasmid, pBEST*luc*™ which synthesises firefly luciferase which migrates to 60 kDa and it also contains the gene for ampicillin resistance, which can appear as a faint band migrating to 31.5 kDa. A negative control was also used which consisted of the same reaction components as the positive control with the omission of

the plasmid DNA. Once the S30 extract reaction was complete, a 5 µl aliquot was removed and added to 20 µl of acetone in a MCC tube and placed on ice for 15 min. The unused portion of the reaction was stored at -20°C. The sample was then centrifuged at 12,000 x g for 5 min. The supernatant was removed and the pellet was dried for 15 min. under vacuum. Afterwards, 20 µl of SDS-PAGE Loading buffer was added and heated to 100°C for 2-5 min. A sample of this was then analysed by SDS-PAGE (2.6.1) using a 13% gel and staining with coomassie blue or used in a Western blotting procedure (2.6.2). Western blotting was carried out using Tris-buffered saline (TBS) with 0.5% Tween® 20 (TBST) as a blocking buffer. Biotinylated proteins were visualised by binding Streptavidin-Alkaline Phosphatase, followed by colorimetric detection with Western Blue® Stabilised Substrate for Alkaline Phosphatase.

5.7.2 Application of the *In Vitro* Transcription and Translation Kit.

The following DNA templates were set up as in 5.7.1 a positive control (pBEST*luc*™), pBPGT2, pUC18 and pCD1-5 and once the reaction was complete these were subject to SDS-PAGE and Western blotting. The positive control developed the expected protein bands of luciferase (60.7 kDa) and β-lactamase (31.5 kDa). However, no novel protein bands were observed in any of the recombinant subclones. This was repeated for a total of three times per sample altering the reaction conditions slightly each time, but the results remained unchanged.

5.8 DISCUSSION.

The method used in this Chapter (5.2) to identify any recombinant novel protein(s) produced by the immunopositive, λBPGT1 was unsuccessful. The technique used would have given an indication of the size of the protein but could not in itself detect low levels of protein expression, or determine whether the protein was in its native form. The immunopositive

clone, λBPGT1 was initially identified as having a "halo" of positive reaction around the plaque and the possible explanations for such an observation were discussed in Chapter 4. It is possible that no protein bands were visualised, as there was a very low level of expression of the cloned product, hence making the immunopositive protein barely detectable with the antisera used, producing just a "halo" of positive reaction. Alternatively, the production of an insoluble protein from λBPGT1 would also be responsible for the lack of a protein band on an SDS-PAGE gel and in Western blotting. As previously discussed (4.7) the "halo" effect observed may have been indicative of degradation of the protein, and indeed if this were occurring, there would be no products observed. For any given protein a variety of factors singly, or in combination may modulate its half-life *in vivo*. Among such factors are the flexibility, accessibility and sequence of the N- and C- termini, the presence of chemically blocking amino-terminal groups such as the acetyl group, and the exposure on the surface of the folded protein of protease cleavage sites (Old and Primrose, 1989). It has been shown that in yeast, for example that proteins derived from β-galactosidase have a half-life that varies from just 3 minutes to 20 hours, depending on the amino acid at the N-terminus (Bachmair *et al.*, 1986). It is not known whether a similar phenomenon occurs in *E.coli* but if it does it may explain the great variation observed in the stability of different proteins that have been made from similar gene constructs.

To facilitate further analysis of the immunopositive clone the foreign fragment was subcloned into a plasmid vector. This facilitated restriction mapping of the fragment. Some digests were inconclusive and occasionally the sizes of the fragments yielded did not add exactly to the estimated size, in some cases due to partial digestion products and in others due to "star" activity. Star activity was observed with single digests using *Sal*I and *Hinc*II (Fig.5.3) and is a phenomenon exhibited by a number of restriction enzymes under non-optimal conditions, and is the loss of sequence specificity (Kessler and Holtke, 1986). In the case of partial digests,

bands appear too light and the addition of all the fragments yields a value too large. Problems with restriction digests are encountered if the G+C content of the fragment is high, which could well have been the case with *B.pseudomallei* DNA, as the genome has a G+C content around 69% (Redfearn *et al.*, 1966). High G+C fragments tend to migrate faster than fragments with lower G+C contents and the problem is particularly common with short DNA fragments. However, subcloning the fragment presented no problems and further subcloning into five smaller fragments was carried out. It was hoped that the gene(s) responsible for the initial immunopositive reaction would be contained within one of the subclones, which would facilitate its expression by isolating the sequence away from other, potentially interfering sequences.

The colony blot carried out with recombinant plasmid subclones was inconclusive as eventually little difference was observed between the reaction of the recombinants to that of the *E.coli* control. However, it was observed that the recombinants developed quicker and slightly darker than the control. It could be possible that there existed a high background reaction from *E.coli* proteins not adsorbed by the acetone powders used and such a level of background reaction did not exist when using the phage clones.

Again, attempts to detect novel proteins produced by the plasmid subclones were unsuccessful, even with the use of *the in vitro* transcription and translation kit. A number of factors again could be responsible for this. Firstly, if the protein were unstable or insoluble it would still be undetectable on an SDS-PAGE gel and in Western blotting even with the use of the kit. If the protein is normally produced at very low levels, which was an explanation to why it was initially undetectable in experiments carried out with λBPGT1, the *in vitro* transcription and translation system should overcome this problem as it increases the level of expression of very such proteins. It could be possible that the immunopositive antigen was in fact a

B.pseudomallei protein produced in-frame with the β-galactosidase system in λGT11 and through subcloning the fragment, such expression has been lost or that a deletion event may have occurred whereby even the loss of one base would disrupt the reading frame of the protein and hence expression would be lost. As no proteins were visualised with the subclones pCD1-5 it is possible that the gene(s) had been disrupted through the subcloning procedure and so expression of the full product did not occur.

As this was the only clone identified from screening with antisera, despite the problems encountered with the identification of any novel protein(s) it was decided to continue work with the immunopositive clone and to sequence the DNA. The lack of expression may have been due to the myriad of reasons discussed and considering that little work has been carried out with cloning and expression of *B.pseudomallei* DNA and proteins in *E.coli* systems there was little information to compare. Sequencing the fragment may provide answer to the problems encountered, as more information regarding the nature of the genes encoded may allude to whether the product is insoluble, easily degraded or perhaps requires post-translational modifications.

CHAPTER 6

ANALYSIS OF THE NUCLEOTIDE SEQUENCE OF AN IMMUNOPOSTIVE CLONE FROM *B.pseudomallei.*

6.1 INTRODUCTION.

Once a fragment of DNA has been cloned that is thought to encode gene(s) responsible for an observed activity, the next step is to sequence the DNA, which gives essential information regarding the structure, function and evolutionary history of the gene(s) encoded. Knowledge of a DNA sequence may often be an end in its own right but the information that DNA sequencing provides can be processed by comparing the nucleotide sequence itself or the deduced polypeptide sequence with a data bank of other sequences. Any similarities highlighted can, in principle arise by two routes: either by convergent evolution, or through their being related by descent from a common ancestral sequence (Old and Primrose, 1989). Convergent evolution implies the two sequences have not descended from a common ancestral sequence, but have instead arisen as a consequence of selection for a particular function. The alternative implies that the two sequences have remained similar in evolutionary time due to a selection pressure, which has limited any scope for divergence. Whatever the route has been, the implication here is that any sequence similarity can be taken to be indicative of similar function. It must also be stressed that selection pressure *per se* does not necessarily act on the protein as a whole but that certain coding regions fulfil in themselves a distinct function, such as DNA binding and so these conserved domains are often the subjects of intense selection pressure. The overall rationale therefore is that when evolution has developed a successful idea it tends to stick with it, which can be paralleled to virulence gene in pathogens. The sequence

data banks are rapidly growing and are at present more than an adequate size to search for similarities in any newly discovered sequence.

The base sequence can also be analysed for structural and physical features such as restriction enzyme sites, base composition and open reading frames, through computer programs. The importance of this information allows additional sequences to be searched for, such as the "machinery of gene expression", this includes ribosome binding sites, promoter and terminator regions. Efficient translation requires that the mRNA bears a ribosome binding site, or Shine and Dalgarno sequence (SD) that varies in length from three to nine bases and precedes the translational start codon by three to nine bases (Steitz, 1979). In addition to this, the transcription of a given coding DNA sequence is initiated by the formation of a stable complex between the RNA polymerase enzyme and a characteristic control sequence, called the promoter. Extensive analysis of these (Hawley and McClure, 1983) has identified two regions that are consistently similar. One of these contains 6 or 7 bases and occurs about 10 bases upstream of the nucleotide at which transcription commences (+1) and is often termed the -10 sequence or "Pribnow Box". The second is 9 nucleotides long and clustered about 35 nucleotides upstream of the of the RNA initiation site, known as the -35 sequence. A consensus sequence for the -35 site has been deemed, 5'-TTGACA, and for the -10 Pribnow Box, 5'-TATAAT. In addition to the transcriptional initiation regions, DNA sequences may also possess transcriptional terminators, consisting of inverted extended repeat sequences distal to the translation termination codon. These sequences induce the 3' ends of RNA transcripts to form "stem-loop" structures which are believed to interfere with RNA polymerase activity and facilitate the disassociation of the synthesised mRNA chain from the template (Adhya and Gottesmann, 1978).

Once a putative gene has been identified it is then possible to analyse the structure and composition of the protein molecule encoded by the gene of interest. Amino acid regions of secondary structure such as disulphide bridges and alpha helices can be detected as well as conserved motifs such as signal peptides and DNA binding regions, and functional properties such as hydrophobicity and antigenicity.

6.2 SEQUENCING OF pBPGT2.

Approximately 30 μg of pBPGT2 DNA was isolated using the Qiagen kit (2.4.2.4), and sent to MWG-BIOTECH as a dried pellet. The fragment was sequenced using the primer-walking service. The complete nucleotide sequence of the *Eco*RI-generated fragment was found to be 4969 nucleotides in length and can be found in Appendix III.

6.3 NUCLEOTIDE SEQUENCE ANALYSIS OF THE 5 kbp FRAGMENT.

The complete 5 kbp sequence received was entered into the DNasis program (Hitachi) installed on a personal computer which consists of a number of programs for the analysis of nucleotide or amino acid sequences.

6.3.1 Restriction Endonuclease Cleavage Site Analysis.

The first analysis carried out on the nucleotide sequence of the 5 kbp fragment was the production of a more detailed restriction enzyme cleavage map than the one previously produced (Fig.5.5) so that further manipulations of the fragment may be more readily performed. The following restriction site table (Table 6.1) was compiled using DNasis software for sequence analysis.

Table 6.1 Computer-prepared Restriction Site Analysis of the 5 kbp *Eco*RI-generated *B.pseudomallei* DNA Fragment.

ENZYME	SEQUENCE	COUNT	CUTTING POSITION
*Aat*II	GACGT↓C	3	800, 2734, 3564
*Afl*II	C↓TTAAG	1	297
*Apa*I	GGGCC↓C	1	917
*Asp*I	GACN↓NNGTC	1	3477
*Avi*II	TGC↓GCA	4	1503, 1979, 3823, 4292
*Bcl*I	T↓GATCA	1	843
*Bfr*I	C↓TTAAG	1	297
*Bgl*II	A↓GATCT	2	619, 1813
*Bse*AI	T↓CCGGA	3	280, 1436, 4010
*Bsi*WI	C↓GTACG	2	2662, 3954
*Bsm*I	G↓CATTC	1	4669
*Dra*II	RG↓GNCCY	3	913, 914, 972
*Eco*47III	AGC↓GCT	1	3916
*Eco*RI	G↓ATTC	2	53, 5016
*Eco*RV	GAT↓ATC	2	517, 592
*Mlu*I	A↓CGCGT	4	2253, 3236, 4021, 4825
*Mun*I	C↓AATTG	1	2522
*Nco*I	C↓CATGG	1	1897
*Nhe*I	G↓CTAGC	1	1983
*Nsi*I	ATGCA↓T	3	3612, 3848, 4552
*Pst*I	CTGCA↓G	2	18, 1062
*Rca*I	T↓CATGA	1	981
*Rsr*II	CG↓GWCCG	4	1019, 3068, 3935, 4244
*Sac*I	GAGCT↓C	3	51, 87, 1705
*Sfu*I	TT↓CGAA	1	3961
*Sgr*AI	CR↓CCGGYG	2	2485, 4679
*Sma*I	CCC↓GGG	2	39, 1617
*Sna*BI	TAC↓GTA	1	1209
*Sty*I	C↓CWWGG	1	1897
*Tru*9I	T↓TAA	2	298, 2391
*Van*91I	CCANNNN↓NTGG	1	4530
*Xho*I	C↓TCGAG	2	2405, 4196
*Xho*II	R↓GATCY	4	32, 619, 1813, 3638

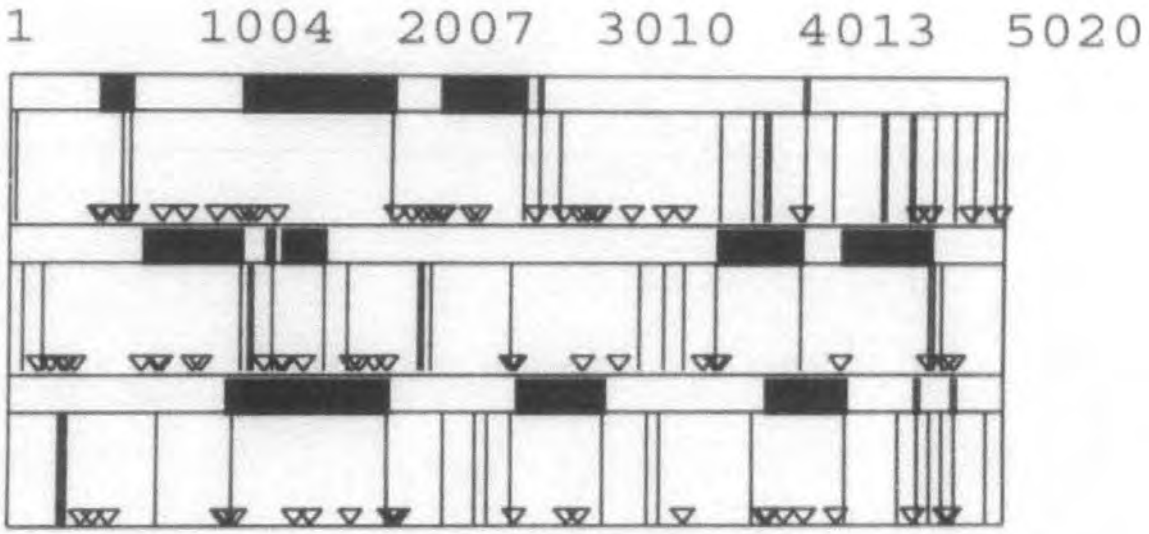
In the restriction enzyme recognition sequences a Y represents a pyrimidine base (C or T), R represents a purine base (G or A) and N represents any base. The values for the *Eco*RI cleavage sites in Table 6.1 (in blue) represents the position of the foreign DNA insert within the sequence, therefore positions in red represent cleavage sites on the vector DNA.

6.3.2 Open Reading Frame (ORF) Analysis.

Following restriction enzyme analysis, the next part of the analysis involved a search for open reading frames of nucleotide bases that when transcribed into mRNA may act as a template for a protein that could be responsible for the immunopositivity previously observed with the clone λBPGT1. The DNasis computer program used translates the nucleotide sequence of interest into all six possible reading frames of base triplets. The program then searches for the translation initiation sequence ATG, followed by a search "in frame" for the presence of one of three possible termination sequences, TAA, TAG and TGA. This is then presented to the user in the format as seen in Fig.6.1, which shows the open reading frame analysis on the 5 kbp fragment carried out with this program.

Prior to nucleotide sequencing, attempts to determine a protein produced by both λBPGT1 and pBPGT2 in which the 5 kbp fragment was cloned were unsuccessful, therefore there were no indications as to the size of protein produced therefore an estimate of the possible ORF size could not be deduced. Thus the strategy used next was to use all putative open reading frames in Web-based BLAST searches to identify possible homologies with other proteins, and to analyse these.

Forward Translation:



Reverse Translation:

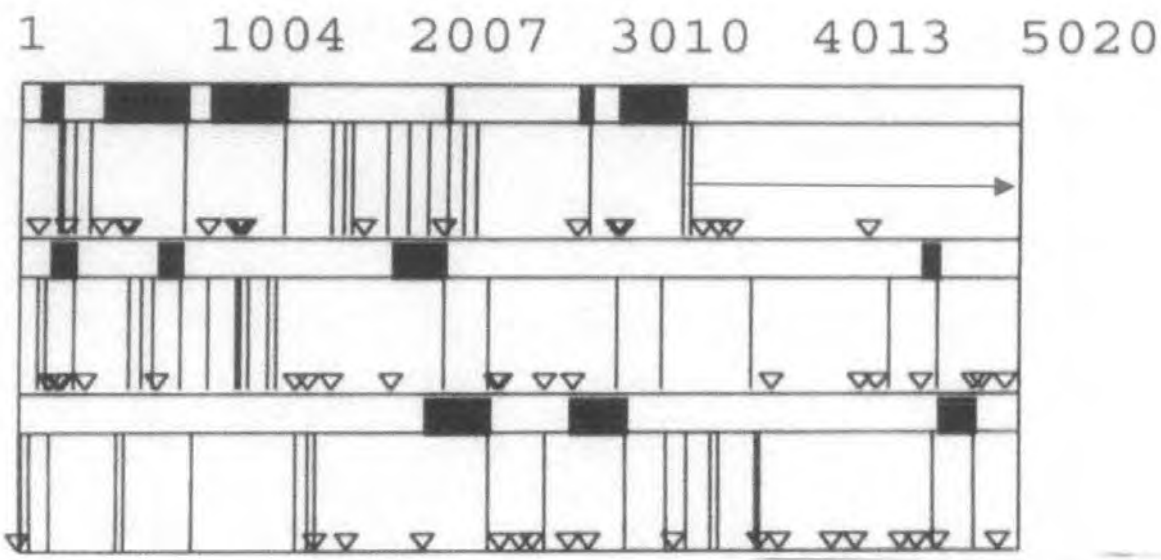


Fig.6.1. Open Reading Frame Analysis of pBPGT2 Forward and Reverse Translations in all Three Reading Frames. The lines represent stop codons and the triangles represent start codons (TAA, TAG, TGA). The red arrow represents an ORF not identified by DNasis.

6.3.3 Web-Based BLAST Searches of Putative Open Reading Frames.

From the ORF analysis carried out (6.3.2) all ORFs identified by DNasis (Fig.6.1) were used in Web-based BLAST searches (Altschul *et al.*, 1997) to identify any homologies with similar sequences with known functions. A number of ORFs identified within the 5 kbp fragment had homologies with proteins from a range of other bacteria and in some cases with other prokaryotes and eukaryotes.

All BLAST results were analysed and eventually four ORFs encoding putative genes, over the whole 5 kbp fragment were identified, the results of this can be seen in Fig.6.2. These four ORFs cover the first 2.5 kbp of the fragment. In the latter 2.5 kbp ORFs did not yield any homologies with proteins from BLAST searches in any of the reading frames. This does not necessarily mean that the sequence contains no coding regions but that perhaps the ORFs code for proteins not yet identified in any other organisms. One of the ORF's identified (ORF1), did not have a termination codon as the sequence was probably disrupted by the *Eco*RI site used to generate the genomic library in Chapter 4 and so the rest of the sequence is probably contained on an adjacent *Eco*RI fragment.

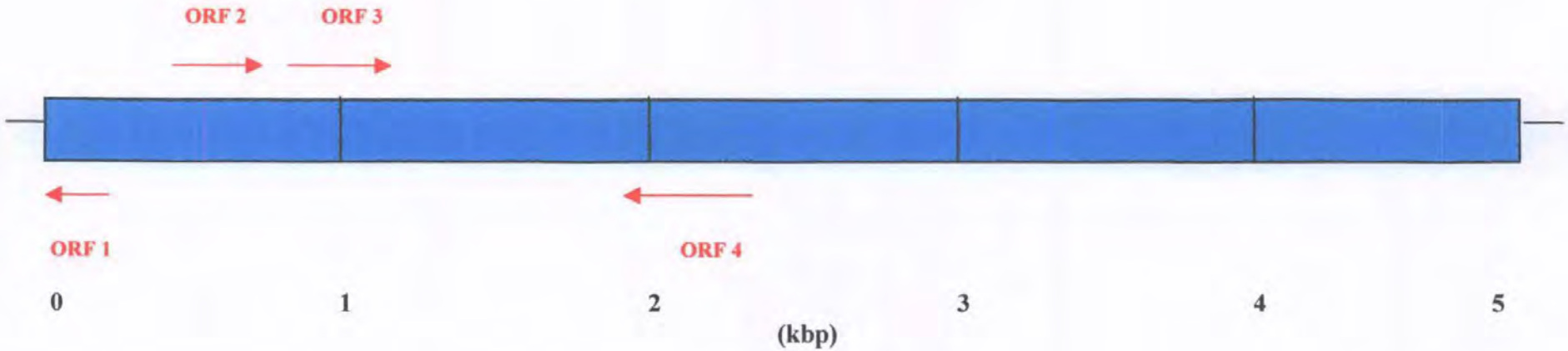


Fig.6.2 Map of the 5 kbp *B.pseudomallei* Fragment Indicating Putative ORFs Encoding Sequences with Homology to Other Bacterial Protein Sequences Deposited in BLAST Databases. Those putative ORFs above the sequence (2 and 3) are in the normal, forward translation, those beneath the sequence (1 and 4) are the reverse complementary.

6.4 ANALYSIS OF ORFs IDENTIFIED ON THE 5 kbp *Eco*RI-GENERATED FRAGMENT FROM *B.pseudomallei.*

6.4.1 Proteins Identified in the BLAST-P Database Having Homology to the ORF Sequences Identified.

The following tables show proteins with homology to the ORFs identified (Fig.6.2) and are given along with their Accession numbers and brief titles describing their function.

Table 6.2 Web-Based BLAST-P Matches for ORF 1.

Acession Number	Match Sequence Title	Percentage Homology (Positives)*
AF075709	*lsf*A *Pseudomonas putida*	96
D90905	Rehydrin *Synechocystis* sp.	82
D14662	Human Antioxidant Protein 2 (24 Kda) *Homo sapiens*	87
AF004670	Mouse Antioxidant Protein 2 *Mus musculus*	85
AF093852	1-Cys peroxiredoxin *Mus musculus*	85
AF014009	Thiol-specific antioxidant protein *Rattus norvegicus*	83
AL033396	Rehydrin-like protein *Candida albicans*	74
Y12089	Peroxiredoxin *Arabidopsis thailana*	67
U40818	Rehydrin *Tortula ruralis*	66
AE000711	Alkyl Hydroperoxide Reductase *A.aeolicus*	63
U31052	Thiol-Specific Antioxidant *Onchocerca volvulus*	73
Y08256	Probable Peroxiredoxin *Sulfolobus solfataricus*	59
X74264	Probable Peroxiredoxin *Methanobacterium thermoautotrophicum*	62
U67520	Probable Peroxiredoxin *Methanococcus jannaschii*	59
AP000005	Alkyl Hydroperoxids Reductase *Pyrococcus horikoshii*	59
AF007757	Probable Peroxiredoxin 2 *Sulfolobus metallicus*	59
L46863	Alkylhydrogenperoxide Reductase *Legionella pnuemophila*	62
AJ222857	*yku*U Protein *Bacillus subtilis*	58
AF016233	Alkylhydrogenperoxide Reductase *Enterococcus faecalis*	52
AB010689	AhpC *Pseudomonas putida*	60
M60116	20 KDa Protein in Rubredoxin Operon *Clostridium pasteurianum*	54
U65510	Alkyl Hydroperoxide Reductase *Rhodospirillum rubrum*	60
AL031798	Thioredoxin Peroxidase *Schizocaccaromyces pombe*	57

* This represents the percentage similarity of the putative amino acid sequence with those deposited in the BLAST-P database.

The amino acid sequence of ORF 1 had a very high homology to *lsf*A of *P.putida* (96% positivity) and with other peroxidases from other bacteria and so could potentially have a peroxidase function in *B.pseudomallei.*

Table 6.3 Web-Based BLAST-P Matches for ORF 2.

Accession Number	Match Sequence Title	Percentage Homology (Positives) *
U18997	Hypothetical 11.1 kDa Protein in EXUR-TDCC Intergenic Region *E.coli*	54
U58768	Putative 10 kDa Protein *E.coli*	47
AE000352	Hypothetical protein in STPA-NRDH Intergenic Region *E.coli*	43
AF073461	Unknown Protein *Eimeria acervulina*	51
X04465	ATP Synthase B Chain *Marchantia polymorpha*	54
U40566	Ubiquitin Activating Enzyme 2 *Arabidopsis thaliana*	46
AE000753	Hypothetical Protein *A.aeolicus*	47
Z99708	Putative Protein *A.thaliana*	49

* This represents the percentage similarity of the putative amino acid sequence with those deposited in the BLAST-P database.

Table 6.4 Web-Based BLAST-P Matches for ORF 3.

Accession Number	Match Sequence Title	Percentage Homology (Positives) *
AE000392	Hypothetical 15.1 kDa Protein in EXUR-TDCC Intergenic region *E.coli*	44
U18997	ORF o157 *E.coli*	44
U75930	Hypothetical Protein 25.6 kDa Orgyia pseudotsugata nuclear polyhedrosis virus	43

* This represents the percentage similarity of the putative amino acid sequence with those deposited in the BLAST-P database.

ORF 2 is immediately followed by ORF 3 in the 5 kbp nucleotide sequence and both are in the same reading frame. Both sequences have homology to a range of putative proteins from other bacteria, most notably from *E.coli*. Interestingly, both have homology to small, intergenic putative proteins from the same region, the EXUR-TDCC in *E.coli*.

Table 6.5 Web-Based BLAST-P Matches for ORF 4.

Accession Number	Match Sequence Title	Percentage Homology (Positives) *
AF027189	Comf *Acinetobacter calcoaceticus*	40
U02552	PilE *P.aeruginosa* (involved in Type IV pilin biogenesis).	36
Z69262	PilE *Neisseria gonorrhoeae*	40
L32968	Prepilin *Moraxella bovis*	35
AF027189	ComE *Acinetobacter calcoaceticus*	43
U32715	Prepilin peptidase dependent protein D *Haemophilus influenzae*	40
M11323	Pilin protein precursor *P.aeruginosa*	37
M11435	Beta pilin *M.bovis*	38
M32345	Q pilin *M.bovis*	38
M59712	Pilin subunit *M.bovis*	38
AF109904	PilA *Vibrio cholera*	35
M14849	Pilin protein *P.aeruginosa*	46
K02080	Pilus protein precursor *N.gonorrhoeae*	38
S68101	Pilin *P.aeruginosa*	44
B31105	Pilin protein precursor *P.aeruginosa*	37
S15265	Pilin protein precursor *Dicelobacter nodosus*	35
U16742	Pilin Monomer *N.gonorrhoeae*	38

* This represents the percentage similarity of the putative amino acid sequence with those deposited in the BLAST-P database.

Perhaps the most interesting homology was identified to this particular ORF, ORF 4, as it has homology to pilin genes from a range of other bacterial pathogens.

6.4.2 Nucleotide Sequences of the ORFs Identified.

The next part of the analysis was to identify possible transcriptional and translational control regions on the ORFs. Analysis of possible -10 and -35 sequences was carried out using a range of *E.coli* promoters and those from similar species. Unfortunately, very little information regarding possible -10 and -35 sequences for different *B.pseudomallei* genes exists. In addition to this possible ribosome binding sites (SD regions) and termination loops were searched for. The results of these can be seen in Figs.6.3, 6.4 and 6.5.

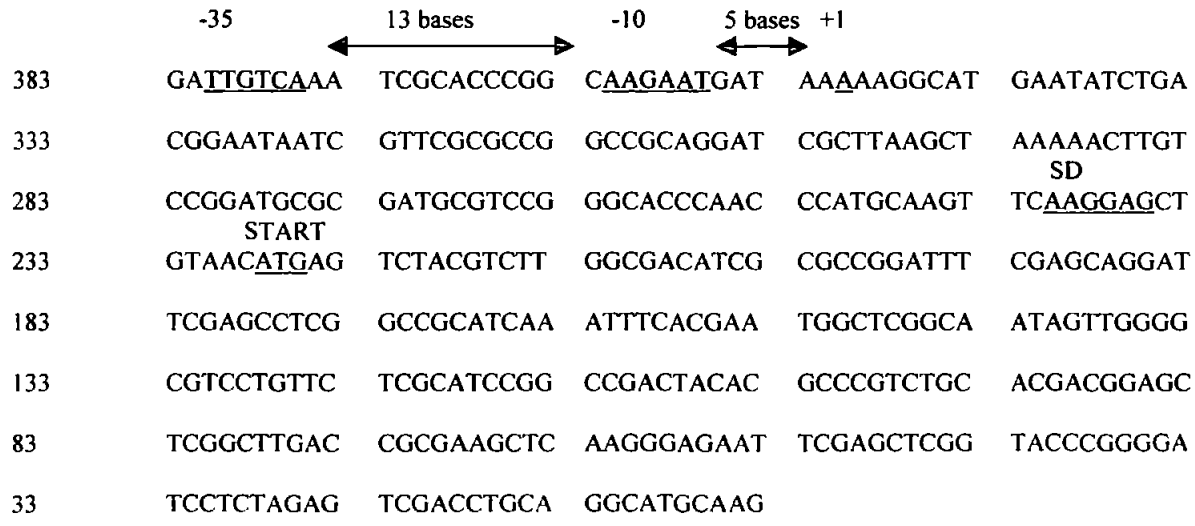


Fig.6.3 Partial Nucleotide Sequence of a Putative Gene Encoded by ORF 1. The open reading frame has been extended to include putative -35 and -10 regions, which have been highlighted on the sequence. A putative ribosome binding site (SD) has also been found and the initiation codon ATG is also highlighted in red. Those sequences in green represent vector DNA as the sequence has probably been disrupted by the *Eco*RI site used to clone the fragment into λGT11.

ORF 2 -35 12 bases -10 6 bases +1

411	TCTCTATTGT	CACGGCAAGT	CGTTTACCAT	GCTGTTCAAG	CGCGATCTTG
461	CGCCACCTGC	GCGCCGTGCC	CGCGCGCCCG	CGGCCAGGCA	CGGGTTTCGC
				SD	START
511	TACGATATCC	GTAGCACAAC	CCGCTTCGGC	AAGAAGGGAG	TCGCAATGTC
561	TGAAGTCAAC	AAGGAGAAAC	TGATGTCGGA	TATCAAAACT	GTTCTCGCGG
611	ACGCGGAAGA	TCTGCTGAAG	CAGGCCGCGA	GCAGCACGGG	CGACCGTGCG
661	ACCGAGCTGC	GCGAGAAGGC	GCTCGCACGC	CTGAAACAGG	CGAAGGAGAA
711	GGCGACCGAT	GTCCAGGTGG	TCGTGGTCGA	AAAAGGCAAG	AAGGCGGCGC
761	GCGCCACCGA	CGACTACGTG	CACGAGCATC	CGTGGACGTC	GATCGGCATC
					STOP
811	GCGGCCGGCG	TCGGCGTATT	GATCGGCCTG	CTGATCAACC	GCAAGTAACG
	ORF 3				
861	CGGCACGCCC	CCGCGAGAGA	CAGCCCCCGG	CTCGCCGCCG	CGTGACGGCC
911	GGGGCCCTGC	GCTCCGGCCC	GCCGCGCCCA	GTTGCGACGC	CGGCCGTTCT
		START			
961	CCGCGCGCAA	CCTCTTTCAT	GACGACAGAA	ACCTCATCGC	ACCAGTCCGG
1011	GCACGGACCG	CTGCGCCGCC	TGCTCGGCTC	GGTGCTCGCG	CTCCTGCAGA
1061	CCCGGCTCGA	ACTCGTCGGC	ATCGAACTCG	CCGAGGAGAA	GGAACGCCTG
1111	ATGGGCGTAC	TCTTTGTCGG	GCTCGCCGCG	ATGATGCTCG	CGACGATGGC
1161	GCTCATCAGC	CTGACGGTGC	TCATTGCGAT	CGCGTTCTGG	GATACGTACC
1211	GCTGGCAATC	GCTCGCCGTC	GTCACCGCGC	TCTACGCGCT	CGGCGCACTC
1261	GCCTGCTGGC	TGAAGGCGCG	CTCGGGACTG	CGCGACGCGC	CCAGCGTGTT
1311	CGAGGCGACG	CTGAACGAAC	TCGAAAAGGA	CCGCGAGCTG	TTCCGCGGCA
			STOP		
1361	AGCCGTGA				

Fig.6.4 The Complete Nucleotide Sequences of Putative Genes Encoded by ORFs 2 and 3. These genes have been included together due to their proximity, but on closer analysis putative -35 and -10 promoters were identified by extending ORF 2, but no clear promoters could be identified for ORF 3, nor a ribosome binding site close to the transcriptional start codon, ATG. A putative SD sequence was identified at position 875. It may be that the two putative proteins are transcribed together especially considering that they are in the same reading frame.

2604	CCGGCGCGCC	CCGACGTGCG	CGCCTTCCTC	GTAACCGCGC	GCGGCGTGGG
		-35		-10 +1	
2554	CGCATCGAAC	GATACGGCCG	AGTGGCTGCA	ATTGCAGGAT	GCGCTCGACG
2504	GCGGGCGCGT	CGAGCCCCGG	TGGCGTCGGG	TCGTGGGACG	GCCGGCATAG
	SD		START		
2454	GCAACGGGAG	GCATGAGCGA	TGCCATGTAC	GGGGCGCAAT	TCGCCCTCGA
2404	GAGGCTTCAC	CTTAATCGAG	GTGGTCGTCG	CGATCGCGAT	CGTGGCGGTG
2354	CTCGCGGCGT	TTGCCGTGCC	GTCGTATCGC	AGTTATGTCG	AACGGGTGAA
2304	CCGATTGACC	GCGGTTGCCG	CGCTCTATCG	CGCCGCGCAA	TACGTGGACG
2254	CGTTCGGCGA	CGCGCCGCCC	ACCGCGTTGC	CGGAAGGGGT	GAACAGGGCG
2204	CCCGAATCCG	GCAAGCTCGT	CTATGTGCTG	CGGATCATGT	TCGACGACGC
2154	GCGCGGCGGA	TACGCACTGG	AGGCGCGTCC	CGCCGCCGAT	GGTGCAATGC
2104	GGGACGACAG	ATGCGGCGTC	TACGTGCTGC	ATGCGGACGG	CACGAGCGAG
2054	AATCGCGTGG	CCGGAGGCGT	CGCGCTCGAT	GGCGGTGCGG	CGGAGGGCGA
			STOP	TERMINATION	LOOP
2004	TGCCTGCTGG	CGAACAGGCT	AGCTGCGCAG	CGCCGGGCGT	GTCGATCTCC
1954	CCGGCGCCGG	CAGCCGGCCG	GCGCGGTCGG	CGCGCCGGCT	CGGTCGACGG

Fig.6.5 The Complete Nucleotide Sequence of an Open Reading Frame (ORF 4) Encoding a Putative Pilin Gene in *B.pseudomallei.* The putative reading frame has been extended to include a putative -10 and -35 sequences. Also shown is the transcriptional start nucleotide (+1) the putative Shine-Dalgarno sequence (SD) and a possible transcription termination loop (LOOP).

6.5 THE PRIMARY AMINO ACID SEQUENCES OF THE PUTATIVE PROTEINS.

6.5.1 ORF 1 Putative Peroxidase.

MSLRLGDIAPDFEQDSSLGRIKFHEWLGNSWGVLFSHPADYTPVCTT ELGLTAKLKGEF (stop codon not found, the reading frame goes beyond the available sequence).

6.5.2 ORF 2 Putative Protein.

MSEVNKEKLMSDIKTVLADAEDLLKQAASSTGDRATELREKALARL KQAKEKATDVQVVVVEKGKKAARATDDYVHEHPWTSIGIAAGVGV LIGLLINRK.

6.5.3 ORF 3 Putative Protein.

MTTETSSHQSGHGPLRRLGSVLALLQTRLELVGIELAEEKERLMGV LFVGLAAMMLATMALISLTVLIAIAFWDTYRWQSLAVVTALYALGA LACWLKARSGLRDAPSVFEATLNELEKDRELFRGKP.

6.5.4 ORF 4 Putative Pilin Gene.

MPCTGRNSPSRGFTLIEVVVAIAVAVLAAFAVPSYRSYVERVNRLTA VAALYRAAQYVDAFGDAPPTALPEGVNRAPESGKLVYVLRIMFDDA RGGYALEARPAADGAMRDDRCGVYVLHADGTRENRAGGVALDGG AAEGDACWRTG.

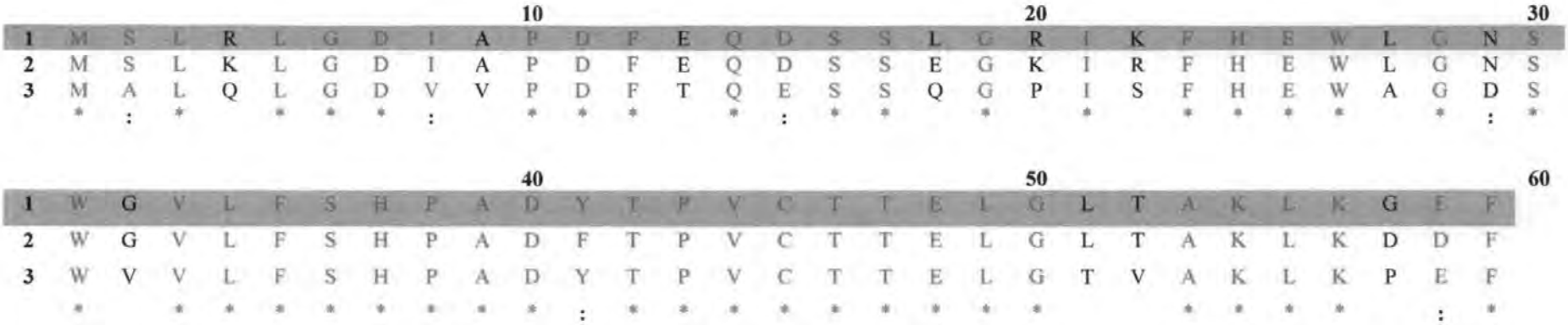
Table 6.6 Codon Usage of the Putative Genes Identified in the 5 kbp *Eco*RI Generated Fragment.

CODON	AMINO ACID	USAGE ORF 1	%	USAGE ORF 2	%	USAGE ORF 3	%	USAGE ORF 4	%	TOTAL USAGE	%
GCA	A	-	-	1	1.0	-	-	2	1.33	3	0.7
GCC	A	1	1.69	3	3.0	5	4.17	7	4.67	16	3.7
GCG	A	2	3.39	10	10.0	9	7.5	18	12.0	39	9.07
GCT	A	-	-	-	-	-	-	-	-	-	-
AGA	R	-	-	-	-	-	-	2	1.33	2	0.47
AGG	R	-	-	-	-	-	-	1	0.67	1	0.23
CGA	R	-	-	-	-	-	-	2	1.33	2	0.47
CGC	R	1	1.69	4	4.0	6	5.0	6	4.0	17	3.95
CGG	R	-	-	-	-	1	0.83	3	2.0	4	0.93
CGT	R	1	1.69	1	1.0	-	-	1	0.67	3	0.7
AAC	N	-	-	2	2.0	1	0.83	2	1.33	5	1.16
AAT	N	1	1.69	-	-	-	-	2	1.33	3	0.7
GAC	D	2	3.39	4	4.0	2	1.67	7	4.67	15	3.49
GAT	D	2	3.39	3	3.0	1	0.83	3	2.0	9	2.09
TGC	C	1	1.69	-	-	-	-	2	1.33	3	0.7
TGT	C	-	-	-	-	-	-	1	0.67	1	0.23
CAA	Q	-	-	-	-	1	0.83	1	0.67	2	0.47
CAG	Q	1	1.69	3	3.0	2	1.67	-	-	6	1.4
GAA	E	2	3.39	3	3.0	4	3.33	3	2.0	12	3.95
GAG	E	2	3.39	5	5.0	3	2.5	4	2.67	14	4.4
GGA	G	1	1.69	-	-	1	0.83	2	1.33	4	0.93
GGC	G	5	8.48	6	6.0	7	5.83	10	6.67	28	6.5
GGG	G	-	-	-	-	2	1.67	2	1.33	4	0.93
GGT	G	-	-	-	-	-	-	2	1.33	2	0.47
CAC	H	1	1.69	1	1.0	3	2.5	-	-	5	1.16
CAT	H	1	1.69	1	1.0	1	0.83	1	0.67	4	0.93
ATA	I	-	-	-	-	-	-	-	-	-	-
ATC	I	2	3.39	5	5.0	7	5.83	4	2.67	18	4.19
ATT	I	-	-	-	-	1	0.83	-	-	1	0.23
CTA	L	1	1.69	-	-	-	-	-	-	1	0.23
CTC	L	4	6.78	2	2.0	12	10.0	4	2.67	22	5.12
CTG	L	1	1.69	7	7.0	7	5.83	3	2.0	18	4.19
CTT	L	1	1.69	-	-	-	-	-	-	1	0.23
TTA	L	-	-	-	-	-	-	1	0.67	1	0.23
TTG	L	1	1.69	1	1.0	1	0.83	2	1.33	5	1.16
AAA	K	1	1.69	4	4.0	-	-	-	-	5	1.16
AAG	K	2	3.39	8	8.0	2	1.67	1	0.67	13	3.0
ATG	M	1	1.69	2	2.0	5	4.17	3	2.0	11	2.56
TTC	F	3	5.01	-	-	1	0.83	3	2.0	7	1.63
TTT	F	1	1.69	-	-	1	0.83	1	0.67	3	0.7
CCA	P	-	-	-	-	-	-	1	0.67	1	0.23
CCC	P	1	1.69	-	-	-	-	4	2.67	5	1.16
CCG	P	2	3.39	1	1.0	2	1.67	3	2.0	8	1.86
CCT	P	-	-	-	-	-	-	-	-	-	-
AGC	S	1	1.69	2	2.0	1	0.83	-	-	4	0.93
AGT	S	2	3.39	-	-	-	-	1	0.67	3	0.7
TCA	S	-	-	-	-	1	0.83	-	-	1	0.23
TCC	S	-	-	-	-	1	0.83	1	0.67	2	0.47
TCG	S	2	3.39	2	2.0	4	3.33	3	2.0	11	2.56
TCT	S	-	-	1	1.0	-	-	-	-	1	0.23
ACA	T	-	-	-	-	1	0.83	1	0.67	2	0.47
ACC	T	1	1.69	3	3.0	4	3.33	3	2.0	11	2.56
ACG	T	3	5.01	2	2.0	5	4.17	1	0.67	11	2.56
ACT	T	-	-	1	1.0	-	-	-	-	1	0.23
TAC	Y	1	1.69	1	1.0	2	1.67	4	2.67	8	1.86
TAT	Y	-	-	-	-	-	-	4	2.67	4	0.93
GTA	V	-	-	1	1.0	2	1.67	-	-	3	0.7
GTC	V	2	3.39	4	4.0	5	4.17	6	4.0	17	3.95
GTG	V	-	-	4	4.0	3	2.5	10	6.67	17	3.95
GTT	V	-	-	1	1.0	-	-	1	0.67	2	0.47
TGG	W	2	3.39	1	1.0	3	2.5	1	0.67	7	1.63
TOTAL		59		100		120		150		429	

6.6 MULTIPLE SEQUENCE ALIGNMENTS.

Once proteins with similar homology have been identified from the Web-based BLAST searches, those regions exhibiting homology can be further analysed by aligning the protein sequences. This can be carried out by most relevant Web-based protein analysis sites. In the case of this project multiple sequence alignments were carried out using the Web-based Network Protein Sequence @nalysis at Pole Bio-Informatique Lyonnais. The program used was CLUSTAL W multiple sequence alignment (Thompson *et al.*, 1994). This program functions to align proteins through sequence weighting, position-specific gap penalties and weight matrix choice (Thompson *et al.*, 1994). For each of the putative genes identified those proteins bearing the greatest homology were used in a sequence alignments. Complete amino acid sequences for the proteins from other bacterial species were derived from searching the BLAST database by their Accession numbers. Multiple alignments can be found in Figs.6.6, 6.7, 6.8 and 6.9.

It was decided to analyse the putative pilin gene in more detail as the other putative genes, although interesting, are not relevant to the aims of the project. That is to clone and characterise putative virulence genes from *B.pseudomallei*.



	10																			20										30
1	M	S	L	R	L	G	D	I	A	P	D	F	E	Q	D	S	S	L	G	R	I	K	F	H	E	W	L	G	N	S
2	M	S	L	K	L	G	D	I	A	P	D	F	E	Q	D	S	S	E	G	K	I	R	F	H	E	W	L	G	N	S
3	M	A	L	Q	L	G	D	V	V	P	D	F	T	Q	E	S	S	Q	G	P	I	S	F	H	E	W	A	G	D	S
	*	:	*		*	*	*	:		*	*	*		*	:	*	*		*		*		*	*	*	*		*	:	*

	40																			50										60
1	W	G	V	L	F	S	H	P	A	D	Y	T	P	V	C	T	T	E	L	G	L	T	A	K	L	K	G	E	F
2	W	G	V	L	F	S	H	P	A	D	F	T	P	V	C	T	T	E	L	G	L	T	A	K	L	K	D	D	F
3	W	V	V	L	F	S	H	P	A	D	Y	T	P	V	C	T	T	E	L	G	T	V	A	K	L	K	P	E	F
	*		*	*	*	*	*	*	*	*	:	*	*	*	*	*	*	*	*	*			*	*	*	*		:	*

Fig.6.6 Alignment of Similar Sequences with the Putative Peroxidase from *B.pseudomallei*. **1**. *B.pseudomallei* putative Peroxidase (59 amino acids available). **2**. *Lsf*A *P.putida* (total sequence 212 amino acids). **3**. Rehydrin *Synecocytis* sp. (total sequence 211 amino acids).

* Asterisks denote identical amino acids; : denotes similar amino acids.

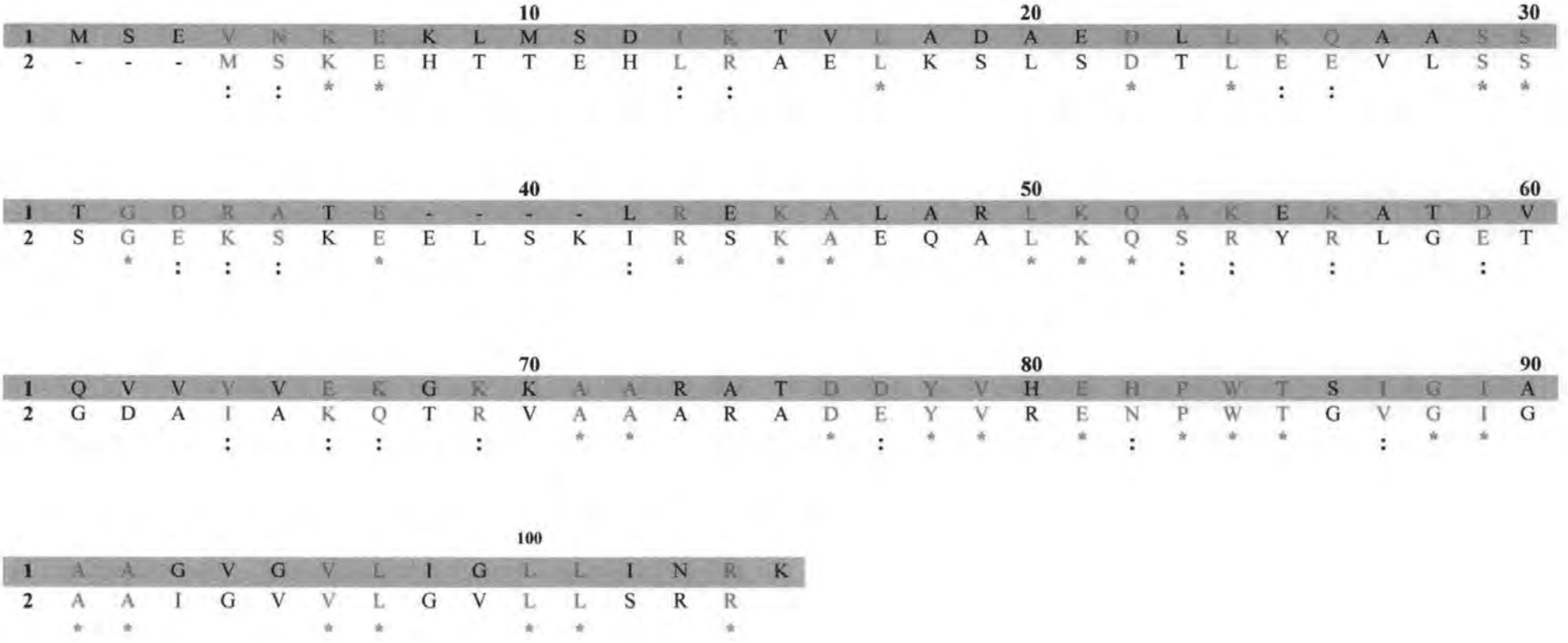


Fig.6.7 Sequence Alignment of the Putative Protein (ORF 2) and a Hypothetical Protein in *E.coli* (Accession AE000352). 1. Putative Protein from *B.pseudomallei* (ORF 2); **2**. Hypothetical Protein *E.coli*. * Asterisks denote identical amino acids; **:** denotes similar amino acids.

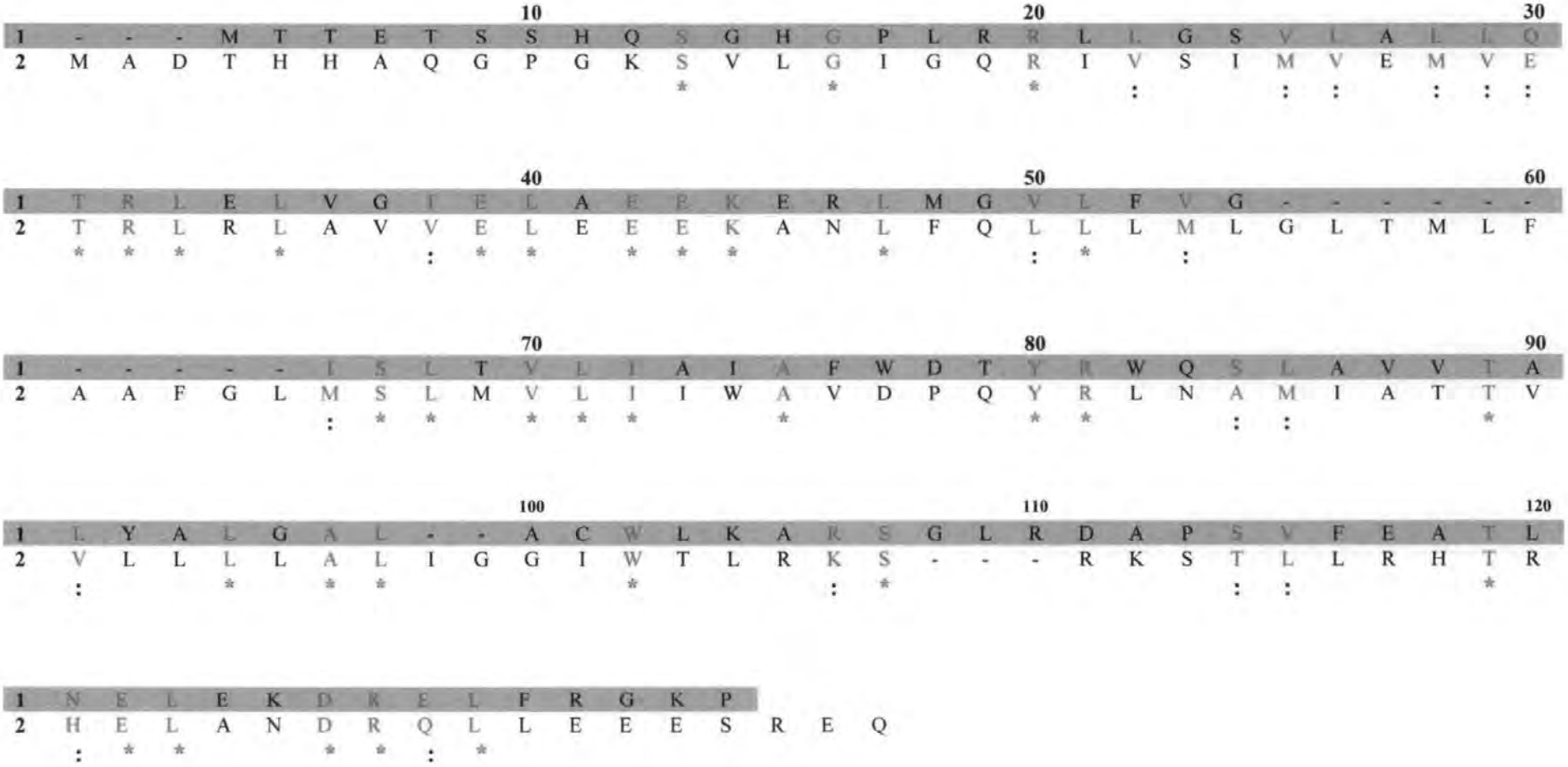


Fig.6.8 Sequence Alignment of the Putative Protein (ORF 3) with a Hypothetical Protein from *E.coli* (Accession AE000392). 1: Putative Protein (ORF 3) from *B.pseudomallei*. **2**: Hypothetical 12.5 kDa Protein (Accession AE000392) from *E.coli*. * Asterisks denote identical amino acids; : denotes similar amino acids.

Fig.6.9 Multiple Sequence Alignment of Proteins with a High Homology to the Putative Pilin Gene Identified on a 5 kbp *B.pseudomallei* DNA Fragment. CLUSTALW Multiple Sequence Alignment (Thompson *et al.*, 1994). **1**: *B.pseudomallei* putative pilin gene. **2**: *Acinetobacter calcoaceticus* (*com*F) (Accession AF027189). **3**: *P.aeruginosa* (*pil*E) (Accession U02552). **4**: *Neisseria gonorrhoeae* (*pil*E) (Accession Z69262). **5**: *Moraxella bovis* (Prepilin) (Accession L32968). **6**: *Haemophilus influenzae* (Prepilin) (Accession U32715).

	10																	20											30	
1	M	P	C	T	G	R	N	S	P	S	R	G	F	T	L	I	E	V	V	V	A	I	A	I	V	A	V	L	A	A
2	-	-	-	-	-	M	R	V	R	H	-	G	F	T	L	I	E	L	M	I	V	V	A	I	I	G	I	L	A	A
3	-	-	-	-	-	M	R	T	R	Q	K	G	F	T	L	L	E	M	V	V	V	V	A	V	I	G	I	L	L	G
4	-	-	-	-	-	M	N	T	L	Q	K	G	F	T	L	I	E	L	M	I	V	I	A	I	V	G	I	L	A	A
5	-	-	-	-	-	M	N	A	Q	K	-	G	F	T	L	I	E	L	M	I	V	I	A	I	I	G	I	L	A	A
6	M	K	L	T	T	L	Q	T	L	K	K	G	F	T	L	I	E	L	M	I	V	I	A	I	I	A	I	L	A	T
												*	*	*	*	:	*	:	:	:		:	*	:	:		:	:		

	40											50																60	
1	F	A	V	P	S	Y	R	S	Y	V	E	R	V	N	R	L	T	A	V	A	A	L	Y	R	A	A	Q	Y	D	A
2	I	A	Y	P	S	Y	Q	N	Y	K	I	R	V	N	R	S	D	V	Q	S	E	L	I	R	V	S	Q	Q	Q	S
3	I	A	I	P	S	Y	Q	N	Y	V	I	R	S	N	R	T	E	G	Q	A	L	L	S	D	A	A	A	R	E	R
4	V	A	L	P	A	Y	Q	D	Y	T	A	R	A	Q	V	S	E	A	I	L	L	A	E	G	Q	K	S	A	T	E
5	I	A	L	P	A	Y	Q	D	Y	I	S	K	S	Q	T	T	R	V	V	G	E	L	A	A	G	K	T	A	D	A
6	I	A	I	P	S	Y	Q	N	Y	T	K	K	A	A	V	S	E	L	L	Q	A	S	A	P	Y	K	A	D	E	L
		*		*	:	*	:		*			:																		

	70									80														90					
1	F	G	D	A	P	P	T	A	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	P	-	-	E	-
2	Y	K	M	V	N	H	G	Y	A	-	-	-	-	-	-	-	-	-	-	-	-	N	A	S	L	A	G	L	G	-
3	Y	Y	S	Q	N	P	G	V	G	Y	-	-	-	-	-	-	-	-	-	-	-	T	K	D	V	A	K	L	G	M
4	Y	Y	L	N	H	G	K	W	P	E	N	N	T	S	A	G	V	A	S	P	P	S	D	I	K	G	K	Y	V	Q
5	A	L	F	E	G	K	K	P	V	L	A	A	P	A	A	G	-	-	-	-	-	N	T	T	T	E	N	I	G	L
6	C	V	Y	S	T	N	E	T	T	-	-	-	-	-	-	-	-	-	-	-	-	S	C	T	G	G	K	N	G	I

	100										110													120					
1	-	-	-	G	V	N	R	A	P	E	S	G	K	L	V	Y	V	L	R	I	M	F	D	D	A	R	G	Y	A	L
2	-	-	N	A	A	S	Y	P	V	-	T	G	T	A	F	Y	T	L	N	L	A	V	D	S	D	N	Q	Y	V	L
3	-	-	S	S	A	N	S	P	-	-	-	-	N	N	L	Y	N	L	T	I	A	T	P	T	S	T	T	Y	T	L
4	K	V	E	V	A	K	G	V	V	T	A	E	M	A	S	T	G	V	N	K	E	I	K	D	K	K	L	L	W	A
5	T	T	D	G	G	A	T	P	R	S	N	L	M	S	A	V	A	L	G	G	F	K	N	N	G	A	G	I	T	G
6	A	A	D	I	K	T	A	K	G	Y	V	A	S	V	I	T	Q	S	G	G	I	T	V	K	G	N	G	L	A	N

	130															140											150			
1	E	A	R	P	A	A	D	G	A	M	R	D	D	R	C	G	V	Y	V	L	H	A	D	G	T	R	E	N	R	V
2	T	A	T	P	I	S	S	-	-	-	-	-	-	T	M	Q	A	G	N	G	I	I	C	L	N	Q	D	G	Q	K
3	T	A	T	P	I	N	S	Q	-	-	-	-	-	T	R	D	K	T	C	G	K	L	T	L	N	Q	L	G	E	R
4	K	R	Q	D	G	S	V	K	W	F	C	G	Q	P	V	T	R	G	A	G	N	A	G	K	A	D	D	V	T	K
5	T	L	G	N	R	A	N	K	D	I	V	G	A	E	I	I	Q	S	R	D	Q	Q	G	V	W	T	C	E	I	D
6	M	E	Y	I	L	Q	A	-	-	-	-	-	-	-	-	-	K	G	N	A	A	A	G	V	T	W	T	T	T	C

	160															170				
1	A	G	G	V	A	L	D	G	G	A	A	E	G	D	A	C	W	R	T	G
2	F	W	S	K	A	A	T	A	C	S	L	S	S	N	S	T	W	D	G	R
3	-	G	A	A	G	K	T	G	N	N	S	T	V	N	D	C	W	R		
4	A	G	N	D	N	E	K	I	N	T	K	H	L	P	S	T	C	R	D	N
5	-	K	K	T	A	P	G	W	K	D	K	F	I	P	T	G	C	K	A	K
6	K	G	T	D	A	S	L	F	P	A	N	F	C	G	S	V	T	K		

6.7 PREDICTION OF THE BIOCHEMICAL NATURE OF THE PUTATIVE PILIN MOLECULE BY WEB-BASED PROTEIN ANALYSIS.

The amino acid sequence of the putative pilin gene was entered into several Web-based protein analysis programs to analyse the biochemical nature of the sequence.

6.7.1 Secondary Structure Prediction.

Using the ProtParam tool at ExPASy on the internet, the molecular weight of the protein was predicted to be 15,806 kDa with a theoretical pI of 5.81. A self-optimised method for protein secondary structure prediction (SOPM) (Geourjon and Deleage, 1994) was carried out by Pole Bio-Informatique Lyonnais (NBIL) on the internet. By analysing the distribution and nature of the amino acids within a sequence the program can make a number of predictions about its secondary structure. Each prediction scheme utilises a different algorithm to calculate the probability that a given sequence of amino acids will be an α-helix, β-sheet structure or a turn. The results of this can be seen in Fig.6.10, which is a consensus sequence of a number of different predictions.

6.7.2 Physio-Chemical Profiles.

The Network Protein Sequence @analysis (NPS@) program available at Pole Bio-Informatique Lyonnais (NBIL) calculates different physio-chemical profiles for the given protein. Amongst these are hydrophilicity and hydrophobicity plots, which relates to the water relationship of different portions of the molecule. Hydropathy plots are calculated by the evaluation of overlapping segments of a given amino acid sequence, with the average score of each segment being plotted at its mid-point (Doolittle, 1986). Such plots are useful in determining whether a portion of a sequence is likely to be on the outside, or conversely on the inside of a protein. It also has uses in determining the presence of certain features, such as membrane spanning domains. NPS@ also utilises the method of Hopp and Woods (1981) to

determine hydrophilicity and that of Kyte and Doolittle (1982) to determine hydrophobicity (Fig. 6.11). As well as a hydropathy plots both flexibility and antigenicity plots were carried out as part of the program. Although interesting, their relative value to in this initial characterisation of the protein product of the putative pilin gene is questionable.

6.7.3 Transmembrane Helices Prediction.

In addition to the other analyses carried out, a program exists on the internet that calculates, based on the sequence of amino acids and their nature, the probability of transmembrane helices present in the protein (TMHMM 1.0). This is based on the method of Soonhammer *et al.* (1998). The computer output result for this can be seen in Fig.6.12.

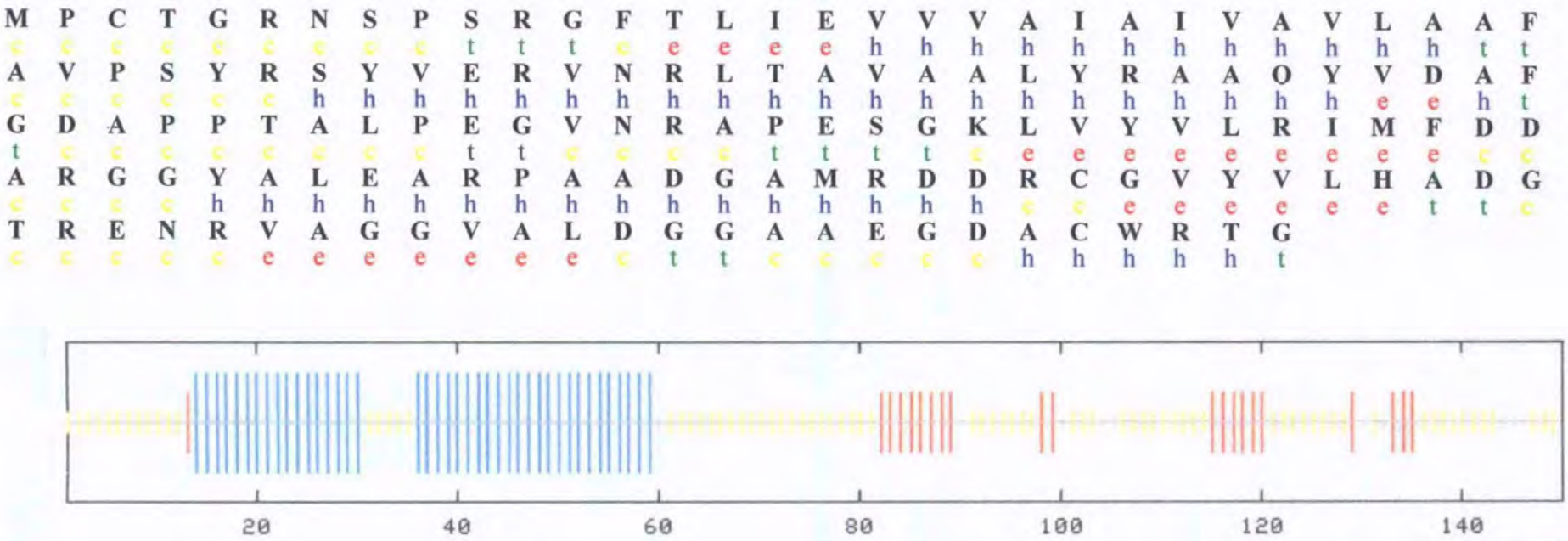


Fig.6.10 Self-Optimised Method (SOPM) for Protein Secondary Structure Prediction of the Putative Pilin Gene Protein Product Carried Out by PBIL SOPM:

Alpha helix	(h):	36.67%
310 helix	(g):	0.00%
Pi helix	(I):	0.00%
Beta bridge	(b):	0.00%
Extended strand	(e):	18.67%
Beta turn	(t):	12.00%
Bend region	(s):	0.00%
Random coil	(c):	32.67%
Ambiguous states	(?):	0.00%

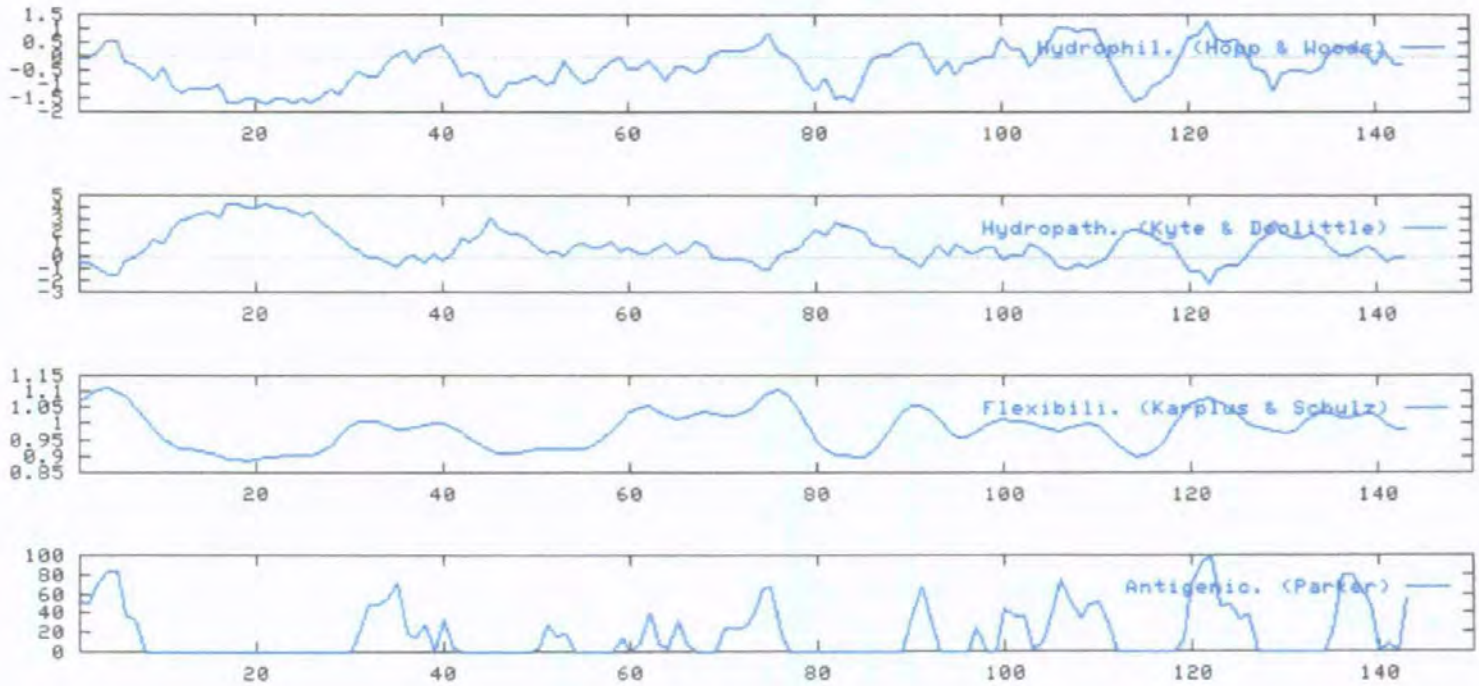


Fig.6.11 Physio-Chemical Profiles of the Putative Pilin Gene Amino Acid Sequence Carried Out By NPS@ at Pole Bio-Informatique Lyonnais (NBIL). Profiles (from top to bottom): Hydrophilicity (Hopp and Woods, 1981), Hydrophobicity (Kyte and Doolittle, 1982), Flexibility (Karplus and Schulz, 1985) and Antigenicity (Parker *et al.*, 1982).

```
Sequence TMHMM1.0 inside      1      12
Sequence TMHMM1.0 TMhelix    13      35
Sequence TMHMM1.0 outside    36     147
```

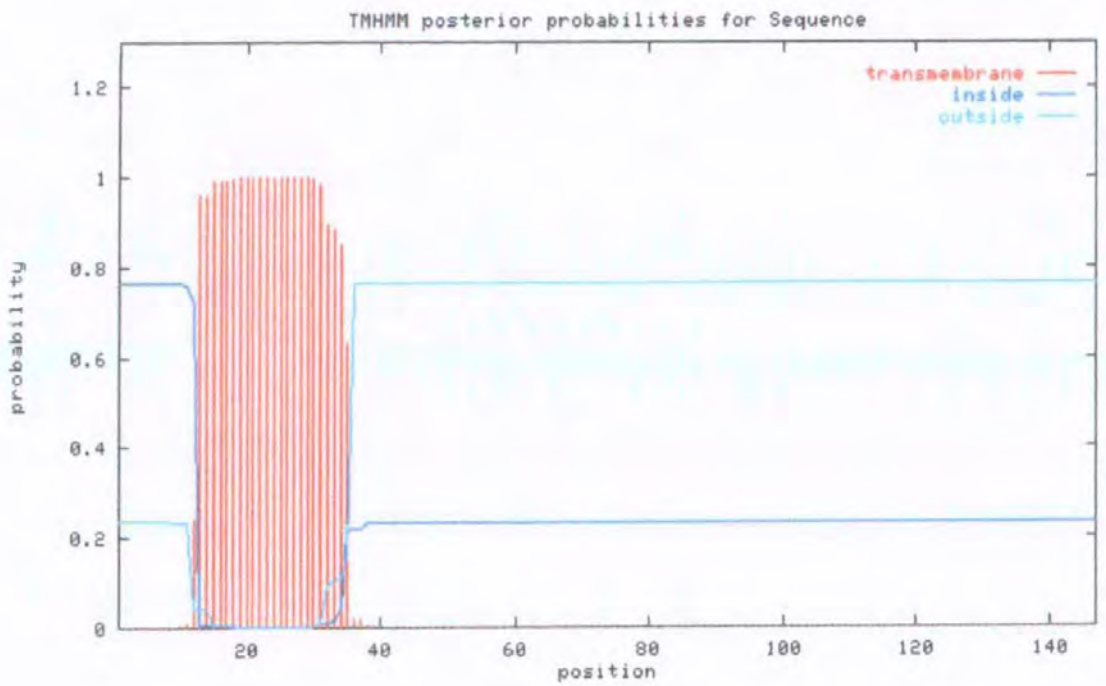


Fig.6.12 Computer Output of a Transmembrane Helices Prediction Program (TMHMM 1.0) (Soonhammer et al., 1998).

6.8 DISCUSSION.

The putative peroxidase encoded by ORF 1 had a very high homology with *lsf*A of *P.putida*, 96% homology at the amino acid level and at the nucleotide level there was 80% identity for the available sequence. *Lsf*A is part of a locus (*ssu*) involved in the assimilation of sulphonate and sulphate ester-sulphate in *P.putida*. The *ssu* locus comprises seven genes of which *lsf*A is the first. All bacteria require sulphur during growth for the synthesis of protein and of a variety of essential enzyme co-factors, however if not present in this form, bacteria must derive it from another form. Sulphonates play an important part in the soil sulphur cycle and Leisenger *et al.* (unpublished) have shown in *P.putida* and *P.aeruginosa* that aromatic sulphonates may be desulphonated by this novel system which is repressed when favoured sulphate sources are present. As *B.pseudomallei* is found in soil it would appear it too might have this novel system, although further work sequencing the DNA up-stream of the putative gene is needed to investigate whether *B.pseudomallei* possesses the other genes found in the *ssu* locus of *P.putida*. Although interesting, it was beyond the aims of the study to investigate this further.

The two putative genes encoded by ORFs 2 and 3 have a high homology to two intergenic regions found in the EXUR-TDCC region of *E.coli*, which encode putative proteins with unknown function. It is unlikely that these were responsible for the immunopositivity observed and so further work on these putative genes was discontinued.

However, a putative type IV pilin gene was identified, encoded by ORF 4. Type IV pili are adhesins expressed by a number of diverse gram-negative pathogens that cause a variety of disease in animals and humans. These include *P.aeruginosa*, *Neisseria gonorrhoeae*, *N.meningitidis*, *Moraxella bovis*, *Dichelobacter nodosus*, *Eikenella corrodens* and the more distantly related type IV producer *Vibrio cholerae* (Ottow, 1975; Paranchych and Frost, 1988; Shaw and Taylor, 1990; Rao and Progulske-Fox, 1993). The contribution of pili to virulence

lies primarily in their ability to promote attachment to various types of receptors during tissue colonisation, although now they have also been deemed responsible for a novel mode of flagella-independent surface translocation called "twitching motility" (Bradley, 1980) and can act as receptors for various bacteriophages (Bradley, 1972). All type IV pili use a similar secretion and assembly machinery, made up of at least a dozen polypeptides which are functionally and evolutionary related to components of the type II secretion pathway or GSP from gram-negative bacteria (Pugsley, 1993; Russel, 1998).

Type IV pili are characterised by the occurrence of N-methylated amino acids (phenylalanine or methionine) as the first amino acid of the mature pilin structural subunit. An extensive region of amino acid homology is present at the N-terminal end, which is characteristically hydrophobic (Patrick and Larkin, 1995). Pilin genes are essentially synthesised as precursors with unique, short, basic, amino-terminal leader peptides. Leader sequences are removed by endoproteolytic cleavage between an invariant glycine residue and a phenylalanine prior to assembly of the pilin monomers into pili. The N-terminal segment is proposed to act as a signal that allows nascent molecules to be recognised for transfer to a specific location in the cell and serves as a structural element that preserves the integrity of the pilus filament (Russell and Darzins, 1994).

The type IV pilin family can further be divided into two groups, type A and type B. Type A consists of pilins from *P.aeruginosa*, *N.gonorrhoeae*, *N.meningitidis*, *M.bovis* and *D. nodosus* and Type B include the subunits of the toxin-coregulated pili (TCP) of *V.cholera* and subunits of the bundle-forming pili (BFP) of *E.coli* (Strom and Lory, 1993). Type B pili are synthesised with longer leader peptides than those of group A, between 13 and 25 amino acids. The putative gene in *B.pseudomallei* has a leader sequence of 12 amino acids, which is fairly long for type A. However, tyrosine residues have been demonstrated in all type A pili, at

positions 24 and 27 in the hydrophobic domain, which have not been identified in type B pili. The putative type IV pilin gene in *B.pseudomallei* possesses such tyrosine residues at positions 24 and 27 (Fig. 6.9) and so is most likely, type A. The tyrosine residues have been shown to be at a dimer/dimer interface in both native pili and reassembled pilin filaments, only present in type A pili.

Type IV pili are related through similarities in the primary amino acid sequences of the structural subunits, a conserved assembly machinery and a similar mechanism of transcriptional regulation (Strom and Lory, 1993). The middle of these sequences is considerably less homologous and contains variable domains that comprise the antigenic epitopes of pili (Strom and Lory, 1993). However, another conserved feature is a pair of cysteine residues that function to form a disulphide loop near the carboxy terminus (C-terminus) of the sequence. Studies concerning pilin structure have found that the C-terminus is exposed at the surface of the filament, whereas the hydrophobic N-terminal domain forms the pilus inner core, and the two are interconnected by a α-helical region (Forest and Tainer, 1997). A pair of cysteine residues are present at the C-terminus in the putative *B.pseudomallei* pilin sequence (Fig.6.9).

It has been estimated that as many as 20-40 chromosomal genes may be required for type IV pilus production in *P.aeruginosa* (Hobbs *et al.*, 1993) and that the genes required for pilus biogenesis in general are not tightly clustered (Russell and Darzins, 1994). The putative pilin gene has homology to *pil*E of *P.aeruginosa*, which was shown by Russell and Darzins (1994) to be required for pilus biogenesis, as *pil*E insertion mutants lacked pili and were twitching-motility deficient.

The N-terminal domain of the *B.pseudomallei* putative pilin gene contains a relatively hydrophobic region (Fig.6.11), similar to other pilin genes, which has been shown in other studies to be important for membrane translocation; and for *pil*A of *P.aeruginosa* has been shown to be important in maintaining the structural integrity of pilin subunit-subunit interactions (Russell and Darzins, 1994).

The putative pilin gene identified in *B.pseudmallei* is a potential virulence gene that may have been responsible for the immunopositive nature of the clone, λBPGT1. Therefore it was decided to continue work with the putative gene to attempt expression of a gene product to facilitate further characterisation.

CHAPTER 7

CHARACTERISATION OF THE PUTATIVE PILIN GENE FROM *B.pseudomallei.*

7.1 INTRODUCTION.

In gene cloning experiments it is necessary to prove that the cloned product is a true representative of that which occurs in the organism from which the cloned DNA originates. As although precautions are taken, there are opportunities throughout the isolation and cloning procedures for contamination to be introduced and for mutations to arise, especially if the product is toxic to the host cells. To assess the cloning fidelity, the basic procedure is to use the foreign DNA sequence as a hybridisation probe against the genomic DNA of the organism from which it was supposedly derived, and to demonstrate that the cloned sequence anneals to a homologous genetic DNA fragment through a Southern blot procedure (Southern, 1975). The detection of homologous DNA sequences through the method of Southern blotting has made a tremendous contribution to molecular biology as it is essential for understanding basic problems such as gene structure, gene expression and genome organisation (Meinkoth and Wahl, 1984). Probes used in these procedures, in general, can be purified cDNA, a cloned fragment of genomic DNA or even a PCR product but in practice the end result of hybridisation is clearer if DNA used as the probe is under 2000 bp in length and separated from any plasmid vector used for cloning (Darbre, 1999).

Previous attempts at determining the protein product responsible for the immunopositivity of the λBPGT1 clone and its subclone pBPGT2 proved unsuccessful (Chapter 5). However, it

was discussed (5.8) that a possible fusion with β-galactosidase in λGT11 may have been responsible for the immunopositive nature of the original clone and on subcloning the fragment such expression was lost. On sequencing the 5 kbp fragment in pBPGT2 a putative peroxidase gene was discovered (Chapter 6) at the beginning of the sequence and so it may have been possible that the part of the gene present was in-frame with the β-galactosidase system. Therefore, the fusion protein created could be responsible for immunopositivity, not through encoding an immunopositive antigen, but through a reaction with the detection system itself. Immunological screening carried out in Chapter 4 used a peroxidase-conjugated secondary antibody and the developing solution therefore included hydrogen peroxide. It is feasible that the immunopositive clone produced a β-galactosidase fusion with the putative peroxidase and this was what reacted with the hydrogen peroxide, perhaps therefore responsible for the unusual "halo" of reaction encircling the plaque. It was therefore necessary to investigate this further.

Unknowingly at the time, subcloning the 5 kbp fragment further into pCD1-5 disrupted the pilin gene sequence. The initial part of the sequence was contained in pCD3 with the remainder of the gene in pCD2. Therefore, to aid expression of this gene a smaller fragment containing just the putative pilin gene was subcloned into a plasmid vector. Often it is the case that genomic DNA fragments used to construct libraries, especially those from bacteriophage vectors are often much larger than the genetic sequence of interest. Consequently, not only will the cloned DNA fragment contain non-essential intergenic regions, it may also contain other functional gene sequences. Where the goal of cloning gene sequences from an organism is to isolate and characterise specific encoded protein molecules, these non-essential regions of DNA should ideally be removed from the initial recombinant molecule, which can be achieved by subcloning procedures (Singer and Berg, 1991). The aim of this was to ascertain whether

the putative pilin gene product was responsible for the immunogenicity of the clone or whether it was just fortuitously stumbled upon as frequently happens in such scientific research.

7.2 ANALYSIS OF CLONING FIDELITY: SOUTHERN BLOT HYBRIDISATION OF *Eco*RI CLEAVED *B.pseudomallei* GENOMIC DNA.

To determine the fidelity of the cloning procedure that generated the 5 kbp *Eco*RI fragment that conferred an immunopositive reaction to antisera (Chapter 4), a Southern blot was carried out.

7.2.1 Source of DNA.

The DNA used in the Southern blotting procedure was that of the *B.pseudomallei* isolate, 204 (Table 2.1). *B.pseudomallei* 204 was chosen because it was the isolate used to create the genomic library in λGT11. However, for use in the Southern blotting procedure the genomic DNA was purified with the RapidPrep™ Genomic DNA Isolation Kit (Pharmacia Biotech) (2.4.4.1).

7.2.2 Creation and DIG-Labelling of the Probe.

As the 5 kbp fragment had been sequenced it was possible to design PCR primers that could amplify a particular region of the sequence. The PCR product could then be labelled with DIG and function as the probe in the Southern blot procedure to check for the presence of this particular sequence in the genomic DNA of *B.pseudomallei*. PCR primers were designed (2.8.1) specifically to the putative pilin gene sequence (Fig.6.5), using the DNasis computer program. The suggested primers given were checked with the use of Amplify software.

The following primers were deemed the most suitable and synthesised at Sigma Genosys:

Forward primer (PilF): 5' TCGAACATGATCCGCAGCACATAG 3'

Reverse primer (PilR): 5' GAGAGGCTTCACCTTAATCGAGGT 3'

These primers were designed to amplify within the 5 kbp target sequence (Appendix III) from bases 2161 to 2405, yielding a PCR product 244 bp in length. The PCR carried out was a standard reaction (2.8.3) with an annealing temperature of 60°C. DIG label (Boehringer Mannhein) was incorporated into the PCR product through the reaction (2.7.1) and the concentration of the PCR product was checked by agarose gel electrophoresis alongside λ DNA standards.

7.2.3 Southern Blotting Procedure.

About 0.3 μg of pBPGT2 DNA and 0.5 μg of *B.pseudomallei* 204 genomic DNA was digested with the restriction enzyme *Eco*RI (2.5.4) and analysed by agarose gel electrophoresis. The agarose gel was used in a Southern blot procedure (2.7.2) using the DIG-labelled probe prepared in 7.2.2. The results of which can be seen in Fig.7.1. The 5 kbp insert of pBPGT2 developed along with a 5 kbp *Eco*RI fragment in the *B.pseudomallei* 204 genomic DNA proving that the cloned fragment indeed originated from *B.pseudomallei*.

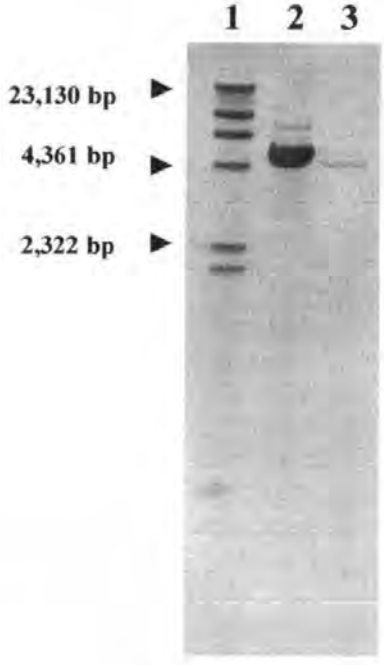


Fig. 7.1 Southern Blot of *Eco*RI-digested Recombinant Plasmid pBPGT2 and *B.pseudomallei* 204 Genomic DNA Using the DIG-Labelled Putative Pilin Gene PCR Product as the Probe to Check the Cloning Fidelity. LANE 1: DIG-labelled λHindIII ladders (from top to bottom) 23,130 bp, 9,416 bp, 6,557 bp, 4,361 bp, 2,322 bp, 2,027 bp, 564 bp and 125 bp. LANE 2: pBPGT2/*Eco*RI. LANE 3: *B.pseudomallei* 204 genomic DNA/*Eco*RI.

7.3 INVESTIGATION OF THE PUTATIVE PEROXIDASE GENE TO DISCOUNT A ROLE IN THE IMMUNOPOSITIVE REACTION.

It was considered possible that if somehow produced as a fusion protein with the β-galactosidase system in λGT11 or indeed from its own promoter, the portion of the putative peroxidase produced may have reacted with the detection system used in the screening process.

7.3.1 Testing of λBPGT1 with the Detection Solution used in the Initial Screening Procedure (Chapter 4).

A loopful of λBPGT1 was streaked out onto NZCYM media supplemented with 0.2% lactose and overlayed with 3 ml of soft NZCYM media also supplemented with 0.2% lactose and containing Y1088 cells to gain single plaques. A negative control of λGT11-infected Y1088 cells was also streaked out. Plates were incubated at 37°C for 18 h. Areas of well-separated plaques were overlayed with nitrocellulose and carefully removed with forceps and placed into Blocking Solution for around 1 h. before placing the nitrocellulose into the developing solution. No positive reaction was observed and it was assumed that the immunopositive reaction had not been due to the putative peroxidase.

7.4 SUB-CLONING OF A FRAGMENT OF *B.pseudomallei* DNA ENCODING A PUTATIVE PILIN GENE.

A smaller fragment containing just the putative pilin gene was subcloned into pUC18. This functioned to separate the putative pilin gene from the rest of the 5 kbp fragment, which may interfere with its expression in an *E.coli* system. Digesting the recombinant plasmid pBPGT2 with *Sal*I cuts the 5 kbp fragment (Appendix III) at positions 1907 and 2726 to generate an 819 bp fragment containing just the putative gene which could then be isolated for subcloning.

Approximately 2 μg of pUC18 DNA and 2 μg of pBPGT2 DNA were similarly digested with *Sal*I. The products were checked on an agarose gel and a fragment of around 800 bp from the digest of pBPGT2 was excised from the gel. This fragment, containing the pilin gene was extracted and purified from the agarose slice using the BIO-RAD Prep-A-Gene kit (2.4.6.2). The digested pUC18 DNA was treated with alkaline-phosphatase (2.5.5) and purified using the BIO-RAD Prep-A-Gene kit. A standard ligation reaction was set up (2.5.6) with the *Sal*I digested pUC18 DNA and the purified *Sal*I 800 bp fragment isolated from pBPGT2 and incubated for 18 h. at 15°C. A transformation using $CaCl_2$-treated DH5α cells was carried out (2.5.1) and the transformation mix plated out onto LB media supplemented with ampicillin, X-gal and IPTG for blue-white screening and incubated at 37°C for 18 h. White colonies obtained were cultured in 10 ml LB broths and incubated at 37°C with shaking for 18 h. After this time, 5 successful cultures (pSAL1-5) were subject to small-scale plasmid extractions (2.4.2.1) and the isolated plasmid DNA digested with *Sal*I to release a fragment. Only one of these released an 800 bp fragment, pSAL2. The recombinant plasmid, pSAL2 was digested with *Mlu*I, as the foreign fragment should contain this restriction site (cuts at position 2343), which is not present in pUC18 hence the result should be to linearise pSAL2. The result of these digests can be seen in Fig.7.2.

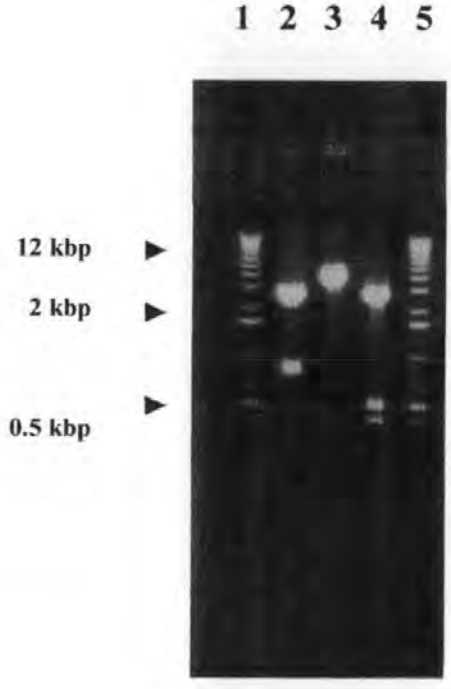


Fig.7.2. Agarose Gel Electrophoresis of a Recombinant Plasmid (pSAL2) Containing a *Sal*I-Generated Fragment from *B.pseudomallei* Encoding a Putative Pilin Gene. LANE 1: 1 kbp ladders. LANE 2: pSAL2/*Sal*I. LANE 3: pSAL2/*Mlu*I. LANE 4: pSAL2/*Sal*I/*Mlu*I. LANE 5: 1 kbp ladders.

7.5 ATTEMPTS TO DETECT THE MOLECULAR DETERMINANT OF THE PUTATIVE PILIN GENE ENCODED ON THE RECOMBINANT PLASMID pSAL2.

Subcloning the fragment further (7.3) functioned to isolate the putative pilin gene away from any interfering sequences contained on the 5 kbp fragment as in the initial subclones created (pCD1-5) the pilin gene was disrupted. As expression of the protein product may be low, the *in vitro* transcription and translation kit was applied to the recombinant clone, pSAL2 and to aid detection of the product, DNA from a DH5α host containing pUC18 (the parent plasmid to pSAL2) was used as a comparative control.

7.5.1 *In vitro* Transcription and Translation of the Recombinant Clone, pSAL2.

Plasmid DNA from DH5α cells containing pSAL2 and DH5α cells containing pUC18 was isolated using the Qiagen Plasmid Isolation Minikit (2.4.2.4) to gain high quality plasmid DNA. The DNA was then used in the *in vitro* transcription and translation system (as in 5.7) using the provided positive control. Products of the reaction were used in SDS-PAGE analysis and then subject to Western blotting. No product was isolated in the recombinant clone that differed to the negative control used. This was repeated, varying the reaction conditions but no product was visualised.

7.6 ATTEMPTS TO CREATE A FUSION PROTEIN.

As attempts to express the protein product of the putative pilin gene using the *in vitro* transcription and translation kit had failed, as well as many earlier attempts in Chapter 5, it was decided to try to produce the pilin gene protein product as a fusion. Problems associated with lack of, or poor expression may be overcome by making fusion proteins between genes, whereby the cloned gene is introduced into an expression vector at the 3' end of a carrier

sequence coding for the amino terminus of a highly expressed carrier protein. The carrier sequence provides the necessary signals for good expression and the expressed fusion protein contains a N-terminal region encoded by the carrier, which may be β-galactosidase, MBP (maltose binding protein), glutathione-S-transferase or protein A fusions. In this case a fusion with β-galactosidase was attempted. In *E.coli* the *lacZ* gene codes for the polypeptide β-galactosidase, which is 1021 amino acids long. However, when expression of a foreign gene is dependent on its fusion to an *E.coli* gene, it is important that the correct translational reading frame is maintained. Therefore for most gene fusions consideration and possible adjustments may need to be made at the gene junction. This method then results in the production of the hybrid protein carrying amino-terminal amino acids from β-galactosidase. Subsequent purification of the recombinant protein from a crude cell lysate can be achieved by immunoprecipitation with antibodies directed against β-galactosidase.

7.6.1 pUC18.

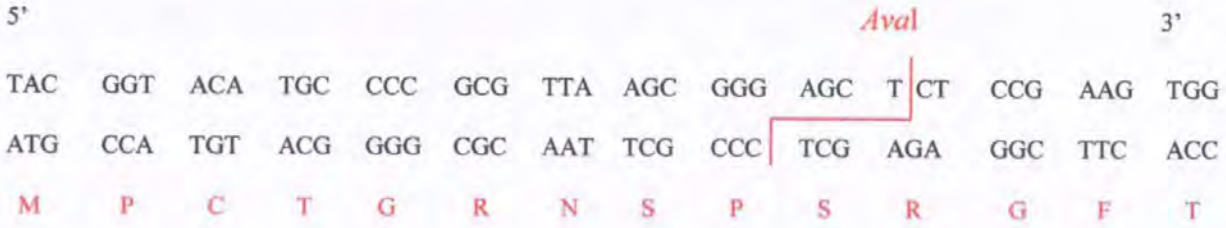
A fusion was attempted with β-galactosidase in pUC18 (Appendix II) by cloning the gene in-frame with the system. The strategy behind creating this fusion protein can be seen in Fig.7.3. Approximately 1 μg of pUC18 DNA was digested with *Sal*I and *Ava*I and the resulting plasmid DNA was treated with calf-intestinal alkaline phophatase (2.5.5). Approximately 2 μg of pBPGT2 DNA was similarly digested with *Sal*I and *Ava*I, which functioned to cut the 5 kbp fragment (Appendix III) at positions 1901 and 2501 respectively to produce a 517 bp fragment containing the putative pilin gene. The products of the digest were run on an agarose gel and the desired band, which was approximately 517 bp was excised from the gel and purified using the BIO-RAD Prep-A-Gene kit (2.4.6.2). A standard ligation reaction was set up with the digested pUC18 DNA and the *B.pseudomallei* DNA fragment (2.5.6) and the resulting ligation was used in a transformation with $CaCl_2$-treated DH5α cells (2.5.1) and plated out onto LB media supplemented with ampicillin, X-gal and IPTG for blue-white screening of recombinants. A very low transformation efficiency was observed and those white colonies

obtained were cultured in 10 ml of LB broth at 37°C for 18 h. Small-scale plasmid extractions (2.4.2.1) were carried out on cultures, but none were found to contain inserts. This was repeated several times with the screening of every colony, including blue ones but none were found to contain inserts.

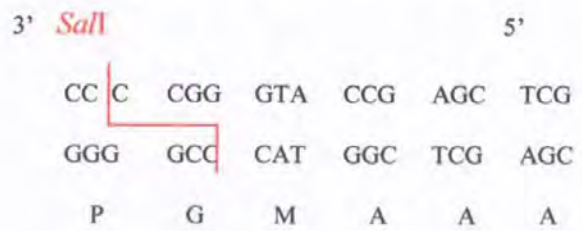
7.6.2 Other Fusions.

Alternative vectors were also used to attempt to create fusions with the putative pilin gene product and β-galactosidase, which were pAX5+ and pEX1 but again after numerous attempts, no recombinants were isolated.

The 5 kbp fragment was digested with restriction enzymes *Ava*I and *Sal*I to yield the following within the pilin gene:



pUC18 was digested with *Sal*I and *Ava*I:



On ligation a fusion protein is created with the following sequence:

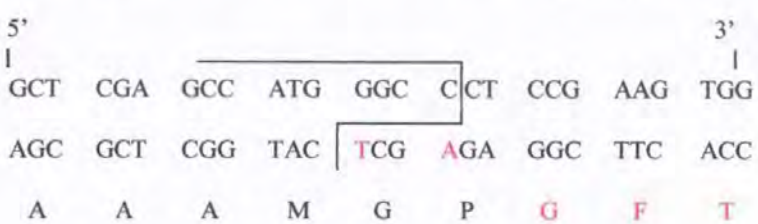


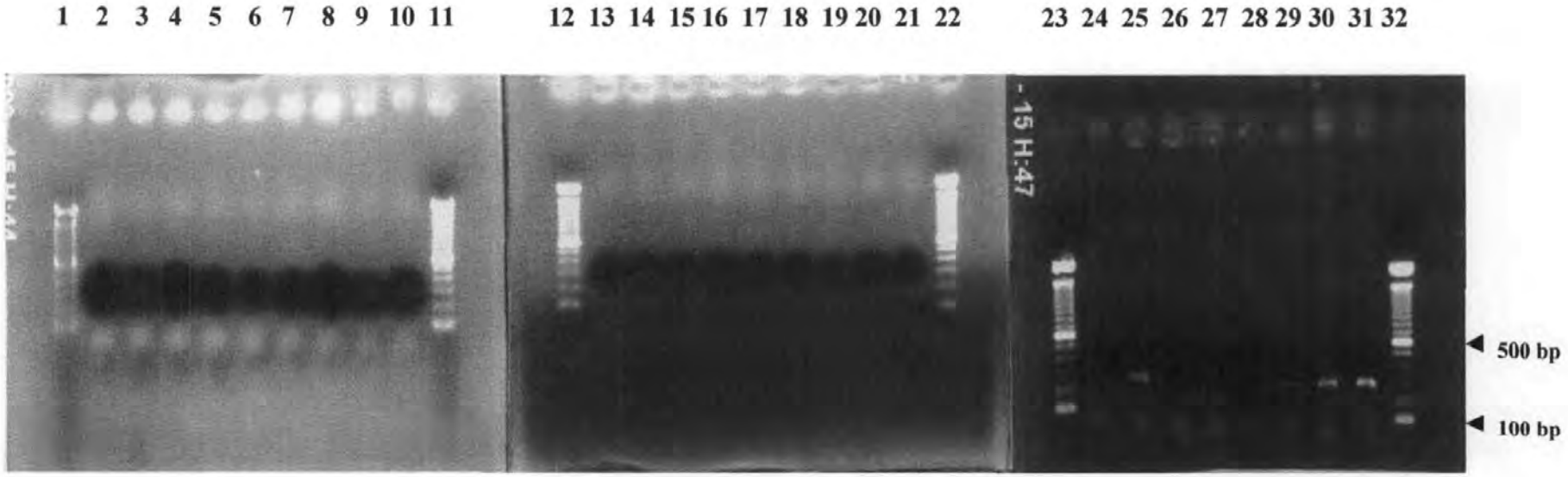
Fig 7.3. Cloning Strategy to Attempt to Produce a Fusion Protein of a Putative Pilin Gene Product from *B.pseudomallei* with the β-galactosidase System in pUC18. The A and T bases highlighted in colour in the fusion protein represent mismatches, but these can often be repaired by the *E.coli* system.

7.7 PCR ANALYSIS TO DETECT THE PRESENCE OF THE PUTATIVE PILIN GENE IN ISOLATES OF *B.pseudomallei* and *B.thailandensis.*

The putative pilin gene was cloned from the *B.pseudomallei* isolate 204 and so to test other isolates for the presence of the fragment, the PCR primers designed in 7.2.2 were used to test isolates of *B.pseudomallei* and *B.thailandensis* (Table 2.1). The PCR primers were specifically designed to the putative pilin gene sequence and so by using genomic DNA from a number of isolates and carrying out a PCR reaction using the primers, a PCR product of the correct size (244 bp) may indicate the presence of the fragment. The PCR reaction does not show whether the sequence is transcribed in the particular isolates, but it was used to observe any differences between the virulent *B.pseudomallei* and the avirulent *B.thailandensis.*

Approximately 20 ng of genomic DNA was used in each PCR reaction with the same reaction conditions used in 7.2.2. The annealing temperature was dropped to 55°C as results tried at 60°C were very faint and even at the lower temperature the reaction products were faint. This was tried a number of times but results could not be improved upon. Results can be seen in Fig.7.4.

It was observed that none of the *B.thailandensis* isolates tested produced a PCR product. However, not all of the *B.pseudomallei* isolates produced a PCR product but this does not necessarily mean that the sequence is not present. It may have been that the genomic DNA used was of a poor quality and contaminants may have inhibited the reaction, or that the PCR itself was unreliable. To investigate this further, it was decided to carry out a Southern blot procedure, which has a higher stringency than a PCR reaction to confirm the results obtained.



LANE	ISOLATE	LANE	ISOLATE	LANE	ISOLATE
1	1 kbpLadders	12	1 kbpLadders	23	1 kbpLadders
2	-ve control	13	22	24	212
3	204	14	25	25	216
4	E27	15	E25	26	217
5	E82	16	33	27	392
6	E255	17	46	28	E254
7	E256	18	53	29	426
8	E260	19	97	30	448
9	E8	20	98	31	576
10	19	21	102	32	1 kbp Ladders
11	1 kbpLadders	22	1 kbpLadders		

Fig.7.4. Agarose Gel Electrophoresis of PCR Products Using Primers Designed to a Putative Pilin Gene Identified in the *B.pseudomallei* Isolate 204 on Isolates of *B.pseudomallei* and *B.thailandensis*. Of the isolates tested, those highlighted had a PCR product of approximately 250 bp.

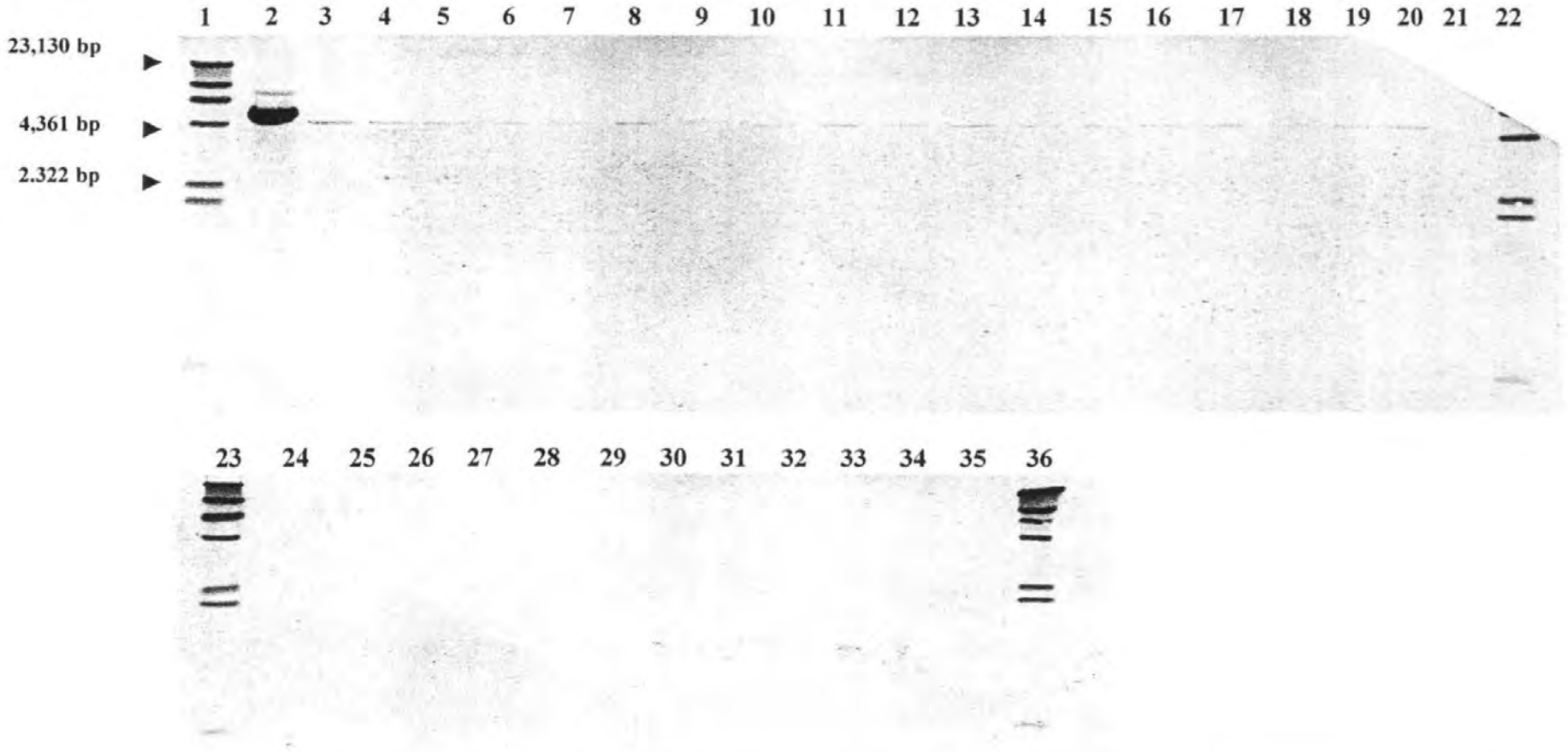
7.8 SOUTHERN BLOTTING ANALYSIS OF ISOLATES OF *B.pseudomallei* AND *B.thailandensis* FOR THE PRESENCE OF THE PILIN GENE FRAGMENT.

7.8.1 Source of DNA.

The DNA used in the Southern blotting procedure was from isolates shown in Tables 2.1 and 2.2. Genomic DNA from the following species was isolated using the Puregene® Genomic DNA Isolation Kit: *B.vietnamensis* (LMG 6998, LMG 6999, LMG 10926), *B.plantarii* (LMG 10908), *B.cocovenans* (LMG 11626), *B.vandii* (LMG 10620) and from the University of Plymouth's culture collection, genomic DNA from strains of *P.aeruginosa*, *P.fluorescens* and *E.coli* DH5α. To a 0.5 ml of overnight culture in a MCC tube, 0.5 ml of Cell Suspension solution was added and the cells pelleted by centrifugation at 13,000 x g for 5 sec. The supernatant was removed and 300 µl of Cell Lysis solution added and the mixture incubated at 80°C for 5 min. A 1.5 µl volume of RNase A solution was added, mixed and incubated at 37°C for 1 h. The samples were cooled to room temperature and 100 µl of Protein Preparation solution was added, vortexed and the preparation centrifuged at 13,000 x g for 3 min. The supernatant was transferred to a new MCC tube and precipitated with isopropanol. Finally, the pellet was resuspended in 50 µl of DNA Hydration solution by incubating for 1h. at 65°C.

7.8.2 Southern Blotting Procedure.

Approximately 1 µg of genomic DNA from *B.pseudomallei*, *B.thailandensis* and other bacterial species used was digested with *Eco*RI. Portions of the digests were run on agarose gel to check that the digests were successful and to give an indication of the concentration of DNA. A suitable volume of the digests were run on an agarose gel for use in the Southern blotting procedure (2.7.2). This was carried out using the DIG-labelled PCR product (7.2.2) as the probe. The results of the Southern blot can be seen in Fig.7.5.



LANE	ISOLATE	LANE	ISOLATE	LANE	ISOLATE	LANE	ISOLATE	LANE	ISOLATE	LANE	ISOLATE
1	*Hind*III MARKERS	8	216	15	98	22	MARKERS	29	*B.vietnamensis* LMG 6999	36	DH5α
2	pBPGT1	9	448	16	102	23	MARKERS	30	*B.plantarii*		
3	204	10	22	17	212	24	E256	31	*B.cocovenans*		
4	E8	11	33	18	217	25	E255	32	*B.vandii*		
5	19	12	46	19	392	26	E260	33	*B.cepacia* 1		
6	25	13	53	20	426	27	E38	34	*P.aeruginosa*		
7	E25	14	97	21	E82	28	MAL6	35	*P.fluorescens*		

Fig.7.6 Southern Blot Hybridisation of *Eco*RI cleaved *B.pseudomallei*, *B.thailandensis* and Closely Related Species' Genomic DNA to Detect for the Presence of a Putative Pilin Gene Fragment.

The Southern blot revealed again that none of the *B.thialandensis* isolates tested appeared to contain the fragment containing the putative pilin gene. The other *Burkholderia* species used similarly did not develop a band on Southern blotting. However, all of the *B.pseudomallei* isolates tested developed a band of approximately 5 kbp and therefore posses the DNA fragment containing the putative pilin gene.

7.9 ATTEMPTS TO CREATE A MUTANT OF THE PUTATIVE PILIN GENE.

Attempts were made to create a knock-out mutant of the putative pilin gene in the suicide vector, pCVD442. The vector was kindly donated by Dr.R.Titball, Porton Down, Salisbury and it was hoped that the mutant created could be tested at Porton Down to ascertain whether the putative pilin gene has a role in the virulence of *B.pseudomallei*. The strategy involves the introduction of a DNA insert into the gene under study to produce defective mutants. However the insertion of DNA in this case provides a new functional element that is resistance to the antibiotic, kanamycin.

7.9.1 Propagation of the Plasmid.

The plasmid pCVD442 can only replicate in strains that have the *pir* gene, usually supplied by a λ lysogen encoding the Pi protein, which is necessary for the replication of this plasmid. A suitable λ lysogen, SY326 was kindly supplied by Mr Derrick Pickard, Imperial College, London and the vector was transformed into this.

7.9.2 Cloning a Kanamycin Cassette into pCVD442 .

It was decided that the most suitable antibiotic to use in such a study with *B.pseudomallei* was kanamycin, which was already the chosen antibiotic for use with mutants at Porton Down. The

kanamycin resistance gene from pGD103 (Appendix II) was removed with the restriction enzymes *Cla*I and *Eco*RI, and attempts were made to clone this into the suicide vector pCVD442. However, the vector proved to be highly unstable and degraded very quickly once isolated from the organism it was transformed into. The plasmid was purified via CsCl/EtBr density ultracentrifugation and re-transformed into a number of hosts, but degradation continued to occur.

7.10 DISCUSSION.

Expression of the putative pilin gene from *B.pseudomallei* in an *E.coli* system proved unsuccessful even when attempts were made to incorporate the protein product as a fusion with β-galactosidase. Synthesis of a functional protein depends upon successful transcription of the appropriate gene, efficient translation of the mRNA and, in many cases, post-translational processing and compartmentalisation of the nascent polypeptide. A failure to perform any one of these processes can result in the failure of a given gene to be expressed. However the use of the *in vitro* transcription and translation kit should overcome these fundamental problems. One factor is that the pilin gene does indeed require post-translational modifications in the form of the cleavage of its signal sequence by a specific protein, which is highly unlikely to be present in the *E.coli* system. This may therefore have prohibited the expression of the protein. Similar problems were encountered by Russell and Darzins (1994) when attempting to express *pil*E from *P.aeruginosa* in *E.coli* as an orientation-dependent production of the estimated 15 kDa polypeptide could not be demonstrated. In this case the authors resorted to expressing *pil*E in a *P.aeruginosa* T7 expression system.

However, there is still the factor of the relative stability of foreign proteins in *E.coli*. The putative gene was initially cloned, and presumed to be expressed if indeed it was responsible for making the clone immunogenic, into λGT11, with the use of a host deficient in the lon protease. Such a host has a reduced complement of intracellular proteases, thus increasing the stability of foreign proteins. However, once the gene was cloned into the plasmid vector pUC18 this did not apply, which could a reason for the lack of expression observed.

The putative gene was placed under the control of an *E.coli* promoter in an attempt to produce a fusion protein. In the design of the fusion with the β-galactosidase system in pUC18 there were a couple of mismatched bases, which would function to decrease the efficiency of ligation and therefore many more clones needed to be screened in order to isolate one in which the *E.coli* system had repaired the mismatches. Hundreds of colonies were screened, irrespective of whether they were white or blue, as often recombinants can still appear a blue colour, without any success. As none of the transformants contained inserts, it could be that the mismatched bases were so rarely repaired by *E.coli* that the screening of hundreds more colonies was necessary in order to find one, or simply that the production of a fusion with the *B.pseudomallei* protein was deleterious to the cell and the fragment was therefore deleted. Alternative vectors were tried that contained no mismatches and that had a lower copy number than pUC18. Numerous attempts at cloning the putative pilin gene in such vectors were carried out again with no success. The reasons for this could be that the fusion protein produced, even at a low level, is deleterious to the *E.coli* host or perhaps that to produce the fusion successfully, the first few amino acids of the foreign protein were needed. These were not incorporated into the fusion protein, as the transcriptional reading frame would have been disrupted.

The experiments used to determine the presence of the putative pilin gene sequence in other isolates of *B.pseudomallei* and in the avirulent species, *B.thailandensis*, showed that the sequence was not present in any of the *B.thailandesis* isolates tested but Southern blotting confirmed its presence in all *B.pseudomallei* isolates tested. Thus indicating that the gene sequence is not present in the avirulent species, but present in the virulent *B.pseudomallei* species, suggesting therefore a potential role for the pilin gene in the virulence of *B.pseudomallei*. There is no information in current literature confirming whether members of the *B.thailandensis* species possess type IV pili, although type IV pili have been confirmed in *B.pseudomallei* isolates (Vorachit *et al.*, 1995; Woods *et al.*, 1999). Interestingly in the study by Vorachit *et al.* (1995) it was observed that some isolates did not produce pili and these adhered to surfaces as dispersed, individual cells rather than biofilms, which was the method of adherence for those possessing pili. It may well be the case therefore that the non-piliated isolates were actually avirulent *B.thailandensis* rather than *B.pseudomallei*.

Unfortunately a mutant could not be constructed with the suicide vector, pCVD442 to elucidate the role of the putative gene in the virulence of *B.pseudomallei*. However, when the suicide vector was donated, it was stated that problems were often encountered when trying to clone with the vector (Petra Oyston, personal communication).

Regrettably, a size and a probable function could not therefore be elucidated for the putative pilin gene. Also, very importantly, it could not be proven that it was this gene product responsible for the immunopositive nature of the original clone λGTBP1. Other type IV pilin genes have allegedly been cloned by Woods *et al.* (1999) with homology to *pil*B, *pil*C and *pil*D of *P.aeruginosa* and the putative gene cloned in this study has homology to *pil*E of *P.aeruginosa*, although it may have a different function in *B.pseudomallei*.

The *pil*E gene of *P.aeruginosa* has been studied in some detail, although a definitive function has yet to be assigned. Russell and Darzins (1994) managed to generate two chromosomal *pil*E insertion mutants, which were broadly resistant to pilus-specific phages and additionally unable to translocate across solid surfaces in a twitching motility assay. The mutants, although they produced unassembled pilin pools, failed to assemble pili. However, Russell and Darzins (1994) have implicated *pil*E in *P.aeruginosa* as a pilot that guides *pil*A subunits to and from the assembly platform during the alternating mechanisms of pilus polymerisation. It has also been suggested that *pil*E functions as a component of the pilus assembly apparatus by capping and uncapping the interactive surfaces of pilin subunits that have been translocated across the cytoplasmic membrane (Russell and Darzins, 1994).

CHAPTER 8

USE OF PCR TO SCREEN FOR THE PRESENCE OF AN ADP-RIBOSYLATING TOXIN IN *B.pseudomallei.*

8.1 INTRODUCTION.

Pathogenic bacteria have evolved mechanisms for evading the defence systems of their hosts and for destabilising the cell's functional system, of which perhaps the commonest and most effective way of achieving this is to produce toxins (Balfanz *et al.*, 1996). In particular, ADP-ribosylating toxins have been the focus of research for more than 30 years, providing invaluable insights into their structure and function, which may be the key to the prevention and treatment of diseases caused by such toxin-producing bacteria. It has been reported that *B.pseudomallei* possesses a toxin belonging to the ADP-ribosylating class of bacterial exotoxins, and more specifically to ADP-ribosylate elongation factor 2 (EF-2) (Mohamed *et al.*, 1989: Ismail *et al.*, 1991). Diphtheria toxin (DT) and *P.aeruginosa* Exotoxin A (ETA) also catalyse the ADP-ribosylation of EF-2 (Van Ness *et al.*, 1980) at a dipthamide residue in a stereochemically identical manner (Patrick and Larkin, 1995). EF-2 is a tRNA translocase responsible for the translocation of the growing polypeptide chain from the aminoacyl-tRNA to the peptidyl-tRNA position at the ribosome, the energy for which is created through the hydrolysis of GTP (Balfanz *et al.*, 1996). The overall effect of the toxin therefore, is to inhibit protein synthesis, leading to inevitable cell death. Diphtheria toxin is synthesised as a protein with an approximate molecular weight of 62 kDa, which is enzymatically cleaved to produce two fragments, A and B, with molecular weights of 21 and 40 kDa respectively (Pappenheimer, 1955; Gill and Dinius, 1971). The A fragment is responsible for the enzymatic

properties of the toxin and the B fragment is thought to be involved in attachment and entry of the toxin into the target cell. ETA is synthesised in the form of a proenzyme with an approximate molecular weight of 66 kDa, which has to undergo conformational change before it becomes toxic. It is Domain III that contains the enzymatic activity of the molecule, that is, catalysing the transfer of ADP-ribose from NAD to EF-2 in eukaryotic cells.

The resemblance in the mode of action of DT and ETA corresponds to similarities in the amino acid sequence of the toxins, especially in the region of the NAD^+ binding site (Domenighini *et al.*, 1991). However, their catalytic and receptor-binding domains are located at opposite ends of the molecules, with the catalytic domain being at the amino terminus in DT and at the carboxyl terminus in ETA (Aktories, 1992). Early studies on these toxins indicated that although the two primary amino-acid sequences showed little overall homology there is homology within the NAD^+ binding site where, for example, Glu-148 of DT and Glu-553 of ETA are homologues within this site (Carroll and Collier, 1988). This strongly suggests that both these enzymatic domains originated from a common ancestral protein and that there is a common evolutionary origin of all ADP-ribosylating enzyme toxins (Patrick and Larkin, 1995).

PCR is a very powerful technique for the molecular biologist and has found many applications in its basic form. The technique can be applied in investigations to compare the sequence homology of conserved genes in different organisms. In the case of this project, although a toxin with ADP-ribosylating activity has been reported and purified in culture filtrates of *B.pseudomallei*, the gene encoding the toxin has yet to be cloned and sequenced. However, other ADP-ribosylating toxins of closely related organisms, such as *P.aeruginosa* ETA have been cloned and sequenced, and so it is possible to use the sequence information available. As the toxin in *B.pseudomallei* has been reported to have the same mode of action as ETA, PCR

primers were designed to the most conserved region of these toxins, the active site. Therefore by using the primers developed in a PCR reaction with *B.pseudomallei* genomic DNA, a corresponding product may be amplified if there is a high homology between the genes of the two toxins. Products of the PCR reaction can be cloned into plasmid vectors for DNA sequencing which may identify the PCR product amplified as part of a gene encoding the ADP-ribosylating toxin in *B.pseudomallei*.

8.2 PRIMER DESIGN AND INITIAL EXPERIMENTATION.

PCR primers were designed in consultation with Dr. M. L Gilpin, based on sequence analysis available for the ETA produced by *P.aeruginosa* (Gray *et al.*, 1984). The primers were designed to regions around the active site of the gene (Fig.8.1). The following primers were synthesised by Sigma-Genosys:

Forward primer (1F) 5' **GGCGACGTCAGCTTCAGC** 3'

Forward primer (2F) 5' **CGAAGCGATCCGCGAGCA** 3'

Forward primer (3F) 5' **CCTGCTGGAGCGCAACTA** 3'

Reverse primer (1R) 5' **GCTCGAGCGCGGCACATA** 3'

Reverse primer (2R) 5' **ATCGCGTCGAGGTCCTGG** 3'

Reverse primer (3R) 5' **AATCACCACGGTGCGCTC** 3'

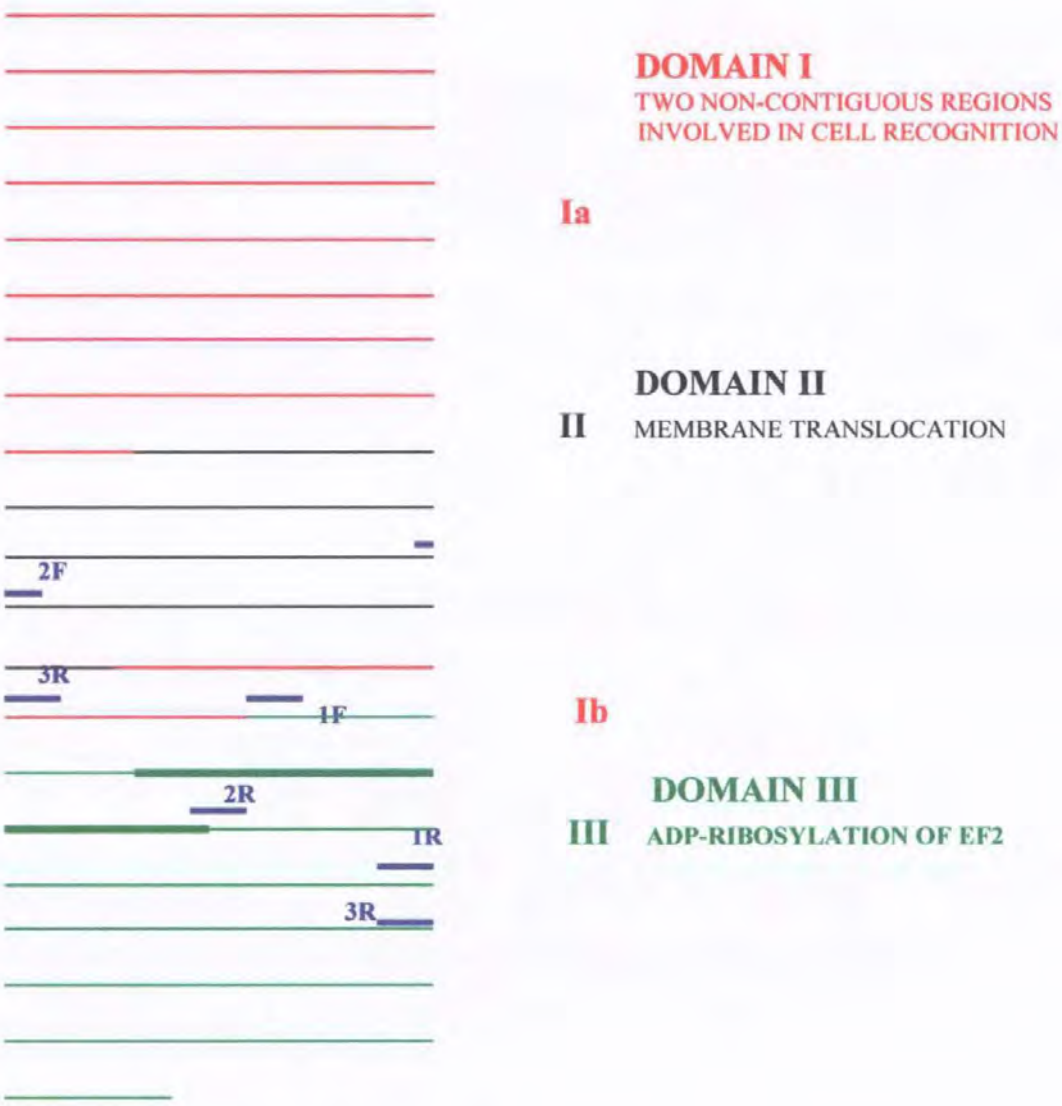


Fig.8.1 Primer Design to the Active Site of the ETA Gene in *P.aeruginosa* (Gray *et al.*, 1984). The figure represents the structure of the ETA gene of *P.aeruginosa*, which includes all three domains of the gene responsible for differing functions. The thick dark green line represents the coding region for the active site of the toxin and the blue lines indicate where the primers were designed.

Synthesised primers were resuspended to give an overall concentration of 100 pmol/µl (2.8.2). A standard PCR reaction was used (2.8.3) with an annealing temperature of 60°C. Genomic DNA from *B.pseudomallei* isolates 204, 576, 4845 and E8 and *B.thailandensis* isolates 82, 27 and 25 (Table 2.1) were used in the PCR reactions. In the initial experiments, different combinations of the forward and reverse primers were used. It was found that using the primer combination, 2F and 1R yielded a PCR product with some of the isolates at around 500 bp (Fig.8.2), which was close to the expected product of 600 bp based on the primers 2F and 1R annealing to the correct sequence in the *P.aeruginosa* ETA gene. It was noted that products were only obtained with *B.pseudomallei* isolates and not with the avirulent *B.thailandensis* isolates used, with the exception of *B.pseudomallei* isolate 4845. Also that isolate 204 had an additional band amplified at around 900 bp.

8.3 CLONING THE PCR PRODUCT INTO A PLASMID VECTOR.

8.3.1 Isolation of PCR Products for Cloning.

From the initial experiments it was decided to clone the PCR products amplified with *B.pseudomallei* isolates 576 and 204 for eventual sequence comparison. A PCR reaction was carried out with both isolates using the same reaction conditions and primers as in 8.2 and the products were run on an agarose gel. The PCR products obtained were of a very low concentration and so the reaction products gained were re-amplified using the same PCR reaction. The re-amplified products were visualised on an agarose gel and bands of approximately 500 bp were excised from the agarose gel and the DNA isolated by electroelution (2.4.6.1), and purified by using the BIO-RAD Prep-A-Gene kit (2.4.6.2). The additional 900 bp PCR product observed in the initial experiment with *B.pseudomallei* 204 DNA failed to be amplified again in a number of PCR reactions carried out.

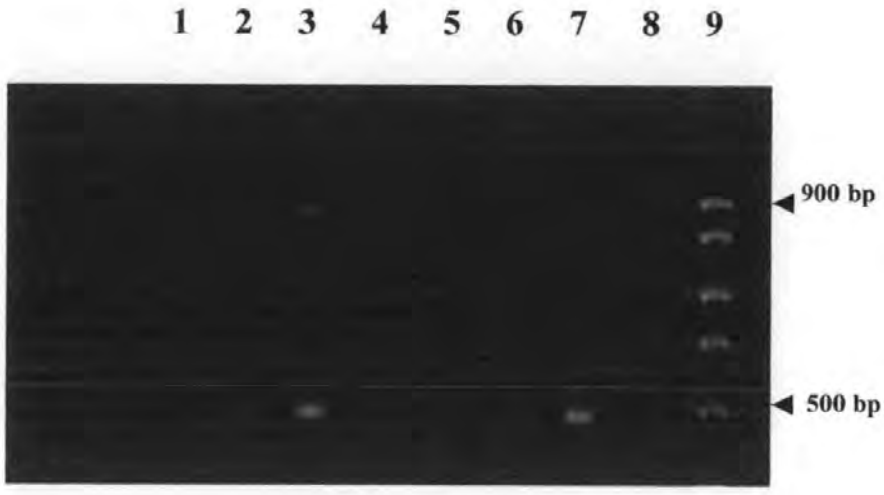


Fig.8.2 PCR Products Obtained in a Reaction Using Isolates of *B.pseudomallei* and *B.thailandensis* to Ascertain the Presence of an ADP-Ribosylating Toxin Sequence. LANE 1: 4845. LANE 2: 576. LANE 3: 204. LANE 4: 82. LANE 5: 27. LANE 6: 25. LANE 7: E8. LANE 8: -ve control. LANE 9: 100 bp ladders.

8.3.2 Cloning the PCR Products.

The PCR products from both isolates 204 and 576 were cloned into pUC18 using the SURECLONE™ Kit (Pharmacia) (2.5.7) and the ligation products were transformed using DH5α competent cells (2.5.3.2). Transformations were plated out onto LB media supplemented with ampicillin, X-gal and IPTG for blue-white screening, and incubated at 37°C for 18 h. Afterwards, 10 white colonies from the 204 transformation (pAD1-10) and 10 from the 576 transformation (pDP1-10) were cultured in 10 ml of LB broth supplemented with ampicillin and incubated at 37°C with shaking for 18 h. Next, small-scale plasmid extractions (2.4.2.1) were carried out on all cultures. Restriction enzyme *Eco*RI and *Pst*I cleavage sites are located either side of the *Sma*I site in pUC18 (Appendix II), which was the blunt-ended site used to clone fragments in the SURECLONE™ kit. Therefore, by digesting recombinant plasmid DNA with these enzymes it functions to release the foreign DNA insert. Therefore, the resulting plasmid DNA was digested with *Eco*RI and *Pst*I and the products of the digests were analysed by agarose gel electrophoresis. Bands of approximately 500 bp were identified in all of the plasmid digests. A recombinant plasmid from each of the isolates 204 and 576, pAD1 and pDP1 respectively, were cultured in 500 ml of LB broth supplemented with ampicillin and the plasmid DNA extracted using the Wizard Maxiprep kits (2.4.2.3). Approximately 0.5 μg of pAD1 and pDP1 DNA was digested with restriction enzymes *Pst*I and *Eco*RI and the results of the digest was visualised by agarose gel electrophoresis to check that the plasmids still contained an insert (Fig.8.3). Both pAD1 and pDP1 released a 500 bp insert and approximately 30 μg the plasmid DNA from both isolates was sent to Dr. R. Titball, Porton Down for DNA sequencing.

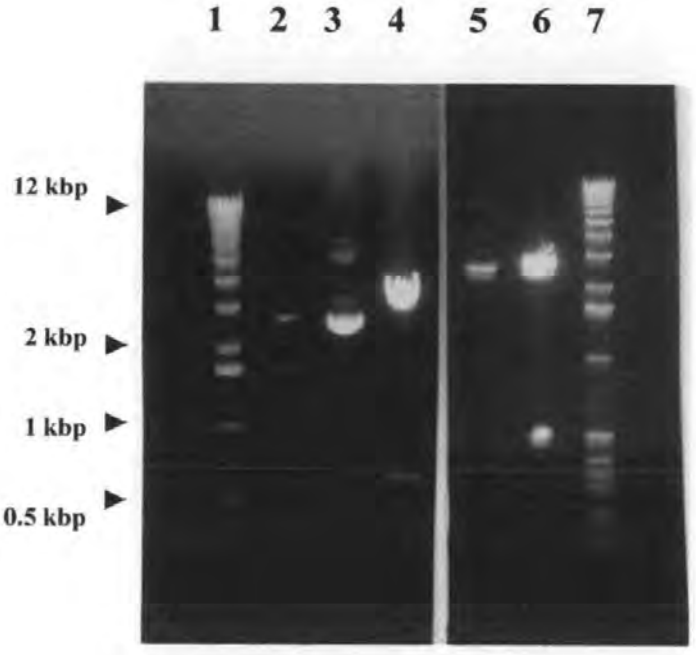


Fig.8.3. Agarose Gel Electrophoresis of Restriction Enzyme Digests of Recombinant Plasmids with *Eco*RI and *Pst*I to Release Cloned DNA Inserts. LANE 1: 1 kbp ladders. LANE 2: pUC18/*Eco*RI. LANE 3: Undigested pDP1. LANE 4: pDP1/*Eco*RI/*Pst*I. LANE 5: Undigested pAD1. LANE 6: pAD1/*Eco*RI/*Pst*I. LANE 7: 1 kbp ladders.

8.4 SEQUENCE ANALYSIS OF pAD1 and pDP1.

8.4.1 The Complete Sequence of pAD1 and pDP1.

The sequence information received from Dr. R. Titball, Porton Down, showed that the two PCR products had identical nucleotide sequences, which can be seen in Fig.8.4. The total length of the sequence was 453 bp, including primer sequences.

8.4.2 Sequence Analysis of the PCR Product.

Firstly an alignment (BLAST 2 sequences program available at Genbank) was carried out between the nucleotide sequence of the PCR product and that of the ETA gene of *P.aeruginosa*. This showed that there was no significant homology at the nucleotide level between the two sequences. The sequence received was then translated into an amino acid sequence in all six possible reading frames, using the DNasis software. Again, no homology could be found at the amino acid level in any of the translations between the two sequences. Investigations into amino acid sequence homologies within the active sites of different ADP-ribosylating toxins have been carried out (Rappuoli and Pizza, 1991) the results of which can be seen in Fig. 8.5. The amino acid translations of the sequence received was searched for any of the patterns in Fig.8.5 without success.

Web-based BLAST-X searches were carried out for the amino acid translations of the sequence, but little information could be derived from this. It was decided that as the PCR reaction failed to amplify a product with any of the *B.thailandensis* isolates tested but gave a result for most of the *B.pseudomallei* isolates tested that the sequence could be of interest and to pursue it further by obtaining sequence information up and down-stream of the fragment. The fragment might encode a toxin slightly different to the other ADP-ribosylating toxins discovered, or perhaps it is part of a sequence encoding another putative virulence factor in *B.pseudomallei*.

```
                              2F
  1   GAGCTCGGTA CCCCGAAGCG ATCTGCGAGC ATCTCAATCA ACATTCGGGC

 51   CCGGGAATCG ACTATGCCGT ACGATGAAAT TCTGCCTATT GAAACGCTGG

101   TTCGACCCAA AGTATCGGAC TTCCGCGAGC ATTACCTCGA AAAAGAGCGC

151   CCGGTGAAAA TCGCGAGGGC GCTCGATGCG TGGCCGGCGA TGCAGAAGTG

201   GTCGCTCGAC TATTTCGAGA ATCGTTTCGG CGACGAGACC ATCGGCGTCG

251   AGAGCTTTCA GCCTGACGAG CGCGGGCCCG GCAACAACAG CCCGCAAGGT

301   TACGTCAAGC ACTTGCGTTT CCAGGAGCTC AAGCTCAAGG AAACTGATCC

351   GGATCTTGCG AACGAAGCCG GACCACATGT ACTACATGGC GTCGCATCCG

401   TTTCGCAAGA GCTTTCCGAA CCTGCGCGCC GATCTCGCGC CGCATCYGTA
                1R
451   TGTGCCGCGC TCGAGCGGG
```

Fig.8.4 Nucleotide Sequence of the PCR Product from *B.pseudomallei* Isolates 204 and 576 Using Primers Designed to the Active Site of ETA of *P.aeruginosa*. The sequence includes both primers sequences (in red) and sequences in green represent vector DNA. Y represents a pyrimidine base (C or T).

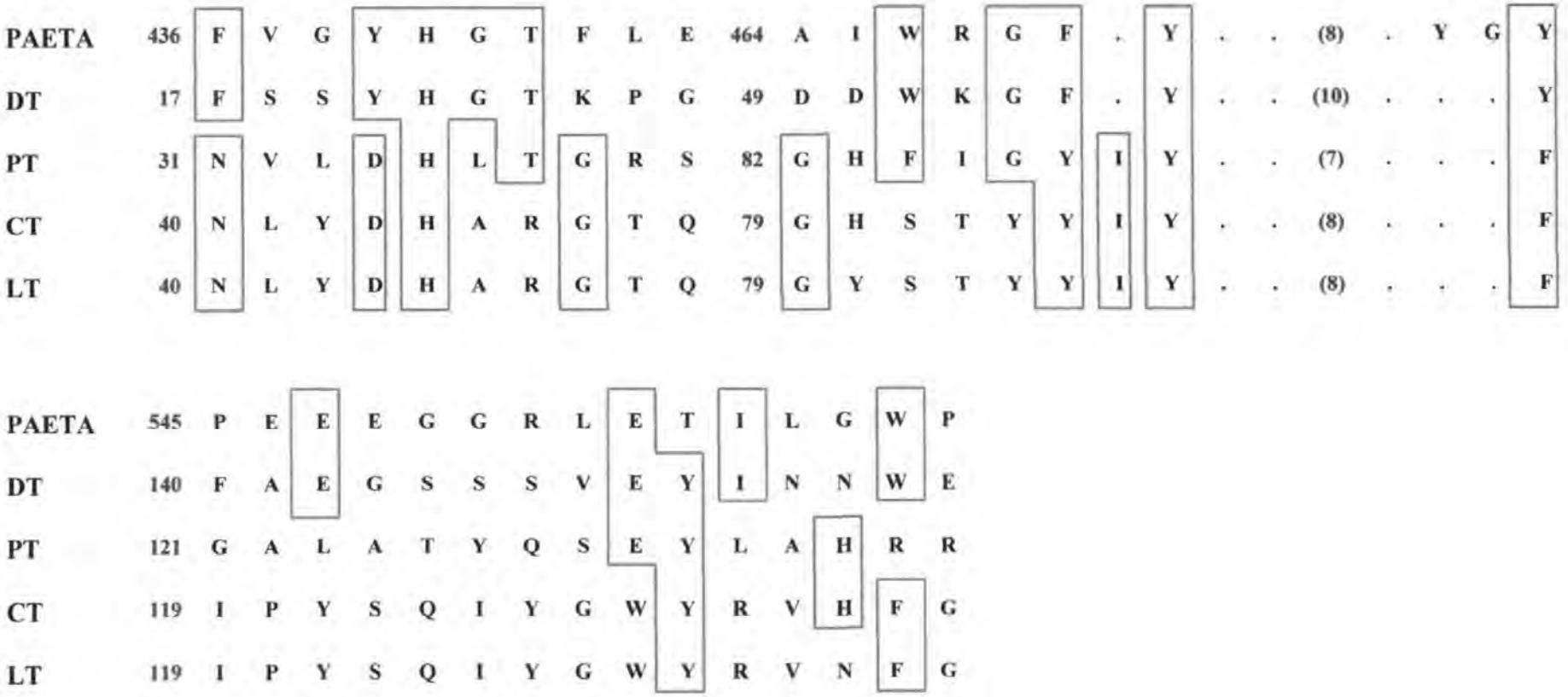


Fig.8.5 Homologies Between the Active Sites of a Number of ADP-Ribosylating Toxins (Rappuoli and Pizza, 1991). *P.aeruginosa* ETA (PAETA), Diphtheria toxin (DT), Pertussis toxin (PT), Cholera toxin (CT) and *E.coli* heat-labile enterotoxin (LT)

8.5 ISOLATION OF A LARGER FRAGMENT FROM *B.pseudomallei* CONTAINING THE PCR PRODUCT.

It was decided to use the PCR product as a probe to isolate a larger fragment of *B.pseudomallei* DNA containing this DNA sequence, so that sequence information up and down-stream could be analysed. The probe was used to screen a *B.pseudomallei* 204 λGEM11 genomic library constructed by Mr M. Mahfouz, University of Plymouth. This was used in preference to the λGT11 library created (4.4) as λGEM11 (Fig.8.6) is not an expression vector and so potentially toxic genes may be more stable within this system.

8.5.1 DIG-Labelling of the PCR Product from *B.pseudomallei* 204 (pAD1).

A DIG label was incorporated into the sequence through a PCR reaction. The PCR was set up as in 8.2 using primers 2F and 1R and *B.pseudomallei* 204 as the template DNA, and 5 µl of DIG-UTP was incorporated into the reaction (2.7.1). The concentration of the DIG-labelled PCR product was checked by running an aliquot of the reaction products on an agarose gel.

8.5.2 Probing a λGEM11 Gene Bank.

A λGEM11 genomic library was constructed with 204 *B.pseudomallei* genomic DNA by Mr M. Mahfouz. This was achieved by partially digesting 204 *B.pseudomallei* genomic DNA with *Sau*3A and cloning fragments into λGEM11 *Bam*HI arms (Promega), using *E.coli* LE392 as the host. Plates were set up using the same method as in 4.4.2, to have around 1000 plaques per plate.

Nitrocellulose filters were placed onto plates as in 4.5. Once removed, filters were washed firstly with distilled water for 2-3 min. followed by 20X SSC for 2-3 min. Washed filters were placed in-between 2 pieces of filter paper, kept in place with paper clips and baked in an oven at 80°C for 2h. in a vacuum (20 Hg/500 mmHg). Afterwards, filters were incubated with

Hybridisation Solution for 2 h. at a volume of 5 ml per 100 cm^2 of membrane, at 68°C in a hybridisation oven. During this time, the probe created in 8.5.1 was denatured by heating to 95°C for 10 min in a PCR machine and placed directly onto ice. The Hybridisation Solution was replaced with fresh Hybridisation Solution and the denatured probe was also added, at a concentration of approximately 120 ng and incubated for 18 h. at 68°C. Next, filters were washed twice with 50 ml of Wash Solution I at 20°C for 5 min., followed by twice with 60 ml of Wash Solution II for 15 min. at 68°C. Afterwards, filters were washed for 1 min. in 20 ml of Detection Buffer I. The antibody conjugate (anti-AP) was diluted 1:5,000 in Detection Buffer I and the filters incubated for 1 h. with 25 ml of the diluted antibody conjugate at 20°C. Following incubation, filters were washed twice in 100 ml of Detection Buffer I for 15 min. at 20°C. and equilibrated for 5 min. in 30 ml of Detection Buffer III before adding the Colour Development Solution (NBT/X-phosphate system). Positive plaques appeared brown in colour after 1 h. and filters were washed in Detection Buffer IV to stop the colour reaction. Positive plaques were identified and picked (2.3.2.1) along with several negative controls. Positive phage were re-screened as before and seven positives in total were selected (λBP1-7).

8.5.3 Isolation of Recombinant Phage DNA.

High titre stocks of positive phage were made from confluent lysis plates (2.3.2.2) and the stock was titred (2.3.1) before extracting the recombinant phage DNA using Wizard® Lambda Preparations (Promega) (2.3.3.1). Resulting phage DNA was purified using the BIO-RAD Prep-A-Gene kit (2.4.6.2).

8.5.4 Restriction Digests of Recombinant Phage DNA to Estimate the Insert Size.

The λGEM11 genomic library was created using a *Sau*3A partial digest of 204 *B.pseudomallei* genomic DNA ligated into the *Bam*HI site of λGEM11. Therefore, only foreign DNA fragments between 9 and 23 kbp were cloned. There are *Sac*I cleavage sites either side of the

*Bam*HI site (Fig.8.6) hence digestion with *Sac*I releases the cloned DNA fragment, and the pattern of bands observed will reflect any internal sites. Approximately 0.3 μg of λBP1-7 DNA was digested with *Sal*I and the results were analysed by agarose gel electrophoresis (Fig. 8.7). Similarly, the recombinant phage DNA was digested with the restriction enzyme *Xho*I (Fig. 8.8), which also has cleavage sites either side of the *Bam*HI site. The restriction enzyme *Eco*RI (Fig 8.9) was also used, which has no cleavage sites on *Bam*HI-digested λGEM11 DNA and so will only cut internally, therefore reflecting any *Eco*RI sites on the foreign DNA fragment. All digests were heated to 65°C for 10 min. before running on an agarose gel. In most digests there appears two bands, one situated above the 12 kbp band at the top of the 1 kbp ladders, which represents the Left Arm of the vector (20 kbp) and one at 9 kbp which is the Right Arm of the vector.

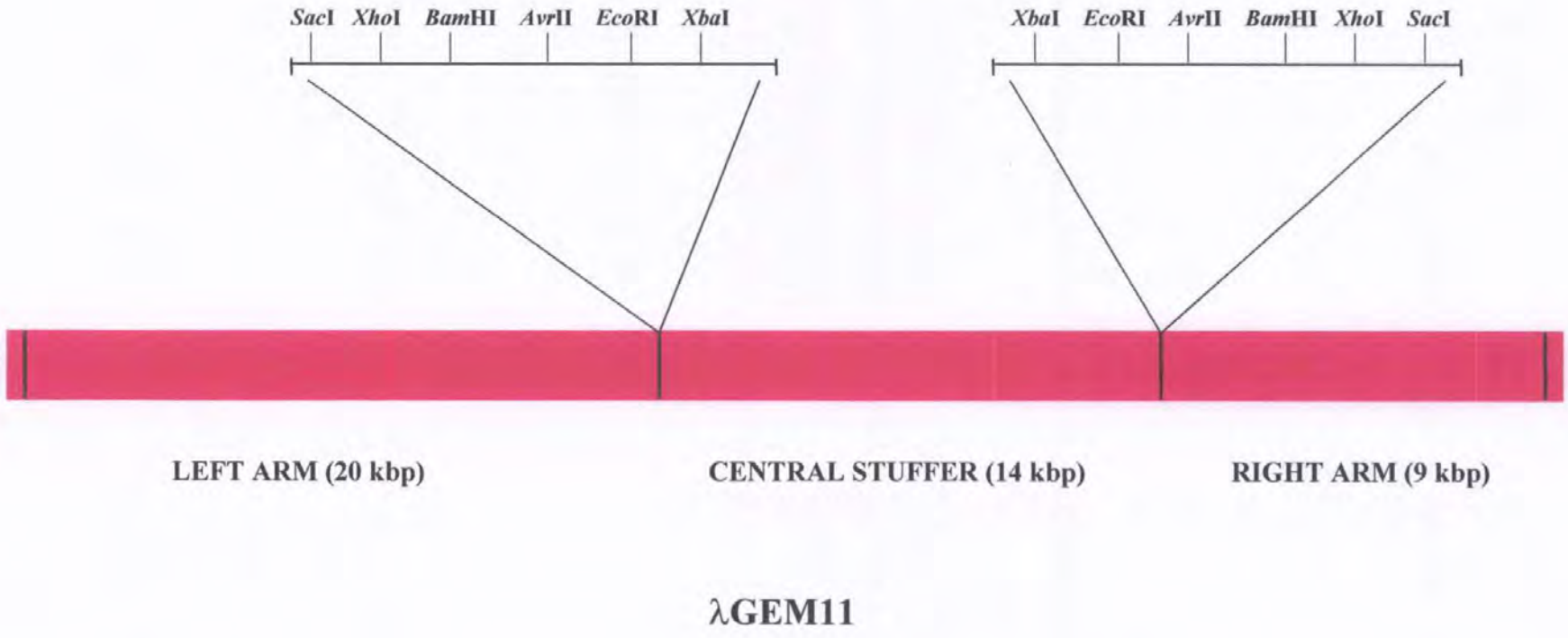
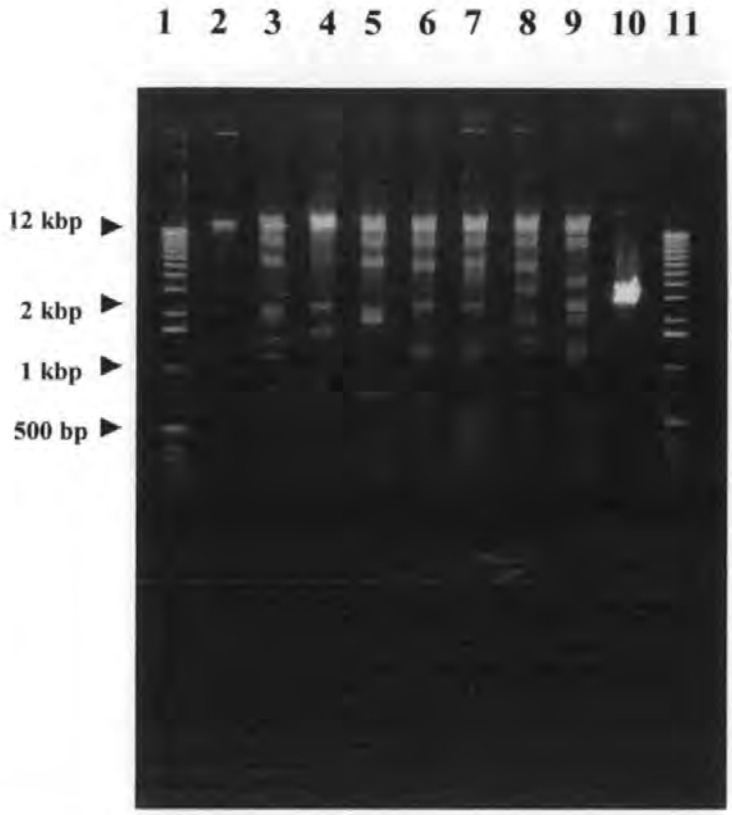


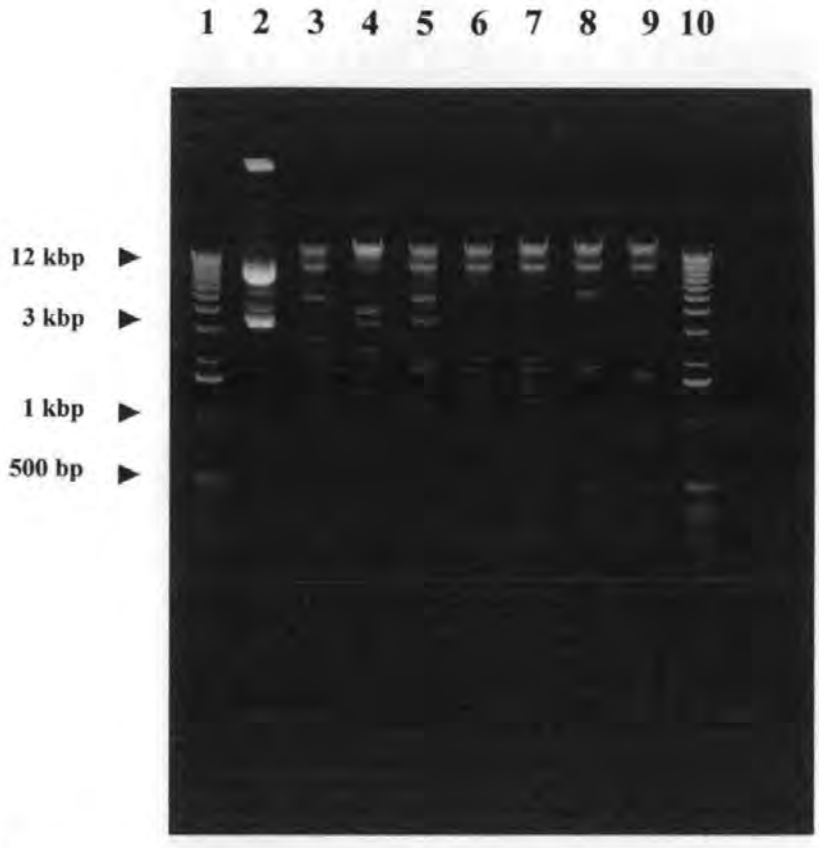
Fig.8.6 The Bacteriophage Cloning Vector, λGEM11. Restriction enzyme cleavage sites are shown for the insertion of foreign DNA fragments either side of the central stuffer region.



LANE	DNA/DIGEST	Number of Bands	Relative Size of Bands (kbp)	Total Size (kbp)
1	1 kbp Ladders	-	-	-
2	Uncut λBP1	-	-	-
3	λBP1/*Sal*I	8	0.6, 0.75, 1.1, 1.5, 2.0, 2.2, 6.0,	14.15
4	λBP2/*Sal*I	*	-	-
5	λBP3/*Sal*I	4	0.75, 2.0, 2.2, 6.0	10.95
6	λBP4/*Sal*I	7	0.55, 0.7, 1.2, 1.3, 2.2, 6.0	11.95
7	λBP5/*Sal*I	*	-	-
8	λBP6/*Sal*I	*	-	-
9	λBP7/*Sal*I	7	0.5, 1.2, 1.4, 1.9, 2.2, 4.0	11.20
10	pUC18/*Sal*I	-	-	-
11	1 kbp Ladders	-	-	-

Those marked with a star (*) indicate that a partial digest was obtained and so analysis of the fragments yielded cannot be carried out accurately.

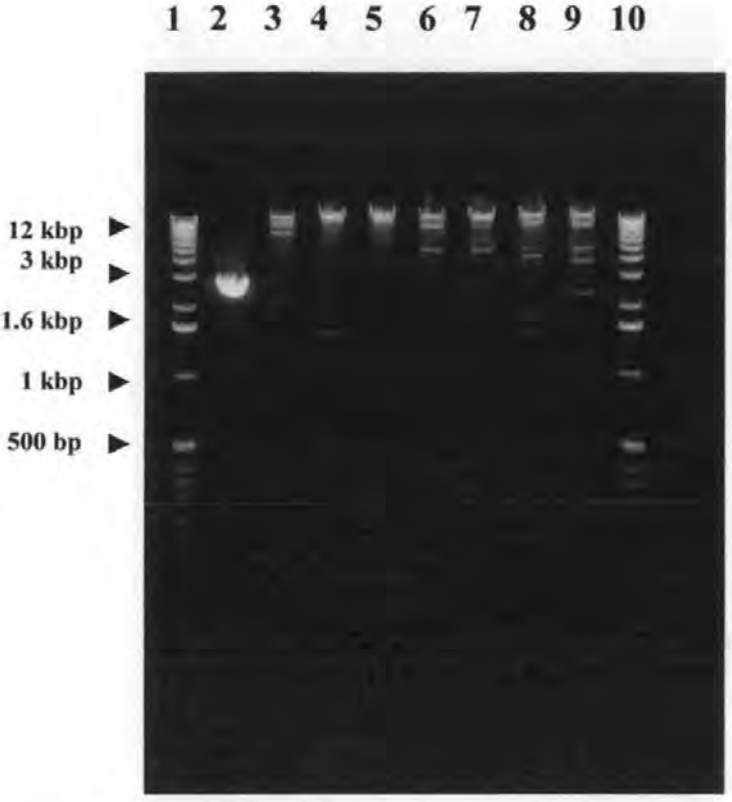
Fig.8.7 Agarose Gel Electrophoresis of Recombinant λGEM11 Phage (λBP1-7) Digested with *Sal*I.



LANE	DNA/DIGEST	Number of Bands	Relative Size of Bands (kbp)	Total Size (kbp)
1	1 kbp Ladders	-	-	-
2	Uncut pUC18	-	-	-
3	λBP1/*Xho*I	5	0.8, 1.3, 1.9, 2.0, 2.6, 5.0	13.7
4	λBP2/*Xho*I	3	2.3, 3.4, 4.0	9.9
5	λBP3/*Xho*I	6	0.6, 1.2, 1.8, 1.9, 3.3, 5.0	13.8
6	λBP4/*Xho*I	6	0.8, 1.0, 1.2, 1.6, 1.9, 2.1	8.6
7	λBP5/*Xho*I	7	0.8, 1.0, 1.2, 1.4, 1.6, 1.9, 2.1	10.0
8	λBP6/*Xho*I	*		
9	λBP7/*Xho*I	4	0.8, 1.0, 1.6, 2.1	5.5
10	1 kbp Ladders	-	-	-

Those marked with a star (*) indicate that a partial digest was obtained and so analysis of the fragments yielded cannot be carried out accurately.

Fig.8.8 Agarose Gel Electrophoresis of Recombinant Phage DNA (λBP1-7) Digested with *Xho*I.



LANE	DNA/DIGEST	Number of Bands	Relative Size of Bands (kbp)	Total Size (kbp)
1	1 kbp ladders	-	-	-
2	pUC18/*Eco*RI	-	-	-
3	λBP1/*Eco*RI	1	8.0	8.0
4	λBP2/*Eco*RI	*	-	-
5	λBP3/*Eco*RI	*	-	-
6	λBP4/*Eco*RI	1	5.0	5.0
7	λBP5/*Eco*RI	*	-	-
8	λBP6/*Eco*RI	*	-	-
9	λBP7/*Eco*RI	3	2.3, 4.0, 5.0	11.3
10	1 kbp Ladders	-	-	-

Those marked with a star (*) indicate that a partial digest was obtained and so analysis of the fragments yielded cannot be carried out accurately.

Fig.8.9 Agarose Gel Electrophoresis of Recombinant Phage DNA (λBP1-7) Digested with the Restriction Enzyme, *Eco*RI.

Analysis of the restriction enzyme digestion patterns visualised on the agarose gels (Figs.8.7, 8.8 and 8.9) was carried out, looking for DNA bands common to a number of clones, indicating that these may represent the same DNA sequence. This was observed in the recombinants, λBP4 and 5, which had very similar banding patterns when digested with *Sal*I and *Xho*I indicating that the cloned sequences were very similar. This is because the recombinants were created from a *Sau*3A partial digest of *B.pseudomallei* genomic DNA, therefore the DNA fragments cloned should all carry the desired DNA sequence, but are all slightly different, depending on the *Sau*3A site cleaved. The other phage DNA restriction digest patterns were all different, with few bands in common. The *Eco*RI digest patterns were all slightly different with the added complication of partial digests. As there was so little phage DNA actually isolated it was decided to firstly clone the 8 kbp fragment created by digesting λBP1 with *Eco*RI and check that it contained the fragment through a Southern blotting procedure. In this way there would be sufficient DNA up and downstream of the PCR product to analyse.

8.6 SUBCLONING AN 8 kbp *Eco*RI FRAGMENT FROM A λGEM11 RECOMBINANT CLONE INTO A PLASMID VECTOR.

8.6.1 Attempts to Subclone into pUC18.

Attempts were made to clone the fragment into the *Eco*RI site of pUC18 but no recombinants could be isolated. Plates were incubated at room temperature with the addition of 0.2% glucose to suppress the *lac* promotor in case a toxic product, deleterious to the *E.coli* was being produced, but recombinants were still not isolated.

8.6.2 Subcloning into pGD103.

As attempts to clone the 8 kbp fragment into the high copy number vector, pUC18 had failed, it was decided to clone the fragment into a lower copy number vector. The plasmid cloning vector, pGD103 (Appendix II) was derived from the low copy number plasmid, pLG339 and permits the propagation of cloned DNA fragments at low gene dosage levels. Hence this vector can be used to clone genes encoding products that are toxic to the host when cloned in a high copy number vector. The selectable phenotype in pGD103 is kanamycin resistance and, unlike its parent pLG339, it contains the *lacZ*α region from pUC8 and so has the advantage of the selection of recombinants by virtue of their *lac*$^-$ phenotype in *E.coli* K12 host strains not possessing the *lacZ*α region.

About 0.5 μg of λBP1 was digested with *Eco*RI and an aliquot of the digestion products was checked on an agarose gel. The products of the digest were then purified by the BIO-RAD Prep-A-Gene kit (2.4.6.2) for use in a ligation reaction. Approximately 0.5 μg of pGD103 was digested with *Eco*RI, treated with alkaline phosphatase (2.5.5) and purified with the BIO-RAD Prep-A-Gene kit. A ligation reaction was set up with pGD103 and the purified digest at a ratio of approximately 1:1, incubated at 15°C for 18 h. and used in a transformation with $CaCl_2$-treated DH5α cells (2.5.1). Transformation mixes were plated out onto LB media supplemented with kanamycin and 0.2% (w/v) glucose. However, half of the transformation plates were kept at room temperature and the other half were incubated at 37°C to investigate whether temperature affected the number and stability of recombinants. Twenty colonies were picked from the room temperature plates (pBPR1-20) and similarly, twenty from the plates incubated at 37°C (pBPT1-20). These were cultured in 10 ml LB broths supplemented with kanamycin and 0.2% (w/v) glucose. Cultures were incubated with shaking according to the temperature used previously on the transformation plates. Therefore subclones pBPR1-20 were incubated at room temperature and subclones pBPT1-20 were incubated at 37°C. When small-

scale plasmid extractions (2.4.2.1) were carried out, it was clear from the size of the plasmids isolated that none of those incubated at 37°C (subclones pBPT1-20) contained an 8 kbp insert and when digested with *Eco*RI were confirmed not to contain an insert. However, of the clones isolated at room temperature, when digested with *Eco*RI, one released an 8 kbp fragment (pBPR1) (Fig.8.10). A 1 L flask of pBPR1 was cultured at room temperature to obtain enough DNA for sequencing, if the fragment was shown by a Southern blot to contain the desired DNA sequence. As pGD103 is a very low copy number vector, large volumes are required to obtain sufficient amounts of DNA for processes such as sequencing. A large-scale plasmid preparation was carried out on the culture (2.4.2.2) and the plasmid DNA was purified by CsCl-EtBr Density Ultracentrifugation (2.4.3.1).

8.6.3 Southern Blot Analysis to Check pBPR1 for the Presence of the Labelled Fragment.

To check that the recombinant plasmid pBPR1 contained the desired sequence (Fig.8.4) a Southern blot was carried out using the previously DIG-labelled PCR product as the probe (8.5.1). Approximately 0.5 µg of *B.pseudomallei* 204 genomic DNA was digested with *Eco*RI and 0.3 µg of pBPR1 was similarly digested with *Eco*RI. These were subject to agarose gel electrophoresis for use in a Southern blotting procedure (2.7.2). The results of the Southern blot can be seen in Fig.8.11. The blot showed a strong positive reaction for the cloned 8 kbp *Eco*RI fragment as well as a corresponding 8 kbp *Eco*RI fragment from *B.pseudomallei* 204 genomic DNA. Proving not only that the 8 kbp fragment contained the desired PCR product sequence, but that the 8 kbp fragment was derived from *B.pseudomallei.*

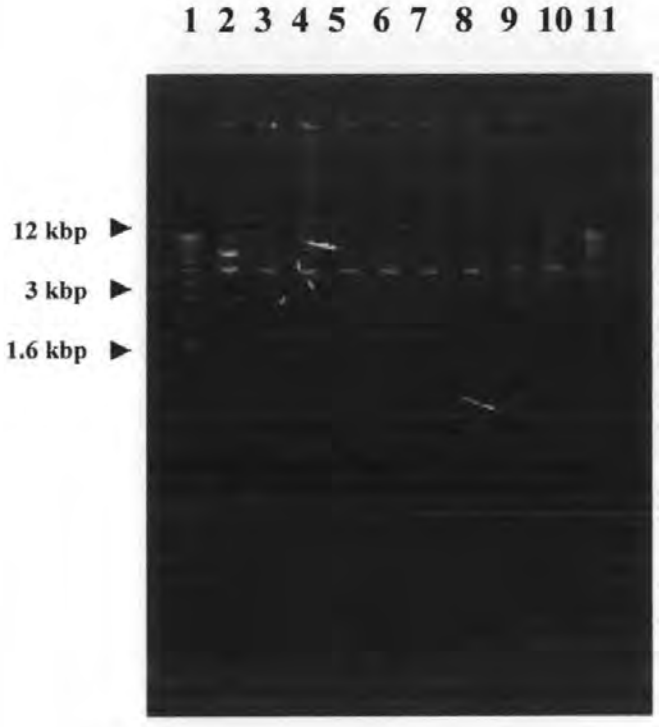


Fig.8.10 Agarose Gel Electrophoresis of Recombinant Plasmids Digested with *Eco*RI. LANE 1: 1 kbp ladders. LANE 2: pBPR1/*Eco*RI. LANE 3: pBPR2/*Eco*RI. LANE 4: pBPR3/*Eco*RI. LANE 5: pBPR4/*Eco*RI. LANE 6: pBPR5/*Eco*RI. LANE 7: pBPR6/*Eco*RI. LANE 8: pBPR7/*Eco*RI. LANE 9: pBPR8/*Eco*RI. LANE 10: pBPR9/*Eco*RI. LANE 11:1 kbp ladders.

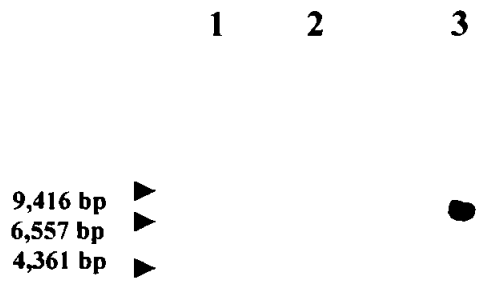


Fig.8.11 Southern Blot to Check for the Presence of a Specific DNA Fragment Contained in a Recombinant Plasmid. LANE 1: λHindIII Markers (DIG-labelled). LANE 2: *B.pseudomallei* 204 genomic DNA/*Eco*RI. LANE 3: pBPR1/*Eco*RI.

8.7 NUCLEOTIDE SEQUENCE ANALYSIS OF THE 8 kbp FRAGMENT.

The recombinant plasmid pBPR1 was DNA sequenced by MWG Biotech Ltd., which showed the cloned fragment to be 7.7 kbp in length and can be found in Appendix III.

8.7.1 Alignment of the PCR Product with the Full Length Sequence.

The first analysis carried out was a BLAST program available at Genbank to align the PCR product sequence (Fig.8.4) used as the probe with the full length 7.7 kbp sequence to ascertain its relative position on the fragment, which can be seen in Fig.8.12. The PCR product was situated near the beginning of the sequence from 434 bp to 863 bp.

8.7.2 Search of the Fragment for any Homologies with ADP-Ribosylating Toxins.

The nucleotide sequence of the 7.7 kbp fragment was searched, firstly at the nucleotide level, using the Web-based BLAST-N service for homology with the nucleotide sequence of the active site of ETA (Gray *et al.*, 1984), which revealed no significant homology. Following this, using the DNasis computer software the 7.7 kbp nucleotide sequence was translated into all six possible reading frames and searched for any of the conserved amino acid sequences common to the active sites of a number ADP-ribosylating toxins (Fig.8.5). This again yielded no observed homologies with the active sites of known ADP-ribosylating toxins. However, as the initial PCR only amplified a product in *B.pseudomallei* isolates, it was decided to search the translated sequences for homologies with any other virulence factors.

8.7.3 Analysis of Web-based BLAST Searches of the 7.7 kbp Fragment for Homologies to Potential Virulence Factors.

As in Chapter 6, the translated sequences were analysed for open reading frames, and each of these were subjected to a BLAST-X search to search the database for homologies with protein sequences of a known or putative function. Over the whole 7.7 kbp sequence four potential open reading frames with homology to known proteins were identified (Fig.8.12).

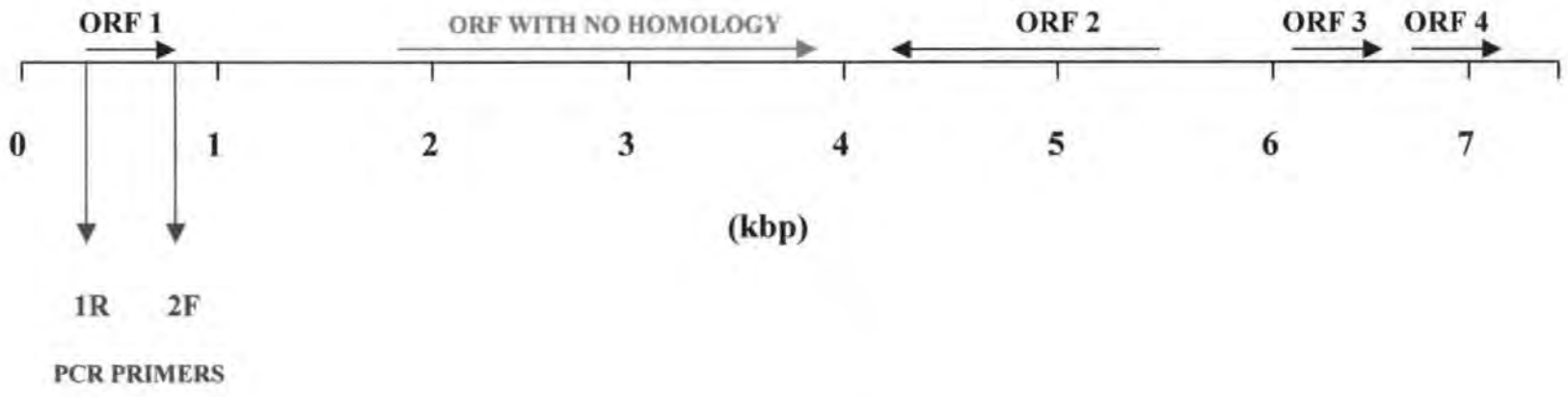


Fig.8.12 Map of the 7.7 kbp *B.pseudomallei* Fragment Showing the Locations of the PCR Primers and Putative OFRs (ORF1-4) with Homology to Proteins in the BLAST Database. The arrow in red denotes a large ORF, spanning over 2 kbp, which had no homology to any proteins in the Database.

8.7.4 Protein Sequences Identified from BLAST Searches with Homology to the Putative ORFs Identified.

The following tables (Tables 8.1-8.4) show the accession numbers and functions of those protein sequences with homology to the sequences encoded by the ORFs identified (Fig.8.12).

Table 8.1 Web-Based BLAST Matches for ORF 1.

Accession Number	Match Sequence Title	Percentage Similarity
AF047828	Syringomycin synthetase *Pseudomonas syringae*	44
AF027770	*fxbC Mycobacterium smegmatis*	36
AL035707	CDA peptide synthetase III *Streptomyces coelicolor*	48
X98690	Pristinamycin I synthase 3 and 4 *Streptomyces pristinaespiralis*	37
Z99120	Similar to antibiotic synthetase *Bacillus subtilis*	40
Z99113	Polyketide synthase *Bacillus subtilis*	40
U08223	Thioesterase *Streptomyces fradiae*	36
U07359	Pyoverdine synthetase D *Pseudomonas aeruginosa*	37
U95370	Thioesterase *Bacillus lichenoformis*	35
AL022076	*pks*13 *Mycobacterium tuberculosis*	36
AJ005061	LchA-TE protein *Bacillus lichenoformis*	36
Y12527	HMWP1 protein *Yersinia enterocolitica*	37
AL021409	Thioesterase *Streptomyces coelicolor*	36
AF074705	Pyochelin synthetase *P.aeruginosa*	38
L42523	Fengycin synthetase *Bacillus subtilis*	38

The putative ORF 1 has homology to a number of peptide synthetase genes most notably, those with an involvement in the acquisition of iron. For example, pyochelin synthetase (AF074705) and pyoverdine synthetase (U07359) from *P.aeruginosa* are involved with the synthesis of the siderophores, pyochelin and pyoverdine, respectively. A number of others are involved with the synthesis of antibiotics. Thus the *B.pseudomallei* putative peptide synthetase gene may therefore have a role in the virulence of the pathogen.

Table 8.2 Web-Based BLAST Matches for ORF 2.

Accession Number	Sequence Match Title	Percentage Similarity
D638223	*opc*P1 *Burkholderia cepacia*	51
L36817	ORF9 *Ralstonia eutropha*	49
X58488	Outer Membrane Porin Protein Precursor *Bordetella pertussis*	49
U16266	ompQ Protein *Bordetella pertussis*	37
A38528	Outer Membrane Porin Protein 32 Precursor *Comamonas acidovorans*	41
X65530	Major Outer Membrane Protein Precusor (Porin) *Neisseria meningitidis*	41
X79464	Class 3 Outer Membrane Protein *N.meningitidis*	41
X65531	Class 3 Major Outer membrane Protein precursor (serotype 4) *N.meningitidis*	40
U07190	Class 3 Outer Membrane Porin *N.meningitidis*	40
J03017	Major Outer Membrane Protein Precursor (Porin) *N.gonorrhoeae*	38
AF051531	PorB Protein *N.meningitidis*	39
AF060527	Porin Protein *N.meningitidis*	40

The sequence encoded by ORF 2 has homology to a number of porin genes from a range of bacteria.

Table 8.3 Web-Based BLAST Matches for ORF 3

Accession Number	Match Sequence Title	Percentage Similarity
Z99105	Similar to Transcriptional Regulator (AraC/Xyls) *B.subtilis*	50
U79580	*chp*D *P.aeruginosa*	48
AL031031	Putative Transcriptional Regulator *Streptomyces coelicolor*	38
AF128627	*mtr*A *N.gonorrhoeae*	49
U35231	Probable Transcriptional Regulator *lum*Q *Photobacterium leiognatha*	49
AE000137	Putative AraC-type Regulatory Protein *E.coli*	42
AF145724	AraC-like Transcriptional Regulator Homologue *Streptomyces albus*	52
M96551	Putative AraC-like Transcriptional Regulator *Streptomyces antibioticus*	39
U50452	AraC-like Regulator *ybt*A *Yersinia pestis*	41
AJ132668	Yersiniabactin Transcriptional Regulator (*ybt*A) *Y.enterocolitica*	41
U59485	*att*O *Agrobacterium tumefaciens*	41
Z48959	AraC-like Protein *Azorhizobium caulinodans*	40
D90899	Regulatory Protein PchR *Synechocystis* sp.	36
Z11519	Putative AraC-like Transcription Regulator *Streptomyces lividans*	38

ORF 3 encodes a sequence that has homology to a number of transcriptional regulators belonging to the Xyls/AraC family of prokaryotic transcriptional regulators.

Table 8.4 Web-Based BLAST Matches for ORF 4.

Accession Number	Match Sequence Title	Percentage Similarity
Z99107	Similar to hypothetical proteins from *Bacillus subtilis*	39
D26185	Hypothetical 32.9 kDa protein in TETB-EXOA Intergenic Region *Bacillus subtilis*	35
Z19055	Hypothetical 34.4 kDa protein in TRPA 3' region *Buchnera aphidicola*	53
AF027868	Hypothetical 33.0 kDa protein in PELB-PENP Intergenic Region *Bacillus subtilis*	45
AJ248287	Hypothetical protein *Pyrococcus abyssi*	46
AF056092	Unknown *Yersinia enterocolitica*	51
AF072874	Putative Membrane protein DcsA *Mycobacterium smegmatis*	44
AF013775	PagO *Salmonella typhimurium*	52
AB001488	Hypothetical 33.6 kDa protein in CSPC-NAP Intergenic Region *Bacillus subtilis*	45
X13330	Hypothetical 32.2 kDa protein in DRSA-VSR Intergenic Region *Escherichia coli*	45
AL031124	Putative integral membrane protein *Streptomyces coelicolor*	38
AE000861	Hypothetical protein *Methanobacterium thermoautotrophicum*	40
X73124	Hypothetical 31.3 kDa protein in PTA 3' Region *Bacillus subtilis*	36

The sequence encoded by ORF 4 has homology to a number of hypothetical integral transmembrane proteins with an unknown function, although many of these are found in intergenic regions.

A putative ORF was identified (Fig.8.12) which was over 2 kbp long but had no homology to protein sequences in the database. It may be possible that it encodes a novel protein that has not been found in any other bacterial species.

8.8 DISCUSSION.

The PCR primers designed to the active site of the ETA gene of *P.aeruginosa* failed to identify a similar active site in *B.pseudomallei*. However, the primers used amplified a 453 bp product in the *B.pseudomallei* isolates tested but failed to amplify a similar product in the avirulent *B.thailandensis* isolates. The exception to this was *B.pseudomallei* isolate 4845. This isolate was donated by Dr. R Titball, Porton Down and is possibly the oldest isolate in the collection (Table 2.1) and therefore has been subcultured frequently, and so it is possible that it may have

lost some of its virulence determinants. However, it could be simply that the genomic DNA was of a poor quality and contaminated with complex carbohydrates, which was a common problem with genomic DNA preparations received. However, no products were amplified with the *B.thailandensis* isolates used, and so the sequence could be part of a gene or genes involved in virulence due to this association. Thus the sequence was used as a probe to isolate sequence information up and down-stream for detailed sequence analysis.

When using this sequence as a probe to isolate a larger sequence, a number of positive clones were identified in the λGEM11 library of *B.pseudomallei* 240 DNA. This genomic library was used in preference to the λGT11 library created in Chapter 4 because it is not an expression library. If indeed the probe isolated larger sequences encoding toxic products then the sequence would be more stable in the λGEM11 library. Attempts to extract the recombinant phage DNA from the positives yielded very little DNA, which hampered restriction enzyme analysis of the clones, as barely enough DNA was isolated to carry out the three digests (8.5.4). Moreover, the quality of some of the phage DNA isolated was not of a standard for efficient restriction digestion and so some could not even be analysed. The only recombinant phage from which sufficient DNA was extracted was λBP1, which gave a single, approximately 8 kbp product on digestion with the restriction enzyme, *Eco*RI. Therefore it was decided to attempt to clone this fragment, as firstly, it was not necessary to isolate the 8 kbp *Eco*RI fragment before cloning, as it could be cloned directly from the purified digest, hence less DNA needed to be used and secondly, it was a large sized fragment and likely not only to contain the sequence, but sufficient DNA up and down-stream for sequence analysis. However, the only way the 8 kbp *Eco*RI DNA fragment could be successfully subcloned was by using a low copy number plasmid (pGD103) and by keeping all incubation temperatures at room temperature. This indicated that the fragment might encode genes toxic to the *E.coli* host when in a high copy number vector, such as pUC18 and if incubated at 37°C. In human

bacterial pathogens it is often the case that certain virulence genes are switched on at 37°C on entry to the human body. Therefore, keeping the subclone at room temperature may have functioned to suppress the expression of such genes. As the subclone was particularly difficult to clone it was decided to DNA sequence the fragment to determine if it did encode a toxin or indeed other putative virulence genes.

Sequencing of the fragment revealed that here was no homology on the 7.7 kbp fragment with the sequence at the nucleotide or amino-acid level with the ETA gene of *P.aeruginosa* nor with any active sites from other ADP-ribosylating toxin genes. This does not discount the presence of an ADP-ribosylating toxin in *B.pseudomallei* for the active site may not possess the sequence encoded by the primers used or the primers may have failed to amplify the site. A PCR product of approximately 900 bp was amplified once with *B.pseudomallei* 204 DNA which may have been the active site, but the product could not be amplified again even after numerous attempts. However, the 7.7 kbp sequence revealed putative open reading frames with homology to a number of interesting and genes from other bacteria some of which may have roles in virulence.

The first ORF (ORF 1) encoded a sequence with homology to a number of peptide synthetase involved in the biosynthesis of a number of important genes from a range of bacteria. These included iron acquisition genes such as pyoverdin and pyochelin in *P.aeruginosa*, and *fxb*C involved in exochelin biosynthesis in *M.smegmatis* (Yu *et al.*, 1998). Also, the biosynthesis of a number of antibiotic genes in bacteria and the biosynthesis of syringomycin, a phytotoxin from *P.syringae*, which has a role in the virulence of the bacterium (Bender *et al.*, 1999). Obviously more work is required characterising this putative gene in *B.pseudomallei* but it could be an important gene in the biosynthesis of a possible virulence factor.

A putative ORF (ORF 2) was identified with homology to porins from different bacteria, including pathogenic species of *Neisseria*, in particular *N.meningitides* and *N.gonorrhoeae*. The porin of *N.gonorrhoeae* is postulated to have a role in virulence by allowing the bacterium to survive inside phagocytes, through two activities. Firstly, when added to cultured epithelial cells it inserts into the membrane and collapses the membrane potential, probably by making pores, and secondly it is able to bind the eukaryotic calcium-binding protein calmodulin (Salyers and Whitt, 1994). More recently the porin has been shown to cause rapid calcium influx in target cells and to induce apoptosis by the activation of cysteine proteases (Muller *et al.*, 1999) and also that the porin is itself able to arrest phagosome maturation within macrophages (Mosleh *et al.*, 1998). *B.pseudomallei* has been shown to be able to survive within human phagocytes (Pruksachartvuthi *et al.*, 1990; Jones *et al.*, 1996) and, moreover, has been postulated by Harley *et al.* (1994) to escape, once phagocytosed, through damage to the vacuolar membrane. If indeed the putative porin of *B.pseudomallei* has a similar function to that of the species of *Neisseria* the porin may contribute to the intracellular survival of the pathogen.

Another ORF (ORF 3) was shown to have homology to putative transcriptional regulators, belonging to the AraC/Xyls family of transcriptional regulators. Most proteins within this family are positive transcriptional regulators and have three main regulatory functions:

1. Carbon metabolism,
2. Stress response and,
3. Pathogenesis (Gallegos *et al.*, 1997).

The putative gene in *B.pseudomallei* has homology to *ybt*A, which is the transcriptional regulator of the siderophore, yersiniabactin, receptor (*psn*) in *Y.pestis* and *Y.enterocolitica*. Interestingly, this forms part of an inorganic iron transport system that functions at 37°C (Fetherston *et al.*, 1996) and the sequence in *B.pseudomallei* could not be subcloned when

incubated at 37°C. A putative peptide synthetase (ORF 1) also had homology to a number of genes involved in the biosynthesis of iron in other bacteria. Perhaps these genes are therefore part of an iron-regulated operon or perhaps are involved in iron acquisition themselves. Obviously characterisation of these putative genes is necessary and more sequence information would be required up and down-stream to investigate this.

In addition to homology to *ybt*A of the *Yersinia* species, the putative transcriptional regulator also had homology to *mtr*A of *N.gonorrhoeae*. The *mtr* (multiple transferable resistance) gene complex in *Neisseria gonorrhoeae* encodes an energy-dependent efflux pump that serves to export structurally diverse antimicrobial, hydrophobic agents (HAs) (Rouquette *et al.*, 1999) and *mtr*A is involved with the regulation of this. Again, ORF 1 had homology to genes responsible for the biosynthesis of antibiotic resistance factors in other bacteria. It is not known exactly what the functions are of these putative genes in *B.pseudomallei* but it is possible that they encode products with a possible function in antibiotic resistance, or iron acquisition or perhaps even another putative virulence factor, as a part of the sequence had no homology to proteins in the database.

CHAPTER 9

*Sau*3A UNDIGESTIBLE *B.pseudomallei* 4845 GENOMIC DNA FRAGMENTS.

9.1 INTRODUCTION.

When creating gene banks the genomic DNA of the bacterium under study is digested with restriction enzymes in order to create fragments for cloning. Quite often this involves partially digesting the DNA with *Sau*3A in order to create fragments of a size suitable for cloning. The restriction endonuclease, *Sau*3A is a four-base cutter with the recognition sequence: ↓GATC and therefore cuts DNA relatively frequently in comparison to six-base cutters such as *Eco*RI. It would be expected that any particular tetranucleotide sequence to occur about once every 4^4 (256) nucleotide pairs in a long random DNA sequence assuming that all the bases are included equally.

*Sau*3A also recognises a tetranucleotide sequence that is included within the hexanucleotide sequence recognised by the enzyme, *Bam*HI. In this case, therefore, the cohesive termini produced by these enzymes means that fragments produced by *Sau*3A will cohere to those produced by *Bam*HI.

As *Sau*3A does digest DNA so rapidly often the digest is set up in a series of partial digests with the enzyme. This then produces a random set of DNA fragments. A series of tubes are set up with a standard amount of genomic DNA but with varying amounts of enzyme produced by a doubling dilution series. These are then left at 37°C for 20 to 30 minutes after which time the

reactions are terminated via inactivation of the enzyme by placing the tubes at 65°C for 10 minutes. A sample of the digestion products are subjected to agarose gel electrophoresis and the enzyme concentration producing DNA fragments of an appropriate size for cloning can be determined.

9.2 RESTRICTION DIGESTION OF *B.pseudomallei* GENOMIC DNA WITH *Sau*3A.

9.2.1 *B.pseudomallei* Isolate 4845.

Serial digestions using the restriction enzyme *Sau*3A were carried out with 4845 *B.pseudomallei* genomic DNA. A series of MCC tubes containing *B.pseudomallei* 4845 genomic DNA were set up whereby the first tube contained double the amount of genomic DNA than the following tubes. Approximately 4 µg of 4845 DNA was contained in the first tube in a total of 39 µl, including a 0.1 volume of *Sau*3A restriction enzyme buffer and the following tubes contained half this amount of DNA in a volume of 20 µl, again including a 0.1 volume of buffer. One µl of *Sau*3A restriction enzyme was added to the first tube, mixed and 20 µl of this preparation transferred into the second tube, mixed and 20 µl of this preparation transferred to the third tube, and so on for a total of 8 tubes. The MCC tubes were incubated at 37°C for 20 min. before placing at 70°C for 15 min. to stop the digestion reaction. A 5 µl sample of each tube was then subject to agarose gel electrophoresis (Fig.9.1).

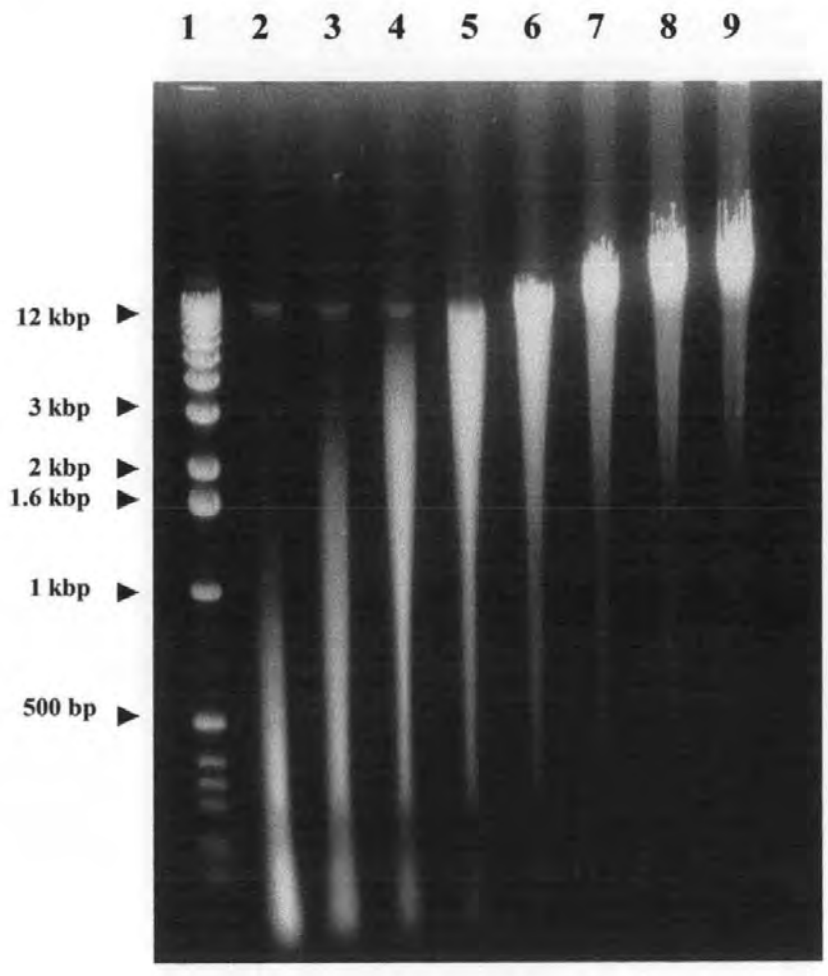


Fig.9.1 Agarose Gel Electrophoresis of a Serial *Sau*3A Digest of 4845 *B.pseudomallei* Genomic DNA. LANE 1:1 kbp ladders. LANE 2: Tube 1. LANE 3: Tube 2. LANE 4: Tube 3. LANE 5: Tube 4. LANE 6: Tube 5. LANE 7: Tube 6. LANE 8: Tube 7. LANE 9: Tube 8. LANE 11: 1 kbp ladders.

There appears to be two bands of the order of 10 kbp in size that have not been digested by *Sau*3A. The restriction enzyme *Sau*3A theoretically cuts DNA every 256 bp and so over 10 kbp it should have cut the fragment at least 39 times. No plasmids have been reported in any isolates of *B.pseudomallei* and so it is unlikely to be two plasmids. As this was observed with one isolate it was tried with other isolates of which there was a sufficient quantity of genomic DNA available.

9.2.2 *B.pseudomallei* Isolates E8, 25, 204 and 576 and *B.thailandensis* 27 and 82.

Approximately 1 μg of genomic DNA from all isolates was digested with 1 μl of *Sau*3A and incubated at 37°C for 1 h. the reaction was stopped by placing the tubes at 65°C for 10 min. Products of the digests were visualised by agarose gel electrophoresis (Fig. 9.2). Interestingly, genomic DNA from *B.pseudomallei* isolates E8 and 576 have completely digested but 25 has a band the same size as that of 4845, and 204 has a faint, smaller band than the others. *B.thailandensis* isolates 27 and 82 contain undigested DNA bands also. Isolate 27 has two, one slightly smaller than the 10 kbp band of 4845 and another at around 6 kbp, whereas 82 has three at approximately 5, 9 and 10 kbp. It was decided to investigate the undigested band of *B.pseudomallei* isolate 4845 as there was a surplus of genomic DNA available for this isolate and it would be necessary to digest high quantities of DNA in order to isolate useable amounts of the undigested fragments of DNA.

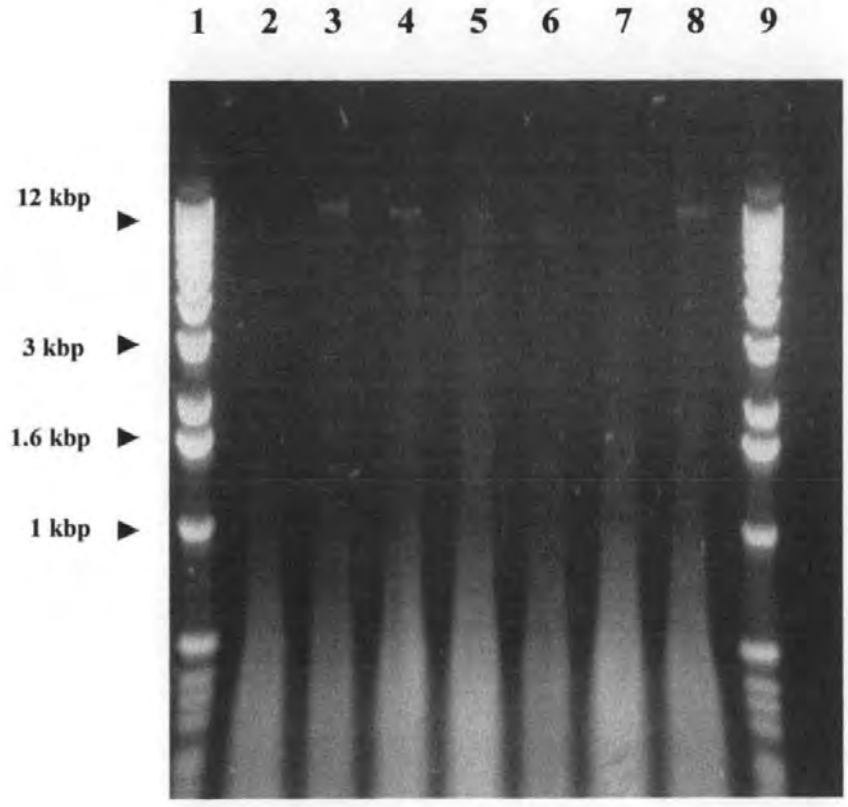


Fig.9.2 Agarose Gel Electrophoresis of *Sau*3A Digested Genomic DNA from a Number of Isolates of *B.pseudomallei* and *B.thailandensis*. LANE 1: 1 kbp ladders. LANE 2: E8. LANE 3: 25. LANE 4: E27 LANE 5: 82. LANE 6: 204. LANE 7: 576. LANE 8: 4845. LANE 9: 1 kbp ladders.

9.3 ISOLATION OF THE *Sau*3A UNDIGESTIBLE *B.pseudomallei* 4845 GENOMIC DNA FRAGMENTS.

Approximately 25 μg of *B.pseudomallei* 4845 genomic DNA was digested with an excess of *Sau*3A for 1 h. at 37°C. The products of digestion were subject to agarose gel electrophoresis and the 10/12 kbp undigested bands were excised from the agarose gel and collected by electroelution (2.4.6.1) and purified using the BIO-RAD Prep-A-Gene Purification Kit (2.4.6.2).

9.4 ATTEMPTS TO CLONE THE *Sau*3A UNDIGESTIBLE FRAGMENTS INTO PLASMID VECTORS.

9.4.1 Attempts to Clone the Fragments into *Bam*HI Sites of Plasmid Vectors.

The nature of the termini of the *Sau*3A undigestible fragment was unknown and so firstly it was assumed that they possessed *Sau*3A termini. Several attempts were made to clone the fragments into pUC18, pBR328 and pGD103 digested with *Bam*HI. However, none of these yielded any results and so it was assumed that the fragments did not have *Sau*3A termini.

9.4.2 Cloning the Fragments Using the SURECLONE™ Kit

As the nature of the termini were still unknown it was decided to perform a blunting and kinasing reaction on the fragments to be able to clone them into a blunt-ended site. The *Sau*3A undigestible fragments were used in the SURECLONE™ kit to attempt to clone into the *Sma*I site of pUC18 (2.5.7). The product of ligation was transformed using DH5α competent cells (2.5.3.2) and two transformations were carried out, one incubated at room temperature and the other at 37°C. It was noted that the transformation put at 37°C, after the incubation time (90 min.) the transformation mixture appeared clear, indicating that the bacterial cells had lysed.

The mixture was still plated out onto LB plates supplemented with ampicillin and 0.2% glucose to repress the *lac* promoter, so that if indeed the fragments encoded toxic products this would limit the expression of them, but no colonies grew on these plates. The transformation mixture kept at room temperature overnight was plated out onto the same media and left to incubate at room temperature. Eight colonies grew on these plates (pSAU1-8) and so these were transferred to 10 ml LB broths containing ampicillin and 0.2% glucose and incubated with shaking at room temperature. Small-scale plasmid extractions (2.4.2.1) were carried out on the cultures and the plasmid DNA was digested with *Eco*RI and *Pst*I to release any cloned fragment(s), the results of which can be seen in Fig.9.3.

The results of this are somewhat ambiguous, as in some of the clones the size of the linearised pUC18 appears to vary. Clones pSAU5-8 have what appears to be a smaller sized linearised pUC18 DNA band than the control. The plasmid pSAU1 has no insert, which is also the case with pSAU5-7. Only pSAU2, pSAU4 and pSAU8 have inserts, which all very in size and none of which have a single 10/12 kbp band. The recombinant, pSAU2 appears to have two inserts, which although add up to around 12 kbp are slightly strange in appearance as the upper band is not as bright as the two lower bands and should appear brighter as it contains more DNA.

Due to the ambiguous nature of the clones in pUC18, attempts were made to clone the fragment into the *Eco*RV site of the lower copy vector, pGD103. However no recombinants were isolated.

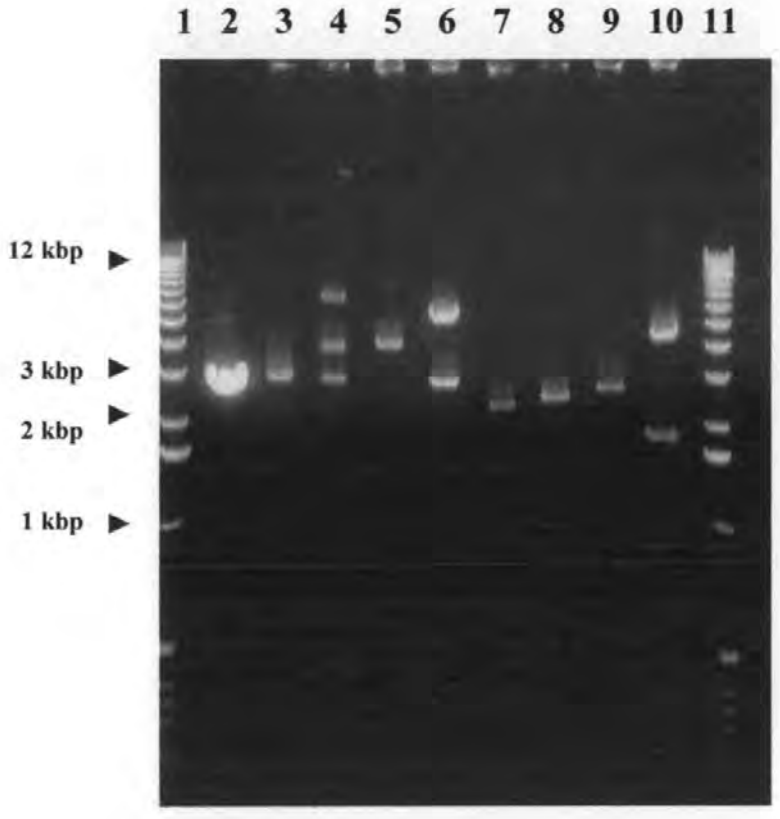


Fig.9.3 Agarose Gel Electrophoresis of Recombinant Plasmids Digested with the Restriction Enzymes *Eco*RI and *Pst*I to Release any Cloned Fragments. LANE 1: 1 kbp ladders. LANE 2: pUC18/*Eco*RI. LANE 3: pSAU1/*Eco*RI/*Pst*I. LANE 4: pSAU2/*Eco*RI/*Pst*I. LANE 5: pSAU3/*Eco*RI/*Pst*I. LANE 6: pSAU4/*Eco*RI/*Pst*I. LANE 7: pSAU5/*Eco*RI/*Pst*I. LANE 8: pSAU6/*Eco*RI/*Pst*I. LANE 9: pSAU7/*Eco*RI/*Pst*I. LANE 10: pSAU8/*Eco*RI/*Pst*I. LANE 11: 1 kbp ladders.

9.4.3 PCR Reactions Using pUC Derived Primers.

Due to the ambiguous nature of the results of the digests with the recombinant plasmids, it was decided to use pUC-based primers in a PCR reaction to ascertain the size of the insert from clones pSAU2, 4 and 8 all of which appeared to have inserts. The pUC primers were designed to either side of the multiple cloning site by Dr. M. Kiernan thus the PCR should amplify whatever DNA is between these.

Forward primer 5' **GTAAACGACGGCCAGT** 3'

Reverse primer 5' **GTACCAGTATCGACAA** 5'

A PCR was carried out with an annealing temperature of 50°C and an extension time of 10 min. for 35 cycles, with twice the normal amount of dNTPs as the inserts could possibly be several kbp in length. The results of which can be seen in Fig.9.4. Again the results were somewhat ambiguous but an insert could be visualised and what appeared to be a ladder of DNA up from the main band at approximately 700 bp. Different annealing temperatures and condition were attempted but yielded no differing products to the ones gained in Fig.9.4. Due to the interesting nature of these clones and the fact that probable deletional events had occurred it was decided to send them for DNA sequencing at Porton Down.

9.4.4 Cultures of pSAU2, 4 and 8 Grown for Sequencing.

One litre cultures of LB containing ampicillin and 0.2% glucose (w/v) of pSAU2, 4 and 8 were set up and incubated with shaking at room temperature. The plasmids were extracted using the maxiprep plasmid preparation (2.4.2.2) and purified by CsCl/EtBr density ultracentrifugation (2.4.3.1). The plasmid DNA gained was analysed by restriction digestion with *Eco*RI and *Pst*I to check again the insert size. The recombinant plasmids pSAU4 and 8 still contained fragments of the same. However pSAU2 contained what appeared to be two fragments in the Fig.9.3, but after a large-scale plasmid preparation it now only contained the smaller fragment (Fig.9.5).

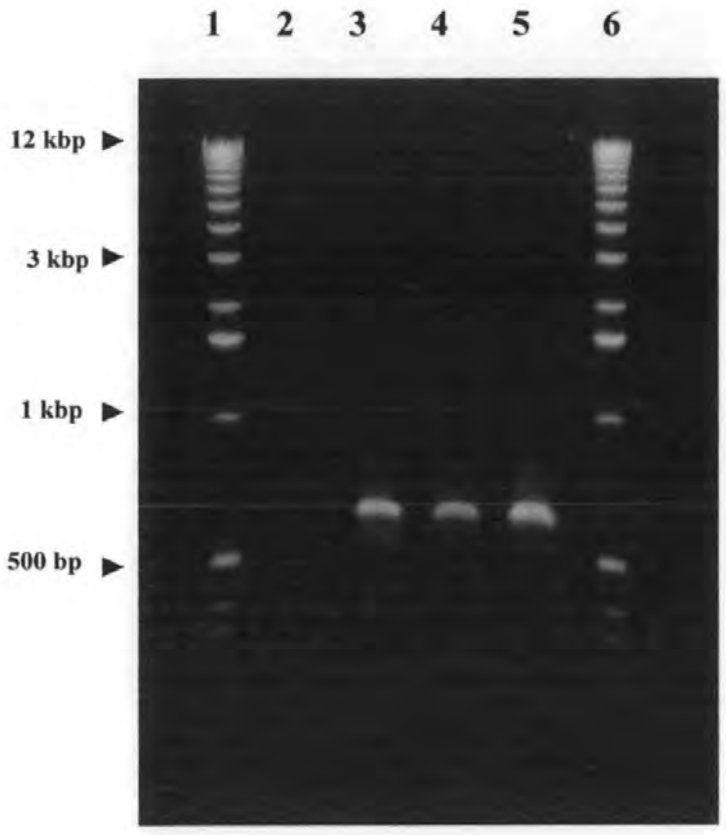


Fig.9.4 Agarose Gel Electrophoresis of the PCR Products from Clones pSAU2, 4 and 8 using pUC Derived Primers. LANE 1: 1 kbp ladders. LANE 2: Negative control. LANE 3: pSAU2, LANE 4: pSAU4. LANE 5: pSAU8. LANE 6: 1 kbp ladders.

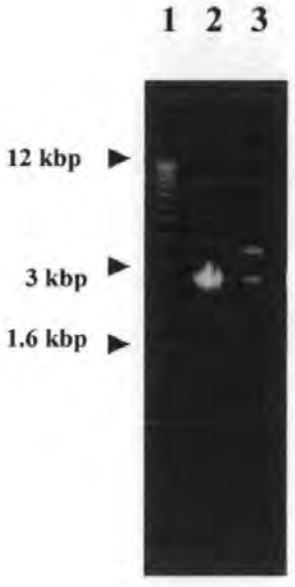


Fig.9.5 Agarose Gel Electrophoresis of Recombinant Clone pSAU2 Isolated from a 1 Litre Culture and Digested with Restriction Enzymes *Eco*RI and *Pst*I. LANE 1: 1 kbp ladders. LANE 2: pUC18/*Eco*RI. LANE 3: pSAU2/*Eco*RI/*Pst*I

9.5 TRANSFECTION PROCEDURE.

In a previous experiment (9.4.2) it was observed that when the transformation mix was placed at 37°C it was noted that the preparation went clear and therefore the bacteria had appeared to lyse. This may have been due to toxic products encoded by the fragments, which had a detrimental effect on the host cells, but given the 90 min. incubation time such products may not have had time to be produced. However, lysis may have been because the 10/12 kbp fragments of DNA were bacteriophage in origin, hence causing the bacteria to lyse. To investigate this, a transfection procedure was attempted with the DNA fragments, which is based on the calcium shock method of Mandel and Higa (1970). Transfection in this case is the uptake of viral DNA (or RNA) by competent cells. A 10 ml culture of *E.coli* C600 cells was incubated at 37°C for 18 h. in NZCYM broth. A 0.5 ml aliquot was removed and used to inoculate a 20 ml volume of NZCYM broth, which was incubated at 37°C for 100 min. After which time, the cells were pelleted by centrifugation at 5,000 x g, for 10 min. at 4°C and resuspended in 20 ml of ice-cold 0.1 M $MgCl_2$. Cells were again pelleted as before and resuspended in 10 ml of ice-cold 0.1 M $MgCl_2$. Finally cells were pelleted and resuspended as before but in 1 ml of ice-cold 0.1 M $CaCl_2$. Approximately 50 ng of the 10/12 kbp DNA fragments isolated was mixed with a 0.5 ml volume of the competent cells, heat shocked at 42°C for 2 min., cooled and incubated on ice for 30 min. The preparation was added to 10 ml of molten NZCYM medium and overlayed onto NZCYM plates and incubated for 18 h. at 37°C. However no plaques were observed on the plates.

9.6 ATTEMPTS TO USE A LONG TEMPLATE PCR SYSTEM TO ISOLATE LARGE QUANTITIES OF THE 10/12 kbp DNA FRAGMENTS.

Problems are often encountered when trying to clone large fragments into plasmid vectors and as so little of the undigestible DNA from the *Sau*3A digest could be isolated, restriction analysis was not possible. It was therefore decided to try an experimental procedure by using the fragments in a ligation with pUC18, but not transforming the resulting ligation products. In this way no selection pressure is put on the sequence by introducing it into a host organism. Instead, as there may be some successful ligations of the whole 10/12 kbp fragments in the ligation mixture, by using commercial systems that allow the PCR amplification of such large fragments, large quantities of the fragments may be isolated for restriction mapping and subcloning. As PCR is highly sensitive it was hoped that it might be able to amplify the successfully ligated product in the reaction, even if the ligation was not highly efficient. The system used was the Boehringer Mannheim, Expand™ Long Template PCR System and is designed to amplify fragments up to 27 kbp from human genomic DNA and up to 40 kbp from λDNA. A ligation was firstly set up using the SURECLONE™ kit as in 9.4.2 and incubated at 15°C for 18 h. The ligation products were precipitated with ethanol and resuspended in analar water. The following reaction was set up firstly as two mastermixes.

Mastermix 1:

Component	Volume (µl)
dNTPs	3
Forward Primer	1
Reverse Primer	2
Template	4
Analar Water	To make a total of 25 µl

Mastermix 2:

Component	Volume (µl)
Buffer 1	5
Enzyme	0.75
Analar Water	To make a total of 25 µl

The forward and reverse primers used were the pUC primers used in 9.4.3. The two mastermixes were added, overlaid with 30 µl of Mineral Oil and cycled through the following program:

Function	Temperature °C	Time	Number of Cycles
Denaturation	94	2 min.	1
Denaturation	94	10 sec.	
Annealing	55	30 sec.	10
Elongation	68	10 min.	
Denaturation	94	10 sec.	
Annealing	55	30 sec.	20*
Elongation	68	10 min.	
Prolonged Elongation	68	7 min.	1

* The elongation time was increased by 20 sec. for each cycle to increase the yield of the product. For example, cycle number 11 had an additional 20 sec. and cycle number 12 had an additional 40 sec. and so on. This was carried out altering the annealing temperature and the elongation time, but no product was visualised on an agarose gel.

9.7 NUCLEOTIDE SEQUENCE ANALYSIS OF CLONES pSAU2 and pSAU8.

Both clones were partially sequenced by Dr.R.Titball, Porton Down, Salisbury. However the clones were very difficult to sequence and only a small amount of sequence information was derived from them. For pSAU2 only the first 308 bp could be sequenced in the forward direction and for pSAU8 only 181 bp could be sequenced in the forward direction and 209 bp in the reverse. The other clone pSAU4, which had a 5 kbp insert was in fact a contaminant.

9.7.1 Partial Nucleotide Sequence of pSAU2.

GGGTCGGCGAGACGNTGACCATTNNGCATCAGATGNGAGAANAGACCACCTGTN

GNGTTGNTTNATTGATGCNTGCCTGTNGGCNTGTGATCTGGTTACGNGCACGAGCT

AACCGAGAGGACTGTAGAACCCATTCCGGGGTACCTGANTATGTCGTGAAGAANG

GNGACNGCAGCATNGACNTGTCTCNGATNCAATCACAGTTGAAACGACAGTCGNG

AGGGGTGGATGTAGTCATNGATGGATACAGGCAGGGTCGCCTTACCNTAGGCCTT

TGTGCGCAGTTGCTTGGACGTTCGCCTGTAGAC (308 bp).

N represents any nucleotide base.

10.8.2 Nucleotide Sequence of pSAU8.

10.8.2.1 Forward Sequence.

CAGGNCGACTCTAGAGGATCCCCGTTTCAATAACAACCGACAGATTCCAATCGCC

CAATACGAGGAATTAAAGCTGACTTCTGCACAGGGTATACAAGAGTGCTTCCAGT

TCTCAAGAGGCCGGCGCTGCAGAGGCTTTCGTCGGCTCATTGATTCAACTCGGCC

AATACGGAGCAAATAC (181 bp).

10.8.2.2 Reverse Sequence.

GATCGTGCGGCGCACGGGGCTCGCAAGCGCTTTGTGAATGGCGTCGGCATCAAAG

GTCATGGGCGGCGCGGTGAAATCGCGAATCGTCGTTTGCCGAAGCTTAAATCGGC

CGATGACGATATAAACAGCAACACTCTATTGACTAGGGGCGATGGTTGATGTCCA

ACAATTAAGTCTGTCGGGCGACGATAAACTCTGTCGGACGGCAG (209 bp).

9.8 DISCUSSION.

Most bacteria have defence systems to guard against invasion by foreign DNA, which involve the production of specific endonucleases. Bacteria protect their own DNA from the potentially lethal effect of their restriction enzymes by previously modifying it, usually by an appropriate DNA methylase. The methylation of certain bases at a limited number of sequences is carried out so that the enzymes no longer recognise them as sites. It would appear that the undigestible 10/12 kbp fragments from *B.pseudomallei* 4845 have been protected in some way against the restriction enzyme, *Sau*3A.

It was found that the DNA fragments deleted when attempts were made to clone them into the plasmid vector pUC18. The undesirable property of this vector, with regards to this experiment, is that it has a high copy number. Hence if the DNA encodes a toxic gene product it may be produced at a level that is deleterious to the host, and so it would be better to clone such a fragment into a lower copy number vector. Many attempts were made to clone the fragment into the lower copy number vector, pGD103, without any success. During attempts to clone the fragments into pUC18, the procedure was carried out using a range of temperatures and it was found that lowering the incubation temperature to room temperature at every step permitted the cloning, albeit deletions of the original fragments. A similar finding was observed in Chapter 8 when subcloning a 7.7 kbp *B.pseudomallei* fragment into pGD103, which could only be successfully subcloned at room temperature. It was discussed that as *B.pseudomallei* is a human pathogen, the expression of certain virulence genes may be triggered on entry to the human body, by the change in temperature to 37°C. Therefore, incubating *B.pseudomallei* clones encoding such thermoregulated genes at a lower temperature may function to repress the expression of virulence gene products that have a deleterious effect upon *E.coli* host cells. In the case of the *Sau*3A undigestible fragments even incubation at room temperature could not facilitate the cloning of the whole fragment and so only clones

containing deletions of the fragments were successful. It must also be considered that the fragments are of the order of 10/12 kbp and large fragments such as these are not cloned as faithfully into plasmids as smaller fragments and that an increased metabolic load imposed upon the cell can instigate plasmid instability itself.

When the recombinant clones pSAU1-7 were initially isolated in Fig.9.3 there were some ambiguous findings. The recombinant plasmid pSAU2 appeared to contain a cloned fragment of around 10 kbp when calculating the combined sizes of the two bands (7 kbp and 3 kbp) produced on digestion of the plasmid DNA with *Pst*I and *Eco*RI. However, after a large-scale plasmid DNA isolation procedure was carried out on a litre culture, there was only one, 3 kbp band cloned (Fig.9.4). It is likely that deletion events were occurring and on close examination of the agarose gel in Fig.9.3, the top band of around 7 kbp was not as bright as the bands beneath. The upper band should appear brighter than the bands beneath it because there would be more DNA fluorescing. What may have happened in this case is that there were two clones present in the preparation, and the one containing the 3 kbp band was present in higher numbers than the clone containing a 7 kbp band, and therefore the former was unwittingly picked for large-scale plasmid DNA isolation. Or perhaps the smaller sized clone was under less selection pressure and when cultured on a larger scale, the clone containing the 7 kbp fragment became present at a much smaller proportion or even selected out completely. Alternatively, the clone may have originally contained a 10 kbp fragment and was undergoing a deletion event due to it's instability in the *E.coli* host, whereupon DNA was perhaps being looped out to leave just 3 kbp and what was observed on the agarose gel represented a culture containing recombinant plasmids at different stages in this process. Therefore, by the time a large culture was produced, all of the clones were deletions of the original 10 kbp fragment.

Another ambiguous result in Fig.9.3 is that pSAU5-8 have what appears to be smaller sized linearised pUC18 DNA bands than the control of 2.7 kbp. When deletion events occur, a piece of DNA is looped out and with pSAU5-8 it is possible that this process has actually removed a part of the vector DNA, which could be why they appear smaller than the control. This might also have occurred in pSAU3, as the linearised plasmid, instead of being around 2.7 kbp, was around 4 kbp. However, in this case digesting with *Eco*RI and *Pst*I functioned to linearise the recombinant plasmid, as one of the sites was lost, due to a deletion but the other still existed.

A myriad of attempts were tried to clone the full-length fragments all of which failed. Interestingly, it was discovered that a similar finding was observed with *Francisella tularensis* genomic DNA and the undigestible fragment also proved impossible to clone intact (R.Titball, personal communication). In one transformation procedure the bacterial cells appeared to lyse after incubation at 37°C (9.4.2), which indicated that the DNA might be phage-encoded. In many bacterial pathogens, virulence determinants are often encoded on a phage, for example, diphtheria toxin in *Corynebacteria diptheriae* is encoded on a lysogenic phage. It is possible that the undigestible DNA may be bacteriophage in origin and due to the difficulties encountered during attempts to clone the full-length fragments it is possible that the DNA encodes toxic products or perhaps other virulence-encoding genes. There exists pathogenicity islands (PAIs) which have now been detected in a number of bacteria, which are large pieces of chromosomal DNA that carry virulence genes, which were most probably acquired by phage-mediated horizontal transfer (Buchrieser *et al.*, 1998). In addition to this, in a study carried out on the lysogenic bacteriophage, MAV1 from *Mycoplasma arthritidis*, which is associated with the arthritogenicity of the bacterium (Voelker and Dybvig, 1998), it was found that the MAV1 genome contained no sites for the restriction enzymes, *Sau*3A and *Mbo*I which both recognise the sequence GATC.

Due to the interesting nature of these DNA fragments it was decided to sequence the clones, despite the fact that they contain deleted DNA, to ascertain any information regarding the nature of the fragment. Unfortunately the DNA proved very difficult to sequence and only a small amount of information could be derived, which was riddled with ambiguities. However, there appeared a *Sau*3A site in the sequence of pSAU2 and so the DNA may therefore be protected in some way, as opposed to containing no cleavage sites for *Sau*3A. This protection though, would not be through the host cells having GATC-specific modifications or GATC-specific DNA methyltransferase activity, because, as with the *M.arthritidis* genomic DNA, the rest of the *B.pseudomallei* genomic DNA is readily digestible with the enzyme. Then again, due to the ambiguities contained within the sequence and that a deletion event had occurred, it is possible that the *Sau*3A site does not actually exist. On the whole, it is possible that this DNA may be bacteriophage in origin and may encode virulence factors.

CHAPTER 10

USE OF THE *B.pseudomallei* GENOME SEQUENCING PROJECT

10.1 INTRODUCTION.

At the International Congress for Melioidosis held in Bangkok, November 1997, it was announced that the Sanger Centre had been funded by Beowulf Genomics to sequence the genome of *B.pseudomallei*. Genome sequencing began early in the year 2000 in collaboration with Dr. R.Titball of the Defence Evaluation Research Agency, Porton Down and Dr. T. Pitt of the Central Public Health Laboratory. The strain being sequenced K96243, was chosen as a typical example of the species and is a clinical isolate from a melioidosis patient in Thailand, supplied by Dr. S. Songsivilai of Mahidol University.

The method in use by the Sanger Centre is a whole genome shotgun approach making use of several libraries. The results of the sequencing project are available on the internet along a with a BLAST facility to search the sequenced data from the shotgun sequences for homologies at the nucleotide level. At the time of writing there were 58, 859 reads in the database giving a total of 26.574 Mb and a theoretical coverage of 98.3% of the genome. The sequences contained in the shotgun database, however, are unedited single reads from ABI sequencers and so contain errors as well as *E.coli* and vector contamination. Work has begun on assembling the shotgun sequences into contiguous fragments and this information is also available to search.

This facility is therefore an invaluable tool to search the genome for sequences from other bacterial species, which may indicate the presence of homologous sequences in *B.pseudomallei*. In this research project there are a number of lines of experimental work where the genome sequencing project could not only be applied, but serve as a very useful tool. Firstly, when testing isolates of *B.pseudomallei* and *B.thailandensis* for activities (Chapter 3) it was discovered that the isolates of *B.pseudomallei* had hexosaminidase activity. However, on screening the λGT11 expression library for a clone with this activity, one could not be found. However, by searching the shotgun and assembled databases with sequence information encoding hexosaminidase genes in other bacteria, a corresponding gene in *B.pseudomallei* may be found.

Similarly, PCR primers were designed to the active site of ETA in *P.aeruginosa* and isolates of *B.pseudomallei* and *B.thilandensis* were tested for the presence of a similar sequence (Chapter 8). The sequence information derived from the reaction products indicated no obvious homology with the active site of the toxin. It is possible to use the same sequence information as a probe to screen the databases to determine if there is a similar sequence, which was not amplified by the PCR primers designed.

Additionally, two 10/12 kbp *Sau*3A undigestible fragments were isolated in Chapter 9 but cloning of the full-length fragments proved impossible. Recombinants containing deletions of the original fragments were isolated but DNA sequencing proved difficult and only a small amount of sequence information was gleaned. However, it was possible to use this sequence information to search for its possible presence in the clinical isolate sequenced and to analyse any assembled sequence information containing the sequence in WEB-based BLAST searches for homologies with any other sequences with known or putative functions in other bacteria.

10.2 PROBING THE SHOTGUN DATABASE FOR THE PRESENCE OF A HEXOSAMINIDASE GENE.

A number of bacterial hexosaminidase gene sequences were deposited in the Genbank database and so several of hexosaminidase amino acid sequences were used to search the *B.pseudomallei* genome project databases to identify any shotgun or assembled sequences with homology (Table 10.1).

Table 10.1 Bacterial Hexosaminidase Amino Acid Sequences Used to Search the *B.pseudomallei* Genome Sequencing Project Databases for Homologous Sequences.

Accession Number	Organism	Sequence Title	Contig/Shotgun Reference	Similarity (%)
U24658	*Vibrio parahaemolyticus*	Beta-hexosaminidase	Contig 1720	52
D42078	*Staphylococcus aureus*	N-acetyl-glucosaminidase	NONE	-
AL162754	*Neisseria meningitidis*	Putative hexosaminidase	Burk258f03.p1c	74
AE004725	*P.aeruginosa*	Beta-N-acetyl-D-glusosaminidase	Burk258f03.p1c	69
AE004155	*V.cholera*	Beta-hexosaminidase	Burk258f03.p1c	60
AE004294	*V.cholera*	Beta-N-acetyl hexosaminidase	Contig 1720	64
L43594	*Serratia marcesans*	N-acetyl-β-glucosaminidase	Contig 1720	62
AB022786	*Enterobacter* sp.	N-acetyl-β-D-glucosaminidase	Contig 1720	61

From the preliminary searches carried out, two different hexosaminidase sequences were identified. One of these was contained on a shotgun sequence, Burk258f03.p1c and another on an assembled contiguous sequence, Contig. 1720.

10.2.1 Identification of a Putative Hexosaminidase Gene Using the Shotgun Fragment Burk258f03.p1c.

The shotgun fragment Burk258f03.p1c, a 612 bp sequence, had a high homology to hexosaminidase sequences from a number of bacteria. This sequence was used to search the shotgun database for more information up and down-stream of the fragment (Table

10.2). The sequence was used on the assembled database but the sequence of this shotgun fragment has yet to be assembled.

Table 10.2 Shotgun Sequences Identified Up and Down-Stream of Burk258f03.p1c.

FRAGMENT ISOLATED UPSTREAM	SEQUENCE INFORMATION (bp)	FRAGMENT ISOLATED DOWNSTREAM	SEQUENCE INFORMATION (bp)
Burk281d11.q1ca	331	Burk211c06.q1ca	332
Burk170h1.q1c	210		

A 1485 bp sequence was compiled from the shotgun fragments and used in a Web-based BLAST-X search. There exists an open reading frame with a high homology to a range of hexosaminidase genes (Table 10.3). The complete nucleotide sequence for the putative hexosaminidase gene can be seen in Fig.10.1.

Table 10.3 Results of a Web-Based BLAST-X Search Using the Putative Hexosaminidase Sequence Compiled from the Shotgun Fragment Burk258f03.p1c.

Accession Number	Match Sequence Title	Similarity (%)
AL162754	Putative hexosaminidase *N.meningitidis*	69
AE002408	Glycosyl hydrolase *N.meningitidis*	68
AE004725	Beta-N-acetyl-D-glucosaminidase *P.aeruginosa*	63
U32777	Beta-hexosaminidase *Haemophilus influenzae*	61
AE000211	Probable glucosidase *E.coli*	59
AE004045	N-acetyl-beta-glucosaminidase *Xylella fastidiosa*	59
U52818	Beta-hexosaminidase *V.furnisssii*	59
AE004155	Beta-hexosaminidase *V.cholera*	59
AF124757	Unknown *Zymomonas mobilis*	50
AE001979	Glycosyl hydrolase *Dinococcus radiodurans*	50
AL359214	Putative sugar hydrolase *Streptomyces coelicolor*	46
AP001509	Beta-hexosaminidase *Bacillus halodurans*	47

1	CACGTTCGTC	GTCGCCGAGG	CGCCGGACGA SD	CGTCGCCGCC	GCGCGCTCGG START
51	GCGCCGCTTC	CTGATCCATT	TTCCGCAAGG	ACGCTGATTC	GATGAAACTG
101	CCCCCCGGTC	CGGTGATGCT	CGACGTCGCC	GGCACGACGC	TCACGCGCGA
151	CGACGCGCGC	CGCCTCGCGC	ATCCGCACAC	GGGCGGCGTG	ATCCTGTTCG
201	CGCGCCACTT	CGAGAGCCGC	GCGCAACTCG	TCGCGCTGAC	CGAGGCGATC
251	CGGGCGATCC	GCGACGGCAT	CCTGATCGCG	GTCGATCACG	AGGGCGGCCG
301	CGTGCAGCGC	TTTCGCACCG	ACGGCTTCAC	CGTGCTGCCG	GCGATGCGCC
351	GGCTCGGCGA	GCTGTGGGAC	AAGGACGTGC	TGCACGCGAC	GAAGGCGGCG
401	ACCGCGCTCG	GCTATGTGCT	CGCTTCCGAG	CTGCGCGCGT	GCGGCATCGA
451	CATGAGCTTC	ACGCCCGTGC	TCGATCTCGA	TTACGGCCGC	TCGAAGGTGA
501	TCGGCGATCG	CGCGTTCCAT	CGCGATCCGC	GCGTCGTCGC	GTTGCTCGCG
551	AAGAGCGTCA	ACCACGGGCT	CGCGCTCGCC	GGGATGGCGA	ACTGCGGCAA
601	GCATTTTCCC	GGCCACGGCT	TCGCGCAGGC	CGATTCGCAC	GTCGCGCTGC
651	CGACCGACGA	TCGTCCGCTC	GACGAGATCC	TCGCGAACGA	CGCGGCGCCG
701	TACGACTGGC	TCGGGCTGTC	GTTGTCGGCC	GTCATTCCGG	CGCACGTGAT
751	CTACACGCAA	GTCGATTCGA	AGCCGGCCGG	CTTCTCGCGC	GTGTGGTTGC
801	AGGACGTGCT	GCGCGGCCGG	CTGCGCTTTG	CGGGCGCCGT	GTTCAGCGAC
851	GATCTGTCGA	TGGAGGCCGC	GCGCGAGGGC	GGCACGCTCG	CGCAGTCGGC
901	GCAGGCCGCG	CTCGAGGCGG	ACTGCGACAT	GGTGCTCGTG	TGCAACCAGC
951	CGGATGCGGC	GGAGCGGGTG	CTCGACGAGC	TGCGCACGAC	GGCGTCGCGC
1001	GAATCGTCAC	GGCGGATCAA	GCAAATGCGG	CCGCGCGGCA	AGGTGCTCGA
1051	GTGGCGCAAG	CTGATGCGCG STOP	AGCCGCGCTA	TCTGAATGCG	CAGGGTCTGT
1101	TGCGCAGCAC	GTTCGCCTGA	CGCGAAGCGA	AAACGGCAGC	CGAAGGCGGG
1151	AGAGGGCGCA	GCGAAGGCGC	GATGAAGGCG	AAACGTAGCC	GACGCACGGG
1201	GCGGCGCGGC	GGCTCGTTCG	CCGCGTCCGG	GCGTCAGCCG	ATCTT

Fig.10.1. Nucleotide Sequence of a Putative Hexosaminidase Gene from *B.pseudomallei* Compiled from the Shotgun Sequence, Burk258f03.p1c. The sequence includes a putative SD sequence but no further sequence information was available up-stream to facilitate the identification of putative –10 and –35 promoters.

10.2.2 Complete Amino Acid Sequence of a Putative Hexosaminidase Gene.

The nucleotide sequence (Fig.10.1) when translated, encodes a 342 amino acid sequence and a predicted molecular weight of 37.5 kDa. The hexosaminidase genes from other bacteria which had homology to the *B.pseudomallei* putative hexosaminidase were between 314 and 361 amino acids in length.

MKLPPGPVMLDVAGTTLTRDDARRLAHPHTGGVILFARHFESRAQLVALTEA
IRAIRDGILIAVDHEGGRVQRFRTDGFTVLPAMRRLGELWDKDVLHATKAAT
ALGYVLASELRACGIDMSFTPVLDLDYGRSKVIGDRAFHRDPRVVALLAKSV
NHGLALAGMANCGKHFPGHGFAQADSHVALPTDDRPLPEILANDAAPYDWL
GLSLSAVIPAHVIYTQVDSKPAGFSRVWLQDVLRGRLRFAGAVFSDDLSMEAA
REGGTLAQSAQAALEADCDMVLVCNQPDAAERVLDELRTTASRESSRRIKQM
RPRGKVLEWRKLMREPRYLNAQGLLRSTFA.

10.2.3 Isolation of a Putative Hexosaminidase Gene on Contig. 1720.

Another region of the genome, Contig. 1720 was identified as having homology to the hexosaminidase genes used to screen the database. Contig. 1720 was a 23,000 bp assembled fragment and the region with homology to the hexosaminidase genes was at the end of the sequence, starting at around 20,000 bp. The end 3,000 bp of the nucleotide sequence was used in a Web-based BLAST-X search to identify protein sequences in the database with homology to the translated nucleotide sequence (Table 10.4).

Table 10.4 Results of a Web-Based BLAST Search using a Putative Hexosaminidase Sequence Located on an Assembled Fragment from the *B.pseudomallei* Genome Sequencing Database, Contig 1720.

Accession Number	Sequence Title	Similarity (%)
L43594	Chitobiase Precursor (N-acetyl-beta-glucosaminidase) *Serratia marcesens*	62
AB022786	N-acetyl-beta-D-glucosaminidase *Enterobacter* sp.	61
AE004294	Beta-N-acetylhexosaminidase *Vibrio cholera*	60
J05004	N-acetyl-beta-glucosaminidase (Chitobiase) *Vibrio harveyi*	61
AF196349	Chitobiase *Aeromonas hydrophila*	59
L04544	Beta-N-acetylhexosaminidase (Chitobiase) *Vibrio vulnificus*	49
AB031320	Beta-N-acetylglucosaminidase *Aeromonas* sp.	45
AB026053	Transglycosylative enzyme *Altermonas* sp.	42
U24658	Beta-N-hexosaminidase *Vibrio parahaemolyticus*	40
AF063001	Beta-N-acetylhexosaminidase *Streptomyces plicatus*	38
AE003924	Beta-hexosaminidase precursor *Xylella fastidiosa*	43

An open reading frame was identified in the sequence and the complete nucleotide sequence for the putative hexosaminidase gene can be found in Fig.10.2.

```
1     CCTTGAAAGA  CCCGATCGCA  GCGACGTATT  TCCACGGCGT  CGTATCGCCG
51    TTTTTTTGCC  AAGGAACGCC  CATGAACCGA  ATCTCGCATT  CCCTGTGCGC
101   CGCGCTATTG  GCCGCCGCGA  CGCTGTTGCC  CACGGCCTCG  CGCGCGCAAC
151   TGCCCGCGCG  CCCCACGGCC  GGCGCGGCCG  CGCCCGCGAC  GGCCGCGCCC
201   GTGCGGCCCG  CGTCCACGCC  GGCCGAGCTC  GCCGCGCGGC  TCGCCAACGG
251   CCTCGCGGTG  CGCGTGGCCG  TCGACAACAA  TCACGCGGCA  TCGGCCGGCG
301   TGCCGTGCGC  CGACCTCGGC  GCGGACTGGG  CGAGCTGCGC  GACGGGCCGC
351   CTGATCCTGC  AGAATCGCGG  CCACTCGCCC  CTCACCGACG  GCGGCTGGAA
401   GCTCTATCTG  CACAGCATCC  GCCGGCTGCT  CCGAATCGAC  CGCCCCGGCT
451   TCACGCTGCG  CCATCTGACG  GGCGATCTGT  ACGAGCTGAC  GCCGCAGCCC
501   GGCACGGTAA  GGCTCGCGCA  GGGCGAGCGC  ATCGAGCTGC  CGTTCGTCGC
551   CGAATACTGG  CTGCGCCGCT  ACAGCGACGT  GATCCCGCGC  CCGTACGTGG
601   TCGTCGACGG  CGCGGCGCCC  GCGGTGCTGC  GCTACGACGA  TACCGACGAC
651   GAGCTGCGCT  ACGTCGAAAC  GCTGCCCGCC  GACGCGCAGA  ACAACTCGCC
701   CGGCAATGCG  CCGCCCGCCG  CCGCGCAGCC  GGTGGCGAAC  CGCGCGCTGC
751   CGAGCGTGAA  GCGGCAGCGC  GCGCTGCCCG  GCGCGCTCGA  TCTGCGCGGC
801   GTCGAGCTGA  CGCTGCCGGA  GCTGCCGTCC  GCGCAGGTCG  CGGCGCTGCG
851   CGAACGCGCG  GGCACGCTCG  GCCTGGACGG  CGCGCGCGTG  CCGGTGTGGG
901   GCGTCGTCGC  GCCGCGCCGG  CTGCCCGCCG  ACATCGCGGT  GCCGGGCGGC
951   TACCGGCTCG  CGATCGGCCC  GCGCGGCGCG  TTCATCGAGG  GGGCCGATCG
1001  CGCGGGCCTC  TACTACGGCG  TGCAGACGCT  CTTCTCGCTC  GTGCCGGCCG
1051  GCGGCGCGAC  GGTGCCCGCG  ATGCTGATCG  AAGACGCGCC  GCGCTTCACG
1101  CACCGCGGGA  TGCACGTCGA  TCTCGCGCGC  AACTTCAAGC  CGCCCGCCAC
1151  GCTGCGCCGG  CTGATCGACC  AGATGAGCGC  GTACAAGCTC  AACCGGCTGC
1201  ATCTGCACCT  GTCCGACGAC  GAGGGCTGGC  GCATCGAGAT  TCCCGGCCTG
1251  CCCGAGCTGA  CCGACGTCGG  CGCGCGCCGC  TGCCACGACC  CGAGCGAGAC
1301  GCGCTGCCTG  CTGCCGCAGC  TCGGCTCGGG  GCCCGACGAT  CGTTCGGGCG
1351  GCGGCTACCT  GACGCGCGAC  GACTACGTCG  CGCTGCTGCG  CTACGCGGCC
1401  GAGCGCTTCG  TCGAAGTGAT  CCCCGAGATC  GACATGCCCG  CGCACTCGCG
1451  CGCGGCCGTC  GTATCGATGG  AGGCGCGCTA  TCGCCGCCTG  CACGCGGCGG
1501  GCCGCGAGCG  GGAAGCGAAC  GCGTATCGGC  TGCTCGATGC  GCAGGACACG
1551  TCGAACCTGC  TGACCGTGCA  GTTCTACGAC  CGGCGCAGCG  ATCTGAACCC
1601  GTGCATGCCG  GGCGCGCTGA  ACTTCGCGTC  GAAGGTGATC  CGCGAGATCG
1651  CGTCGATGCA  CGCGGACGCG  CAAGCGCCGC  TGCGGATCTG  GCACTTCGGC
1701  GGCGACGAGG  CGAAGAACAT  CCTGCTCGGC  GCGGGCTTCC  AGCCGCTCGA
1751  CGGCGCCGAT  CCCGGCAAGG  GCCGCGTCGA  TCTCGCCGCG  CAGGACAAGC
1801  CGTGGGCGCG  CTCGCCCGCC  TGTACGGCGC  TGCTTCGGCG  CGGCGAGATC
1851  AAATCGATCG  ACGAATTGCC  GACGCGCTTC  GCGAAGCAGG  TCAGCGCGAT
1901  CGTGAACGCG  AACGGAATCG  GCACGATGGC  CGCGTGGCAG  GACGGCATCA
1951  AGCACGCGAG  CGGGCCGCGG  GAGTTCAGCA  CGCGGCACGT  GATGGTGTCG
2001  CTGTGGGACA  CCATCTTCTG  GGGCGCGTCC  GACAGCGCGC  GCGATCTGAG
2051  CGCGAAGGGC  TACCGGACCG  TGCTCGCGCT  GCCCGATTAC  CTGTACTTCG
2101  ATTTCCCGTA  CACGCGCAAT  CCGCGCGAGC  GCGGCTATTA  CTGGGGCTCG
2151  CAGGCGACGG  ACGAGTACAA  GGTGTTCTCG  CTCGCGCCGG  AGAACCTGCC
2201  GCAGAACGCC  GAGGTGTTCG  GCGATCGCGA  CGGCAACCCG  TTCGAGGTGA
2251  CGAGCGCGGG  CGCGGCGCCG  AGCATCGAGG  GCATCCAGGG  GCAGGCGTGG
2301  GGCGAGGTGA  TGCGCAACGG  GCAACTGCTC  GAATACATGG  TGTATCCTCG
2351  CCTTCTCGCG  CTCGCCGAGC  GCGCGTGGCA  CAAGGCCGAC  TGGGAACTGC
2401  CCTACGCGGC  CGGCGTGCGC  TACAAGCTCG  GCGACACGCA  TCACGTCGAC
2451  ACGGCCGCGC  TCGAGCGCGA  CTGGGCGGGC  TTCGCGACGG  TGCTCAAGCA
2501  GCGCGAACTG  CCGAAGCTCG  AGCGTGCGGG  CATCGGGTAT  CGCAAGCCCA
2551  CGTTTACGCT  GACGGGCGAA  TGACGGGACG  ATGATGGGCC  GATCGCATCG
```

Fig.10.2 Nucleotide Sequence of a Putative Hexosaminidase Gene Identified on Contig. 1720. The nucleotide sequence has been extended to include putative –35 and –10 promoter sequences and a putative ribosome binding site.

10.2.4 Complete Amino Acid Sequence of a Putative Hexosaminidase in *B.pseudomallei.*

The nucleotide sequence in Fig.10.2 encodes an 833 amino acid sequence, with a predicted molecular weight of 91 kDa. Most of the hexosaminidase amino acid sequences with homology to the putative *B.pseudomallei* hexosaminidase were between 858 and 885 amino acids in length.

MNRISHSLCAALLAAATLLPTASRAQLPARPTAGAAAPATAAPVRPASTPAEL
AARLANGLAVRVAVDNNHAASAGVPCADLGADWASCATGRLILQNRGHSPL
TDGGWKLYLHSIRRLLRIDRPGFTLRHLTGDLYELTPQPGTVRLAQGERIELP
FVAEYWLRRYSDVIPRPYVVVVDGAAPAVLRYDDTDDELRYVETLPADARNNS
PGNAPPAAAQPVANRALPSVKRQRALPGALDLRGVELTLPELPSAQVAALRE
RAGTLGLDGARVPVWGVVAPRRLPADIAVPGGYRLAIGPRGAFIEGADRAGL
YYGVQTLFSLVPAGGATVPAMLIEDAPRFTHRGMHVDLARNFKPPATLRRLI
DQMSAYKLNRLHLHLSDDEGWRIEIPGLPELTDVGARRCHDPSETRCLLPQL
GSGPDDRSGGGYLTRDDYVALLRYAAERFVEVIPEIDMPAHSRAAVVSMEAR
YRRLHAAGREREANAYRLLDAQDTSNLLTVQFYDRRSDLNPCMPGALNFASK
VIREIASMHADAQAPLRIWHFGGDEAKNILLGAGFQPLDGADPGKGRVDLAA
QDKPWARSPACTALLRRGEIKSIDELPTRFAKQVSAIVNANGIGTMAAWQDGI
KHASGPREFSTRHVMVSLWDTIFWGASDSARDLSAKGYRTVLALPDYLYFDF
PYTRNPRERGYYWGSQATDEYKVFSLAPENLPQNAEVFGDRDGNPFEVTSAG
AAPSIEGIQGQAWGEVMRNGQLLEYMVYPRLLALAERAWHKADWELPYAA
GVRYKLGDTHHVDTAALERDWAGFATVLKQRELPKLERAGIGYRKPTFTLT
GE.

10.3 ATTEMPTS TO PROBE THE SHOTGUN DATABASE FOR THE PRESENCE OF AN ADP-RIBOSYLATING TOXIN.

The shotgun and assembled databases of the *B.pseudomallei* genome sequencing project were searched for any homologies with the ETA gene of *P.aeruginosa* (Gray *et al.*, 1984). No significant homology was observed using the sequence of ETA, or the active site of the gene either at the amino acid or nucleotide level. Various conserved amino acid sequences belonging to ADP-ribosylating toxins (Chapter 8) were searched for and the sequences isolated were analysed for homology, but none could be found.

10.4 SEARCHING THE SHOTGUN AND ASSEMBLED DATABASES FOR THE 10/12 kbp *Sau*3A UNDIGESTIBLE FRAGMENTS.

The sequence information received for the recombinant clone, pSA2 (9.7.1) was used in a BLAST-search of the assembled *B.pseudomallei* genome sequencing project database for a homologous nucleotide sequence. An assembled fragment, Contig. 1748, contained a sequence with a high homology (80%) to that of pSA2. The assembled fragment, Contig. 1748 was 38 kbp (APPENDIX III) in length and the homologous sequence to pSAU2 was located from 34,835 bp to 35,143bp. The nucleotide sequence of Contig. 1748 was searched for any homology with the nucleotide sequence of the other recombinant clone, pSAU8, however no homology could be found. It is possible that the two clones pSAU2 and pSAU8 represent the two undigestible *Sau*3A fragments. Therefore a similar search was carried out with the two sequences from the recombinant clone pSAU8, which comprised a small amount of information at the beginning of the cloned sequence and at the end (9.7.2). Shotgun sequences revealed had a lower homology than the previous sequence for pSAU2, at around 60% and when analysed using the assembled database, the

two sequences were not found on a contiguous DNA fragment. For this reason, it was decided to only pursue the pSAU2 sequence. Web-based BLAST-X searches were then carried out on the 38 kbp sequence to reveal any areas with homology to protein sequences deposited in the Genbank database, the results of which can be seen in Fig.10.3. The sequence encoded 11 regions with homology to protein sequences deposited in the BLAST database (Table 10.5). However, the end 8 kbp of the fragment contained repeated DNA and the same proteins were identified in this region as those located at approximately 20-25 kbp on the fragment.

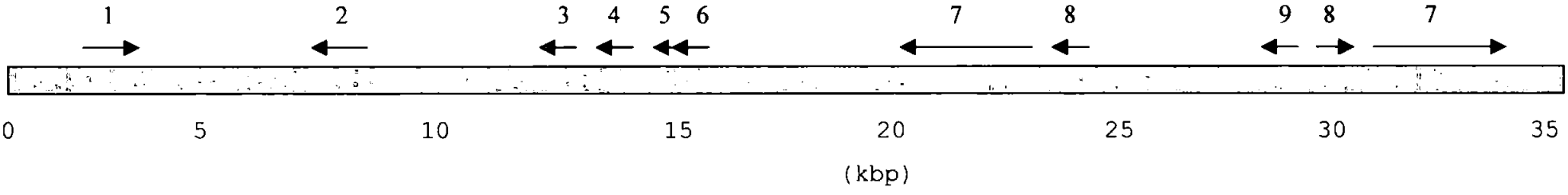


Fig. 10.3 Map of the 38 kbp Fragment Showing the Locations of Sequences with Homology to Other Protein Sequences Deposited in the BLAST Databases.

Table 10.5 Web-Based BLAST Results Showing Protein Sequences that have Homology to Regions on the 38 kbp Fragment.

Sequence No. On Fragment	Accession Number	Sequence Title	Similarity (%)
1	AP001512	Integrase/Recombinase *Bacillus halodurans*	39
1	D86934	Transposase A *Staphylococcus aureus*	45
1	X03216	Gene *tnp*A *Staphylococcus aureus*	45
1	AL445067	Site-specific integrase/recombinase *Thermoplasma acidophilum*	43
1	Z29084	Product similar to *tnp*A of *S.aureus Clostridium butyricum*	45
1	L29642	Methyltransferase *Pseudomonas fluorescens*	37
1	Z74024	Probable integrase/recombinase *Mycobacterium tuberculosis*	40
2	U31759	Thermostable alkaline protease *Thermoactinomyces* sp.	42
2	AP001512	Protease *Bacillus halodurans*	40
2	D64006	Serine protease *Synechocystis* sp.	37
2	AL031541	Putative secreted serine protease *Streptomyces coelicolor*	37
2	AL031231	Serine protease *Streptomyces coelicolor*	42
3	AF224059	Putative variable cytadhesin protein *Mycoplasma gallisepticum*	36
3	AB009049	Similarity to carbonic anhydrase gene *Arabidopsis thaliana*	38
3	L22858	Hypothetical 24.1 kDa protein Autographa californica polyhydrosis virus	52
3	M34651	Probable nuclear antigen *Pseudorabies virus*	34
4	AL445064	Hypothetical protein *Thermoplasma acidophilum*	49
4	AB035945	Serine/threonine protein kinase *Chlamydophila pneumoniae*	49
4	AL079038	Putative serine/threonine kinase *Streptomyces coelicolor*	42
4	Y10256	Serine/threonine protein kinase *Homo sapiens*	44
5	AL445064	Hypothetical protein *Thermoplasma acidophilum*	61
6	AL445064	Hypothetical protein *Thermoplasma acidophilum*	46
7	AE004565	Probable transcriptional regulator *Pseudomonas aeruginosa*	59
7	AB016282	Immunity repressor protein Bacteriophage phi-105	61
7	AF062070	Thermosensitive mutant immunity repressor Bacteriophage phi-105	59
8	AB029393	Streptococcal hemagglutinin *Streptococcus gondii*	31
8	U96166	*srp*A *Streptococcus cristatus*	33
8	U30626	Cell surface flocculin with structure similar to serine/threonine rich GPI-anchored cell wall proteins *Saccharomyces cerevisiae*	37
9	AF250878	Hypothetical protein *Salmonella typhi*	43

Perhaps the most interesting putative gene found on the 38 kbp fragment was an open reading frame with homology to a range of proteases from other bacterial species. Protease activity has been published for *B.pseudomallei* and indeed from the activity experiments carried out in Chapter 3, strong protease activity was observed for both *B.pseudomallei* and *B.thailandensis*. However the gene(s) responsible for the proteolytic activity observed have yet to be cloned and sequenced. The putative protease gene is located from its start codon at 8755 bp to the stop codon at 7019 bp in the 38 kbp sequence and is complementary to the sequence in Appendix III. The nucleotide sequence can be seen in Fig. 10.4. The nucleotide sequence encodes a gene product of 578 amino acids as follows and a theoretical molecular weight of 63 kDa:

MADLNEMKLLIRLPNQHVLNFRSPTRRSAADLIGITGGKVEALFP
DTSVQFVGGQNDWVVFQPDTVEDIHPWDRAHKQALSIVNSASGV
GDVYVEPNIIHKRTLDSSDAGVRSSITPLSAQEDVTLAPPYPPTESL
NPNYPPSEGASFSPAWHLQKAGFPRAWQTTKGEGIRIAHLDIGW
WPNHYSAPLKVRKDLGYNFVEGNSNTVDPGVGPNKGHGTATLAL
LAGNAVSLCGKAGQSEGDQVYRGFIGGAPSAEIVPVRIAGVDGSV
VYLYGETMARGLAYAINPGDGRRCDVVSLSHGGLPMKSWAHAV
NMLYDAGVVVVAAAGDSYWAVLTDIATHFTVYPSAFYRVVTSTG
VTFDDGPYKRDRLGVMQGCWGPDKVMKKAVGAYTPNVPWMC
YNTKYGWDMNGAGTSASTPQMAAACALWLAKYGSVFPNDWRR
VAACRAALGRSVADAEKDFSEIGLGRLDVSAMIEQTLADEVKHLC
DENRLQNIDPDDVSFPFLRLLFGLAPPGNGIEEMYEVEALQIFYQT
TNERLFRAVEDYENERLQSNEDLSELRAQFLQEKGMSDALRGYL
TSHA

```
1       CAGGCAATTT  GAGGTCGATT  GCATCTTCCG  ATATCTAGCA  TCTCGGAATT
51      AGATATATTG  TTGCGCTTGA  AGCTGTGACC  AACAGCCTAC  ATGATAGGCT
101     GAAGCTATAC  CTCGGGGAGT  AGAGCATGGC  TGACTTGAAT  GAGATGAAGT
151     TGTTAATCCG  TTTGCCTAAT  CAGCATGTTT  TGAATTTCAG  AAGCCCCACG
201     AGAAGAAGTG  CGGCGGACCT  TATTGGTATC  ACGGGCGGTA  AGGTTGAGGC
251     GCTTTTTCCT  GATACGAGCG  TTCAGTTTGT  TGGTGGCCAA  AATGATTGGG
301     TCGTATTTCA  ACCAGATACC  GTAGAGGATA  TTCATCCATG  GGACCGCGCA
351     CATAAGCAGG  CGTTGTCTAT  CGTAAACAGT  GCGAGCGGTG  TCGGTGATGT
401     GTATGTTGAG  CCCAATATTA  TTCACAAGCG  AACATTGGAC  AGCTCAGATG
451     CGGGAGTAAG  AAGCTCAATA  ACGCCTCTCA  GCGCTCAGGA  AGATGTCACA
501     TTAGCGCCCC  CCTATCCCCC  CACCGAATCG  CTCAACCCAA  ATTACCCTCC
551     GAGTGAAGGC  GCGTCGTTTT  CGCCTGCATG  GCACTTGCAG  AAGGCGGGTT
601     TTCCTCGAGC  ATGGCAGACT  ACAAAGGGAG  AAGGTATCCG  CATCGCACAT
651     CTGGATATCG  GATGGTGGCC  AAATCACTAT  TCTGCACCTC  TAAAAGTTAG
701     AAAAGACCTG  GGATACAACT  TCGTCGAAGG  AAATTCCAAT  ACCGTTGACC
751     CCGGAGTTGG  GCCGAATAAA  GGACATGGAA  CAGCAACTCT  CGCTCTGCTC
801     GCTGGTAATG  CCGTGAGCCT  ATGCGGAAAA  GCTGGCCAAA  GTGAGGGCGA
851     CCAGGTATAT  CGCGGATTCA  TTGGCGGCGC  GCCAAGCGCG  GAAATTGTCC
901     CTGTGAGAAT  CGCTGGTGTG  GACGGGTCCG  TTGTTTATCT  TTACGGTGAG
951     ACTATGGCTC  GTGGTCTTGC  TTATGCGATT  AATCCCGGCG  ATGGGCGACG
1001    TTGTGATGTG  GTGAGTCTGA  GTCACGGTGG  TTTGCCAATG  AAGTCATGGG
1051    CGCACGCTGT  GAACATGCTT  TATGACGCTG  GAGTAGTAGT  GGTGGCGGCT
1101    GCCGGCGATA  GCTACTGGGC  GGTGCTTACA  GACATCGCAA  CCCATTTCAC
1151    GGTTTATCCG  TCAGCGTTCT  ACCGCGTTGT  TACATCGACA  GGCGTAACAT
1201    TTGACGATGG  CCCTTACAAG  CGAGATAGAC  TCGGGGTAAT  GCAAGGCTGT
1251    TGGGGGCCCG  ATAAGGTGAT  GAAGAAAGCG  GTGGGAGCCT  ATACGCCGAA
1301    CGTTCCCTGG  ATGTGCTACA  ACACGAAGTA  CGGTTGGGAC  ATGAACGGTG
1351    CCGGAACCTC  CGCGAGTACA  CCGCAGATGG  CGGCTGCCTG  CGCTCTATGG
1401    CTGGCTAAGT  ACGGCAGCGT  ATTTCCAAAT  GACTGGCGCA  GAGTTGCGGC
1451    GTGTCGAGCT  GCGCTCGGTA  GGTCGGTTGC  GGACGCGGAG  AAGGATTTCA
1501    GCGAAATTGG  CCTGGGGCGG  CTAGATGTCA  GCGCGATGAT  AGAACAGACG
1551    CTTGCGGATG  AAGTTAAGCA  TCTTTGTGAC  GAAAATAGAC  TCCAAAACAT
1601    AGACCCAGAC  GACGTTTCAT  TTCCATTTCT  TCGACTTCTA  TTCGGGCTCG
1651    CCCCTCCGGG  AAATGGTATC  GAGGAAATGT  ACGAGGTAGA  GGCTTTGCAG
1701    ATTTTCTATC  AAACCACCAA  TGAACGCCTT  TTTCGCGCGG  TCGAAGATTA
1751    TGAAAACGAG  CGACTGCAAA  GCAATGAGGA  TTTGAGCGAG  CTGCGAGCAC
1801    AATTTCTACA  GGAAAAGGGT  ATGTCCGATG  CGTTGCGTGG  ATACCTGACC
1851    TCTCATGCCT  GA
```

Fig.10.4 The Complete Nucleotide Sequence of a Putative Protease Gene in *B.pseudomallei.* The sequence has been extended to include putative –35 and –10 promoter sequences and a putative ribosome binding site.

10.5 DISCUSSION.

The extracellular products from isolates of *B.pseudomallei* and *B.thailandensis* had been tested for hexosaminidase activity in Chapter 3, and it was revealed that all *B.pseudomallei* isolates had a positive activity with three different substrates, whereas the *B.thailandensis* isolates had no activity. However, construction and screening of genomic libraries in Chapter 4 yielded no clones with such activity. By searching the information available on the genome sequencing project with other bacterial hexosaminidase sequences available, two putative genes were identified. The first, isolated on a shotgun fragment was smaller, and yielded a 342 amino acid sequence with a predicted molecular weight of 37.5 kDa, with a high homology to beta-hexosaminidases from a range of other bacterial species. Beta-hexosaminidases catalyse the removal of beta-1,4 linked N-acetylhexosaminidase residues from oligosaccarides and their conjugates.

The second, isolated from an assembled fragment, Contig. 1720, yielded a putative 2,453 bp gene with high homology to chitobiases (N-acetyl-beta-glucosaminidases) from a range of other bacterial species. The nucleotide sequence predicts a gene product of 833 amino acids and a predicted molecular mass of 91 kDa. Chitobiases break down chitobiose, a disaccharide sugar unit found in chitin.

Both of these enzymes would possibly be extracellularly produced in *B.pseudomallei* and although the predicted sizes are given, they may well be processed to a smaller, active molecule. Both protein sequences were analysed by the Centre for Biological Sequence Analysis (CBS) on the internet using the Signal PV1.1 program (Neilsen *et al.*, 1997) to predict whether there were likely signal peptide cleavage sites in the amino acid sequences. A likely signal peptide cleavage site could not be found for the smaller putative hexosaminidase (37.5 kDa) but one was identified in the larger, putative chitobiase

between amino acids 25 and 26 (Fig. 10.2). Thus, the mature enzyme would be around 88.5 kDa.

In Chapter 3 extracellular products isolated from a representative group of *B.pseudomallei* and *B.thailandensis* isolates were subjected to SDS-PAGE and immunoblotting and the relative molecular weights of protein bands were calculated. These were analysed for proteins around 37.5 and 88.5 kDa. Results of immunoblotting yielded protein bands around 85.1 kDa but this was also found in one of the *B.thailandensis* isolates and hexosaminidase activity was not found in this species (Chapter 3). However, a similar sized band such as this may not actually represent the same protein. In addition to this a protein band of 36.3 kDa was only found in *B.pseudomallei* isolate 576 but it must be remembered that the protein bands were overshadowed by the ladder of reaction to LPS and so it was difficult to analyse protein bands. However, SDS-PAGE analysis revealed protein bands of 87.1 and 38.9 kDa found only in *B.pseudomallei* isolates, which possibly represent the two putative proteins identified, but cloning and characterisation of the putative genes and their respective products would be necessary to prove this.

The *B.pseudomallei* genome project was searched for homologies to ETA of *P.aeruginosa* both at the nucleotide and amino acid level but no obvious similarity was found. Sequences were also analysed for amino acid patterns found in the active sites of a number of ADP-ribosylating toxins but again no obvious similarities were identified. This does not discount the presence of an ADP-ribosylating toxin in *B.pseudomallei* but perhaps implies that it is different from ETA of *P.aeruginosa* and perhaps those from other bacterial species.

When the *B.pseudomallei* genome sequencing database was searched for the sequence information possessed on one of the *Sau*3A undigestible clones (pSAU2) (Chapter 9) the

sequence was present on an assembled fragment totalling 38 kbp (Contig. 1748). Previous attempts in Chapter 9 to clone the full-length 10/12 kbp sequences had failed but clearly the process of cloning fragments in the genome sequencing project was successful, presumably because different vector systems and enzymes were used.

Interestingly, the 38 kbp region of this fragment was found to contain no *Sau*3A cleavage sites (GATC). Additionally, the G+C content of this fragment was 56%, which is different from the rest of the genome, which is 69% (Redfearn *et al.*, 1966). Such findings are consistent with Pathogenicity Islands (Pais), which have been found in a wide variety of bacterial species including *E.coli*, *Salmonella*, *Helicobacter pylori*, *Dichelobacter nodosus* and *Vibrio cholera*. Pais are large pieces of chromosomal DNA that carry virulence genes, which were most probably acquired by phage-mediated horizontal transfer (Buchrieser *et al.*, 1998). A Pai can be defined according to the following criteria (Hacker *et al.*, 1997):

(i) Carriage of (often many) virulence genes.
(ii) Presence in pathogenic strains, and absence or sporadic distribution in less-pathogenic strains of one species or related species.
(iii) Different G+C content in comparison to DNA of host bacteria.
(iv) Occupation of large chromosomal regions (often more than 30 kbp).
(v) Represent compact, distinct genetic units, often flanked by direct repeats.
(vi) Association with tRNA genes and/or insertion sequence (IS) elements at their boundries.
(vii) Presence of (often cryptic) "mobility" genes (IS elements, integrases, transposases, origins of plasmid replication).
(viii) Instability.

One of the main features of Pais is the carriage of virulence genes and a putative gene encoding an alkaline protease was found within the sequence. Sexton and Jones (1994) believed that the extracellular protease they isolated belonged to the family of alkaline

proteases sensitive to metal chelators. The authors also found the protease capable of digesting transferrin and haemoglobin and so postulated its role in iron acquisition in the host. The putative protease gene product was calculated as having a theoretical molecular weight of 63 kDa. However, as with the putative hexosaminidases, it is likely that the protease contains a signal peptide cleavage site and so the sequence was subjected to analysis by the same program as the hexosaminidase sequences. A possible peptide cleavage site was recognised between amino acid 29 and 30 (RSA-AD). Thus giving the mature protease a size of 59.5 kDa. Conflicting reports exist in literature regarding the size of the protease whereby Sexton and Jones (1994) confirmed the presence of a 36 kDa protease and Percheron *et al.* (1995) a 42 kDa protease. Despite this, a 59.5 kDa protein band was searched for in the SDS-PAGE gels and immunoblotting in Chapter 3. Microtitre plate tests carried out on *B.pseudomallei* and *B.thailandensis* ECPs revealed that proteolytic activity was common to both species. A protein band of 60.3 kDa was found in *B.pseudomallei* isolates and a *B.thailandensis* isolate on SDS-PAGE gels. On immunoblotting a protein band of this size was not found, only ones at 57.5 and 63.1 kDa, which were, incidentally, found in both species.

As the protease may be produced in both species its role in virulence is questionable and recent published literature suggests that the expression of a high proteolytic activity is neither sufficient nor necessary for virulence when bacteria are injected via the intraperitoneal route (Gauthier *et al.*, 2000). However, a pathogenic role for the protease cannot be excluded, and it has been suggested that its role in the physiopathology of melioidosis is probably dependent on the route of entry of the bacteria into the organism (Gauthier *et al.*, 2000).

In addition to the putative protease another region encoded homology to a *Streptococcal*

haemagglutinin from *Streptococcus gondii*, which was present in two copies along with a putative transcriptional regulator. It has been postulated that the haemagglutinating adhesins of bovine group B *streptococcal* isolates are directly involved in the adherence mechanisms of these organisms (Wibawan *et al.*, 1993) and so if it has a similar function in *B.pseudomallei* is another extracellularly produced protein with a putative role in virulence.

The 38 kbp fragment also contained a putative integrase/recombinase gene at the beginning of the sequence, which are often associated with Pais at their boundaries. In addition to this the 10/12 kbp fragment isolated from *B.pseudomallei* 4845 was unstable and only deletions of the fragment could be cloned. All of the discussed factors fulfil much of the criteria for the fragment being a Pais, although its presence in non-pathogenic strains has yet to be tested. In addition to this, it is also possible that the fragment is bacteriophage in origin due to lack of *Sau*3A sites as was found in the genome of the lysogenic phage, MAV1 of *Mycoplasma arthritidis*, discussed in Chapter 9. Not only this, but a number of the regions identified on the fragment had homology to various phage genes (Table 10.5). It is possible therefore that the 38 kbp fragment may be a Pais, bacteriophage in origin and encodes a number of excreted virulence genes in *B.pseudomallei*.

CHAPTER 11

GENERAL DISCUSSION

Following each experimental chapter within this thesis there are individual discussions, which focus on the findings of that particular practical investigation. The purpose of this final chapter is to tie together all the information in a general discussion.

This research project was concerned with the molecular characterisation of the pathogen, *B.pseudomallei*, of which there is a paucity of knowledge in published literature. As stated in the aims of this thesis the main aim was to apply the techniques of molecular biology, in particular those of gene cloning, to isolate and characterise extracellular virulence factors from the pathogen. With respect to this initial aim, although no firm information regarding the nature of extracellular virulence factors nor the pathogenic processes of the bacterium has been gained through this research, previously uncharacterised genes and biochemical activities of *B.pseudomallei*, which may have a role in the pathogenic processes of the bacterium have been discovered and are available for future study. In addition to this, fundamental differences between *B.pseudomallei* and the avirulent species, *B.thailandensis* have been elucidated, which was another aim of the research project.

As previously stated in the aims, the order in which experimental work appears within the thesis is not representative of the chronological order in which the investigations were carried out. Thus some of the investigative work may appear illogical to the reader. However, the initial part of this thesis deals with the initial characterisation of extracellular products from *B.pseudomallei* and *B.thailandensis* followed by the construction and screening of *B.pseudomallei* genomic libraries.

When culturing facilities became available, at the end stages of the research project, a direct comparison of the ECPs isolated from *B.pseudomallei* and *B.thailandensis* isolates was carried out through SDS-PAGE profiles and immunoblotting with convalescent human antisera. As there was little opportunity to experiment with culture media, coupled with a lack of published literature regarding the evaluation of different media with respect to the isolation of extracellular products, it may have been the case that the full range of ECPs were not produced, or produced at a very low level. Indeed, the concentration of ECPs derived from the cultures was very low, which prohibited an extensive comparison between all of the isolates used. Future studies may benefit from a better-defined media for use in isolating ECPs, as the isolates in this study were cultured simply in LB media. This may be achieved by simply assaying isolates on a range of media to test for enhanced expression of certain activities.

However, a representative group consisting of both species, and including *B.pseudomallei* isolates from clinical and environmental sources could be compared and it became evident that the two species did exhibit different extracellular protein profiles on SDS-PAGE and immunoblotting. Isolates of *B.thailandesis* invariably contained fewer ECPs than those of *B.pseudomallei*. However, the two species did share a number of common bands, although obviously it cannot be concluded that similar sized bands represent the same protein.

ECP protein profiles gained by SDS-PAGE and immunoblotting also revealed differences between isolates of *B.pseudomallei*. Not all strains of a virulent bacterial species are equally pathogenic and considering the wide range of clinical features and outcomes associated with infection with *B.pseudomallei* it is likely that the isolates of *B.pseudomallei* differ with respect to the virulence products they produce, which includes ECPs. Certain protein bands present in clinical isolates were not found in those isolated from the environment and may be indicative of virulence products elicited on infection. It

was not known what were the functions of the proteins identified, although suggested functions based on the relative sizes compared to those in published literature was discussed in Chapter 3.

The results of the immunoblotting using convalescent human antisera were overshadowed in most isolates by the reaction to what was considered to be LPS, producing a ladder of bands. The exceptions to this were *B.pseudomallei* isolates 576 and 217. It has been reported that 5% of *B.pseudomallei* isolates possess an atypical LPS pattern (Dr. T. Pitt, personal communication), which includes *B.pseudomallei* 576, and on immunoblotting with the human sera there was no evidence of the ladder of LPS reaction. It was not known if *B.pseudomallei* 217 was included in the 5% of isolates with atypical LPS patterns but as there was no reaction of LPS with the sera it is possible that it is. The convalescent human antisera probably contained antibodies raised to the typical LPS pattern, found in 95% of *B.pseudomallei* isolates and so did not react with LPS from those isolates with an atypical pattern. If this was indeed the case, in order to analyse the other protein profiles, which were overshadowed by the LPS reaction, the study may have benefited from using human sera from an infection with an isolate of *B.pseudomallei* that has an atypical LPS pattern, such as *B.pseudomallei* 576. In this case antibodies would be raised to the atypical LPS pattern, which would possibly not react with LPS from isolates with the typical pattern therefore making the bands visible. Experiments were carried out testing isolates of both species for a number of activities with the aim of identifying novel targets for future gene cloning studies. Some of the activities identified had been previously reported for the bacterium, but not confirmed in the newly identified species, *B.thailandensis*. Protease and lipase activities, already identified in *B.pseudomallei* (Ashdown and Koehler, 1990), were shown to be similarly possessed by *B.thailandensis* isolates. This was also the case with phospholipase C activity, previously identified by Korbsrisate *et al.* (1999).

Of those tested, only a hexosaminidase activity, previously unreported, was found in isolates of *B.pseudomallei* and not in *B.thailandensis*, which represented therefore an ideal target for future cloning work as it might have a role in the pathogenicity of *B.pseudomallei*. Furthermore, the ECPs from the *B.pseudomallei* isolates tested yielded a positive reaction with three different substrates all testing for hexosaminidase activity: MU-diacetyl-chitobioside, MU-N-acetyl-galactosaminidine and MU-N-acetyl-glucosaminidine, and it was postulated that *B.pseudomallei* may possess three separate enzymes, or a single hexosaminidase enzyme responsible for the activities observed. However, as this was carried out at the end of the research project there was little opportunity to investigate this. Although the *B.pseudomallei* genomic library, constructed in λGT11 was tested with one of the substrates, MU-N-acetyl-glucosaminidine, no recombinants were identified exhibiting this activity. However, it must be remembered that genomic library construction and screening was carried out well in advance of the experiments testing the ECPs for various activities, due to the lack of culturing facilities. The logical pattern would have been to test the ECPs first and from the results of this to identify activities such as the hexosaminidase activity to screen the genomic library with. In this case all three hexosaminidase substrates with which the positive reaction was observed would have been used. The logical next step if a clone exhibiting the desired activity was not isolated on screening the genomic library would have been to develop PCR primers based on other hexosaminidase gene sequences deposited in the Genbank database in the hope that these would amplify a similar sequence in *B.pseudomallei* that could then be cloned and sequenced.

Towards the end stages of the research project, the *B.pseudomallei* genome sequencing project became available on the internet to search sequence information that had been determined in the form of shotgun sequences and assembly of these sequences into larger, contiguous DNA sequences. It was a relatively quick and straightforward process to use

hexosaminidase sequences already deposited in the database to search the genome sequencing project for homologous sequences. This process yielded the isolation of two putative hexosaminidase genes, which, if their gene products are produced in *B.pseudomallei* may have been responsible for the activities previously observed in the ECPs isolated. Of the two putative hexosaminidase genes, one had homology to beta-hexosaminidases from a range of bacterial species and the other had homology to chitobiases (N-acetyl-beta-glucosaminidases) of which a putative signal peptide cleavage site was recognised in the latter. The next step would possibly be to develop PCR primers to both sequences in order to amplify the sequences from *B.pseudomallei* isolates for cloning and characterisation of the putative genes and their respective gene products with the ultimate aim of determining a role of the enzymes in the virulence of *B.pseudomallei* through mutational studies.

An anomalous result in testing the ECPs was the positive reaction of only *B.pseudomallei* isolate 576 in the panel of isolates used, to the chromogenic substrate for phosphatase activity, 5-bromo-4-chloro-3-indolyl phosphate. It was postulated that as this isolate has a different LPS pattern than the others in the panel of isolates tested, that this in some way enabled other components, in addition to ECPs, to be isolated during the ECP isolation and concentration procedure, such as the phosphatase not normally extracellularly produced by *B.pseudomallei*. Testing of other isolates with atypical LPS patterns may reveal the same result.

One of the major virulence factors identified in *B.pseudomallei* is the presence of a thermolabile toxin (Nigg *et al.*, 1955), which has been postulated to possess ADP-ribosylating activity (Ismail *et al.*, 1991). However, no simple, precise and highly sensitive test exists for the exotoxin (Ismail *et al.*, 1987). Again, as culturing facilities only became available at the end of the research project there was little opportunity to investigate or

design an assay for use to test the ECPs of *B.pseudomallei* and *B.thailandensis* for the presence of the toxin. Attempts were made throughout the research project to acquire antisera raised to the exotoxin, as well as other specific extracellular products of *B.pseudomallei* but these were unsuccessful.

The application of recombinant DNA techniques to *B.pseudomallei*, through cloning fragments of the organism's genome in *E.coli* host/vector systems, to facilitate the construction of representative genomic libraries was the very first experimental procedure carried out. It was hoped that by subsequent screening of the genomic libraries for antigens or specific activities of *B.pseudomallei*, that the corresponding gene sequences which may be of importance in the pathogenicity of the bacterium would be isolated. However, at this time there existed no Category 3 culturing facilities at the University of Plymouth and the research project was dependent upon donated sources of genomic DNA. As the genomic DNA used to construct genomic libraries was donated there was no control over the genomic DNA isolation procedures, or any subsequent purification steps that were carried out prior to receipt of the DNA. This indeed proved to hamper the construction of various genomic libraries. The genomic DNA received often appeared to be contaminated with what was postulated to be complex carbohydrate material. When the Containment Level 3 unit became available at the University of Plymouth and it was possible to culture the organism, it was noticeable that growth of both *B.pseudomallei* and *B.thailandensis* was accompanied with the mass production of such material and complete separation of bacterial cells with this material proved difficult. It was therefore likely that a cell lysis procedure, followed by repeated phenol-chloroforms, the process used by the donators of the genomic DNA, would not have eliminated this carbohydrate material and so restriction enzyme digestion and ligation procedures would be seriously hampered as a result. This was observed in the numerous attempts to construct both plasmid and cosmid genomic

libraries with the eventual outcome of just one successful genomic library in the bacteriophage vector, λGT11.

Screening of the *B.pseudomallei* genomic library in λGT11 with various substrates yielded very little in the way of results. There was little published work on the application of gene cloning to *B.pseudomallei* and so it was assumed that considering such a large number of genes from a wide range of bacterial taxa have been successfully cloned and expressed in *E.coli* systems that it would be possible to do the same with *B.pseudomallei*. As screening procedures yielded only two immunopositive clones with antiserum raised to *B.pseudomallei* 576 formalised cells, it could be the case that the *E.coli* expression system differs somewhat to that of *B.pseudomallei*, hence the relative success of expressing genes is limited due to such incompatibilities. Additionally, incompatibilities may also exist between a product encoded by *B.pseudomallei* and the metabolic activities of the *E.coli* cell or indeed deficiencies in the gene library screening procedure may exist. The genomic library in λGT11 was shown to be composed of 99% recombinants and so it was unlikely that problems existed from the library being composed of a high proportion of vector-only ligations.

Incompatibilities at the level of gene expression can be the result of a number of attributes of the cloned gene, including promoter strength (Deuschle *et al.*, 1986), the structure of the ribosome-binding site (De Boer *et al.*, 1983) and the choice of codons for amino acid incorporation (Kurland, 1987). As very few genes, particularly those involved in virulence of *B.pseudomallei* have been cloned and sequenced there was little information to compare. Although the phosphatidylcholine-hydrolysing phospholipase C (PC-PLC) gene was successfully cloned in an *E.coli* host/vector system and expressed even without IPTG induction, indicating that the promoter was recognised by the *E.coli* RNA polymerase (Korbsrisate *et al.*, 1999).

Incompatibilities between the *B.pseudomallei* product and the metabolic activities of the cell may be due to the fact that the *E.coli* host may not possess certain factors for any post-translational modification of *B.pseudomallei* proteins. Or cloned products may be inactivated or degraded by enzymes in *E.coli* or conversely, it could be that the *B.pseudomallei* components may be lethal to *E.coli* host. This is especially relevant because the genomic library was constructed in λGT11, which is an expression vector and so the expression of any clones that are encoding toxic products may have been lethal to *E.coli.*

Any of the reasons discussed may have been why only two clones were identified from all of the screening procedures carried out. However, problems may also have arisen from inadequacies in the screening procedures themselves. The antiserum used to screen the bacteriophage genomic libraries was donated from an outside source and so again there was no control over its preparation or purity. It was already evident that extracellular carbohydrate material was a possible source of contamination in the genomic DNA preparations that had been received and considering that the antisera was raised to whole and formalised cells, it was likely that most of the immune response would be to this material. This became even more likely when immunoblots were carried out on ECPs in Chapter 3, as a ladder of what was most likely LPS material was the overwhelming response. Thus, screening the *B.pseudomallei* genomic library with the donated antisera would have not have yielded many positive results.

Alternative cloning systems or techniques perhaps need to be considered, for example utilising a more closely related species, such as *P.aeruginosa* may aid expression of *B,pseudomallei* proteins, although care would have to be taken considering that *B.pseudomallei* is a Category 3 pathogen, and altering the host from *E.coli* may increase the containment level at which cloning experiments must be carried out. Transposon

mutagenesis is also a powerful tool to study bacterial pathogens and has been used successfully by DeShazer and Woods (1999) to identify a 12 kbp locus with a high degree of homology to the general secretory, or type II secretion pathway proteins and the authors have claimed to have cloned a number of putative virulence genes using this system. Future studies may also be improved by more selective methods of screening perhaps using monoclonal antibodies raised to certain molecules of interest produced by *B.pseudomallei*.

Two immunopositive plaques were identified, which developed 'halos' of positive reaction encircling the plaques. It was not known what this unusual reaction indicated, but it was postulated that it could be indicative of an insoluble product, or one that was being degraded or the very low level of production of the immunopositive product. However, only one of the clones retained its immunopositivity after isolation and re-screening, and so the stable immunopositive λGT11 clone (λBPGT1) was used for further analysis.

Numerous attempts were made to detect a protein product being expressed by the immunopositive clone (λBPGT1), responsible for the activity observed, through SDS-PAGE and Western blotting, none of which were successful. Analysis continued by attempts to subclone the immunopositive-conferring sequence of DNA contained in the recombinant clone (λBPGT1). The size of the cloned insert was determined and the 5 kbp *Eco*RI fragment was successfully subcloned into the plasmid vector pUC18. However, attempts to determine whether the bacterial subclones were producing an immunopositive antigen, by carrying out a colony blot, was inconclusive. The recombinant colonies did develop quicker then the negative control, but not as a 'halo' of positive reaction and after a few minutes no difference could be observed between the negative control and the recombinants. Acetone powders were used to pre-adsorb any anti-*E.coli* antibodies that may have been present in the antiserum, used as the primary antibody. However, it may

have been the case that this was not sufficient in the colony blots to prevent the high *E.coli* background reaction.

A restriction endonuclease cleavage map was constructed using the recombinant plasmid pBPGT2 in order to further analyse the cloned DNA. Further subclones of the original 5 kbp fragment were created (pCD1-5) in an attempt to isolate the immunopositive-encoding sequence away from other potentially interfering sequences encoded on the 5 kbp fragment. However, attempts to express the immunoreactive product using a commercial *in vitro* transcription and translation kit on the subclones pCD1-5 and pBPGT2 were unsuccessful.

Unfortunately, a protein product responsible for the immunopositivity of the original clone could not be determined. It is not clear exactly what the 'halo' of immunopositive product indicated or indeed what exactly was responsible for it's production and so explanations can only be postulated. The immunopositive 'halo' of reaction, as previously discussed, may have been indicative of degradation or insolubility of the product, and if any of these were the case then a product may not have been visualised by SDS-PAGE or Western blotting of the original phage lysate. Moreover, a second immunopositive clone was identified, which possessed the same 'halo' of reaction but failed to retain its immunopositivity, which may indicate degradation of the product, or perhaps an instability of the cloned fragment encoding the immunopositive antigen. If the immunopositive antigen was initially being produced at a very low level, it too may not have been visualised on an SDS-PAGE gel or perhaps on Western blotting. It was hoped that the *in vitro* transcription and translation kit would overcome problems such as a low level of expression. However, it is possible that subcloning the fragment into a plasmid vector resulted in a deletion of nucleotide(s) and thus a loss of the correct reading frame for the immunopositive-encoding sequence, and therefore a product was not detectable with the

use of the *in vitro* transcription and translation kit. Alternatively, when the kit was used on pBPGT2 there may have been other, interfering sequences contained on the 5 kbp fragment that prevented the successful transcription and translation of the immunopositive product, and in further subcloning the fragment in order to alleviate such a problem, the sequence was unwittingly disrupted.

Sequence analysis of the 5 kbp *Eco*RI fragment revealed four ORFs with homology to protein sequences deposited in the BLAST database. One of these was of great interest as it had homology to type IV pilin genes from a range of bacterial pathogens. Multiple sequence alignments revealed that the N-terminus of the putative pilin protein contained significant homology to type IV pilin genes over the first 40 amino acids. This included, most notably, *pil*E of the closely related species *P.aeruginosa*, which is postulated to have a role in pilus biogenesis (Russell and Darzins, 1994). Pilus biogenesis includes the concerted synthesis of major and minor subunits, various post-translational modifications and assembly into a functional organelle (Strom and Lory, 1993). Other type IV pilin genes have allegedly been cloned by Woods *et al.* (1999) in *B.pseudomallei* that have homology to *pil*B, *pil*C and *pil*D of *P.aeruginosa*, but none with homology to *pil*E.

There are three major conserved features of type IV pilin genes all of which were identified in the putative pilin gene sequence of *B.pseudomallei*. Firstly, they are synthesised as precursors with unique, short, basic amino-terminal leader peptides, which are removed by endoproteolytic cleavage between an invariant glycine residue and a phenylalanine (Strom and Lory, 1993). The prepilin leader peptidase cleavage site is conserved in all members of the type IV pilin family, (Gly-Phe(Met)-Thr-Leu-Ile(Leu)-Glu-) (Strom and Lory, 1993). However, many similarities exist between the biogenesis of type IV pili and the assembly systems required for the export of extracellular proteins or the import of DNA in a large range of gram-negative bacteria of which one of the

characteristics is the possession of multiple genes with prepilin-like signal sequences (Bally *et al.*, 1992). However the putative pilin gene sequence in *B.pseudomallei* contained tyrosine residues (Try-24 and Tyr-27), which are thought to be involved in pilin subunit-subunit interactions (Watts *et al.*, 1983) not found in genes required for the export of extracellular proteins. The final conserved feature is that the carboxy terminus contains a characteristic pair of cysteines, forming a disulphide loop.

Reports have stated that genes required for pilus biogenesis are not always tightly clustered (Russell and Darzins, 1994), which may be why no other pilin genes were identified on the 5 kbp fragment. In addition to this, *pil*E of *P.aeruginosa* was mapped away from any other pilin genes 5 kbp downstream from *pil*V another gene involved in the biogenesis of pili (Alm and Mattick, 1996). Perhaps sequencing up and down-stream of the 5 kbp *Eco*RI fragment may facilitate the isolation of other genes involved with the biogenesis of pili in *B.pseudomallei*.

Numerous attempts at expressing the protein product of the putative pilin gene were unsuccessful. Therefore it could not be determined whether the immunopositive reaction was due to the pilin gene product, or whether this interesting gene has fortuitously been stumbled upon. As previously discussed, very little exists in published literature concerning the cloning and expression of *B.pseudomallei* genes in *E.coli* host/vector systems and based on the relative difficulty encountered during the research project on the expression of any *B.pseudomallei* products it may be that incompatibilities do exist. Putative –10 and –35 promoter sequences were identified for the pilin gene sequence, but these did differ from the consensus *E.coli* promoter sequences and so perhaps expression failed as the promoter sequences were not recognised by the *E.coli* system. However, attempts to place the gene under the control of an *E.coli* promoter to produce a fusion protein with the β-galactosidase system in *E.coli* were also unsuccessful.

Again it is not known why expression of the putative gene product failed but the sequence does require post-translational modifications, which are probably not present in the *E.coli* system and this may therefore have prohibited the expression of the protein. Additionally, the product may have been unstable or deleterious to the *E.coli* host. Similar problems were also encountered when trying to express the *pil*E gene product of *P.aeruginosa* in an *E.coli* system (Russell and Darzins, 1994). In this case the authors resorted to expressing *pil*E in a *P.aeruginosa* T7 expression system. This may be the step required to express the putative pilin gene product from *B.pseudomallei* in any future work carried out, as expression of the product is invaluable for its further characterisation.

Attempts were made to further characterise putative pilin gene product through Web-based protein analysis programs to analyse the biochemical nature of the sequence. The information generated on the prediction of secondary structure, different physio-chemical profiles, a prediction on transmembrane helices and a table of codon usage of this putative gene and the others identified in the sequence are displayed in the thesis, but are included mainly for reference purposes. Although the hydropathy prediction proved that the N-terminal domain of the *B.pseudomallei* putative pilin gene contains a relatively hydrophobic region similar to other pilin genes, which has been shown in other studies to be important for membrane translocation. However, in future studies of the putative pilin gene when hopefully it has been characterised further and a mode of action has been elucidated for the gene product, such secondary information may prove be valuable.

Southern blot hybridisations using the putative pilin gene sequence as a DNA probe revealed the cloned 5 kbp *Eco*RI-generated fragment of DNA was homologous to a sequence found in the genome of *B.pseudomallei* 204, therefore establishing the fidelity of the cloning procedure. The presence of the gene sequence in other *B.pseudomallei* isolates and its absence in the *B.thailandensis* isolates tested was confirmed through Southern

blotting analysis. This indicates a possible role of the gene in virulence and the possibility of the absence of pili in the *B.thailandensis* species, which has yet to be established. Other species of the genus *Burkholderia* were also tested for the presence of the sequence, but a homologous fragment was not hybridised in any of the other species tested.

Mutational analysis of the putative pilin gene could not be carried out due to the instability of the donated suicide vector, pCVD442. Mutational analysis may have alluded to the role of the gene product, if indeed it possesses one, in the biogenesis of pili in *B.pseudomallei*. Russell and Darzins (1994) constructed a gene-replacement mutant of *pil*E in *P.aeruginosa* and carried out expression studies with the T7 RNA polymerase promoter system to show that the gene was required in pilus biogenesis because the mutants were pilus- and twitching-motility deficient. A similar study of the putative pilin gene of *B.pseudomallei* could possibly establish a role in pilus biogenesis and virulence.

One of the aims of the research project was to use sequences encoding extracellular virulence factors available from other bacterial pathogens to construct PCR primers to test for the presence of homologous sequences in *B.pseudomallei* and *B.thialandensis*. It was decided that the target of such a procedure was the toxin of *B.pseudomallei*, which has yet to be cloned and sequenced. As previously discussed, the toxin is postulated to be of the ADP-ribosylating class of exotoxin with the same mode of action as ETA of the closely related species, *P.aeruginosa* (Ismail *et al.*, 1991). This made the proposed ADP-ribosylating toxin of *B.pseudomallei* an ideal target because the ETA gene of *P.aeruginosa* had been cloned and sequenced (Gray *et al.*, 1984) and so PCR primers could be designed to test for the presence of a similar toxin in *B.pseudomallei*.

PCR primers were designed to the most conserved region of ADP-ribosylating toxins, the active site of ETA. The primers amplified a 453 bp product in the *B.pseudomallei* isolates

tested but failed to amplify a product in the *B.thailandensis* isolates. The PCR products from *B.pseudomallei* isolates 204 and 576 were cloned into pUC18 and DNA sequenced. These two isolates were chosen for comparative purposes, both are clinical isolates but have different LPS patterns. Both sequences were identical but no obvious homology was observed with the ETA active site either at the nucleotide or amino acid level nor with published consensus sequences for active sites of ADP-ribosylating toxins (Rappuoli and Pizza, 1991). Despite this finding, the sequence was used as a DNA probe to isolate a larger fragment to analyse sequence information up and down-stream. It was hoped that as the PCR product was only amplified in *B.pseudomallei* isolates it might encode other putative genes with a role in virulence or a toxin that did not possess obvious sequence similarity over the 453 bp sequence isolated.

The 453 bp PCR product was DIG-labelled and used to screen a *B.pseudomallei* genomic library constructed in the bacteriophage, λGEM11, chosen in preference to the λGT11 genomic library as λGEM11 is not an expression vector and so may facilitate the isolation of clones encoding toxic products to the *E.coli* host. Screening yielded six stable positive clones (λBP1-6). Restriction endonuclease analysis was carried out on the clones and a 7.7 kbp *Eco*RI fragment was isolated from one of the clones for further analysis. The first stage of analysis was to subclone the fragment into a plasmid vector. Difficulties were experienced with this procedure and attempts to subclone the fragment into the high-copy number vector, pUC18 were unsuccessful and so a lower copy number vector, pGD103 was used. However, subcloning of the 7.7 kbp fragment could only be achieved by keeping all incubation temperatures at room temperature. It was then proven by Southern blotting analysis that the cloned 7.7 kbp sequence contained the sequence of the PCR product before DNA sequencing.

Subsequent DNA sequencing of the 7.7 kbp *Eco*RI fragment revealed again that the sequence had no obvious homology either at the nucleotide or amino acid level with ETA, or indeed the consensus amino acid sequences for active sites of ADP-ribosylating toxins. This finding in itself did not discount the presence of an ADP-ribosylating toxin in *B.pseudomallei* for the active site may not possess the sequence encoded by the primers used, or the primers may have failed to amplify the active site. The *B.pseudomallei* genome sequencing project facilitated the opportunity to search the information available for homologies with ETA at the nucleotide and amino acid level as well as DT toxin, which has the same mode of action. Both BLAST-N and BLAST-X searches of the database yielded again no obvious similarity.

However, the presence of a toxin with ADP-ribosylating activity cannot be discounted in *B.pseudomallei*. The ADP-ribosylating toxin in *B.pseudomallei* may have a completely different active site and possess unique biochemical properties to such toxins. Indeed ExoS of *P.aeruginosa*, in relation to members of the family of ADP-ribosylating toxins possesses several unique biochemical properties. Firstly, while most bacterial ADP-ribosylating toxins target specific host proteins, ExoS has been shown to ADP-ribosylate several host proteins *in vitro*, including vimentin (Coburn *et al.*, 1989) although the *in vivo* target protein(s) of ExoS is undefined. Secondly, ExoS possesses an absolute requirement for a eukaryotic protein termed FAS (for factor activating exoenzyme S) to catalyse the ADP-ribosylation reaction (Coburn *et al.*, 1991). Thirdly, while most bacterial ADP-ribosyltransferases possess a recognised A:B structure-function organisation (Kruger and Barbieri, 1995) the A:B organisation of ExoS is not apparent. Thus, the toxin of *B.pseudomallei* may too differ from other members of the family of ADP-ribosylating toxins.

The identification of the toxin in culture supernatants of *B.pseudomallei* would be of huge value in future studies as the N-terminal sequence of a purified protein can be determined, and so by using the amino acid sequence derived, a specific oligonucleotide may be synthesised for gene library screening by DNA hybridisation. The ETA gene of *P.aeruginosa* was cloned in this manner (Gray *et al.*, 1984), by purifying ETA and subjecting it to N-terminal amino acid sequencing. As previously discussed attempts were made to obtain a monoclonal antibody to the exotoxin but these were unsuccessful. This may have been of value as immunoblotting using the ECPs isolated from culture supernatants of *B.pseudomallei* with the monoclonal antibody to the exotoxin may have facilitated its isolation.

However, the PCR product was only amplified in *B.pseudomallei* isolates and difficulties were encountered with subcloning it into the high copy number vector, pUC18, which may be indicative of the fragment encoding products toxic to the *E.coli* host. Sequence analysis carried out on the 7.7 kbp fragment, despite it not having homology to known ADP-ribosylating toxins, showed the 7.7 kbp sequence to encode other putative genes that could possibly have a role in virulence. Although it must be remembered that although such BLAST searches may give clues to the characteristics of a protein, there is no substitution for actual biochemical and molecular biological experimentation. Indeed, information derived from BLAST searches in this case functions as a useful starting point for further characterisation of putative genes and their gene products.

The sequence was shown to encoded an ORF with homology to porins from different bacterial species, most notably porins from *N.meningitidis* and *N.gonorrhoea*, which function to aid intracellular survival in these bacteria. Other interesting homologies included an ORF with homology to peptide synthetase genes involved in the biosynthesis of a number of important genes such as those involved in iron acquisition and antibiotic

resistance. Additionally, another ORF was shown to have homology to genes from the AraC/Xyls family of transcriptional regulators, some of which regulated genes with a putative role in virulence. In addition to these there was an ORF of over 2 kbp which had no homology to any proteins deposited in the Genbank database and so may encode a novel protein, perhaps with an important role in *B.pseudomallei* pathogenicity. This fragment may prove to be an important operon containing genes involved with iron acquisition, antibiotic resistance or intracellular survival but much more work is required to characterise the putative genes identified and their respective gene products.

Whilst carrying out a *Sau*3A restriction digest of *B.pseudomallei* 4845 for the purposes of creating a genomic library, it was noted that there existed two bands of around 10/12 kbp in size that were resistant to the endonuclease action. *Sau*3A digests were then carried out on a number of other *B.pseudomallei* and *B.thailandensis* isolates, with similar results. Although the *Sau*3A resistant DNA fragments varied in size, and number with isolates there existed such bands in *B.thailandensis*.

This finding occurred at the beginning of the research project when there was no facility to isolate genomic DNA and work was dependent on the donation of genomic DNA from outside sources. Moreover at the time there was only sufficient genomic DNA of *B.pseudomallei* isolate 4845 to continue work with the *Sau*3A undigestible bands. It was not known at the time exactly what this finding indicated although there was published information on MAV1 a lysogenic bacteriophage from *Mycoplasma arthritidis*, where it was found that the MAV1 genome contained no sites for the restriction enzymes, *Sau*3A and *Mbo*I which both recognise the sequence GATC (Voelker and Dybvig, 1998). It was postulated that these fragments were therefore bacteriophage in origin and as virulence determinants are often encoded on a phage in bacterial pathogens the fragments may encode such genes in *B.pseudomallei*. This theory was somewhat strengthened when in a

transformation procedure the bacterial cells appeared to lyse after incubation at 37°C, which indicated that the DNA might be phage-encoded, although a transfection procedure using the DNA fragments yielded nothing. Additionally, despite numerous attempts, the full-length fragments could not be cloned and only clones containing deletions of the original fragments were isolated, pSAU2 and pSAU8. DNA sequencing proved difficult and only a small amount of sequence information was derived from the two clones.

However, using the small amount of sequence information from one of the clones, pSAU2, to search the *B.pseudomallei* genome sequencing project for any homologies with the sequence facilitated the isolation of a 38 kbp assembled fragment containing homology to the sequence from pSAU2. The sequence information from the clone pSAU8 was similarly used to search the genome sequencing project database but shotgun sequences were identified with only around 60% homology to the forward and reverse sequences of pSAU8 and moreover they were found not to be contained on the 38 kbp fragment. It may have been that the two clones pSAU2 and pSAU8 contained DNA from the two separate fragments, which may have different sequences. However, the homologous sequence to pSAU2 was located from 34,835 bp to 35,143 bp and it is possible that the sequence encoded by pSAU8 is not actually contained on this assembled fragment, as the two sequences may be as much as 12 kbp apart. Despite this the two shotgun sequences with homology to the forward and reverse sequences of pSAU8 were found not to be contained on the same assembled fragment. It may be the case that the isolate used in the genome sequencing project possesses only one of these fragments or that the deletion event that occurred in pSAU8 resulted in non-contiguous DNA fragments lying next to each other. However analysis of the 38 kbp fragment continued.

Sequence analysis of the 38 kbp fragment was carried out in the form of BLAST searches to identify any regions containing homology to proteins with known or putative functions

deposited in the Genbank database. The 38 kbp fragment appeared to fulfil some of the criteria defined by Hacker *et al.* (1997) to indicate that it could be a Pathogenicity island (Pai), which have been found in a wide variety of bacterial species. Firstly, such fragments are often associated with the carriage of virulence genes, and the fragment was shown to encode a region with homology to serine proteases. Sexton and Jones (1994) believed the protease they isolated from *B.pseudomallei* to belong to the family of alkaline proteases, sensitive to metal chelators. An ORF containing a putative gene sequence for the putative serine protease was identified along with putative –10 and –35 promoter sequences and a putative SD region. The nucleotide sequence encoded a putative product with a theoretical molecular weight of 63 kDa. However, a putative signal peptide cleavage site was recognised and the mature protease is estimated to be 59.5 kDa. This conflicts with the size of the protease isolated by Sexton and Jones (1994) (36 kDa) and Percheron *et al.* (1995) (42 kDa), although these two published reports also disagree with the size of the protease. The role of the protease in the virulence of *B.pseudomallei* has not been fully established. Although the expression of a high proteolytic activity by any *B.pseudomallei* strain was shown to be neither sufficient nor necessary for virulence when bacteria were injected via the intraperitoneal route (Gauthier *et al.*, 2000) a pathogenic role for the protease cannot be excluded. Moreover, microbial pathogenesis is usually complex and multifactorial. Bacterial pathogens often have a number of biochemical mechanisms which may act individually or in concert to produce infection and disease, thus the elimination of one of the components may or may not render the organism avirulent. In addition to the putative protease another region encoded homology to a *Streptococcal* haemagglutinin from *Streptococcus gondii* which is postulated to have a role in virulence.

The 38 kbp fragment was of a different G+C content to the rest of the genome, and was calculated to be 56%, which is very different from the genome value of 69% (Redfearn *et al.*, 1966). Pais often occupy large chromosomal regions, often more than 30 kbp and the

size of fragment identified in *B.pseudomallei* was 38 kbp. In addition to this, portions of the fragment contained repeated DNA. At around 15 kbp two regions encoded homology to the same hypothetical protein from *Thermoplasma acidophilum* and the end 8 kbp of the fragment was a repeated DNA region from immediately upstream of this, encoding homology to the *Streptococcal* haemagglutinin and a transcriptional regulator. There was also the presence of a region with homology to integrase genes at the beginning of the sequence which are often associated with Pais at their boundaries.

Much more work is needed to ascertain whether this 38 kbp fragment is a putative Pai. It needs to be established whether this fragment or, perhaps part(s) of this fragment is present in other isolates of *B.pseudomallei* and, importantly in *B.thailandensis*. Pais found in other pathogens are characterised by their absence, or sporadic distribution in less pathogenic strains of one species or related species (Hacker *et al.*, 1997). Pathogenic *Yersinia* spp. can be subdivided into highly pathogenic (high-pathogenic) and low-pathogenicity strains by the presence of a 'high-pathogenicity island' (HPI) (Carniel *et al.*, 1996). Initial experiments carried out showed the presence of *Sau*3A undigestible fragments in a number of *B.pseudomallei* and *B.thailandensis* isolates. It is possible that a region of the 38 kbp fragment is present in *B.thailandensis* isolates but excludes several genes essential for virulence. *B.pseudomallei* infections range greatly in their severity and outcome, which possibly reflects the possession of different virulence genes and perhaps even more so with *B.thailandensis* isolates, which do not cause illness. Therefore investigations to analyse the presence of the 38 kbp fragment or parts of it in different isolates could possibly highlight specific genes found only in highly pathogenic isolates.

In addition to this, the assembly of shotgun fragments was in its infancy at the time of writing and so the fragment may actually be much larger, as Pais can encompass chromosomal DNA segments of up to 200 kbp (Hacker *et al.*, 1997) and possibly encode

other secreted virulence products. In addition to this a number of Pais have been integrated into, or close to tRNA genes (Blum *et al.*, 1994) and it would be useful to search up and down-stream of the 38 kbp fragment, to analyse for the presence of such genes or other insertion elements.

A common theme when trying to clone both the 7.7 kbp fragment and the *Sau*3A undigestible fragments was that the incubation temperature was a very important factor in determining the relative success of the cloning procedure. Bacteria are, in general, conservative in their gene expression and will only express genes that they require in a particular environment (Harb and Abu Kwaik, 1999). Intracellular pathogens, such as *B.pseudomallei* should therefore possess genes that are only expressed in the intracellular environment. Being a human pathogen, the virulence genes of *B.pseudomallei* expressed in the intracellular environment of a host cell may be thermoregulated. All the fragments that could not be cloned at 37°C do possibly encode putative virulence genes and so perhaps these genes are thermoregulated in *B.pseudomallei*. Moreover, if the 38 kbp fragment is a pathogenicity island, it may resemble other such Pais, for example that of *S.typhimurium* that have been shown to be induced during intracellular infection of macrophages (Groisman and Ochman, 1996). More work is obviously needed to characterise these genes and to apply techniques such as β-galactosidase gene fusions, which has been successfully carried out with *lac*Z-listeriolysin (*lis*A) gene fusions and *lac*Z fusions to other virulence genes of *L.monocytogenes* to show that these genes are thermoregulated and therefore triggered by thermal cues, such as entry to the host (Leimeister-Wachter *et al.*, 1992).

The ability to sequence entire microbial genomes has generated vast amounts of information that leads to the understanding of gene structure and function, evolution, microbial diversity and pathology. *B.pseudomallei* is a medically important bacterium in the tropics and it is hoped that genome sequence information may offer new insights for

understanding disease-pathogenesis and host-microbial interactions, and for the development of new drugs and vaccine strategies.

It is clear from the information gathered in this research project that the molecular characterisation of this bacterium although a major task in itself will eventually provide invaluable insights into how the bacterium causes disease and the weapons with which it uses to do so. When a fuller understanding of the pathogenic processes utilised by this pathogen have been established through recombinant DNA techniques, new drugs and hopefully a vaccine can be developed.

APPENDIX I

Formulation and preparation of reagents, solutions and growth media.

BACTERIOLOGICAL MEDIA.

LURIA BERTANI (LB) MEDIUM.

per litre of ddH_2O:

Tryptone	10 g
Yeast extract	5 g
NaCl	10 g

The pH adjusted to between 7.2-7.4 and sterilised by autoclaving 15 min. at 15 lb/in^2.
LB agar made as above, with the addition of 1.5% (w/v) Agar No.2.

ANTIBIOTIC STOCK SOLUTIONS.

Ampicillin (sodium salt)	50 mg/ml in ddH_2O
Chloramphenicol	25 mg/ml in ethanol
Kanamycin	25 mg/ml in ddH_2O
Tetracycline	10 mg/ml in 50% (v/v) ddH_2O/ethanol

All antibiotic stock solutions were filter sterilised (0.22 μm) and used at 1000 times less concentrated as a working standard.

BACTERIOPHAGE METHODS.

SM BUFFER.

Per litre of ddH_2O:

NaCl	5.8 g
$MgSO_4.7H_2O$	2 g
1 M TRIS (pH 7.5)	50 ml
2% Gelatin solution	5 ml

Made up in 100 ml batches and sterilised by autoclaving for 15 min. at 15 lb/in^2.

NZCYM MEDIUM.

Per litre of ddH_2O:

NZ Amine	10 g
NaCl	5 g
Yeast Extract	5 g
Casamino Acids	1 g
MgSO4	2 g

The pH adjusted to 7.0 with 5 M NaOH and sterilised by autoclaving for 15 min. at 15 lb/in^2.
NZCYM agar made up as above, with the addition of 1.5% (w/v) Agar No. 2.
NZCYM soft agar for overlays was made up as above with the addition of 0.75% (w/v) Agar No.2

M9 SALTS (X10).
Per litre of ddH_2O:

Na_2HPO_4	60 g
KH_2PO_4	30 g
NaCl	5 g
NH_4Cl	10 g

Made up in 100 ml batches and sterilised by autoclaving for 15 min. at 15 lb/in^2.

PHAGE PROPAGATION MEDIA.
Per litre of ddH_2O:

M9 Salts (x10)	100 ml
20% glucose	20 ml
1 M $MgCl_2$	5 ml
1 M $CaCl_2$	0.1 ml
Casamino acids	10 g

DNA ANALYSIS METHODS.

CHROMOSOMAL DNA ISOLATION .

B.pseudomallei
Solution A: Sucrose (75% w/v) in 0.15 M Tris-HCl (pH8.0)
Solution B: Lysozyme (15 mg/ml) in 0.05 M Tris-HCl (pH 8.0)
Solution C: Na_2EDTA (0.5 M) (pH 8.0)
Solution D: Triton X100 (8% v/v) in 0.2 M Tris-HCl (pH 8.0)

PLASMID DNA ISOLATION FROM *E.coli*.

All solutions made in ddH_2O.

B & D SOLUTION I.

Glucose	50 mM
TRIS/HCl	25 mM
EDTA (sodium salt)	10 mM

The pH adjusted to 8.0, made up in 100 ml volumes and sterilised by autoclaving for 15 min. at 15 lb/in^2.
Prior to use only in the maxiprep method add 5 mg/ml lysozyme to sterilised solution.

B & D SOLUTION II.

ddH_2O	4.3 ml
NaOH (5 M)	0.2 ml
Sodium dodecyl sulphate (10% solution)	0.5 ml

B & D SOLUTION III.
Sodium acetate 3 M
The pH adjusted to 4.8 with glacial acetic acid, stored at 4°C.

RNase SOLUTION.
Ribonuclease A (pancreatic RNase) 1 mM in a 50 mM TRIS buffer (pH 8.0). Before use, this solution is incubated at 95°C for 15 min. to remove any contaminating deoxyribonuclease.

TE BUFFER.

TRIS	10 mM
EDTA (sodium salt)	1 mM

Made into 100 ml volumes, sterilised by autoclaving for 15 min. at 15 lb/in^2.

AGAROSE GEL ELECTROPHORESIS OF DNA.

TBE BUFFER (X10).

TRIS	0.9 M
Boric acid	0.9 M
EDTA (sodium salt)	0.025 M

pH is adjusted to 8.5, sterilised by autoclaving for 15 min. at 15lb/in^2. This X10 stock is diluted as required.

FICOLL LOADING BUFFER.
Per 10 ml ddH_2O:

Bromophenol blue	0.025 g
Ficoll 400	1.5 g

ETHIDIUM BROMIDE SOLUTION.
Ethidium bromide 10 mg/ml in ddH_2O
Stored at 4°C in a dark bottle.

DNA MANIPULATION METHODS.

TRANSFORMATION OF *E.coli* K12 STRAINS.

RF1 MEDIUM.
Per litre ddH_2O:

RbCl	12 g
$MnCl_2.4H_2O$	9.9 g
Potassium acetate (1 M, pH 7.5)	30 ml
$CaCl_2.2H_2O$	1.5 g
Glycerol	150 g

The pH adjusted to 5.8 with 0.2 M acetic acid and filter sterilised (0.22 μm).

RF2 MEDIUM.

Per litre ddH_2O:

MOPS (0.5 M, pH 6.8)	20 ml
RbCl	1.2 g
$CaCl_2.2H_2O$	11 g
Glycerol	150 ml

BLUE/WHITE SCREENING.

IPTG.

Isopropyl-β-D-thiogalactopyranoside. Prepared as a stock solution of 24 mg/ml in Distilled water.

X-gal.

5-bromo-4-chloro-3-indolyl-β-D-galactoside. Perpared as a 20 mg/ml solution in dimethylformamide.

PROTEIN ANALYSIS METHODS.

SDS-PAGE.

MAIN GEL.

Acrylamide (40%) (SIGMA)	3.2 ml
Main gel buffer	5.2 ml
ddH_2O	1.72 ml
APS (10 mg/ml)	380 µl
TEMED	30 µl

STACKING GEL.

Acrylamide (40%) (SIGMA)	0.7 ml
Stacking gel buffer	2.0 ml
ddH_2O	1.3 ml
APS (10 mg/ml)	100 µl
TEMED	16 µl

SDS-PAGE SOLUTIONS.

All solutions made in ddH_2O.

MAIN GEL BUFFER.

TRIS	0.75 M
SDS	0.2% (w/v)

The pH adjusted to 8.8 with HCl.

STACKING GEL BUFFER.

TRIS	0.25 M
SDS	0.2% (w/v)

The pH adjusted to 6.8 with HCl.

AMMONIUM PERSULPHATE (APS).
10 mg/ml solution made immediately prior to use.

SDS SOLUTION.
10% (w/v) solution.

ELECTROPHORESIS BUFFER.
Per litre of ddH_2O:

TRIS	3.025 g
Glycine	14.4 g
SDS	1 g

The pH adjusted to 8.3 using 5 M NaOH.

SDS-PAGE LOADING BUFFER.

TRIS	1.51 g
Glycerol	20 ml

Components dissolved in 35 ml of ddH_2O and pH adjusted to 6.75 followed by addition of:

SDS	4.0 g
Bromophenol blue	0.002 g

Diluted to 90 ml with addition of ddH_2O and stored at ambient temperature for no more than 12 weeks. Prior to use addition of 2-β-Mercaptoethanol to a final concentration of 10% (v/v).

STAINING OF POLYACRYLAMIDE GELS.

COOMASSIE BLUE STAINING SOLUTION.
Per litre volume:

Methanol	400 ml
Glacial Acetic Acid	70 ml
ddH_2O	530 ml
Coomassie brilliant blue-R	1.25 g

DESTAINING SOLUTION.
Per litre volume:

Methanol	400 ml
Glacial Acetic Acid	70 ml
ddH_2O	530 ml

IMMUNOBLOTTING (WESTERN BLOTTING).

TRANSFER BUFFER.
Per 2 litre volume:

TRIS	6.1 g
Glycine	28.8 g
Methanol	400 ml
ddH_2O	1600 ml

The pH adjusted to 8.3.

IMMUNOSTAINING.

PBS.
1 tablet PBS (Oxoid, Basingstoke) dissolved in 100 ml of ddH_2O. Sterilised by autoclaving for 15 min. at 15lb/in^2.

BLOCKING SOLUTION.
1% (w/v) casein in PBS. Sterilised by autoclaving for 15 min. at 15lb/in^2.

DAB/NICKEL CHLORIDE DEVELOPMENT SOLUTION.

PBS	100 ml
DAB	50 mg
$NiCl_2$	300 ml

Solution made fresh immediately prior to use in blot development and the chromogenic reaction initiated by addition of 100 µl H_2O_2 solution.

SOLUTIONS FOR SOUTHERN BLOTTING/PROBING λGEM11 GENOMIC LIBRARY.

DEPURINATION SOLUTION.
HCl 0.2 M

DENATURATION SOLUTION.
NaCl 1.5 M
NaOH 0.5 M

Solutions made separately and sterilised by autoclaving for 15 min. at 15lb/in^2 and added together prior to use.

NEUTRALISATION SOLUTION.

TRIS	1 M
NaCl	1.5 M

The pH adjusted to 8.0 with HCl sterilised by autoclaving for 15 min. at 15lb/in^2

SSC (X20).
Per litre of ddH_2O:

NaCl	175.32 g
Na citrate	88.24 g

Sterilised by autoclaving for 15 min. at 15lb/in^2.

TRANSFER BUFFER.
Per litre of solution:

SSC (X20)	500 ml
ddH_2O	500 ml

Made prior to use.

HYBRIDISATION SOLUTION.
Per 100 ml of SSC (X5) solution:

Blocking reagent (Boehringer Mannheim)	0.5% (w/v)
N-laurylsarcosine, Na-salt	0.1% (w/v)
SDS	0.02% (w/v)

Solution made prior to use and heated to 65°C for 30 min. to dissolve blocking reagent.

WASH BUFFER I.
Per litre of solution:

SSC (X20)	100 ml
SDS (10%)	10 ml
ddH_2O	890 ml

Sterilised by autoclaving for 15 min. at 15lb/in^2

WASH BUFFER II.
Per litre of solution:

SSC (X20)	5 ml
SDS (10%)	10 ml
ddH_2O	985 ml

Sterilised by autoclaving for 15 min. at 15lb/in^2

DETECTION BUFFER I.

TRIS	100 mM
NaCl	100 mM

The pH adjusted to 7.5 with HCl and sterilised by autoclaving for 15 min. at 15lb/in^2.

DETECTION BUFFER II.
Blocking reagent (Boehringer Mannheim) 0.5% (w/v)

Dissolved in Detection Buffer I, solution made prior to use and heated to 65°C for 30 min. to dissolve blocking reagent.

DETECTION BUFFER III.

TRIS	0.1 M
NaCl	0.1 M
$MgCl_2$	50 mM

The pH adjusted to 9.5 with HCl and sterilised by autoclaving for 15 min. at 15lb/in^2.

DETECTION BUFFER IV.

TRIS	10 mM
EDTA	1 mM

The pH adjusted to 8.0 with HCl and sterilised by autoclaving for 15 min. at 15lb/in^2

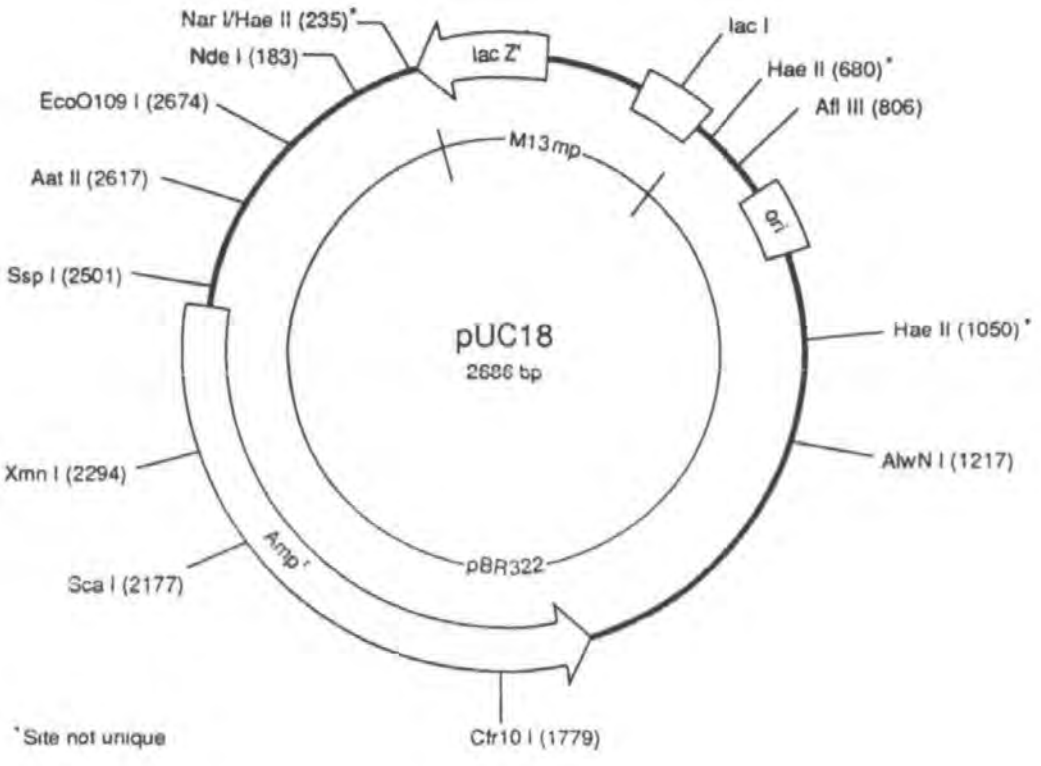
COLOUR DEVELOPMENT SOLUTION.

Per 100 ml of solution the following are added prior to use:

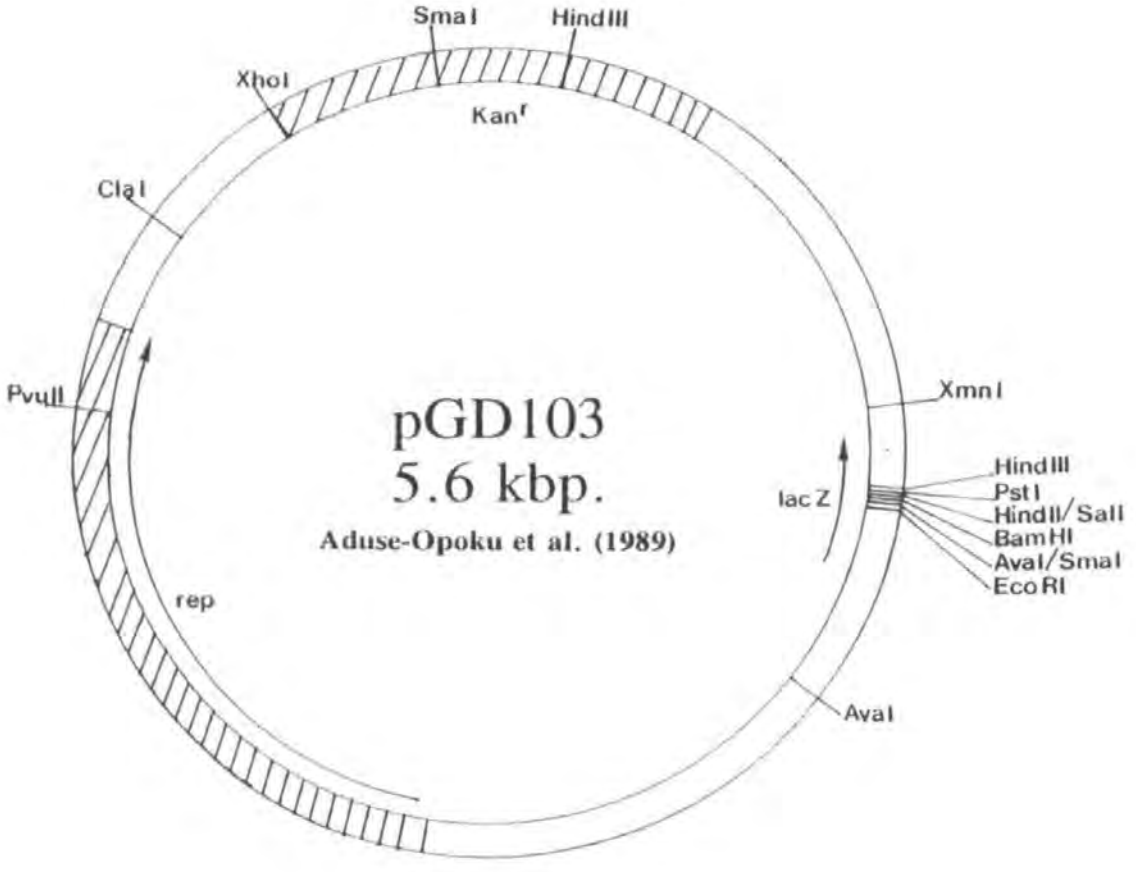
Detection Buffer III	100 ml
NBT	0.033 g
X-phosphate	0.0165 g

APPENDIX II

Plasmid Restriction Maps.



The plasmid cloning vector, pUC18 (Yanisch-Perron *et al.*, 1985).



The plasmid cloning vector, pGD103 (Aduse-Opoku *et al.*, 1989).

APPENDIX III

SDS-PAGE gels, Nucleotide Sequencing Data.

SDS-PAGE Gels of ECPs Isolated from Isolates of *B.pseudomallei*, *B.thailandensis* and other species of *Burkholderia* (Chapter 3).

GEL	LANE No.									
	1	2	3	4	5	6	7	8	9	10
1	MARKERS	19	22	33	25	46	53	97	98	MARKERS
2	MARKERS	102	112	212	216	217	392	204	426	MARKERS
3	MARKERS	448	576	441398	Hainan 1	Hainan 2	Zhan I	E8	MARKERS	E25
4	E27	E82	Jie187	Hainan 706	Hainan 106	Hainan 93	Hainan 55	Hainan 4	MARKERS	-
5	MARKERS	LMG 6998	LMG 11626	LMG 6999	LMG 10908	E260	E256	E255	E254	MARKERS
6	MARKERS	*B.cepacia* I	*B.cepacia* II	*B.cepacia* IIIaepi	*B.cepacia* IIIa non-epi	*B.cepacia* IIIb	*B.cepacia* IIIc	*B.cepacia* IV	MARKERS	LMG 10620

SDS-PAGE Gel 1

1 2 3 4 5 6 7 8 9 10

SDS-PAGE Gel 2

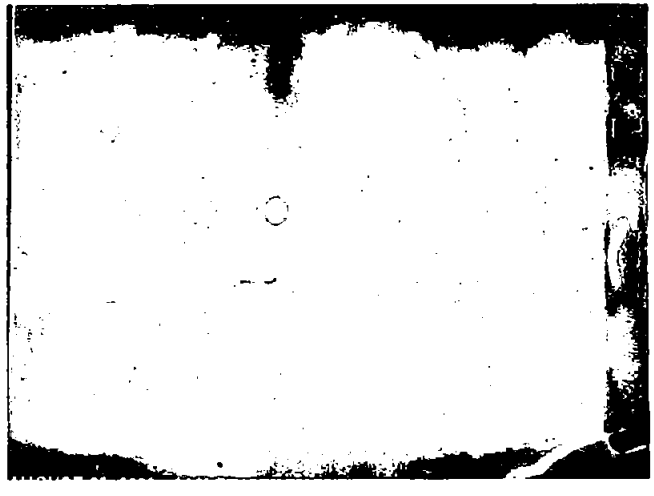
1 2 3 4 5 6 7 8 9 10

SDS-PAGE Gel 3

1 2 3 4 5 6 7 8 9 10

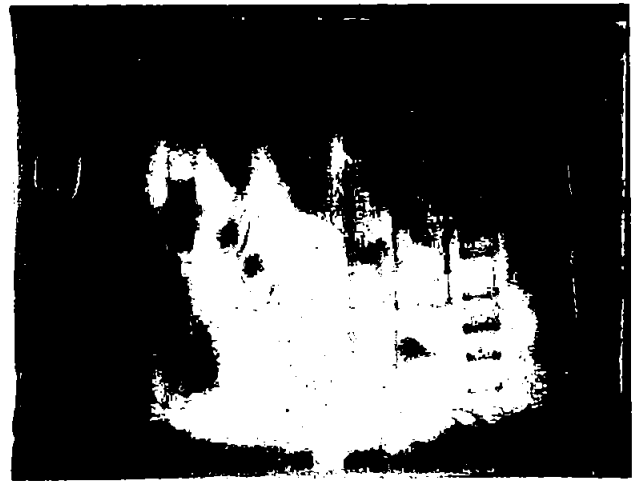
SDS-PAGE Gel 4

1 2 3 4 5 6 7 8 9



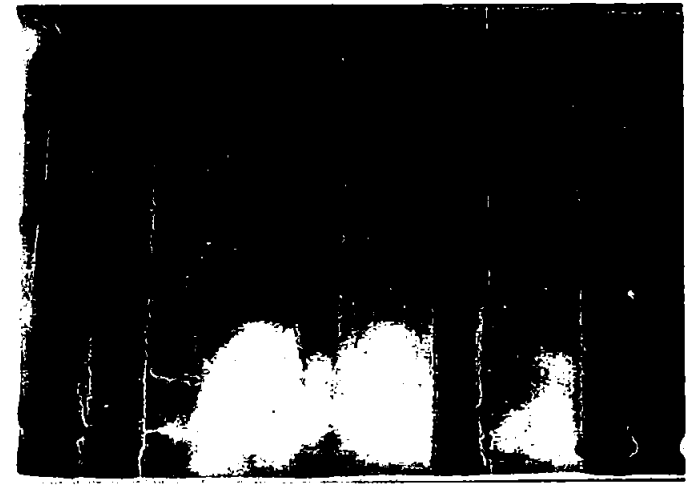
SDS-PAGE Gel 5

1 2 3 4 5 6 7 8 9 10



SDS-PAGE Gel 6

1 2 3 4 5 6 7 8 9 10



The complete 5 kbp nucleotide sequence of the immunopositive clone pBPGT2 (Chapter 6).

```
   1  AAGCTTGCAT GCCTGCAGGT CGACTCTAGA GGATCCCCGG GTACCGAGCT
  51  CGAATTCTCC CTTGAGCTTC GCGGTCAAGC CGAGCTCCGT CGTGCAGACG
 101  GGCGTGTAGT CGGCCGGATG CGAGAACAGG ACGCCCCAAC TATTGCCGAG
 151  CCATTCGTGA AATTTGATGC GGCCGAGGCT CGAATCCTGC TCGAAATCCG
 201  GCGCGATGTC GCCAAGACGT AGACTCATGT TACAGCTCCT TGAACTTGCA
 251  TGGGTTGGGT GCCCGGACGC ATCGCGCTAC CGGACAAGTT TTTAGCTTAA
 301  GCGATCCTGC GGCCGGCGCG AACGATTATT CCGTCAGATA TTCATGCCTT
 351  TTTATCATTC TTGCCGGGTG CGATTTGACA ATCCTGGAAA CTTTTTCGTT
 401  CCGCATCGGC TCTCTATTGT CACGGCAAGT CGTTTACCAT GCTGTTCAAG
 451  CGCGATCTTG CGCCACCTGC GCGCCGTGCC CGCGCGCCCG CGGCCAGGCA
 501  CGGGTTTCGC TACGATATCC GTAGCACAAC CCGCTTCGGC AAGAAGGGAG
 551  TCGCAATGTC TGAAGTCAAC AAGGAGAAAC TGATGTCGGA TATCAAAACT
 601  GTTCTCGCGG ACGCGGAAGA TCTGCTGAAG CAGGCCGCGA GCAGCACGGG
 651  CGACCGTGCG ACCGAGCTGC GCGAGAAGGC GCTCGCACGC CTGAAACAGG
 701  CGAAGGAGAA GGCGACCGAT GTCCAGGTGG TCGTGGTCGA AAAAGGCAAG
 751  AAGGCGGCGC GCGCCACCGA CGACTACGTG CACGAGCATC CGTGGACGTC
 801  GATCGGCATC GCGGCCGGCG TCGGCGTATT GATCGGCCTG CTGATCAACC
 851  GCAAGTAACG CGGCACGCCC CCGCGAGAGA CAGCCCCCGG CTCGCCGCCG
 901  CGTGACGGCC GGGGCCCTGC GCTCCGGCCC GCCGCGCCCA GTTGCGACGC
 951  CGGCCGTTCT CCGCGCGCAA GGCCCTCTTT CATGACGACA GAAACCTCAT
1001  CGCACCAGTC CGGGCACGGA CCGCTGCGCC GCCTGCTCGG CTCGGTGCTC
1051  GCGCTCCTGC AGACCCGGCT CGAACTCGTC GGCATCGAAC TCGCCGAGGA
1101  GAAGGAACGC CTGATGGGCG TACTCTTTGT CGGGCTCGCC GCGATGATGC
1151  TCGCGACGAT GGCGCTCATC AGCCTGACGG TGCTCATTGC GATCGCGTTC
1201  TGGGATACGT ACCGCTGGCA ATCGCTCGCC GTCGTCACCG CGCTCTACGC
1251  GCTCGGCGCA CTCGCCTGCT GGCTGAAGGC GCGCTCGGGA CTGCGCGACG
1301  CGCCCAGCGT GTTCGAGGCG ACGCTGAACG AACTCGAAAA GGACCGCGAG
1351  CTGTTCCGCG GCAAGCCGTG AACGCGCGGC TGCACTCCTT CCCTTCCCGC
1401  CCATCTCACG CCGACGCGCC ATGAGTTACT CCGCTCCGGA ACAACGATTC
1451  CACACCGCTT CGCAGAGGCG GCACATGAGC GCGGCCCAGT ACCGCGCGCT
1501  GCGCAAGGAA CTGCTGATCC TGCGCTCGGA TGTCGAGCGC CTCGGAACTC
1551  GCGCAGGCCG GCGCCGAGCT GCGCCAATCC ATCACCCATT TCAAATGGCT
1601  GAAGCTCCTC GTGCCCGGGG TGTCGGGTGG GTCGCTCGGC AAATCGGCGA
1651  GGAACGTGAA CGCGACGCTC GGCATGCTCG TCAGCCAATA TCCGCTATTG
1701  AGCTCGCTGG CGTCGGTCGT GCTCGCCAAG CCCGTGCGCG CACTGCTGCG
1751  CGCAAGCCGC GCCCCGCGCT CAAATGGGGC GCTGTCGGCT TCGCGCTCTG
1801  GGAGGGCTAC CAGATCTGGA AACAGGCGAA CACGACGACG CATCCGACGC
1851  CGGAGCGCGG CAACGCACGC GCCCGTAGCG ACACGCCCAG GACGCCCATG
1901  GCGCCCGTCG ACCGAGCCGG CGCGCCGACC GCGCCGGCCG GCTGCCGGCG
1951  CCGGGGAGAT CGACACGCCC GGCGCTGCGC AGCTAGCCTG TTCGCCAGCA
2001  GGCATCGCCC TCCGCCGCAC CGCCATCGAG CGCGACGCCT CCGGCCACGC
2051  GATTCTCGCG CGTGCCGTCC GCATGCAGCA CGTAGACGCC CGATCTGTCG
2101  TCCCGCATTG CACCATCGGC GGCGGGACGC GCCTCCAGTG CGTATCCGCC
2151  GCGCGCGTCG TCGAACATGA TCCGCAGCAC ATAGACGAGC TTGCCGGATT
2201  CGGGCGCCCT GTTCACCCCT TCCGGCAACG CGGTGGGCGG CGCGTCGCCG
2251  AACGCGTCCA CGTATTGCGC GCCGCGATAG AGCGCGGCAA CCGCGGTCAA
2301  TCGGTTCACC CGTTCGACAT AACTGCGATA CGACGGCACG GCAAACGCCG
2351  CGAGCACCGC CACGATCGCG ATCGCGACGA CCACCTCGAT TAAGGTGAAG
2401  CCTCTCGAGG GCGAATTGCG CCCCGTACAT GGCATCGCTC ATGCCTCCCG
2451  TTGCCTATGC CGGCCGTCCC ACGACCCGAC GCCACCGCCC CTCGACGCGC
2501  CCGCCGTCGA GCGCAACCTG CAATTGCAGC CACTCGGCCG TATCGTTCGA
```

2551 TGCGCCCACG CCGCGCGCGG TTACGAGGAA GGCGCGCACG TCGGGGCGCG
2601 CCGGCAGCCT CCATGCTTCG ATCAAGCAAC TCGGCGCCTG CGCCGCGCCC
2651 GGCCATCCGG CGTACGGCCT GAACGCACCG ATGCCGTCGA ACGCGCCCGG
2701 TTCGCGCCAG CCGCCCGGCT CGCGCGTCGA CGTCGTACCG TCGCCCGCCG
2751 ACGGGTGTGG CGCCCTCCCG CTCTCCAGCG CGATCACGCA TGCTTCGAGC
2801 CCGCATCCGC CGCACGAAAT GCAATGGATC GGCTGCGACG CTCGTCATAC
2851 GCCGGACCTC CGTCTTCGCG ATCTCGAACC AAGCGAACGA CATCACGATG
2901 AGCGAAGCCG TTACGGCAAG TACGCCCGGC AGCGCGACGC CCCGCGAACC
2951 GCGCATCCGC GCTCCGCCGT CGCGGCGCGT CGCACGGCGT CGGGCGCGCC
3001 CGGAACATCG GCAGTCGCTC CACATCCACC CCGACCATGA CGCGCATCAT
3051 GTGCCGGCTC GGTTACGGAC CGCCACATAG CGCCGCCACG CCAGTCTCGC
3101 GCGGGTGTCG GGTGCGCTCG CCACTCGAAA TTCGCAGTCG ACATACCGCG
3151 CCGGGCGCCC CGTGCGCCTG CCTCGCACTT GCACGCAGAC GGAAACGACC
3201 GCCACATTGG CCCAGTCGTC GACAGACAAG GCCGACGCGT CGGCCCATCG
3251 TGCAGCACCA TATAGTCGAT AGCGCAAACT CAAGCGTTCG ACGCCTTCGA
3301 CAAGCGGCTG CGCGATGCCG GGTCTGCCGC TGCCTTCGCA GTACAGTTCC
3351 GGTTCGCCCG TCGACCTGAC GCGCGCGTAA AAGCGGTTGA CGATCAGCGG
3401 CTGCGTATCG GCCGCGCCGA CGCCCTGGCC GAGGCAGTCC GTCGGCTGCC
3451 CGCTCGCCGT CGGCCATGTG GAGACGCCGT CTCCGACATA GCGGACGACG
3501 AGGCCGTCGG AGCGGCTCGC GAGCGGATCG CATGCGGTCT GACCGTCCGC
3551 GCCGACCGGA CGTCCCGCCG CACAGCCGAA CAGCCCCGGC ACCGCGCGGG
3601 CATCGTATGC ATCGAGCGGC ACGAAGCCCG CCATCTGGAT CTGCTGCGAG
3651 ATCAGCGCGA GCGCCGCCTG CGCCGCGTCC CGCAAGCGCG CGGCATCCAC
3701 CGCCGCTGAA TAAGCAGCCC GCTGCACGCG ATAGAGCGAC ATGGCGGCAA
3751 GCAAAATCAG CAAGCCGAGC GCCATCGCGA TCATCACCTC GACCAGCGTA
3801 TGCGCGCGCC AACGGGTCGT GCGCATCATC GGCGCTGCGT CGATGCATAC
3851 GCGAGCGTCA CGCACGACTT GGCGGCATCG CGCGCAGCAG CGGCATCGCC
3901 GTCGCAACGC AAAGCGCTTG CAGCCAATCC CGCGGACCGT TCCCATTCGA
3951 CCCGTACGTT CGAAATATCC GCACCCGCGT TCGTCGCGAC GAGCTTGCCG
4001 TGGCGCAATC CGGACATCGA CGCGTCCCAG TCGACGCCCG ACGGCCCACC
4051 GGCTCGCGCC ATCTCCGCCC ACGAATCGGC GACGAGCACC GCCGGCTCGC
4101 ATGCCGCGGT GTCGCGCTCG CTGCGTATCG CCGCGAGCAA CGCCGCCACC
4151 ACACCGAGCG ACGCAACCGC GAGCAGCGCG CACGCGATCG CCGCCTCGAG
4201 CAACGTGCTC CCACGCATCG AGCGGCAACA TGTCATGTCG ACCGTCCGCA
4251 ACCGCCCTCG GCGAATCGCA CGCGACCACC CGCCGCGATG CGCAGGCAGC
4301 GGCGCCAGCG TTCCCCGCGC CATGCGCCCG ATGCGTCTGC CGGCGCAAAC
4351 TCGAAACTTC GGAAGCCGCC GATCACCTGG CCCGCGGGCG GCGTGAACAC
4401 GACATCGACA GTCGTGCCCG TCACGGCCAC TCGGGCATCG GCGCAATGCG
4451 CCGCAACAGG GCGTACCGCG TTCGCCATCC GCGGCGGCGA CGGCCCAGCC
4501 GCACGACCAA TCGGCCGCCC CGCCACCGCA TGGCCGCCCC GCCGCAATGC
4551 ATTTCGCGGT CGCATCGCTG CGGCACAGCA CCACGCGCGC GCCGAGCCTG
4601 ACCGCTTCGC CGCGCGCATA GGCGAGCAAT GCGCCGAACA CGCGCGCCCG
4651 CGCGTCCACT CGATCGCGCA TTCGAGCGCC GGCGAGCGTC GGCGCGGTCA
4701 GCATCGCGGC AAGCGCCACG AGCATCAGCG CAGCCATCAC TTCGACGAGC
4751 ACGAAGCCCG CCGTTGCCCA CCTGTCCAGC TCCGCTCGCA TCGCGCTCTC
4801 CTTGAATCCG ATGCGCCGAC TCTACGCGTG CAGGCAAGCG CCGACCATCC
4851 GCCGAATGGC CGAGTCGCGC GCCGCGACCG GGCACACGAC AATCGCCCAG
4901 GCGAGCACGC ACACCGCGTG TTACGCGGGA GAAACGAATG GAACGGAAAA
4951 CGAAACGGGA AACGGAATGG CAAGCTCAAC GCTTGCGAGG CGTCTTCGCG
5001 GGCGCCGCGC GGCGGAATC

The complete nucleotide sequence of the 7.7 kbp *Eco*RI *B.pseudomallei* fragment (Chapter 8).

```
   1  AAGCTTGGCT GCAGGTCGAC GGATCCCCGG GAATTCGGGC ACCAGGCGAA
  51  ACAGCTGCGG ATACGTGATG AAGAACGTGC TGATGGAAAT CGAGACATCC
 101  ATCGAGATCA CGCAGTCCCC CAGTTGCGCG GAATGTGCAG CGTCTCGCCC
 151  GGCTCGATGA TCGCCTGGTA CGGCGTGCAG TGCCGGTAGT TCGGATAGCG
 201  CTCGAGATCC GGCTTCTCCA AATCCACCGG ACTGAACGTC TCGCGCTCGA
 251  ACTGGCCGAT ATAGAGGTTC TTGCTGGCCT GATCCGGCGG GAACAGATAC
 301  ACCATCTTCC GGCCCACGAT CTGCGTGAGG AAATTCGGCA TCGGATCGGT
 351  GTGGATCGGC GTATGCGTGC CGGCCGGCCC GATCCACAGG TACGAATCCA
 401  TGTGCGCACC CGGGATGTGC TCGATATGCC CCTGCACATA CGGATGCGGC
 451  GCGAGATCGG CGCGCAGGTT CGGAAAGCTC TTGCGAAACG GATGCGACGC
 501  CATGTAGTAC ATGTGGTCCG GCTTCGTTCG CAGGATCCGG ATCAGTTCCT
 551  TGAGCTTGAG CTCCTGGAAA CGCAAGTGCT TGACGTAACC TTGCGGGCTG
 601  TTGTTGCCGG GCCCGCGCTC GTCAGGCTGA AAGCTCTCGA CGCCGATGGT
 651  CTCGTCGCCG AAACGATTCT CGAAATAGTC GAGCGACCAT TTCTGCATCG
 701  CCGGCCACGC ATCGAGCGCC CTCGCGATTT TCACCGGGCG CTCTTTTTCG
 751  AGGTAATGCT CGCGGAAGTC CGATACTTTG GGTCGAACCA GCGTTTCAAT
 801  AGGCAGAATT TCATCGTACG GCATAGTCGA TTCCCGGGCC CGAATGTTGA
 851  TTGAGATGCT CGCGGACGCT GGCAGGCACC ATGTCCGTCC AGAGTTGCTC
 901  GATGCGTTGC AGACACTGTT GCTTCGTCGC AGGTTCGCCG ACCGTTTCCC
 951  AACCGGCGGG AAGGCTTCGC CGTCCATCCA GACCGAATAC TGGCCCTCAT
1001  TATTTCGCAC TACGAGTACA ATTCGTCGTC CATAGCTGGT TCCCAGGGAA
1051  AGCAAGGTGA TGTGCCGGGT TTTTGCCGTG GCGGCGCGCA GGTGAAATCG
1101  CGCAAGACCA TCGGCTGGAA AGACGCGCCG GATGTACCGC CTCGGCATCG
1151  CTCGGATGCA GGCTCGCGAC TACGCCCAAC GCAGTGAATG CATCAGAGAA
1201  TTGAGTATTT CTTCTGCGGG AATTGTCCTT TGAACTAACC TGATATTCGG
1251  TTATTGGGTA TGTATTTGTC AACGAACTCT CATCTTCAAG GCACGATTAG
1301  AATTAGCCAT AACCTCTCGC CTGTCAAGCT ATTTTGCGGC AAGGTACGCC
1351  GCTTTATTTC AAACAACAAT GCCTCCCTGA TAAAGATTAC AGGCGAAAAA
1401  GCTTTATTTC AAACAACAAT GCCTCCCTGA TAAAGATTAC AGGCGAAAAA
1451  TGCCCCGCGC GCGGGGGGGC GATGCCCTGT GACTGGCCGG CCGCTGCGAG
1501  GATTCGATGG GCGCTCGCTC GCCGCCCTTC ATCCGCAGCC GCTCGAAACG
1551  GATCCAGGCA CATGAATCCA AACAATATTT TTCCCGGCTC AAGCGGCAAG
1601  GAAAGCGAGC CGATTGAACT CGGAATGCTT GCGGCTGACC ATTAATTAAG
1651  CGCGCACATT ATAGAAAATA TATTAATCAT CCATATTCGT GAGGCACGGA
1701  AATTTTTACC GTCGATGCGT TTTTGCAAAG AAGCCTTGCC GACCCCAAAC
1751  GAGTGCGCAA ATCTCATAAA CTTCCCCGGA TCATCCCGCC AGTCATTTTC
1801  ATGAAATAAT TTCGGTGGAG AGAATAGATC AATCGATTAT TTTCAGTCAG
1851  CCTGATCATT GCACCGAAGG CAACTAATTG CGCGTTCGAA ATACGGCGGC
1901  GGCGCTTGGC ACCCGGTGCC GTCCGATCGC AACATGGAAT GCGCGCTACG
1951  CGACATCGAA CCCCGGGAGA CCTGACCGCC ATGCATGTTT CGACCGATAC
2001  GTCCGCCGAA GCCCATTCAC CGGCGCCGAA TGCTCGCGAC ACCGACTCAC
2051  GCGCAGACTC ACTCGATATG TCAGACAGTG ACGCACTGCG CCGGATTGCC
2101  GAAGCCGTCG ACGGCGGTGC GGCGAACATC GAGCGAATCG TTCCGCTCGC
2151  CCGTGCGCGC GAGCGTATGC CGACGCGGCC TCGGCTCGAG CGGTGCGGCG
2201  GCGGGCGAGT GACGGCGGCG CACATCACGC TCGACTCGCG TGCGCGTCTC
2251  GATGCGTTGC TGCACGCATT GCAACGCGCG ATCGACCAGA ACGCGGACCT
2301  GCGAACGTGC ATTTTGGGGG CGTGCCTGCG GCGGCCGATG CAAGTCACGC
2351  TTCGCGAGGT TCGCCTGCGA GTGCACGCCG CGACGCTCGA CCCCGACCTC
2401  GATCCCGCCG CGCAGTTGGC CGCGCTGAGC ACCGGGCCCG GCATGCGCAT
2451  CGACATGCAA CGCCCGCCGT GGGTGCTCGC GTGCATCGCG CGCATTCCGG
2501  GCAGCGGGCA ATGGCTGCTG CGGCTCGTGG CAGCCCCGAT CGCGGCCGGA
2551  TTCGACGCGC TCGACGCGCT GCTTCGCGAG GCGGTGATTC ACGGCGACCG
```

2601 GGAGCCCGGG CCGACGCCGT TTCACTGGAC TGTGGAAACG GCTGTTGAAT
2651 CGTGCGGAGG CGAACCTGCG TCGTTGCCGA CCGCGGGCGC GGTTTGGCCG
2701 TCGAACGACG TATCACGCGC TTGCGATCCG GATGCCGCGT CGTGCGTCGA
2751 GGCGCGCATC GCCGCGATCG CGTCCGATCT GCCGGGCGTC GTGCATGGCG
2801 GACCACGAGA CGATTTGCGC GCGCTCGGAC GAACGCCGTT GCAGGCGCTT
2851 CGACTCGCGC GCCGTATCCG CGACGAGCTG GGCGTGACCG TACCGGTCGA
2901 GTCGATCCTC GCGAGTCCGA CCATCGTCGA GCTTGCCGGA TACGTCGAGC
2951 AATTGCGCTC GCGGGACGTC CGCGACGGCG CTGCGCCCGT GTCGATCGGC
3001 GAAACACCGG CGGACGCGGA TGCTCGGGCG CAGGCGCAGG CGGATACGGA
3051 TACGGCGCAC ACCGATTGCC TGATCGTCAT TCAAGCAGGA GGCGCCGAAC
3101 AAGCGCCGGT GTTCTGCATC CCGGGCGCGG GGGGCAGCGT CGCGTCGTTC
3151 GTTGCGCTTG CGAGCATGCT GCGCGCCGAC ATACCGGTAT ACGGCTTGCA
3201 GCCTCGCGGG CTGGACGGCC TGGGGCCGCC GGACCGGTCC GTCGAAGCGG
3251 CTGCGCGCCG GTACGCGCGG GCCATTCTGG ATGCCGCCCC GCCCGGGCCG
3301 CCGCGCATCG TCGGCCACTC GTTCGGCGGC TGGATCGCGC TCGAGACAGC
3351 GCGGCTGCTG GACGGCATGG GAACGCGCTG CGCCCCGCTC GTCCTGCTCG
3401 ATTCGAATCC ACCGCCCGCG TCGCAGGCCT GGCGCGCGCC TTCCGAGGCA
3451 GACATGCTGC GCACGCTCGT CGGCCTGCTC GAGCAGGCGC CGGGCGGCGC
3501 CCCATCCGGG ATCGGCGACG AAGAAATCGC CCGTTGCGCG CGACGCCCGC
3551 AGGATCGCGC GGATGCGCTC GTCCACGCCT GCATGGTGAG GACCGAAATG
3601 CTGCCGCCGC GCGCGCCGGT CGAAGCGGTG GCATGGTGAG GACCGCCCTG
3651 CGAAGCCCAT TCGAGCACCC GCTACGCGCC GGGCGGCCGA TACGCGGGCG
3701 ACGCAACGGT GATCGTCGCC AACGGCGATC GCGACGCGGG CGAGATGGTG
3751 CCGGCGTTCG GATGGGCCGC GCTGATCGAG CGAGTCGAGG TGGCCGTGAC
3801 GCCGGGCAAT CACATGAGCA TGCTCGCGGC GCCGTATGTT CGTCACGTCG
3851 CGCTGACGAT GAAGGCGGTA TGGCGCATGA TCTGAATCAG CGCGGCGGCC
3901 GCCGCCGCGC TGCGATGGCC GTTCGTTGCG CCCCGTGCGC CGACCGATGA
3951 AGCCGGCCGG GCGCCGGCAC GGCGGCGATC GGTCGTGGCG CGGCCGGCCG
4001 CGTCGCGCTT CCGGCCGAGC TGTCTTCGCG GCGAAACCTC GCGCGCTTCC
4051 GATCGATCGC GCCGCGCAAC GGCCAACGCC TGCGCATCCG GCGCGCATGC
4101 TTCGCGGCCC GACACGAGGC GTTCATGAAC AGCGCCGCCG GTATCGTTTG
4151 CAAGCCGGTT GGAGCCGACG GCGCGCCTCC CCCGCCGTGG ATCGAACGCC
4201 CCATCGCGAA ACGGCAGCCG GCATCGGGCC GGTCATTCCA ACGCCGCCGT
4251 CGGAGCGCCG CGCCGAGCCG GCCGGCCGCC CTCGCCTGCC GCGCGCCGTC
4301 TAGAACCGAT GCCGCAGCCC CAGCGCGACA ACGACCTGAT TGCCGTTCGC
4351 GGACGGCGTC AGCGTCCAGA TCGTCGCGTT GAACGCCGGA ATGCCGTTGC
4401 CGCCGCTCAC GCGCTGATAG ACGGCCGCGA GATACGCGTC CGTGCGGATC
4451 GAGAACGCGT AATCGGCTTG CGCGACCATG TGATTCCACT TCGGACGGGT
4501 TTCGCCCGAG CGCGCGTCGA AGCGGCCCAT CGTGTACGTG TACGCCGCCG
4551 CGAGGCTCAA CCTCGGCGTG ACGACGTAGC GCCCGTCGAG CGTGAAATTG
4601 TCGAAGACGA GCGAATTGCC GTCGAGCTTC GCGATGCTGC CGCCCTGGAG
4651 CACGCCGCTC ACGCGGTCGG TGGCCGAATG GGACCACGCG GCGCCGATCG
4701 AATGCGGCCC GAACGCATAG CGGCCCGCGA CGGCCCAGAT CTGCTGGCTG
4751 CCGCCCGTGA TCGTCGCCGA GCCGTCCGTC GTGCTCAGCG CGCCGTTCGC
4801 GTTCGCGGCG TTCGGATTCC GATTGATTCT CAGATAGCCG GCGCCGAGTT
4851 TCAGCGGGCC GTTCGCATAG GACAGGCCCG CGCTCCACGC CGCGTTGTTG
4901 CCGAACGGGC CCGCCGTGTT CGAGAAGCCG TACATTGCGC CGAGGGTCCA
4951 TCCGCGGTAT GTCGGGCTCG TGTACTTGAC CGCGTTGTTG ATGCGGATGT
5001 TGCGGTTCGA GTCGTCGTTG TCGTACGGGT GAACGGCGAG GTTGCCGCCC
5051 CAGCCGGTCC CCGACGCGCC GAGCGGCGTG ACGAAATCCA GGATGAGGTC
5101 GTACTGGCGG CCGAGCGTGA GCGCGCCCCC TTCCTTCGAT TTCAGTCCGA
5151 TCCATGCCTG CCGGCCGAAC ATGTCGACGC CCTTTTGCGA TAGCGCGCCG
5201 CTCGCCCCCG AGAAGCCGTT CTCGAGCGCG AAGATCGCCG ACACGCCGTC
5251 GCCGAGGTCT TCGCGGCCGC GCAATCCCCA GCGTGACGCA TTGAGCGCGC
5301 CGCTCGTCAC GGCCCAGCCG GGGCTTCCCG GCGAGCCGGC TCCGTGGGTG

```
5351  CGCTGGTTGT TCGCGTAGGT GATCGACGTA TCGATCAGGC CGTACAGCAC
5401  GACGCTGCTT TGCGCGTGAG CCGTGCTCGC GCCGAAGCCG GCGGCGCATG
5451  CCGCGGCGGC GGCGCCGGCG AGGCGCGCGA GCCGGCGGTT GCCGGGAAGC
5501  TGCCGGGCCG CGCGAGCCGA CATCCCGGAG CGGAGGGTCT TGTGTCGTTG
5551  GTTATCCATA TTAATTGTTA TGTTGGCGGG CGCGATTGGC GGTGCGTGCC
5601  CGGGCGCCGA TTGGCGGACG GAGCGCCATT ATTGGAGACG ACGCGCGTGT
5651  CGATTAACCC TCGGGTTCGG AAAGGTCCTC TGCGAAATCG CGATCAATCG
5701  AATCCAATCG CGTTCAAGCA CATTGCGCGC GAATGCGATC GAGCGCGCGC
5751  CGGGCGTTCG GGCGCGAGCG CGCCGCGGGT CGCGCGCCGC AAACCCGGGC
5801  GGACCGGACC CGCGCCGCTT CCGGCGCATA ATGGCGTCTC GGCCGGCCAG
5851  CGGCCGTTCG AGGAGCGCGC ATGCACCGTG TCACCCATTA CCGCCGCGCC
5901  GACGACGGCA TCGAAGCGAT CAGTCTCGAT ACCGATCGTG CGTTTCCGCG
5951  TCACGCGCAT GACGAGTTCG GCGTCGGCGT GATCGTCAGC GGCGCGCATC
6001  GTTCGTGGAG CGGATGCGGA ACAGTCGATG CGCGCGCCGG CGACACGATC
6051  ATGGTCAATC CGGGCGAAGT CCATGACGGC CTGCCGCTCG GCGCACGCAC
6101  GGCGCGCAGC TGGCGCATGC TCTATCTCAC ACCCGCGCTC GTGGCGGGCT
6151  TCGCCGCCGA AGAAGGATGC GGCGGCGTCG AGCTCGCGCA TCCGGCCGTC
6201  CGCGACGCGC GACTCGCGGC CGCCGTCGAG CGGCTGTTCG CGCGTGTCGT
6251  GGCGGGCGCG CAAACGCCTG CGGCGTCCAT CGCGTGCGAC GAGGCGCTCG
6301  CCCTGCTCGT CGCCGCGCTG CTCGCGCGGC ATGCGAGCGG GCGCGCGCCG
6351  CGCCGGGAGC GCGCCCGCGA TCCGGATCGT GCGCGAGCGG CTCGACGCGT
6401  CGCCGGCCGA TGCCGTGTCG CTCGCCGAGC TCGCGAGTCT GGCGGGCGTC
6451  AGCCGTTTTC AACTGCTGCG CGGTTTCGCG CGCGAGCTCG GCATCACGCC
6501  GCATGCCTAT CTGATTCAGT CGCGCGCACG CTTCGCCCGT GCGTTGCTTG
6551  CGCGCGGCCT GTCGATCGCG CAAGCCGCGG CCGAGGCGGG TTTCGCCGAC
6601  CAGAGCCATC TGACGCGCGC GTTCGTTCGC CAGTTCGGCA TCACGCCGGG
6651  ACGCTTCGCG TCACGCTGAA CGCGTGCTCG CTGCCGCCGC GCGAGCTCGG
6701  GCCGTGCGCG TCGCATGTCC GCAATAGTCC GAGATGGTCC TCGCCATCGC
6751  CCGCTCCCCC CGCTTCCCGA TCGCCGCCGC GCCGTTGCAA TTTTGTTCAA
6801  GACGCGCTCC GCCCGCGAGG CGACGATGCG GGCATGAAAA CGAAACTCAT
6851  CGGTTATCTG TCTCTCGCCG CGGCGATGAT GGGCGTCGGC AGCACGGTCG
6901  TCGCGAGCCG CATCGCGGGC GATGGCCTGC CGCCCTTCAC GGCCGCCGCG
6951  CTGCGCTTTC TGATCGCTTC GCCGCTGCTG TACGCGCTGA TGCGCGCGCA
7001  GCGGCTGCGC TGGCCGCGCC TCGCGCCGCG CGAGCTCGGG CTGCTCGTCG
7051  TGCAGGCGGC GTGCGGCGGC GTCGGCTATA CGGTGCTGCT GATCGTCGGC
7101  ACGCGCCTGT CGTCGCCCGT CGACGCGGGC GTGATGCTCG GCACGCTGCC
7151  GGCGATGTCG ACGCTGATCG CGGCCGTCGT GCTGCGCGAG CGGCAGACGC
7201  CGCGCGACTG GGGCGCGGCC GCGCTCGCGA CGGCCGGCGT GCTGCTCGTC
7251  ACGCTCACGC CGGGCCGCAC GACGATGTCG ACGCGGGCGC TCGCGGGCGA
7301  CGCGCTCGTG CTCGCGGCGG TCGCGTGCGA GGCTGTCTTC ATTTTGCTGA
7351  ACCGCCGGCT TGCCGTGCCG CTCGCGCCGC TCACCCAGTC GAGCGCGATG
7401  TCGGCGGCCG GCTGCGTGCT CGCGCTCGCG CCGGCGGCGT TCGAATGGCG
7451  CGCGGCGACG GCCGCCGCGT GGCAACCCGC CGCGCTCGCG GCGATCGCGT
7501  ACTACGCGCT CGTGCCGACC GTGCTCGGCT ATCTGTGCTG GTACGCCGGT
7551  TCCGCGCGCA CGAGCGGCAC CGAGGCGGCG TTGTTCACGG CCGTGGCGCC
7601  GGCCTCGGCG GTGCTGTTCG CGGCGGCCGC GTTCGGCGAG CCGCTCGACC
7651  GAACCCGGCT CGCCGGCATC GCGATGGTGG TGGCGGGGGT GCTGGTGGGC
7701  GCGATCCGGC GCGCTGCGGC GCGCATCGTC CACGCGAAGC CGCTGCGGCG
7751  AATTC
```

The Nucleotide Sequence of the 38 kbp Sequence Containing the Cloned pSA2 Sequence (Chapter 10).

```
1       CCGCGTTCGCGGCGGAAATGCAGGAAAAGCAGGTGGACTGGCAGCTTCATCTCTTCGGCG
61      GCGTCGGCCATGCCTACACGAATCCCGACGCCGACGGGTGGAACAAGCCCGGCTACGGCT
121     ACAGCCGCGCCGCGGACCAGCGGGCATGGACGATGATGCTCGCGCTGTTCGACGAAGTGT
181     TCGGCGGCGCAGCCGCCGTCGGCAAGGCGTGAATGTCGGCGCGTTGAAGCGCGGTCGAAT
241     TCGCGCCCTCCGTCTGAACGACGGAGAAAAGAAAAACCCCGCGCAAGGGCGGGGTTGATT
301     TCTGCATGCCGCGGTTGCGCGACTCAGAACGGAAGTGGGCATCGATGAGAGCGGAAGAAG
361     GGGAGGAAGATTTGGGCTCGAAGCCTTGTCCCTTGGGTTGAGGGGAGCTTCCTGCATCAA
421     CGCACGAACCTGTAGCTGCCCCCGCTCCGTATCTCAGTGGTGCCGTTATCACTTCGCGCA
481     CAACGTCCCTTCCGAAGAAGGAAGGCCCCCGTTAGGGGGCCAAAATCAGGCTGGCGCGCA
541     CGCAAAGCTCTTGCGGACAAATTCAGAGAAAACGTCGCTGATAGCGTTACCGGCCTTCCT
601     GTAGCAGTGTGCGCACCGGCACCACGACCGTCTCTGCACCGATTGCGTCCAGGCATTCGA
661     ACAAGAACGCTGCCTGGAACGAACCTCGCGACATCTTCAGTTTGATTGTGGTTTCGGTCT
721     CAGCACCGGGCCGTCGTTCATTCAATCGCTCCGCCAGCACCTTGTACGTTATTCCTCTTT
781     GAGCCAGCTCCGAACGAAGCAATGCGGCAGCCGCCCGACGATAGCGGTCCTTGTTAGGAC
841     TACTTGCATCAGCCATTTGTGTCCTGTAGAGTTAGGTATCGCCCGACGATACCTTTGAGT
901     CGCAATGTGAAACGTCTGAATAGGAGAATGATGTCAACAACCGACTTGGAAGCACGTAGC
961     GTCCGCCGCATGAGTATAGCAACCAGCCTCGAACGCGTCGGCGCCGGGACAACGCCTCCG
1021    CTTCGTAAGCGTAGGCGAGCCAAGTCACGCCCGCGCTATCGCCAGTGCCGCGTAATCAGT
1081    CGGGGGGTCGTACTAGACTCTTGGGGGAAGCCAGGCATGCACCCCGCTTTCAGCGCGCGA
1141    ATTGATAAGCCAATTCGACGGACAAAATCGATTATCAATTAATTGATAATCACGAGCATC
1201    ACATCTCTTTTCAATTGACACACACAAATAATGCCACCGACTTTACCTCGACCTCTTGAA
1261    AAATTATATTTCCGTGATACGATGCGCTGGTTTGACATGAAGCACATCATCAACGAAAGA
1321    GGAAATCAACAAACATGGAAGGATAGGCCAACACCACATTCGCGGGGATATCCCGTTAGC
1381    GCCCCAAAAAAACAAGGGCGCTAACGGGATTCCCCGAGAGCGCCCTTGCGCCGACAGTCC
1441    ATAAGACACGCCGGACCTCAGAAACCGCAGTGTCCTATGAAACTCCCTCGGTGTCAAGGA
1501    TGGAATCGTTCGTCCATCGATTTTGTGATGGACGAACGATTAATATCGTCCGTAGCAAAA
1561    CACCGAAATGAGCGAAACCGATAATCCCGGCAACAATCTCCCGGCAACTGATATCCAGCA
1621    GCAGCCCGGCACACCGCCCCCGCCAATGCCGCAGGGATATCACGTTAAGATTCTAACGCT
1681    CCAGCGAGGCATTAATCGCGGGGAGCGCATCTCCATATTGCTAGATGGTTCTCGTGAACC
1741    GGTGACTCTCGCAAATGAGCACATGCTTTGCGGGAAGCGATTACTCCTCTCACCCAACAC
1801    TTATTTCAATCACCTTTCACACATCGCCCTGATATACAAGTGGGCGGATGAGCACAACAT
1861    ATCGGTAGCCCAGAGGTTGCTTGGGGGCATCGGTTTTGCCAATTATGAGGTGAGCTCAAT
1921    CTTCGCAACAGTCCGAAAGCGTCAACGGCAAGCCAAGCCCGGCACGATGATTGTCGGCAA
1981    AACGGAGCAAAGAGCGCGCATGGTCCGAATAAAGAACTTCGTGATTAGCGGGATGGAAAC
2041    CGCCACGGCAAATCTCGACCCACGAACCGACACGGAAAAGCTGAACGCGCTTGAATATCG
2101    CGTGGAAACGACACGAACACTGTTTGAGCAAACAACGCCGAGCAAGGAACGCTCGCCATC
2161    GAAGAAAGGACTCACGGCAGAGTCGGTGCAATTGCTGGCAGATACACTTTCTCCAGGGAG
2221    CGCGAATAATCCCTGGAGAAAGCCGCTAGTCCAGCATCGCAATTTCATACTCGTCATGCT
2281    CTATTTGGTCACCGGCTGTCGCCGAGGTGACTTGGCCAAGGTAAAACTATCCGACATCGT
2341    GCATGAACAAACACCCTACATCCGCTTTGACGCACACGTCAGTGACCCTAACGACAGACG
2401    GCCTGCTGAGCCGAGGCTGAAAACGCTGTCTCGCGAGTACCCCATTCATACGACGGTTGC
2461    CGATGCCGTTAGCCGGTACATCGCCCAACATCGGGTGGAAATCCCGAACGCGGCCGACAG
2521    CGAGTACCTATTCCTCGAAACCAGCGAAGGGCGTGAACTCGCTATGCGGACGGTCAATGC
2581    CATCTTCGAGACGTTACAAAGCGTCATTCCTGACCTCACGCCGCACATTTTGCGGCACAC
2641    GAAAACCGAAAGCCTGCTCCGAGACGCAGAAGCGATGGGGCTGACGGAAACACAAGTTCT
2701    CGAGACCGTCATGTATCTCAACGGCTGGCGAACAGACAACCGGGCAACATACACGGCCCG
2761    CAAACGTGAGGAAGTCGCACGCCGCGTTGCCACCGCCCGACAAGAGGAGTTCTTCGTCAA
2821    ATGATGGACTACCAGAAGAACTTTCTCGACGTCTATGGCAATCCGGTCATCGTTGAGCCT
2881    CACCGATGGACCATTGCTACATCGCGTGGAATCGAGGGGCTGCCTTGGAAGGGAATTGAC
2941    AATGCCGGCGAGGGTTTCTCGTCCCTCTGTCGTGCGTACGTCCAAAACTTCCTGCTCCGC
3001    AAATCCCCCATCGAGGCACTCAACGTGCATCGCGAGCTATGCAGTTTCGCTCGGTTCTAT
3061    AAGACCACACCCGGAGCTTCTATGCTCGCAGCTGTATTGGGATACCTCGCACAGTTGACA
3121    TCGAAGGGCAGCCGCTGGCGATGGTCACGAATTCGTGCCTTCTATCGCTTTTTTGCTGAA
3181    CAGCGGCATCCGGAGTTCGACCGGAGGGCTCTCGTGGATATCAATCGAATAGTAGTTGGT
3241    GGCAATCCCAAGGGAGTAGCCGTCCGAACCGGCGACCCAGTCTCCGGGCCGATGAGTCGC
3301    GAGGAAGCAACCGTGTGGAGAGATGCCATCCTTGATGACGACGACGATAGCTATTCGTCA
3361    ATCAGGGAGCGCGCCATTGCGACCCTAGTCACGGTACTAGGACTACGGCCGACAAGCATC
3421    GTCCATCTGATGGAGTGCGATTTCGTTCAAAACGTGACAATCGATGGCACGCCGGTCGCG
3481    TTTGCTCTTGATATACCGCGCGCCAAGAAAAGGGCTGCGCCTAGGACACATCGCCGACGC
3541    ATTCCCATTCATCCCCAGCTAGCCGCCACGGTTCAGAGGTTGATTGAGCTCAACCGCAAG
3601    CTGCATCCTGAGCCGCAGTCGGCCGACGTGCGACCAATAATCTACGCCACGAAGCCCAAT
```

3661 TCCACCTTCGGCACACTGGGTTGGAAGGAAAGCTCCGACGGGGTGTCTGCCATTCTGAAC
3721 CGGGCGGTCAACCGCTATGGCGTCGTATCGCCCAGAACCCAGCAGGCGCTGATAGTTTTC
3781 CCAACACGTCTCCGGCGCAGCGCGGCGACAGCGCTAGCGAGAGAGGGTTATTCGGCCGAG
3841 CACATTGCATCTTTCCTCGACCAAGAAAGCTTGCAAAACGTCACCGTGTACACCGACGCT
3901 GCTCGGGGCACTATCGAACACTTGGATGTGTCTCTTGCAGATGCCTATGCGCCGTTTATC
3961 GACGCATTCACGGGCGCTATGCCCAAACCTTCATCGAAACGGAGAAGGAGAAAAACGGTC
4021 TCGTACGTGACGGACGACGGCCGTCTTCTTACGCTCGGCGAGTGCGGTCAGGAGCGGGTG
4081 CCTTGCCGGCGTCATGCTCCCTATTCGTGTTACCGCTGCGAGCTTTTCGGGCCGTTCAGT
4141 GATTCAGAGCATGTCGCTCTGCAGAGCGACATGCTCGCTCGCCGCGCGCGCCTCGGAAAA
4201 TCCGGCTCGCTAGCAGAGCGACGAATGGCCGCTCTTATTGACCGCACCATTTTCGGTGTT
4261 TCGCACGTGAAGCTGTTGATTGAGGAAATGAAGAGTGAGCAAGAGTTCAGTTGAGCGGAC
4321 GGACGTGGCTCTAGCAGCCCTCCAACACGTGGCTACCGTGCGACCGGTAGCGAGCGTCGT
4381 ACGAATCGACCCGCGAACAGACAGCTGCATTGTGGCTCGACGTTCGCTGATTGAGTTGTA
4441 TAAGACCTCACAGCCGTTTGGCGAAATCAACTGGCTCTCGAGTTGCTGGGACATAACCGA
4501 CGCTTTGCGAGACAAAATTAAGGGATACACGGCGCGAAAGCGTGCACTCGTTCACTTTAC
4561 CCGTCACCGTGAAGTCGGGCAGCGCATAGGACTGCCATTTCCCGAGGAGGACGACCTCGC
4621 CAGCGTAATCAAAGCATTCGCGTGCCGCCAACATGTCCTCAATGGCGTGACCGCCGCGAG
4681 GCATATGGTCTACATCCGCGCGTGGCGGTACATAGTTGGCGCAATGCGCGGGACTGACAT
4741 CGCACAGGTCACCCCCGCAACATTCAATTCCGCCGCACTAGCCGCAACTAGGGAAAAGTC
4801 ATCGAGCGCATATACGGTCCACGGTGCCCTTGAGGCAATTGCGGATATGCTCGACGAACT
4861 GCGCCTTACCAAAGTACCACTCCGCTGGAGGTGGGTGAAGAAAAAGCGGGGAAAGGGGTA
4921 CGGTGGAAGTAAGCAGATACGACTGGGACAAGCACCTGACACCGACCGGCGCGCATCCGA
4981 TGACGTCGTCTACGCTGTCGCGCAGTTGTATCGCATCGTTCCTAAGACCAAGTGGGCAGA
5041 CCGAATCTGTGTGCTGCTGGCGACGATTCTCGTCTGTACGGGTCTCCGCCTCGGGCAAGT
5101 GCTCTGCTTGAGAGCTGAAATGCCCGCTTTCGATGAGGCAACCGGTGAGCACTTCATTCG
5161 GCTCGTCGCATTTAAGCGAACGGACGCGACGCGAAAAACGCTGCTCTCCGAAACCGTAGG
5221 CCTTCTGAATGATGTCTTCCGGGAGCTGTTGGAAATTACAGAGCCGGTTCGTGAGGTCGC
5281 TCGTTGGCTCGATAAAAACCCGGGGAGAGTCAACCTGCCCAAGCCGGATAACTCCGACGG
5341 AACTGTGACGGCTGCGAGATTGCGAACCTGGTTCGGACTCTCCACTGGCGCCTTTAAGCG
5401 TCGTCTCCAGACATGGGGCTTCGCAGGAGCCGCGATACCGCTGCACGAACTCAACGAACG
5461 CCTACTTCGGGATAGATTCGACAAGCCAGTTGTTCCCGGCACGAGAGGCGAACAACTGCT
5521 GCTCAAAGACTGCCTTTCAATATCATTCGTTTCGGCAATGAAGCAAACGACGGCTGCCCT
5581 GAAACATGCTGTTCGCCCCATCGGCGAGCAACACGTGTCGGACCGGTTACGCGGTCGTCA
5641 GTTGGGAGACGGCCGCTCACGAGCGACCATCTTCGACCGCTATTCTTTGACAGATAGCCA
5701 AGGTAACCCCCTGAGCACATATACCCATGGCTTTCGCCATAAGCTCAACGACGCTCTGGA
5761 TAAGGGCGGTGCACCGGACGTCATTCAGGCACAGTGGTTCGGGCGTGCGAATCCGCGCGA
5821 CAACAAGGCATATCAGTATCGCACTCATGCAGAAATGCGAGAGCGAGCGCGCGACATGCT
5881 CATCAAGGGCCAACTCAATGGGCGATTTGCAACGTTGCTTAGCCAGGTCACGCCGGAGTG
5941 CCGAAAGGAGGCTGCTGAGTCCATGGTGCAAGTCGCGCATCCTATATCCGGGGGCTACTG
6001 CTTGCAGAATTTCGCACAAGTCGACTGCGAGCACTGTGGTCAGTGTCTCGATGAGTGCAT
6061 GTCTTTTCACTGGGTACCGGGCGAAACCGACCGACAAGACGAAGTCATCGCCATCCGAGA
6121 CACCATTTCGCGAAAGCTGAATGTCATGCTCGACAAGATGTCCCCTACCGAAAGACAGAA
6181 CGACGAGTCGTTCTCACGACTACGTCATCAGCAAAACTTCGTGAACCAAATCCTGAAATC
6241 TATGGAGGAGAGCAAGGTTGGTGAACCGAACAAAACCTAAGAAGGGTGCAAGAGCTAGAG
6301 CCGTACCCAAGATGGGGCCTCGCGACACCGAACGCGTAATGAAGGAGCTTCTTGCTTTCG
6361 AGCAAGGAAAGCGAGCTACCATCACGTTGACTTGGAAACATCTCGCTCGGGTAGCCGGCT
6421 TTACCGACCGGTCTTTGCGTTCGAAGCCCAAGCTTCGCGCACGGTTCGATGAGGTAAGAG
6481 ATGCCCTGCTGGGGAAGGGACGTATTCGCCGCTCCACCGGCATGGTTCACGAAGACCTTC
6541 AGAAAGCGAATTCTCGGCTTCGACAACAGCTCAAGATGTACGAGAAACTCGAAAACGAAT
6601 GGCTTGAAAGATGGATTCGCATCGCCGCGTCACTCAAAGCACACGGGATGAGCATCGACC
6661 AGTTCGATGTCGACCTTGAGCTGAGGAGTGAGCGGCGCTCAGGCTGATTCTCGCGCACAT
6721 CAAGACAGGACGCAAAAAGCATGGGGGGGCAAAGCATCACGCAGCGGCAAGCGATGTAAT
6781 TCCATCGATGACCGTTTAGCGGAAAAGAAGCGTTACGGCGGCCGTGAAATTAATTGACGA
6841 ATTCCAACCAAAAAAAAGATGAATTGACGTGATTATCACTTCGCGGCCGTCGGCTACCAC
6901 GACCGTCTTTTTCCATTTCACCTCGACGATTATAGGTGCCCCACACGTGGGCTGGCCTCA
6961 CCAACGAAGACAAATGGCATGGCGAGCAAAAAAGGCACGTCACGCCGAGGCGGTTCTGTC
7021 AGGCATGAGAGGTCAGGTATCCACGCAACGCATCGGACATACCCTTTTCCTGTAGAAATT
7081 GTGCTCGCAGCTCGCTCAAATCCTCATTGCTTTGCAGTCGCTCGTTTTCATAATCTTCGA
7141 CCGCGCGAAAAAGGCGTTCATTGGTGGTTTGATAGAAAATCTGCAAAGCCTCTACCTCGT
7201 ACATTTCCTCGATACCATTTCCCGGAGGGGCGAGCCCGAATAGAAGTCGAAGAAATGGAA
7261 ATGAAACGTCGTCTGGGTCTATGTTTTGGAGTCTATTTTCGTCACAAAGATGCTTAACTT
7321 CATCCGCAAGCGTCTGTTCTATCATCGCGCTGACATCTAGCCGCCCCAGGCCAATTTCGC
7381 TGAAATCCTTCTCCGCGTCCGCAACCGACCTACCGAGCGCAGCTCGACACGCCGCAACTC
7441 TGCGCCAGTCATTTGGAAATACGCTGCCGTACTTAGCCAGCCATAGAGCGCAGGCAGCCG
7501 CCATCTGCGGTGTACTCGCGGAGGTTCCGGCACCGTTCATGTCCCAACCGTACTTCGTGT

7561 TGTAGCACATCCAGGGAACGTTCGGCGTATAGGCTCCCACCGCTTTCTTCATCACCTTAT
7621 CGGGCCCCCAACAGCCTTGCATTACCCCGAGTCTATCTCGCTTGTAAGGGCCATCGTCAA
7681 ATGTTACGCCTGTCGATGTAACAACGCGGTAGAACGCTGACGGATAAACCGTGAAATGGG
7741 TTGCGATGTCTGTAAGCACCGCCCAGTAGCTATCGCCGGCAGCCGCCACCACTACTACTC
7801 CAGCGTCATAAAGCATGTTCACAGCGTGCGCCCATGACTTCATTGGCAAACCACCGTGAC
7861 TCAGACTCACCACATCACAACGTCGCCCATCGCCGGGATTAATCGCATAAGCAAGACCAC
7921 GAGCCATAGTCTCACCGTAAAGATAAACAACGGACCCGTCCACACCAGCGATTCTCACAG
7981 GGACAATTTCCGCGCTTGGCGCGCCGCCAATGAATCCGCGATATACCTGGTCGCCCTCAC
8041 TTTGGCCAGCTTTTCCGCATAGGCTCACGGCATTACCAGCGAGCAGAGCGAGAGTTGCTG
8101 TTCCATGTCCTTTATTCGGCCCAACTCCGGGGTCAACGGTATTGGAATTTCCTTCGACGA
8161 AGTTGTATCCCAGGTCTTTTCTAACTTTTAGAGGTGCAGAATAGTGATTTGGCCACCATC
8221 CGATATCCAGATGTGCGATGCGGATACCTTCTCCCTTTGTAGTCTGCCATGCTCGAGGAA
8281 AACCCGCCTTCTGCAAGTGCCATGCAGGCGAAAACGACGCGCCTTCACTCGGAGGGTAAT
8341 TTGGGTTGAGCGATTCGGTGGGGGGATAGGGGGGCGCTAATGTGACATCTTCCTGAGCGC
8401 TGAGAGGCGTTATTGAGCTTCTTACTCCCGCATCTGAGCTGTCCAATGTTCGCTTGTGAA
8461 TAATATTGGGCTCAACATACACATCACCGACACCGCTCGCACTGTTTACGATAGACAACG
8521 CCTGCTTATGTGCGCGGTCCCATGGATGAATATCCTCTACGGTATCTGGTTGAAATACGA
8581 CCCAATCATTTTGGCCACCAACAAACTGAACGCTCGTATCAGGAAAAAGCGCCTCAACCT
8641 TACCGCCCGTGATACCAATAAGGTCCGCCGCACTTCTTCTCGTGGGGCTTCTGAAATTCA
8701 AAACATGCTGATTAGGCAAACGGATTAACAACTTCATCTCATTCAAGTCAGCCATGCTCT
8761 ACTCCCCGAGGTATAGCTTCAGCCTATCATGTAGGCTGTTGGTCACAGCTTCAAGCGCAA
8821 CAATATATCTAATTCCGAGATGCTAGATATCGGAAGATGCAATCGACCTCAAATTGCCTG
8881 CCAACTCAAGTGCGTTGCACCCGGTCGAGCGAATGCTGTACCACGACGGCGCTCTGTCGT
8941 CAGCTTGTCATTGCCGCAGTGTAAGTACACCTCATTACAGATAGGACATTGTCGAGAACA
9001 TTTCAAGTCGTTGTTGGCGTAGAAGCGGGTGCCGGCTCAGCCTCACCTGGGACCAGCCGT
9061 ACAGTACTGACGCGCACAACATGGCTCCTCGAAGGGCATGATTTGGGCCGCGCCGATTGC
9121 CGGAGCGATGCTCATTGCGGCGTTCGCCAAAGCTCTTTGATGTCGATGCGGATTGAGGGG
9181 CAAACAGTATCCGTGACCGCTGCATCCGGTTCTGCAGGCTGAGGAGGCGAGCGACAAGCG
9241 TTACTGGATAGGGAGCACGCACCCTGCACCTTTTCAATGAATTGTAGCCAACGTGCCTTC
9301 CGCGTGTTTATATCCCTCGATGCGCATTGCCGAATCTATCCGCCTCGGTTCAATGCCCGA
9361 AGAACCGGGTCGCACACGCAGAGAGACTTTCCACTCCGCCAAATGACGAGAATGCGTGCG
9421 CCCCTAAGTCGTCCACCTCGCTGGGCTCATAACTGGTTGCATGTCCCAGACGTGCAGCGA
9481 GCGCGAATACGAGGTACGGGCCTTCCTCGCTACTCGACACCCTGATTCGGCTGGTCGTAG
9541 CGCTTCTAGTATACGGATATGGAGGCGTGCCTTGTCCTCTCAGTTCTACTTCCGTCACGG
9601 GTCTGGGCTTGCCCAGAGCGGGGTCATGAATGAGTTGCGCTGCGGCAACCGAAATTTCAG
9661 CGGGAGATTCGTTTTCATCGGCCACGAGAGCGTAGACCTTATGGTTGTCTCTCAAGCGAA
9721 CAATCGCGCCGGCATTCAAATATCCGGTCGGCCCGTTCTTGCCTGTAAGGCTTGACGGCA
9781 TCACCAAAAAGGGGAGTCTGTCAGAATCTGCGTATTGAACCCCGTGCCCATCGGGCATGC
9841 TCACTGGGTGGGCGGTCACCGAGTTGAAGAAGCCCGGAGCCGGGTCACCCTCTTTCTGCA
9901 CGGGTAGCGGACCGATACCGAGGCCAGTGCCCCAGACAGGCTTGTCTCCAACCTTCCTCG
9961 CGCTGTTGGCGGCAAAGCCATAGACGCAGAACGCCGGGGACTTCACAGGGTCCCACTGGA
10021 CAGCCTGCGCCTGGCGAACTGCAGCCGTGCAGGCGGTGTCGTCGCGAACGCATCTGCCGC
10081 CCCGGTAGATGTCGAGCCCATCACAAAGGCTGGAACTACCGATATTCTCTGGGTGATAGG
10141 CAGTGGATGCACCGTCAGTGTCGGTGCTCAGGCTGCCGGCCCATACGACGAAGCCACCTG
10201 GGAAGGTAGCAAACCTGTTTCCCTTCGGCTGTTTCCCTGGAAGATGAATCTCGGTAAAAG
10261 TGGGACAGGCCCTTTCAGCAAGGGCTGTCCCGGAACCCGACATGAAGCCTCCTACGAGGA
10321 GCACAGCAATCGCAGAACGTACGTGGAAAGTCACAATGCCTCCGACGCAAGTATAGAGAG
10381 TAAAAGTGTCGACGCGCTGAGTGGCTAGCGACGCATTCGCCGCTCAACTAACATAGTCCA
10441 CCATAACCTCCAGAAAACCAACTTGCGACTTAGGCAGCGGCGCAAGATGTACCGGAAACT
10501 CGAAAGCGAACGGCTCGCGTGACGGTTTGCATCGCGTCGTCGCTCAAAACCTATGCGATG
10561 AGCATTGCCAATTCGCTGTCCCGCCGGTATTCAGGCGCGAGCGACTTCCATATGGGGCAG
10621 CGGCTACGGGTGCGGACATGCGATATATTGATGCAGTCATCATCGACGGAAAAAACTATG
10681 TCCCGCCTCGGTATCGAACACGACCACGGCCTCGTCTATGAGGGCACCGACCATCCAAGC
10741 CACCTCACTGTCCCTGCGCCAATAATTTCACAGTGCGTACTCATTGAGTCCCCGTCTAAC
10801 CTCGACGATTTGCCTAGCGGGATGCTGAGTGACCCGTTTCGATGGATTTTTCGCGAGGAC
10861 TCGTTCGACCCGGTGTCGCGCGTTCGGCGAGGTCGCCTTTTCCAGCCGTTTGAAAGAGCC
10921 AATCGCGTATCGTATTTCGTTGAAGCACATCCCAACTTGCTATCGGATAGTCGCCGGAGG
10981 ACCGAAGACGGTCGGGTCCAGAAGGAATTGAATGTGTTCATTCACTGCACCGAACTGTTG
11041 CGACGTGCCAATCGAGGGGAGGGGCTCCAACTGGCGATAGGCAACGCGCAGGCGCACTCA
11101 CTATGGCGAATCCTGCAAACAGAGCAGACGGTGAGTCAGGATGTTTTGGTGACGCTTCGA
11161 GCAGAGTCTGCGTTTGGTGTCCTTCCCGCTCTCAACATTGACCAAGCTCCGGACGAAACG
11221 CGCAAGTCGGTGTCGGATGCCTACGAGCGCGTCATGCAAGTGGCATATAGGGACTCGCCA
11281 ACATCGGTCGTCGACCAATGCCGTAACCTGTGTGCCGTCGTCAGCGCGCGGTGGCTGTAC
11341 AAGTTGACAGGAAACAGAGCAATGCTCGAAGAAGACCTTGGCGATTGCATTACAGCGGTC
11401 AAGAAGCACTTCGGCGAGAAGGAGCAACGGCTCATTCGCGCGTCCCTCGAAACGGTTAAT

11461 CTCCTCCATCCGAGGGGAAAAGATAACGAACGCGAACGCTACAAACTTCGCGAGGTATCA
11521 AAGGAGGACGCCGAACTTGCCTTGCACGCAACGGGCTTCTTGCTGAGAGAGCTCGGATGG
11581 GGATGCTGAGACTCTCCCATTTTGGTGATTGGCGGTACATCGGTCGTAGACTTTGCCGGC
11641 GCCTTGCAGATGGTAGCGCTGATGTACCGTTGTCTATGGCGAAGGCAACGGCCGAATTTC
11701 CGAGCGAGCTAGTGCAGCACTAGCATCACTAGGGCGAGGAGGAACAGCCCCCCTGCCACA
11761 TAGACCCAGCCCGGGACGGCCTGTCTACCCGACTTGGAGTTGAAAGGCGCTTGCTGATAC
11821 GTCGGTGTTGTGGTTGCCGCAGGAGACGCTGCTGCCGGTGGTTCGCGATGCCCCCGTTCG
11881 ACCGGCTGCGGAATTGGGCGCGGCGCCGACATCGGTTGTGGAACAACGGGTCGTGGCGCT
11941 GGGGTCGACGACGGTGGATTCTGCGTTATTTTGGTCGGTGGCGACGGACGTGGATTGCGC
12001 GCTGCCATGGCTACCGTAGGCGCCTGCTGTTGTGCCTTCATCTTCATTTGCGCTAAACGC
12061 TGCGCGTTCGCAGCTGTCGCACCGGTCGAGATGTTAGTCTGTACTGGTGATGGTCGTTGA
12121 GCCTTCGGCTCTTGCGGCGCGGGCTGGTTGGCGCTCATCCCAGCAAGCAATTTCCGATTG
12181 TCAGTCGAGGAGACACGAGGGCTCGGACTATCCTTCATTGCATCCGGACGCTCGGAATCC
12241 ATGCGATACCGTGCCTCATGCTCGTCGATGGTGTGAATCTGTCCCGGAGCAGAAACGACG
12301 AGCTGGACGTATACGCGTCCGTCTCGCTCTCTCCTCCTTCGAATTTGGCCCTTCACGCTC
12361 AACCACCGCCCCTTCCATGAGCTTTGCGGCTTGATTTTCATCTTGCGTAGGACGTCATCC
12421 CAAATCGTAAGCCGCACGCTGTTTCCTGGTCGCCAATCGCCGAAATTGATGAAGGCATAC
12481 GGACGGTTGTTACCATCTTTCGTCTTCTTGTTCGCAACGTCCACGACCATACCCACCAAT
12541 TCGACCACCTCGTCCAGTTCATCGAGAGCCGCAGCAAAATTCGACGCATCGACGACGGGA
12601 TGGGAGGGGAGATAGCGTCCGCGCGCGAATTTGGCGCCCAATTCCGGAGCAAGATATCTC
12661 GGTCCCGTAGGGATACCCTGACCCAAAATGAAGTCAGCCAGGGAAGGCACGCGGTCGGCG
12721 CCAGCAAGGCATATGGTCGCGAAGCTCGCGACGTCCCGTGCGAGCTCACCAATCCGCGAC
12781 AGCGACGCGAACACACGAGAACTCTGCGGCGAGGCAAAATCATTTGCCCGGAATACAAAT
12841 GCGTCTGCGTCGCTGTTTGTATGGCCCCATAGCTCGGCTGCAGCGCTAATCGCTCGGAGT
12901 GCGATGTCGAGCGAAATAAACGCAAATCTATCCAGCTTCGCGTTGAACAGCGTATGGTCT
12961 CGTCCCGGGTGCTGGAAGTTTCGGTGACCTATCTCAGCGCTATGCATGCCCGCAAGCGCC
13021 GGAACAAACATGCCGTCGTAATCGATAAGCCTTACTGACGCACCGTCCGGTGAAACCATC
13081 AAGTTGCCGGGTTGGATGTCGCCGTGGGCTATACCTTGACTCTCGAGATATTGTGCGAGC
13141 CGACGAAGCGCTTCGATTAGCCTAGCCAGCCTCGTCCGGTCGTCGTAGTACTGTTCAACG
13201 AATTCGCCTAGAGTGACGCCGCTCGCCCACTCCATCTTCACGATAGGGTAGGCAGTGCCA
13261 TCCACCCATATCCCGTTCCTCGCGAAGTTGAACCGCACGAAGTACGGCGACCTCAACGTC
13321 TCAAGATGGCGCGCGACGGCGGTGTAACGGTCTTCGAGGTCCTTCGATTCTTTGTGGAAG
13381 CACCGAAGAGCGACTTTCCGATTTTGCGCCAGTGTGAACGTATAGGTGAGTGCAAAGCCT
13441 CCGCACATGACGAGCGGAAGTCCGAAGCCATCTCTTGCAACGGAGCCTTGTGCAAGCTCA
13501 GCGTCCCCACGAAATGCCGTCTTGGGAAATTGCAGTACCTCTTGGTACCGTTCCAGAGCC
13561 GGGTACGTCATTTCGCCTGGGGTCGGCTACGCACTGCAAGGACGAGTAGTGAGGTATCGT
13621 CGATACGCAGGACTCCTTCGCGTCTTTCGGACAGAACAAGCTCGGTGAACGCAGCATCCG
13681 TATTCACCGACAGCAGCCTTTCCAGGCGCGAGCCTGCGGGATTTTCCGACAATAGCCACT
13741 GGCCCAATGCGTCCGTCATGAGAAGGACAACGCTGCCTTCGCTGAGGTGCCAGACCGTAC
13801 CGGAATCGTTGTTAAACGACGCCGCACTCACGAACTCATTCGCTTCGGCGCGCGTAGAGA
13861 GCAGCCGTGGACGCTGGTCGAACTCGGCGGCGGTCGTGTACGGGTGCGACTGAACGACTT
13921 TTGCGGCCTCCACGTGTACCGCGACGCTATCGCCCACGGAGACAATCTCGAGGTTCCCTA
13981 AATGTTCTTGCCATTGGCAGCCCAGTAGGGTTGAGAAACTGCCACGGTCGAAAGCCATCT
14041 GTTCGGACCAGCCGAGACTGTTGTAGTCCCAGGTGCGGTCATAGCTTTGCCTTGCGGCGT
14101 TGACCCATTCAGGGCTCACGGCGGGAAGCTCGACGAATCGAGAAGCGAGCAGGTGTGCCC
14161 AGGCACGTGAATCGAACGATTCCGAAGCGCCGTCTGCAACGGCGAAGACGCCACGACCCT
14221 CGTCGATGCAAACTGCGTCCTCGTTCCGCTCAGGCTCGTCGGTGACCTTTGCCACAGTCG
14281 ACTGAAACGCTAGAACCAGCCGCATAAATTATCGGAGCTGCGACGGGCGGGTGCCGATGT
14341 CGAAGAAGTTCACGATTTGTTCAGCCTCGGCGTTGAACATGAAACCTCGGCTCTCGAACG
14401 AAGCCGCAATGCCCTTCTCCTGGGCAACGCGCACCAGATGCTCTGGAAGCTGGCTCGACA
14461 TACGGAAAAGCAGCTTGGCGTACTGGTCTGCAAGGGCGTTCTCTGATGCCGGAAATTGCA
14521 CTGGTTGTCCCGCTTGAGAGCTTACGTGCAGGTTCAGCAATAGTACGTTGCCATCGTCCG
14581 TCTGAAACTGGCGCAGCGTATCGGCAAGCTCTTCCGGGTCCCCGTCAGTTGAATCACCAT
14641 CGGTAACGTGTAGAACCGTTGGCGGATAGCTCCTGGGATGCGCGTCACACCATCCAGCGA
14701 GCGCTTCGGCCGCCGCCATGATTGCCGCTCTCATAGGCGTACCACCATCGGCCTTGGGCT
14761 CGAACCACACGGGAAACTTCGTCAACTGCTCGACGAGTCCGCCGGCACCATCATCCACCT
14821 TTCGTTTTCGTTCTTCGATACGGAGCGGATTGGCCTCGACCACCGAAATTGGCTGCAAGG
14881 TCGTCTGCGAGAGTACACCGTGAAAGCCGTTCGTTACCGAGTCCCCACCATACCCGAGAA
14941 CCCCAACATTGAAATAATCTCGAACACCGTCTGACTTCCTGCAGCGAACGATAAGGTTGG
15001 CCAGAGTCCGATTCAAGACGTCGGCAACGAACTGCGCTTTTGTGCGTTCGCTTCCCGCCA
15061 TCCTGTCCTCCATCGACCGTGACTGGTCAACAAGAAACAGCAATGCGGTAGGGTCACTGC
15121 GCTGAATGGCAGCCTGGTAAGCCATATTGCTCTCCGTGTGGGGGTACCGTTGCCTAATCT
15181 TATTTAACCGCAAACATACCAGCAGCCATCATGATATGCCAGCGACGGCCGCTATTGCTG
15241 GAGGTTGCCTGGCACTGCAGTACAGAAACTACAGCGCATTTCTCTCGGTCCTTGCGTAGA
15301 TATGCCATATGCATGGTGATGGCGATTGAAAACGATTTTCCTTCGAGTTCACCTCAAAGG

15361 GCATACGGTCGCGGAACGGAAGTCGTGTGTCGGCTCCCCGGACGAACCGCTCAAGCTGCG
15421 GCCATCGAAAGCAAAATGTCAGGTTTGTGCAGGGCTGATACCGTTTCGGCTTCGCTCCCC
15481 GGCGCTACTATCCCGCCAACGAGTTTTAGGTGGAGTCGTGCCCGGGCAGGGAAACCGTCG
15541 TTCCATGCCAATGAAAATGCTTGTTTGAAATCTGCTCGAAAGCCTCTATGCCCGATATTT
15601 TCGACTCGTAATCGTTGATGCCACGCCCGGCCACCAATACAATCGCGCTAACAACTCAAA
15661 AATTTTCGCACTTCGGGGGCACCATGAACTTGTGGACGCACATCGCCACGTTTCTGGCCG
15721 GCTTAGCGACTGGCTGGACGCTAAAGGTCGTGATTTCCAATCGTTCGAACCGTTCGCACC
15781 GCGCTACCTTTGTTTCACAGAAGGGCAACATGGCAGGCGGCGATATCGTCGGCGGCGACA
15841 AGAGCACCCGCGAGTAGTATTCGTGTTCACGAACCAAAGCAATAACGAGGTCAATGGCGG
15901 AGACCTCGTTGCCGGCGACAAGGTCGATAACAGCACGACTACAGTCATCCACGTCACGTC
15961 GAATGGCTCGGCCAGCGAAGAGTTGTCAAAGCTCTACGAGAAGTTAAAGACTGACGGCAT
16021 CGGTGACCCGTCGGGCGGAAAATTTTCCGAGAAGCTCGAGCACTACATGTCGTCGGCGAC
16081 TGATGGGGACGTGCGAGGGCTGGAGAAAAAGTTGTGCGAAAGCGGAAGACAAGACATCTT
16141 GCATGCCGCTCTCAGCCTCAAAGAGCGCGCCGCCAAGTCCATCATGAAGCATCAGACGTC
16201 GAGAACGGCGCAGCGAATCTTCACCATTGTCCTCACCGAGCTGCACGCGCGATTCACTCT
16261 CACTGTCACGCCGGTTATCCAGCAGGGTCTCGACCGCGTCGCAGTCGACCAGTGCATCAA
16321 TACGATTATTCAAGGGACTCAAAACGTGTTGGGCGAGAACGTTCTTGAGTTGGATGCCAT
16381 TGACCTGCTCGCGTTGCTCTATTACCTGGGCGGCAACTGCCATATCCGGTGGGATAAATG
16441 CTGATATACCACCCTGCCTATGACGCCTATCACTGTGTGTTCCGGATGCTGCTCATTGCC
16501 GAGCGAGTCTCTGGCATCGAGGTGGAGAAAGCGCGAATCTTGGACTTCTATCTGGTTTTC
16561 CCGGGCATGGTGAAGAAGATACGTTTGCCGGACACACTCAAAGTACTCCGGGCCCATGCA
16621 AAAGGCGCTGGTAACGTGTACCGCGACCCTGTCAGCAGTACATCCACATTCATGGAGATG
16681 CGGCATATACAGGAAGCTGCACTCAAGTCAATCGCCGCGTCTGGCTTGATTGACCTCGAG
16741 AGATTCTCGGCCGGATATGTTGTGCGCTCTCAAGCGCCGGTTCCGCCTTCGCTCGCGGCA
16801 AAGCTGGACGAATTTGTGGCCGCAAACTCGCCTATGGCCGAGAGTATCCTTAACGGCCTC
16861 TCACAGATTCCACTGCTTGGGAGCAACGGGCTGAAGCATCGTTCCGAACTCATGGAGTAC
16921 CGGTATGACTATGCTTGACCCGACGCTGCGCGTAAGCCGACTCGTCGTTTTCCAGGGCGG
16981 TCACGTCGCATTCGACTGCTCGTTCCACAGCGGAGTCAATGTCATTCGCGGCCGGAATAG
17041 CTCTGGCAAGACAACCATCATGGACCTTCTCGCGTTTTCGCTGGGCGCAGAAAATATCCG
17101 ATGGAAGCCAGAGGCACTGCGTTGCACCAATACGTTCGTTGAGGTACAGCTTAACGGCGC
17161 CGTCGCGTGTCTTCGTCGTGAGATTGATACGGCACTTCAACGCCCAATGTCGATATTTTG
17221 GGGCGCACTCGACGCATCACTGAAGGCCGGCCCGCAACAGTGGGAGCTGTATCCCTTCAA
17281 GCGAAGCGCGCATCGCATCAGCTTCTCTCAAGCATTGTTCGATGCACTTGACATGCCGCA
17341 AGCGCAAGGTGACGGAGCATCGAACCTAACGATGCATCAGCTTTTGCGCGTCCTCTATGC
17401 AGACCAACCATCGGTACACAGCCCCATCTTCCGCCTCGACAACTTTGATAGTCCTCTGAC
17461 GCGCGACATGGTCGGCGGGTACCTCAGCGGCGTATATGACGACCGGCTTTATTCAGCGCA
17521 ACTGCGACTTCGCGAACTTGACGGGCAGATTGCGAAGCAGGAGACGGAGCTCCGAGGCAT
17581 CTTTCACGTACTTGGACGTTCTGGTCAAACCCCGGACCTTCAGTTTCTGACCGATAAAAT
17641 TTCCGAACTGGAAGCGGCTCGCAGAGCGCTTTCCGATACGCTGACAAAACTGAAGCAGGA
17701 CCGAACGCTGCCTCGCAAAGATGCGAACAAGGCGCGCGCCTCCGCCGACACTTTGCGTAG
17761 CACTCTGAACGCTGCGCGCCAGGCCGAATTCGCAGCCAAGGACCGTCTGGCGGCGCTCGA
17821 ACTCGACATCGAAGATTCTCGCCTGTTCGTCGGCGAACTCAAGTCCAGACTGCAGAGCCT
17881 GGGCGAATCGAAAGAAACCCGGACTTATTTTAGTAATTTGCAGTTCCAGTTTTGCCCGAG
17941 CTGCCTCGCCGAGCTTCCAAAAGCAACGGCGGACTCCCACGTTTGTCATCTTTGTACCGG
18001 CTCGCTCGGTGACGGCCGGGCAGACACGCAGCTTCTCCGGATGAAAAACGAACTGAACAT
18061 CCAACTTAAAGAATCGACGACTCTCATCGAAACTCGTCTGAAGGATGCCGAAAAGCTTCG
18121 TGTCGAAGTCCCCCACCTTTCACAACGAGTTAGAAAGCTGGAGCAGGACTATGCTGCGGC
18181 AGCTACGTCGTGGTCTAGCGAACTCGAGATTGCCCTTGAGGAACACGCACGCAAGCTCGG
18241 TGCAGTAGACGAAGAAATACGTCAAGCGTACGAGCAACAAAAGCTCGCTACCGTCATCAC
18301 GGACCTGCAAAAACGCCGTGACGCGCTAACCGCCGAAGGCAAGCGGCTGCAAGAGAGCAT
18361 CATACGTTTGGAGCAACAGCAAGCCGAGCGGAAGGTCGAAGTCGCGAATACGATTACCGC
18421 TGCAATGGTTAGGCTTCTCAAGCTCGACCTGCCCCTGACGCCTGAATTCATCGATGCGCA
18481 GGTAGCGAACTTCGACTTCGTGGCAAACGAAGTCTACGTTAACGGCTCCAAAAATTTTTC
18541 TGAGAGCTCGGCTGTAGTGTTGCGCCATATCTTCCACCTCTCACTGCTTACGGCAAGCGT
18601 GCAGAAGCCGTTCATGCGCGTACCGCGCTTCATGATGCTTGACGGAATCGACGACGGTGG
18661 CATGGAAAAGGAACGCAGCCACCGCCTGCAAGAAATCATCGTTGAGGAATGCGCGAGCTA
18721 CGATGTCGACTATCAGCTCATCTTCGCGACTTCCGACATAAACCCGAAGTTCGAAGAGAC
18781 TGACTTGGTGGTGGGCAGATTCTTTACACCCCAGGCTCGCTCGCTGGACGTGCGAGAAGC
18841 CTAGCCTGGCAAATACCATCGCGATTGAAGTCGGCATCTTTAGCGCACAGCAAAACCGAG
18901 AGGCATCGTGTCAGTCATCAAGCAGCAGTATAAAAGCCTAGCCCCAAAGGGCGAGTGTCT
18961 GTCGGACATCGTTGTTCTCGCGCAGGCCTGGAAGAAATCGCATACCTTCATTCGAAGACA
19021 CAATTGGTACGCCGATGTTCTCGAGCTCGATTGTTCGACTGTCGACCTCGAGGAGCGCCT
19081 CAATCAATGGGCAGACGAAGTCGCGCAGCCGGAGTTTTCGACCGCTGAGTTGTTGCTCGT
19141 GCCTGCCCCGAAGAACGCGCGATGGGAGTTCAAGGCGGCGCCAGACCTACGAGCCGTTCC
19201 CGACATTCTCGACGTCAACTTGGACGATTTCGGTTCCGAACCATCGTTTGGGGACTGGGT

19261 GCCAATCGAAAATGGCGATGCTCCATCTCAGAATTCATCGGACGCCGGTCAAAAGCTACG
19321 CCCTCTAGCTCATGTCGCGATACGAGACCAGACGCTCGCGACTGCCGTGATGATGTGTCT
19381 CGCTGAGGCCATCGAATCCCACCAAGGGGACACCACCGAAGGTGATTTGACGAAGCTACG
19441 CAAATTGGGGGTCGTTAGCTATGGCAACCGCCTGCAGTGTCGATGGGTTGAGAAGGGCGA
19501 TGGCACAAAACGCGCGCACTTTTCGTGGGGTAACAGCCGTACCTACAGACAATATTTCGA
19561 GGACTACCGCGCGTTTCTTTCGCGACCGCGAAGGGTCTGCGCGGAGCTTTCACCGCAAAC
19621 CGCGACGAGAAAGGAACTATTCGTCGTGTCACTCGACATCAAGTCGTTCTTTGACTGTGT
19681 CGATAGCAAGGCATTGGTACATGAACTTTACCGACTTCAGACAACCTATCAGAAACATGA
19741 GGGCCTGTCCGAGAACGACTCCACAGACGTGCCGTTCTGGGAATTGGCATCGCTCATTCT
19801 TAGTTGGCGGTGGCGGGAAAGCGACCATCAATCTGCACCGCTAATTCGCGGCCAGGACTT
19861 TGAGCACCTTGCAGTAGGTCTCCCGCAAGGACTCGTCGCATCTGGCTTCTTCGCCAACGC
19921 TTACCTCGTCGGCTTCGACACCGAGATGTGGAGAGTATGTGAAGAGCAGCAGCAACTGCC
19981 AGAAGATATCCGCATACTCGACTATTGCCGCTATGTCGATGACCTTCGTATCGTGGTCGA
20041 GGCGCCGCTTGCAATGGGTGCCGCTGGGAGCGAGAGGACGCTGCAACGAGTCCAACAGTT
20101 TGTGTCGGAGATGTTGGCGAATCACTGCCTGCGTCTACAGGCAAACGTGTCACTGACGCT
20161 TTCCCTGCAAAAAAGCAGCATCACACCATACCGGTCGATTTCGGCGCAGAGCAATGTCTC
20221 ATCGCTGATGGAAGCATTGAACGCAGACCTAAGTGGAACGTTCGACCTTGAATCTTTGGT
20281 ACAGACAGCCGGAGGCCTCGAAGGCCTTTTGTGGATGTCTGAGCAAATAGAAGACGCTCC
20341 CCCATCAAGCCGCTCACGACTGGCGCTCGCCAATGTAGCCGTTCCCAACACCGATGTACG
20401 GGACGATACTGTCAAGCGATTCGTAGCAAGCAGGCTCGCTCAGGTAATGCGTCACAGGTT
20461 GGCGATGACGGATACCGAATCTCCGGCGGACACTGACGTGGCTGTCTCCGACGCCGTGAG
20521 CGCCGGAGCAGCGCTCTCTCACGAATTCGAACAAACCGCTCGAAAGCTCATAAAGTGCTG
20581 GGCCGAGAATCCTTCACTAGCACTGCTGCTGAGGTGCGGGCTAGACCTCTTTCCTCACCC
20641 GAGGCTGCTTTCCCCAGTTACCGAGGCCCTAATATCCAAACTACCAGGAGCCCAAAACGC
20701 CTTCGACAGCGAGCAGATAAAACAGGTGCGCACGGCCGAGTACGTCGCTGCTGATATTTT
20761 GCGCGCGGGTGCCACAGAGACTGGGTATCGCGATAGGGACGAGTACCCGGAGTTCGTCGA
20821 CATCGAAGGTTACCGGGAGAGCCTCGGTGCATTTGCGCGATGGATACTCCTTGAGTACAA
20881 GCATGCGCCCTGGTATCTGCTTCAACAGGTCCGTCTCTACCTTGCGTCCATTGGAGATTT
20941 CTCAAACATTCGGTCTGTGAACGGAAGCATGGAGGAATCGTACGCCTTACTCCATCGCGC
21001 CTCGCTATATCAGCCAGTACGCTCGGGCGAGTTGTTGCGCCTTCTGCCGTTTGCGCTCGT
21061 CGCCCAGCAGCTCAATCCCAATCCGCGCAGATTCGGCGTTTGGCTCGCCGAATGCCTTCG
21121 TCGGACGCGGAGCGAGGGCGAACAGGCCAAAGCCGTTACAACCGTCGCACTCAACCGACC
21181 GGATTTGATGCTCGAAGCTTTGAGGTCGAAGACTATTAAAGCCAACGGGTGGGTAAAGCA
21241 TGTGCCAACGCCGCTCGCCGAAGTCGACCGGAAGCTGCTGAAGAGGGGAAAGACGACCGA
21301 CGGCACGAGAACCCTGATTCGTTGCATGCTCGACAACGACAACCTGTTCGCACAAGAGAA
21361 CGGCTTGCTGGTCCTTGCCCAAGCGCTTCTGCGCAGCCCGGACATCCAAAAGCGCCTATC
21421 TGACGGCTTGGACGCCAGTCAAATTACTCTCCGGTGCGAGGAATGGGAGCAGATTCATTC
21481 CATTCCAACGGCCACGGACTTCATGAAGGTCGAACTTGCGCCCCCTCCAGGCACGCAGAA
21541 TCCATTGTATGAAACGCCACTTTGGGTGGCAGAGGACAAGGCTTGGCTCTACGGCCTCGG
21601 ACGTATACTTCGCGCCGCGTTGACTGGCGAGTTTGATTTTACTAGCCGGCGCTATCTTGT
21661 GACGGAAGAGGTCGGTCGATACTTCGGTCTTCGAAGTACTTGGTATAAGCGTAGGTTCGG
21721 TATGCTGAATTCTGGCACTGGCTTGCTTCACGAACCAACGCCCATCTCCCCCTGGCTTTC
21781 TGGACTGTTGTCCACGTTGCTTCAGTGGCCCGGGGTCGAGTTCAGGGCGAACGACGCCGC
21841 TGCAGCGGGTGACGCTAACACCCCGGAAGAACTGCTGGCCCTAATCGAGAAACGAGCGGC
21901 AGAACAGCGCGCATTGTACGGTTCGCGCAGCCGCACGCCAATGTACGTCATACCGGCATC
21961 AGGGGATGATGCACCATTGAATGAAAGACCATTGCGCGTTGCGATTGTGCAACCGCTTCG
22021 ACCAAGGCGTGACGATTTCGATGCAAAGGACCCTACGCATTGGACAAAGGGAATGCTTGC
22081 CGAGCATCGAAGGCACCTTGCTGCCGTATGCCGACTCGCTGAACAGCAGCTACGAACGTG
22141 GGCGTCGGCCCACAGCAAAGTCGACACGGACAAGGGGCCGCTGGTCGACATAGTGTTGTT
22201 CCCCGAACTCTCAGTGCATCCAGAGCACGTCGGCTTGCTTCGTCGGCTGTCGGACAAACT
22261 GAAAGCGAACGTTTTTTCCGGTCTAACCTTCGTTCATTCTGACAAGGTCGGAGGTCCGAT
22321 AAATCAGGGACTTTGGTTGATTCGCAGCGAAACTGCCGAACATGGTCGAAGCATCCAGTA
22381 CGTTTGGCAAGGCAAGCTTCATCCAATGAAGCTTGAACGAGCCATGGGCATTAAGCCGTA
22441 CCGGCCACACGTGACACTTGTTGAATTGCCAATCGGCACGAAGTCGCCGACCCGGATAGC
22501 AGCGGCGATTTGCTACGATGCGACGGACCTCGATTTGGTTGCAGACCTCCGAACCAAGTC
22561 CGATGTCTTTCTCGTTGCAGCACTCAACCAGGACGTTCAGACATTCGACAATATGGTCGC
22621 TGCTCTGCATTTCCACATGTACCAGCCCGTCATATTGGCGAACTCCGGTGAGTTCGGCGG
22681 CTCGACCGCTCAAGTCCCGTTGCCGAAGCACGAACGCCTCGTGGCGCATGTACATGGCAG
22741 CAATCAAATAGCGGTCAGCGTGTTCGAAGTCGACCCCGCACCGTTCAAGTCCACTGCAAC
22801 ACCGAAAGCCACGAAGGAACTGAAGGCACCTCCCGCCGGCTACAAGGGGCGACCCTGTTA
22861 GAGGCAACACATGAGTGCAAACCAGACTATCGTAATCGATTGCCCGCAATGCGCCGTTCG
22921 CGTTAAGTCCGAAGCCTCCTGCTGGGTCGGTGATGCGCAAGAACAAGCATATTTCCTAGC
22981 GGAATGTCCGTCCTGTCGCCAGGCATTGTTCGGCTCATCTTCCCTCATTCGCGACGAGCG
23041 GGGTGACTGGGAGTGGGACACAGCGGAACGGCTTTGGCCTGCGCCAATGATAGCGGACTT
23101 GGGTCCGTCGGTCCCGCCTGCGGCGCGGCGAGACGTGAGCGCCGCGCAAAAGTGCTTTAG

23161 TCACGGCATCTATCCCGCCGCTGCGGTTATGTGCGGCCGTGCTTTAGAGCGACTCGTCGC
23221 GGAAAAGACGGGGAAACCACAGTCCCTTGCTAGGGGCCTTGCCGAATTGAGAGCGCAAGG
23281 TGTCATCGACGAACGGCTTCACGCGTGGGCAGATGCCTTACGTGTCGAGCGCAATATCGG
23341 CGCTCATGCGTCCAACACGGAAACGACGAAAGACGATGCCGAAGACATCATCGACTTTAC
23401 AGTCGCGATATTCGACTACGTCTACACTCTTGCGGAGCGATATCAAAAGTATCTCAAACG
23461 AAAAGAGGCTGGGGGATAGCCGCCCTCGGTACGAGGAAGACGCCATGCGAGCACCAGATT
23521 GATAGGCTCCATTTCGCAAAGGTCGGCGAGTCGTAGGTTGAAGCTCTCCGAGCGCCCTAA
23581 GTAATTTCTGATTTGATAGAGTGATGAAGATAAAGATTCATTGGACCACCGCATTGACCC
23641 TTCGAGCTGACCCACAGCTCATATATAGCCCGTCTGTACTTGAAAGCGTGGACGACGTGC
23701 CGGGGGTCTATGTATTCGCGCGGAAGTACGGCAGTCGCGTTACCCCGCTGTATGTGGGGA
23761 AAGCTCTAAACCTAAAAGGGAGAATCAAGCAACAGTTAAACAACTTGCGACTGATGCGCC
23821 AAATCCAGGACTGGGAAAAAAACGGCGCTCGGGTCCTTTTGCTTGGGTATCTGAAGGTTC
23881 ACGGTTCGCAAGACGTAGGAACGGCGTTGGACGTCATCGAGCGCGCACTGATTGAACATG
23941 GGCTTGCCGAGGGACATCCACTCGTCAACGTACAAGGTACTCGTACTCCGTTCCACACAC
24001 TAAGCTTCACGGGCAATAGAATGTCCGAACAGGTTGCGCCACGTCAGATGTTCGTTAAAG
24061 CATGAAAGCGAGCTCCCCACGAGCCTCAAACTCGTGGGGCCTGTCGGGGATGGCCTTGGC
24121 CTGAGATGCGTCCGCTTCCGTGCCTTGATTTTTCGCTCAAACCAGCGGTCGATGTGGTCA
24181 GCTTGCACCGCTCGAACGGTTTCGCCGCCTGGCTGGCGGGCATCTTTCAAGCAGTTTACT
24241 AGTCCAGACCTTCATCTGCTCAAGCAGGAACAGTCGGTCGTCTCCGGCGAGCGGCTCTAG
24301 AAGACGCACGATATCGTGCGCGACCTGCTGCGCGTCGCCCGAACTCCCTTCCAACAGGTC
24361 CATCCACGAGCACTAGAGTACCGCACAGAGCCGCTCGAGTGTCGCAACACCCGGCATCAC
24421 CGCCCCTCGCTCGAACCTAGAGACCGTCTCGGCATCGACGTCGATTTGCTCTGCGACGGC
24481 GGCTTGCGTCAGCCCGCGCTGCTTCCGAACCGATGCGATGTTCCGGCCAATGACCGACGC
24541 GAGGCTTGTAGGCATTACTCAACTCAACCTGTGGAGAGACACATAAAAATGTTGAGTTGC
24601 CGGCGGCGGTTTGGTAAGGTTCCAATCTGTAGACTAACAACAGATTAGTCCCATTCATCC
24661 ACGAAAGCATGCTGTCGTCCCACGAACGACGCACCGGTGATATGCCACCGCGAGCCCGTC
24721 TCGGGCACGTGCTGAAAACAAAGGAAGCTGGCTCCTGCCAGCGGACCGGCGATATTTCGC
24781 CCATGCATCGGCGTCCGTGACCGAGCTCCGGACCGCGAGCGGCCGGGCGTCGGTCGAAAT
24841 TTGAAAGACGGGGAAAACGTGAAGAAATCGGGGATTGTCGCGGCTTCTGTGCTCGCCGCA
24901 GCTTTGCTAGCCATCGCCTCGCCGGCGACATACGCATGGACCGACATTGTCTCGGGCCCG
24961 TACGTCGAGACCGCGCCACATTATCTTTCGCACACGTTCTATGTGTCCGGCGTTTGGCCA
25021 ACCAAGGGGAAAGCATCGGCCCAAGCACATATGGAGTGCGTACAGGACAACAGCTTCGAG
25081 GACCGCTTTCGCGGTAGCTGCCAATCAATCGCAACGTTCGACGGCCCGGCTGTCATTGCT
25141 CTTGTGTCCGCTCAGAATCACTACTGGTCAGCATATGCAGACCCTGACACGAACACCGCA
25201 GTGCGAAAGGCAATGGCACTTTGTCATAAGGACAAAGCGGAGGAGTGCCAGCTGACCAAT
25261 ATCTACTGGGATTTCGGGCACAACTTCCTCGATGCCCCGGAGAGCTTCAAGCAGCTCTAC
25321 ATGAAACTTCACGACGGCGAATACTGGCGGCCGCTGAAGGGCTACTGGGTTCCCTGAGTC
25381 GCTTGCCTGAAACCGCCCGAGGCGCTGCCATGTAGCGCCTCCGTTACGGTATGTGCCCGC
25441 CATCGTCTTCGAGGACGTCCCGACTGAGGCGAGCGCACATTCAACGCTAGGTGTGAGCAC
25501 GCATCGCCAGTCATCACCGCGCGTCCTGTCTATGGCATACGAGCACGCCGCAGTCGGGTG
25561 CGCCCCCCCATTCTTAAGGCACCGCTGAGCGGCGTGTTAGCGCTCGGCGGGAAAGACTCA
25621 CGACCATACTTGGGACAACTCAACGTGAAACATCGGCAAATTGCAGCGACGACCGCTTGC
25681 ATCTCTCTCGCGATTCTGACAACCGGCTGCGCGACGGTTACTGGCGGGACCACGCAAAAC
25741 GTATCGGTGAAAACACAAAAGGACGCAGCAGACGTTGCAGGAGCGGACTGCGTCCTGTCG
25801 AACTCAAGGGGTAGCTGGACGGTGACCACGCCCGGGAAGGTCAGCGTCCATCGGGCACAG
25861 GACGACTTGAATGTCCGATGCTCGAAAGATGGCGAACGCGACATCACCACGGCCGTGAAG
25921 TCGAGGACACGCACTGGAACGATAGTTGGAAATGTCGTCCTGTTAGGAGTAGGCGCCTTG
25981 GTCCTGGATGGGGTCGACAAGGCGAATGGCTCGGCCTACGCCTATCCGGAAGCCATCACC
26041 GTGTCGTTTGTAAATGGCAATGAGTCGCAGGGAATCCCCGGTGACGCAAGTACCGGCTCT
26101 ATCAACGGCGCGCCAGCTCAAGCGGCAGCGACGCCGGCGACAAACACCGCGTCGAGTCTG
26161 ACCAGGTAACTTATTTTCTTGCGGAGCTCGCGCATAGTTCTGCACTCGCTGCCGTCATAG
26221 GAGCGGATGACTGCGCCGTTGATGTTCGGCGCCGTCATCGAGACGTCGACGCTCTTGATA
26281 GGCGAGCGTCAACGATGCAGAGGCAGTCTTTCTCGGTTGCCGTACAGAACGGCGATTACA
26341 TGACGTCGGTCCACTCGACCGATTCGAACGTGCGACATTGGCGTGCTCCGTAGCACACGC
26401 TGTTGACGCCGAGTTCGTATGCAGCATCGAACATACAACGCCCGCACAGGGCTGCCATCG
26461 CCGTTTCCTTGCCGGATTCGACCGCATTCGACGCCTGAATCAAGGTCTGCAGCGGGACCT
26521 CGTGAGTTCCTTCAGGCAAATTGTCGTCATGGGGCACATCGAGGAAGGGACCGTCGGCCT
26581 TGCGCCGTTTGGCATGGGCGCGGTCAATGACACGGCGCCCCCAGCCATCAGGCTCAGTAT
26641 CGGCAATGGCACCCGCAAAGGTAGAAGTCTCCTTGCTTCCCGTGCGGTATTGAAAGTCGG
26701 TGCTCAGGGGGAGGTCCGGAGACAGTGCGAAGCAATCTGGATTACTGAGCCAGGCGGGGG
26761 AGTACCGGAATGCCGAGACCTCTTTTTGCCCCTGTCTGGAAAATTGCAGTTCGCCCACTT
26821 CCAGTCCCGTGTCGCCCAGATAGACGAGCAGCGTCCGCGTTGTGCTGGTTTTTGCTTTTT
26881 TTGTCGTCGCCATTACAGGGCTCCACTGTCCGGGTTGGCTTTGCGTCCCCGGATGCGTTT
26941 GGGCACCGCCTTGTTCATGAGCGCCAGGCCGAGCGCATCATTGGTGGTGTTCAGGGCCTT
27001 GTTGAGCTCGGGGAGGAGATTCAGGATGTTCAGGACCCGAATCACAGCGCCCCACGACAC

27061 TCCGGGGTCGCCCTGTTCCAGCCGGTAGAGCGTCTTTTTTGAGATGCCCGCCCGCTCCGC
27121 CATGGATGACGTAGAAAGCCCCCGGCGTCGGCGGGCCAGCGCAAGGTCATCTCCGGCAGC
27181 AACCAGCGCACGCTCGACCGGCAAGGGCATCGGATTTTCGGTGAAGCGGCCGGAAGGTTT
27241 GGGTGTCGCCGCTATTTTTGGTTTATCGGTCATATCTGAACCCTTAATGCGTTAATACGT
27301 CATATTTGACCACTATACAGCTTTTCTCAGGTGGGCGTCAGCGCGCTGATGAAAACGGTC
27361 AAATATGGGTCATAATGTCATTTAATAGGTCATATTTGACTCTTCATCGTGGAGTCAGGC
27421 GCAAGAAGAGTCCAGAGGCTCCCTGTTGAGAAACGCCCACCCAACCGGGAGCGCGACATC
27481 CTGTTCCTGTCGAGGGTGGCCGACCTCGAGTGCTGCCACGCGCTTTCAGGGAACGCCCGT
27541 GACCAGCGTTTCAACGTCGCGGCGAGTCGTGCGCGTGACCGCATGTATCTGGTCAGGCCG
27601 GTAACGGCCGAGTAGCTCTCTCCGAGAAGGACATCCGGATGACGCTGCTGGAGCACTTCA
27661 GCGGGCCTGCAGTTGAGAATGAGGTGGACCAGGCAGGACTCATCAAACTCTGCGAGTCGG
27721 GTTTCGAGCGCGAGTTCTTCGGCATGCTGGTCGAGCGCGGGTTATCGCGTCATCCCGCAG
27781 GTGAAGAGCGGCGCTATCGTCTGGACCTGGTCGTGGAAGGGGCAAACGACGCTCGACTAG
27841 CCTCGAGTGCGATGGCGACGCCCTCCACGGCCCCGACTGGTGGGCCCGTGACATGTCCCG
27901 TCAGCGAATTCTTGAGCGCGCAGGCTGGACCTTCTGGCGATGCTTCGCGTCGACGTGGTG
27961 TCTGCGCAAGGAGGAAGTCTTTGAGGAACTGATTGAGTGACTCACGAGCATGGGCGTCGA
28021 GCCCCTAGGTCGCGAACTAATAAAAAAGACCCATAAATTTCTCGACTTAGCGGTCCTAGA
28081 TTCTTAGGTTATGCGTCTACCTACAATCGAGTGGAGCCTGACGCAGTCAGCCCAATCGGC
28141 ACTAGTGTTGTTGGATACTAGGTACGTATCTCGTTCGATAGGCGACAGGCTATTCCAGAA
28201 AGGGCTCCAAAACTGACGCCACCAGTACGCAACATCGCCTTGCAGCGCGCCCAACCCCAG
28261 CGGGTCCAGGTGTGGGAACACGGTCCAAGGCGGAGGTGGTGAGAACCTACCTTCTGCGAT
28321 ATGGTGAAACACCTCTGCATCGTCTGAAACGAGCGATACGACGACGCCAGCGAACTTCGG
28381 CGCTCGACTCAGGTGGATATGGAAATGCGCGCCTCCGGCGACCTTGCTGTAGGCAGCTCC
28441 GCCCGCTGGCTCATATCCACGCGTCGCCACCTGTACGAGCAGCGCAGCATCCGATTCGCG
28501 AACCTGAAACTGAGCAAAAAATGCCCTGATTCCCGCTTCGGCAAGGTCCTGCACGACGGG
28561 ATATGCAACGTCGGGCTGACGCAGCGACCGGCTCGCTGAGTCACCATAAGTTTTACGCAG
28621 AGCCTTCGTGTCAAGGCCAAGCAATCGCTCGAAGGTATTGCTCCGCTCGATTCGGATTCG
28681 GCGCTCCATAAGCGCATTTTTCAAGCTGAACATGTGGCGAAGATTAAAGTGGGTCGAATT
28741 TGTTGTTGAGCCAGTCCGCGCCAAAGTACCCAACTGCACCGAACAAAAGCCCGCCAATCG
28801 CTCCCGTGATAATAGCCCCAGGGCCGGTCTCAACACCGATGCTCGCGCCCAGCGCTGCCC
28861 CAATCCGCGCCCCCGCCGTGGCCCCGCTCCAGCCGCCCATCTGTCTTACGGCTTCCCTTC
28921 GAATCGGACGGATGCTCTTTGTTTGCACCGATTGCTCGGCCGCCACGTCGAGGTCGTACG
28981 CGGTGAACACGATTCCGAAAATTTGCACGACGCGCGCGTACTTCGTGAATCCGAGCGCAG
29041 TCCGTAGCGACCTTTCGTTGAAGATGGCCGATGCCGGAACCTTGGGACTCTGAACAAGCA
29101 CCTCCTTGTCGATGTCCTTTACATACCCAGCAATTTTTGACACCCGCTTGGCCAGGTGCG
29161 GATTCTCGGCTTTGTATTGCTCAAGAGCCGTCAAGATTTCGTCGGTCTGCACCAGCTTCG
29221 CGCCCGCGCGCTTCGCTTTCGCGATATCGATGACGATGCTGCGGCCTTCGAAGCGAGGTG
29281 AGCCACCGGGGAACAACTCGCTCGTCGAGACGTAAGACGACATATTGTTGCCCATCACGT
29341 GTTCAGCGACGCTCGTTCCGGCTTCGAAACTTTTGGGAGATAGTCCGTACTTCTCTGGAT
29401 GCAACTCCTGCAGCTGCTGGAGCGGAGAAAGCCCCATGTCCCGCAGGAACACGTTGTCCG
29461 GTGCCGCGTAAATTTTGCGAACGAGAAATCCGTCTTTCCCTGGCTCGTCGACAATCGCGA
29521 AGAAGAAGGCGCGTGCCGGCATGAGGTGTCCGGTCACATCGTGGCGCGAACCTTCCTTGT
29581 CGTTGGTCGGTGTCTGGTAGTGATAGGCGGGCATTATTTTTTCCACTTTTCCAAAGCTGC
29641 GACGCCGACTGCAATCGACAAGGTTTCGATGCTGGCCGTCGAAACTCGCGGGCACAGGCC
29701 ATCCTCGTCCGTCAAACCCTGGTATTCGTTACCAGACTGGTCCCGGATGTGATACGGAAC
29761 GTTCGCGAGAGGTTCACCATGCTCATCGAGGATTCGGACTTGCTCGTCAAAGGCATGCAT
29821 GGCTTGGTCGCTCAGTGGGCTACCGTCGGCTGCGTACCCCATCTGTGCGAGTTGATGTCC
29881 CTCGAAGGTCATGAACATATCGTTCTTCGAGGCAATCATCCGAGGGGGAGGGTCGCACTT
29941 ACAGACGCACAAGTCGTTCTCCAAAGCCACCTGCTTGCCGTCGAAAGTCATCGGCCAGGA
30001 CGGACCCGCGCCAACAATGTGGCCATCCGATTTGCAAGTAGGGCAGTGGACAAGCGCGCC
30061 ATGATAGGCGAGAGGCGTTCCATTATTGAAGCCCGCAGCCACTCCCTCTTGCACAATGCC
30121 ACCCGCCGTCGTTCTATCCCCTAACACGAGCATATTGCGTTTCATGCGATACCTCTATCT
30181 AACCATTGCCCGATTCGACGGCCGCTCCTAACGCGCCTCTCGTGCAAATCTCGACTGATT
30241 TCCACTTCCCGCCAGTTGGCGTCAAGCATCCGGTTATACATTCGCGGGAAAACCGGAAAT
30301 CATAGATTAGAACGCAACGAGAGCACAGTGCCGCAGGCAAAATCAGACGAAGTCGCCGTC
30361 TCTTTTATGCACTGCGCGTCACATATAAAGACAGGTCAGTTTTGAGCTAGAGAGCATGCC
30421 ACCGAAATCTATAGGCAACGGGCGGAATCGCCGGTTGCCGCAGGTTACGCGAAGCAACAT
30481 GAGACGTGGTCCATGCCCGCGACGAATCTTCTGGCTCCGGGGTGAATGTAGTCCTATCGG
30541 TGCCAGACTCATTGGCCGCATAAAATCGAAATTTGCCATTTACACGGCCAATTGTTTTAA
30601 ATCAATGACATAGCGCCACCGCCAACGGGGCAGTCGCTCCGTCCCAGCGCAGGCGCGGCC
30661 CGATTCGTCACAACGTGATATTATGTTATATCATCGACTTTGATATAAGCCGTCCGAATC
30721 GGCCGGGGGAGCCGACTGTCTTGCGGGCCTGCATGGCGAAAGTTCGATAGTTGACGGGCT
30781 CGAATCTGAAGCGGCCACTGCTTCGGCACAACGCTGGCGCGTAACGCCGCAGGGCTTCTT
30841 CGGCGCTCACCTGGCAATGTGTTTTCGGACGTCGTCACCCATGCTGACGATGTCGTCCGG
30901 GACATCAAACTTCCTCTTTGCAATACCGATGCGCTTTTCTTCGCGGCTTAGTCGAGCCTC

30961 AGCAGACGCCTTTGCCCGAGATTTTCCGAGCTCAACGGCGTCCCGCACGTTCATCTCGAG
31021 TGCCCCCCGGTCTTTTTCTGGAGACGCCAGTTCGGCGAGAGCGCCATCCCGGCCTATGAG
31081 CCATAAGGCCGTCGCATACAGGCCGATACCCACACCCGGTTCGCCAGCTTCTATACGCAT
31141 CAGAGTCGGTACACTAACCCCGAGCCGGGCCGCCCAGGAAACGAGCGACTCTTTTCTGCG
31201 TAGACGAGCCACAGCCAGGTCTGCGCCCAAACGCTCGAGAGCCGCAACGGTGGCTGGAGG
31261 CATTTGCCAAATCGCGACGGGGGTTTTCGGCGTAAATTCTATTTACGTACGAGACTGTAT
31321 TCCGTGCAAATCTTCTTTACGATTTGCGTTTCCGCTCATCGGCTAGGCAACGTAGTCACT
31381 CTGTTTCGTTGCCTCTTCATCGCTTCCCGACGTCGCTCGTACAGCTCGTCGCGCAACAGC
31441 TGCTTGAACCTGTAGCCGATTTCGAACGGGGACATCGTCCGAAGCGCAACGTACTCTCGC
31501 GGCGTCATCGCATCAAGGACCTGCATAAGACCATGGTCCATAGCATCAGCAAGCATCCGC
31561 TCGAGCGTGTACCGCCCGCCAATGGTTGACTCGGCGCGTTGTCCCTTGGTTCGCATCCTT
31621 CCCCCTCAATACCAGGCTTTTTGCTCTCAATCATGCGGTGCCGCCAACCACTGCTGACTA
31681 TGAGAATCGTTCATGTAGTCTGCGCCGCAGAAGGTCGACGACAGCAGGGTCACCCAACCC
31741 ACATTCGGACTGACAATGGACAAGCCGCTCTCTGCTGCTCTCGAATCCGTCGCTGCGGTC
31801 ATCAATCGCAAGCCATTCTAGGCCCTCACATCCTTGTCGGCGAACGTAGTCGAGGACCTG
31861 TGCGCCTCTCGATAGCATTGGCCAAAATTGCTCGTCGAATTCGCAACCCTGATACGTTGC
31921 GCCGATTACCCGTGCGCGCAACGATGCCGACGGCAGCTGGGAACGGGTGTAGTCAAGCCC
31981 GAATCGATAGGCCCAATCGGACGAAAGCACAACCGGCACGTCTGGATGCGAGTTCAATGC
32041 GTCGTCCAGCGTGGGCACGAACTCAAACAGTTCGATGCGCCCGGGCGCGCTGCTCAGGAT
32101 GCGGCGGCCCGCAACATACGAGTTGCCTCGGTGCAGGACCCCGTCAATGTCGAGAAAGAG
32161 AATTCGCCGCTCGCTCATCGTCACTATCGGTTGCCACGCTACTTGGCGCCGGCCCGCAAT
32221 CTGCAGGCTCGGCGCTAGCACAGACATCACCAGCTACCACACTGAGCCATCCGACAGAGC
32281 GGGCAGCACCGCCGTTGCCTTCGAAATGCGCCAGTAGCTCGCGGTTTCTCGAGGTAACCA
32341 ATTTGGGCATTCAAATGGTTCGCTGCCTGCTATGTGCGAGATAGGTCTGTAAACGCCGCA
32401 CCCACGGGTTGCCAGCAGTTGCTCTCATGATGTCGTCGGTAGAGACGGGGTGCGTAGCGG
32461 CCTCGGCGGCCTTCTCAAGAAGCACTCGCAACCGGTCAGCCGCACCAGCGTCAAAACCGA
32521 ACTGCGAGTCGACGTTCACACGCCTCGCCGCCGTCACCTCGCTGAACTCGAAATCCGTCG
32581 ATGTGACCACGATGTACTCGACCGACCGGGAGGACAAAAAACCCAATATCGCTTCCTCTT
32641 TCGTGCCGCCGAGGATTCTCACTGCGTTCACGATGCGCTGCGCGAACGGCACGCACTCTT
32701 GGCCCAGGAACTCGATGTCGTAGAGCAGTGGCCAGTCGTTTGCAATGACAACAAGGGCGC
32761 CAGGGTATGGCTCGATTGTCTCCGCGAACAGGCGAGCGTCCGATTCGAGCGAACGGGTCG
32821 CTGCAGGCTCGTTTTCGCGAGTCGTTATGGACGGATGTCGGTCAACCGCGCCCAGGACGC
32881 CGCTGATGCTGAAGAAAATCAAAACCATGTATGTCCCATTTCGGAAGGAAGCTACCGGTG
32941 CGTGCCCGGGGTTTTCAATTTGCGTCGGCCGCTCGCCTCGCCGAGACGACGGGCTTATCA
33001 ATAGCGCGCACCTCCCTCGGTCCATTGCTCCGCTGGAGGGTCAATCGAACTGTGCGAAAA
33061 GTGTCCGCGACGGAGGCTCTCGCTCCGTCGGCTCGAGTCGAATTGTAGCTTTGTATGCGT
33121 ATTCTTTTCAAGCGTACACATACGGAGTGGTAGAAAATTTTTCGACTGCAAGGGATGCTG
33181 TGCGGCTATGCCGAAGCTCCATGAAGCGGTCGCCAACGGGCTGTCCGAACGACTTAGAAC
33241 GATTCGAAAACGTGCCGGCCTCTCGCAAGACGAACTCGCTGCGAGGGCGTCGCTGGCTCG
33301 CACGAACCTCGCGAGCGTCGAACAAGGACGGCGGGCAAACCTCCGACTCTCGACCCTTTC
33361 CCGTCTAGCTGATGTACTGGGCGTGGATGTGCTCGACTTCTTCTGTGACCGGCCTGTCGG
33421 CGAACAGGAACCAGCCGGGGAAGACGCCCCGGCTCGGGTAATAGCGAATGTGAAACGGCT
33481 GCGCAGTGAGGCAGGGCTCAGTCAAGAAGCACTCTCGGTCAAAGCCCATCGATTTCGGAC
33541 GTACGTCGGCCGGTTGGAGAACGAAGCGGCCAGCCCAATGGTGGTCGATTTGCAGGACCT
33601 GGCTGACGCCCTTGGCATCTCGATTCCGGCACTCCTCGAGCCTAGCCAGCACAAAGCCAA
33661 ATAAGGAATGAGCGGAACGCGGCCAAGGGCCTAGCCGTCAACGTTCGGCAGTGCCTCTTG
33721 CTTCTCGAAAAGGATTGGGCGACAGAGGCAATAGAGCGAACGCAAACTATGCTTGGTCGG
33781 CGCCGTTTTTCCCTCCGCAACTTCCTTCGCGCGAAACAGCGACGCCGCGATAAATTCCGA
33841 TAGCTCGTCCGCTCTTTGGCGCAGTGCGGCGGTAAGGGTTGGATGCCATGGGGCAGCCTC
33901 GTAAACTTCCTCTACCGTGTGCTGTAGCGCTTCGCGCATCGCCTCGTCCGCCACTGAACG
33961 GAATGCCGGATGCCGGTAAATTGCTGAGGCGAGTCGCTCGACAACGCATGCCAATGCACC
34021 GAACTGAACGTCCAGCAAGTGAAGCGAACGGAACACACCGCACACTACGCGAACAGCAGA
34081 GTCGAACCATGTGTTCGAACTTGCAATCACTTCGGCCCACGCGTCGAGGACAATACTCGA
34141 ATAACCGGCTTGCGACAGCATCCATAGGAAGTCATAAGACGCGATGGAGATGTAGGTGCG
34201 GTTTGAGAGAAGCGACAGCATCAAATACTGACGGTGCTCCGCAAGTGGGAGCTGTCCCGT
34261 CTGCGTCGCGTGTGCAATAACAGCCTGTGGCGACACGGATGAGACCTCGAATCCGCTAAG
34321 CGCTAACGCACGCAGGCGCGCGTCAATCGACACGAGTTCCGCTCGCTTTTCAAGGCAAAG
34381 AGCGATGGCATCGTAGGATTCGTCGTCCAAGACACCCTTGAATTCTAGAAAATTGTTAGG
34441 CACCTCCGCCGGGCCATACGCGGGCGCGACTTCGCAGTAGCGCTGAACCGACTCGATAAT
34501 TGACGTCAGCATCGCCACTCTTTGCTGCCGGTCTTCGGGACCGTACCGAACTATCGCAAG
34561 CTCCCCGTCTATTTCCCTCGCTCGACCTATCGAGCGTTCATCTTTTGCCTCGTCCAAGGC
34621 GATAAGCAGCAAATCATAGGCACGCTGCGAAACAACAGGATGCGGCAACAAGCCCAGAAC
34681 TGACAGAGCCTTAAATAAGCCCAACTCGACCAGCGTGATGCAATCCAGCACGCACGATGA
34741 ACTTGAGCCGGGTGGATTCGACCGCTCAGGCCCACCAGTGACACGTCCCTCGGCAACAAC
34801 CAATGGCGGCTGGCTCGGCCCCCAGGACAAGATAATGTCCACTGGCGAACGTCCAAGCAA

34861 CTGCGCACAAAGGCCTACTGTAAGGCGCCCCTGCCTGTAGCCATCGATGACTACGTCCGC
34921 TGCCCCCGACCATCGTTTCAACTGTGAGTGAATCCGGGACATGTCGATGCCCCCGTCATC
34981 GTTCTTGAAAGAGATGGTCCGAAATCCCGGAATGGGTTCTACGGAGTTTTCGGTTAGCTC
35041 GTGCGCGTAACCAAGCAGGTATCTCAAGGCAGGCATCAATGAAGCAACGCGATAGGTGGT
35101 CTCTTCTCCCAGCTGATGCCGAATGGTCACCACGTCGCCAACGCCCTTGCCCATGAAGTT
35161 CGCGGCGAAACCAGAGTCGTGAGGGAAGTAATTCTGTGCCTGCGGAAGGTTACCGACGTC
35221 TGCTGGGTCGACGACAATGGTGAACGGCGACCCGTCTTCTCGCAGCAAGGTCGCCGCAAC
35281 ACCCGCAGAGACAGCTGTAGGTGCAGCGCCCCATGGAATGTCGTCGGCGGTAAGCAGGGC
35341 GGTGATGAAGCCAGCGGCTGCGTTCGCATTGTCGAGGTTGCATCTGAACAATCGATACAA
35401 CCGGCGCATACCGAGGTCACGGTCGCCGTATTGAATTTCAGCCAACGCTACCCTGCGGAT
35461 AACCTTTGGGTCACCGCTTAGCTCCAGCGGGGCCTCGCCAAACGCACGCCGAATTTCTCC
35521 CGGCCTAAGGGTTTGAAGAGCCACCCTATACCAAAACTCCCAGGCTCGCGCATCCGTAGG
35581 GCAACGGTCCCTTTGTCGCAGCGCCAGCGGCTTTAACCGGTGCCAGTCATTCGCCTGCTG
35641 AGCTGCGAGGATAGCCATGCCGACGAAGGCGTCGTCCTCGAACCAATCTGGTGGAGCGCT
35701 ATCCACGATGGAGCGTGCCTTCGCGCGGTAACCACCCTTCGCATACGCCTCCAGTAGGCG
35761 TGTTTGCAACCCACGACTGAGATTGGCATTCGGAACCGAGGTCAGCAGCGATATCGCCGC
35821 CTTGTAGTTCCCAAGTGCAGTAAGCGTATCGGCCGCCTGAATCTTCACGGGCTGAGCCGA
35881 GGCCGTCGTCACCAAAGCAGCAATCTTCCCGCCTAACATCGCAGCTTTCTCTTGCTGCCC
35941 AACCGCCAGATAAATGCGGGCCGCCGCGTTCAGCAAAATAGGGGCGGTCTCAACCTTGAA
36001 ATCATGCTTATCAAGCAAATCAGTGGCTGATGAACTGTCGCCGCTGTTTCGAAGCGCAAT
36061 CCACATCAGCGCCTGAGCAGTTTTCGCCGCGTCGCTGTCAGTGGCCACAGCGTTCGATTC
36121 CGCGATGTCCCGTACCAACTCGGGCATTCCGAGAATAGCGGCAGACTCGCCGACCATCAT
36181 CTGCGCCGAGATATGCATATCTTTCAGGTGTGGTGCGCCGACCTCAACAACCTTCTTCAG
36241 GTCTCCCAGTTCGTGGTACGCGTATACACGCGGAGCCGCGAGCTTCAACGCAGACTCGCC
36301 AAATACGCCCGCTTCCTCGCAAATCTTCAAGCAAGCCAAGGCGTCGCCGAGCAGTAGCCT
36361 TCCGACCGCAAGGTTAGCCGCTGCATCCGAAACTAGTGGGCCGATTTCCTTGCGCCAGAG
36421 TCTTTCATCGCGTGGTTCAAAACATTCGCATACGGCTTGCAATGTACGTGCGACCTCTGG
36481 GGGCAGGAGCCCCATAGAAGCCAAAAACGGGTCCTTACTGGCGAAATCTGTGGCGATAAA
36541 TAGAACGTTGGTCCACGTGTAGAAGCCTGCACCATCCGCATGTAAAGCACGGTCAGCATA
36601 TGTGAGGGACAACGCTCGGTTCCCCGCATCATGAAAGCTAATTGCCATCATCTGCAACGC
36661 ATCAGCGTTGCCTTCGAACTTCCTTGGCAAATCAGCTTCAGCAAGAAACTGCTTTAGTGC
36721 GATTTTCGTATTGGCCCAAGCAAGCCATACGTGAGCAGAATCGGGAAAACGGGAGCGCGC
36781 TCTCTCAGCGATGTCAAACGCGTCATCAAGATGGTCAGAAAGCAAAAGGCCACGCACCTT
36841 TCCGGACGCAACTCGGTCATCGTCCGGGTAGGTCTCATATGCCGTATCAAACTTTTGCAC
36901 TGCGCCACTGACATCTCCCGCGTGCCAATCCACGATTCCTGAGAGGAGATACCATCTCGC
36961 TTTCTGCCGCTCGTTCAATGACGAGAACTCATCAAAAATCACCGATAGCAGCGCACTTGC
37021 ACCGCTATAGCTCATTTCCTTAATGAACTCGGCCACCTTGTCCAGCTGTCTCGATACCGT
37081 ATTATTCTCGTTTGTTACGAGCGACGACACAAAGGCATGGCCACCCACCTGTGTCATCAG
37141 CGAATCGAGCTTCGACGATAGCGCGCTTACGAGCTCATTGGTTTCAACTACCGTCGAATG
37201 AGTCTTCTCCGACGATGTTCTAATGTCATCATAGACACCGCCTGGAGAGTTAGGGCTGTA
37261 CTTCTTCTGAAGAATGGCATGACGCTTTATGTGATTACAAATGTCGTCCCAAAACTCGAT
37321 GCTTACCTTGAACTTTCCAGACTCAGCCCGAGCGTCAGTCAAATCCTGGGCATACTTGAC
37381 GAGCTTCGAATCACTCGCGGCCGTCGTCGCTACAATTAACTCGCCGATACTCCAGCCAGC
37441 CTTGTCGGCCGCAACCACTTCATCAGCTATAACTTTTTCAGTGACCTTCGTTTCGTAGTA
37501 ACGCTTGCACTGAATTCCTTTACGCTTACCAACGCGAGGCTCAACAAAAATGTCTACACC
37561 GCCCTGACTTTGTCCGGTTCGACCGTTGCGCTTCGGGGTGGGGTCTTCAAAGACTACTGA
37621 GTAGATTTCGGCGCAGTAATCTTCGAAGTCCTCGGCATGCTTGGGCTTGACGATTTCTAA
37681 AGTGAACGTACGCATGTAGATGCTTCTTCGCTCTTCCGAAACCAAAGGATATACATCATC
37741 GAATCCGACGTTCCGGTATCGGCGTATGCATAGTCGTCGGCTTGAACATTGTGCCCGAAA
37801 ACACCCAAAGGAAAATGCGTCCGCCGACGCAGTGAAGGTACTCGACCTGGCTCATCGGCC
37861 GTCATTCGATGTATGAACAGCTTCGCGACAGAACGATGCCGCTCCGAACCGAGAGGAAGA
37921 TGCATCGGCTCCTTCTACAAAGCGAAGCCTGCGGTTTCCCGAATTTTTCCCTTTTTGCTC
37981 AAGCTGTTAGGTGACCCAACGTGCCGCCCATCTTCCTGTTTCACAGCACCACGCCGTGGC
38041 ACAAAAAAGGGCGCCCCGAGAGGACGCCCTAAAGTCACCTTTTTCCGATATCCATCAAGG
38101 AGTTAGAGTGAGAATCTTCCTTCCTCTTGTTGGTATGCTCCCTAGAACGGAATATCGTCG
38161 TCCATCTCATCGAACCCGCCGCCCGCCGGGGCGCTCGGTCGGTTCGCGCCGACGACGACA
38221 CCGCTGCTCGCACCCACACCGACGACACCCGACACGCGGACGACACCGTCGTCGAGCTCC
38281 ATGTTCCGGGCGACG

REFERENCES

Achouak, W., Christen, R., Barakat, M., Martel, M.H. and Heulin, T. (1999). *Burkholderia caribensis* sp. nov., an exopolysaccharide-producing bacterium isolated from vertisol microaggregates in Martinique. *International Journal of Systematic Bacteriology.* 49(2): 787-794.

Adhya, S. and Gottesmann, M. (1978). Control of transcription termination. *Annual Review of Biochemistry.* 14: 967-996.

Ahmed, K., Enciso, H.D.R., Masaki, H., Tao, M., Omori, A., Tharavichikul, P. and Nagatake, T. (1999). Attachment of *Burkholderia pseudomallei* to pharyngeal epithelial cells: a highly pathogenic bacteria with low attachment ability. *American Journal of Tropical Medicine and Hygiene.* 60(1): 90-93.

Aktories, K. (Editor). (1992). ADP-Ribosylating Toxins. *Current Topics in Microbiology and Immunology.* 175.

Alm, R.A. and Mattick, J.S. (1996). Identification of two gene with prepilin-like leader sequences involved in type IV fimbral biogenesis in *Pseudomonas aeruginosa. Journal of Bacteriology.* 178(13): 3809-3817.

AlonsoDeVelasco, E., Dekker, B.A., Verheul, A.F., Feldman, R.G., Verhoef, J. and Snippe, H. (1995). Anti lipopolysaccharide immunoglobulin isotype levels and opsonic activity of antisera: relationships with protection against *Streptococcus pneumoniae* infection in mice. *Journal of Infectious Diseases.* 172(2): 562-565.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research.* 25: 3389-3402.

Anuntagool, N., Intachote, P., Naigowit, P. and Sirisinha, S. (1996). Rapid Antigen Detection Assay for Identification of *Burkholderia (Pseudomonas) pseudomallei* Infection. *Journal of Clinical Microbiology.* 34 (4): 975-976.

Anuntagool, N., Intachote, P., Wuthiekanun, V., White, N.J. and Sirisinha, S. (1998). Lipopolysaccharide from Non-Virulent Ara^+ *Burkholderia pseudomallei* Isolates is Immunologically Distinct from Lipopolysaccharide from Virulent Ara- Clinical Isolates. *Clinical and Diagnostic Laboratory Immunology.* 5: 225-229.

Anuntagool, N., Panichakul, T., Aramsri, P. and Sirisinha, S. (2000). Shedding of lipopolysaccharide and 200 kDa surface antigen during the *in vitro* growth of virulent Ara^- and avirulent Ara^+ *Burkholderia pseudomallei. Acta Tropica.* 74: 221-228.

Arun, S., Neubauer, H., Gurel, A., Ayyildiz, G., Kuscu, B., Yesildere, T., Meyer, H. and Hermanns, W. (1999). Equine glanders in Turkey. *Veterinary Record.* 144: 255-258.

Ashdown, L.R. and Koehler, J.M. (1990). Production of Hemolysin and Other Extracellular Enzymes by Clinical Isolates of *Pseudomonas pseudomallei. Journal of Clinical Microbiology.* 28(10): 2331-2334.

Bachmair, A., Finley, D. and Varshavsky, A. (1986). *In-vivo* half-life of a protein is a function of its amino-terminal residue. *Science*. 234: 179-186.

Balfanz, J., Rautenberg, P. and Ullmann, U. (1996). Molecular Mechanisms of Action of Bacterial Exotoxins. *Zentralblatt fur Bakterologie-International Journal of Medical Microbiology, Virology, Parasitology and Infectious Diseases*. 284: 170-206.

Bally, M., Filloux, A., Akrim, M., Ball, G., Lazdunski, A. and Tommassen, J. (1992). Protein secretion in *Pseudomonas aeruginosa*: characterisation of seven *xcp* genes and processing of secretory apparatus components by prepilin peptidase. *Molecular Microbiology*. 6: 1121-1131.

Bauernfeind, A., Roller, C., Meyer, D., Jungwirth, R. and Schneider, I. (1998). Molecular procedure for rapid detection of *Burkholderia mallei* and *Burkholderia pseudomallei*. *Journal of Clinical Microbiology*. 36 (9): 2737-2741.

Bender, C.L., Alarcon-Chaidez, F. and Gross, D.C. (1999). ***Pseudomonas syringae*** phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. ***Microbiology and Molecular Biology Reviews.*** 63(2): 266-292.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*. 7: 1513-1523.

Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschape, H. and Hacker, J. (1994). Excision of large DNA regions termed pathogenicity islands from t-RNA specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. *Infection and Immunity*. 62: 606-614.

Bradley, D.E. (1972). Shortening of *Pseudomonas aeruginosa* pili after RNA-phage adsorption. *Journal of General Microbiology*. 72: 303-319.

Bradley, D.E. (1980). A function of *Pseudomonas aeruginosa* PAO polar pili: twitching motility. *Canadian Journal of Microbiology*. 26: 146-154.

Brett, P.J. and Woods, D.E. (1996). Structural and immunological characterization of *Burkholderia pseudomallei* O-polysaccharide-flagellin conjugates. *Infection and Immunity*. 64(7): 2824-2828.

Brett, P.J. and Woods, D.E. (2000). Pathogenesis and immunity to melioidosis. *Acta Tropica*. 74: 201-210.

Brett, P.J., DeShazer, D. and Woods, D.E. (1997). Characterisation of *Burkholderia pseudomallei* and *Burkholderia pseudomallei*-like strains. *Epidemiology and Infection*. 118: 137-148.

Brett, P.J., DeShazer, D. and Woods, D.E. (1998). *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. *International Journal of Systematic Bacteriology*. 48: 317-320.

Brett, P.J., Mah, D.C.W. and Woods, D.E. (1994). Isolation and characterization of *Pseudomonas pseudomallei* flagellin proteins. *Infection and Immunity*. 62(5): 1914-1919.

Brook, M.D., Currie, B. and Desmarchelier, P.M. (1997). Isolation and identification of *Burkholderia pseudomallei* from soil using selective culture techniques and the polymerase chain reaction. *Journal of Applied Microbiology*. 82: 589-596.

Brown, T.A. (1993). *Gene Cloning* (Third Edn.). Chapman and Hall.

Buchrieser, C., Brosch, R., Bach, S., Gulyoule, A. and Carniel, E. (1998). The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal asn tRNA genes. *Molecular Microbiology*. 30(5): 965-978.

Burkholder, W.H. (1950). Sour skin, a bacterial rot of onion bulbs. *Phytopathology*. 40: 115-117.

Camilli, A., Beattie, D. and Mekalanos, J.J. (1994). Use of genetic recombination as a reporter of gene expression. *Proceedings of the National Academy of Sciences of the United States of America*. 91: 2634-2638.

Carniel, E., Guilvout, I. And Prentice, M. (1996). Characterisation of a large chromosomal 'high pathogenicity island' in biotype 1B *Yersinia enterocolitica*. *Journal of Bacteriology*. 178(236): 6743-6751.

Carroll, S.F. and Collier, R.J. (1988). Amino acid sequence homology between the enzymatic domains of diphtheria toxin and *Pseudomonas aeruginosa*. *Molecular Microbiology*. 2: 293-296.

Chaiyaroj, S.C., Kotrnon, K., Koonpaew, S., Anuntagool, N. and White, N.J. (1999). Differences in genomic macrorestriction patterns of arabinose-positive (*Burkholderia thailandensis*) and arabinose-negative (*Burkholderia pseudomallei*). *Microbiology and Immunology*. 43(7): 625-630.

Chambon, L. (1955). Isolement du bacilli de Whitmore a partir du milieu exterieur. *Annals of the Institute Pasteur (Paris)*. 89: 229-235.

Chaowagul, W. (1996). Melioidosis; a treatment challenge. *Scandinavian Journal of Infectious Diseases*. 101: 14-16.

Chaowagul, W., Suputtamongkol, Y., Dance, D.A.B., Rajchanuvong, A., Pattara-arechachai, J. and White, N.J. (1993). Relapse in melioidosis: incidence and risk factors. *Journal of Infectious Diseases*. 168: 1181-1185.

Chaowagul, W., Suputtamongkul, Y., Smith, M.D. and White, N.J. (1997). Oral fluoroquinolones for maintenance treatment of melioidosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 91: 599-601.

Chaowagul, W., White, N.J., Dance, D.A.B., Wattanagoon, Y., Naigowit, P., Davis, T.M.E., Looareesuwan, S. and Pitakwatchara, N. (1989). Melioidosis: a major cause of community acquired septicemia in Northeastern Thailand. *Journal of Infectious Diseases*. 159: 890-899.

Charuchaimontri, C., Suputtamongkol, Y., Nilakul, C., Chaowagul, W., Chetchotisakd, P., Lertpatanasuwun, N., Intaranongpai, S., Brett, P.J. and Woods, D.E. (1999). Antilipopolysaccharide II: an antibody protective against fatal melioidosis. *Clinical Infectious Diseases*. 29: 813-8.

Chittivej, C., Buspavanij, S. and Chaovanasai, A. (1955). Melioidosis with case report in a Thai. *Royal Thai. Army Medical Journal.* 68: 11-17.

Chodimella, U., Hoppes, W., Whalen, S., Ogniborne, A.J. and Rutecki, G.W. (1997). Septicemia and suppuration in a Vietnam veteran. *Hospital Practice.* May 15: 219-221.

Choy, J.L., Mayo, M., Janmaat, A. and Currie, B. (2000). Animal melioidosis in Australia. *Acta Tropica.* 74: 153-158.

Clarke, L. and Carbon, J. (1976). A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E.coli* genome. *Cell.* 9: 91-99.

Clayton, A.J., Lisella, R.S. and Martin, D.G. (1973). Melioidosis: a serological survey in military personnel. *Military Medicine.* January: 24-26.

Clegg, S. and Gerlach, G.K. (1987). Enterobacterial fimbriae. *Journal of Bacteriology.* 169: 934-938.

Coburn, J.A., Dillon, S.T., Iglewski, B.H. and Gill, D.M. (1989). Exoenzyme S of *Pseudomonas aeruginosa* ADP-ribosylates the intermediate filament protein vimentin. *Infection and Immunity.* 57: 996-998.

Coburn, J.A., Kane, A.V., Fieg, L. and Gill, D.M. (1991). *Pseudomonas aeruginosa* exoenzyme S requires a eukaryotic protein for ADP-ribosyltransferase activity. *Journal of Biological Chemistry.* 266: 6438-6446.

Cohen, S.N. and Chang, A.C.Y. (1977). Revised interpretation of the origin of the pSC101 plasmid. *Journal of Bacteriology.* 132: 734-737.

Cohen, S.N., Chang, A.C.Y., Boyer, H.W. and Helling, R.B. (1973). Construction of biologically functional bacterial plasmids *in vitro. Proceedings of the National Academy of Sciences of the United States of America.* 70: 3240-3244.

Cohen, S.N., Chang, A.C.Y. and Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences of the United States of America.* 75:4242-4246.

Colling, M., Nigg, C. and Heckley, R. (1958). Toxins of *Pseudomonas pseudomallei. Journal of Bacteriology.* 76: 122-126.

Collins, J. and Hohn, B. (1978). Cosmids: a type of plasmid gene-cloining vector that is packagable *in vitro* in bacteriophage heads. *Proceedings of the National Academy of Sciences of the United States of America.* 75:4242-4246.

Currie, B.J. (1995). *Pseudomonas pseudomallei*-insulin reaction. *Infection and Immunity.* 63(9): 3745.

Currie, B.J., Fisher, D.A., Howard, D.M. and Burrow, J.N.C. (2000a). Neurological melioidosis. *Acta Tropica.* 74: 145-151.

Currie, B.J., Fisher, D.A., Howard, D.M., Burrow, J.N.C., Selvanayagam, S., Snelling, P.L., Anstey, N.M. and Mayo, M.J. (2000b). The epidemiology of melioidosis in Australia and Papua New Guniea. *Acta Tropica*. 74: 121-127.

Currie, B., Smith-Vaughan, H., Golledge, C., Buller, N., Sriprakash, K.S. and Kemp, D.J. (1994). *Pseudomonas pseudomallei* isolates collected over 25 years from a non-tropical endemic focus show clonality on the basis of ribotyping. *Epidemiology and Infection*. 113: 307-312.

Dance, D.A.B. (1990). Melioidosis. *Reviews in Medical Microbiology*. 1: 143-150.

Dance, D.A.B. (1991). Melioidosis: The tip of the iceberg? *Clinical Microbiology Reviews*. 4(1): 52-60.

Dance, D.A.B. (2000a). Melioidosis. In: Tropical and Exotic Infections (Beeching, N.J., Hart, C.A. and Duerden, B.I. eds.) *Journal of Medical Microbiology*. 49: 5-27.

Dance, D.A.B. (2000b). Melioidosis as an Emerging Global Problem. *Acta Tropica*. 74: 115-119.

Dance, D.A.B. (2000c). Ecology of *Burkholderia pseudomallei* and the interactions between environmental *Burkholderia* spp. and human-animal hosts. *Acta Tropica*. 74: 159-168.

Dance, D.A.B., Davis, T.M.E., Wattanagoon, Y., Chaowagul, W., Saiphan, P., Looareesuwan, S., Wuthiekanun, V. and White, N.J. (1989a). Acute suppurative parotitis caused by *Pseudomonas pseudomallei* in children. *Journal of Infectious Diseases*. 159(4): 654-660.

Dance, D.A.B., Smith, M.D. , Aucken, H. and Pitt, T.L. (1999). Imported melioidosis in England and Wales. *Lancet*. 353: 208.

Dance, D.A.B., Wuthiekanun, V., Naigowit, P. and White, N.J. (1989b). Identification of *Pseudomonas pseudomallei* in clinical practice: use of simple screening tests and API20NE. *Journal of Clinical Pathology*. 42: 645-648.

Darbre, P.D. (1999). *Basic Molecular Biology. Essential Techniques*. John Wiley and Sons. Rickwood, D., (Editor).

De Boer, H.A., Hui, A., Comstock, L.J., Wong, E. and Vasser, M. (1983). Portable Shine-Dalgarno regions: a system for systematic study of defined alterations of nucleotide sequences within *E.coli* ribosome binding sites. *DNA*. 2: 231-241.

Desakorn, V., Smith, M.D., Wuthiekanun, V., Dance, D.A.B., Auken, H., Suntharasamai, P., Rajchanuwon, A. and White, N.J. (1994). Detection of *Pseudomonas pseudomallei* Antigens in Urine for the Diagnosis of Melioidosis. *American Journal of Tropical Medicine and Hygiene*. 51(5): 627-633.

DeShazer, D. and Woods, D.E. (1999). Pathogenesis of melioidosis: use of Tn5-OT182 to study the molecular basis of *Burkholderia pseudomallei* virulence. *Journal of Infectious Diseases*. 16: 91-6.

DeShazer, D., Brett, P.J., Carlyon, R. and Woods, D.E. (1997). Mutagenesis of *Burkholderia pseudomallei* with Tn5-OT182: isolation of motility mutants and molecular characterization of the flagellin structural gene. *Journal of Bacteriology.* 179(7): 2116-2125.

DeShazer, D., Brett, P.J. and Woods, D.E. (1998). The type II O-antigenic polysaccharide moiety of *Burkholderia pseudomallei* lipopolysaccharide is required for serum resistance and virulence. *Molecular Microbiology.* 30(5): 1081-1100.

Desmarchelier, P.M., Dance, D.A.B., Chaowagul, W., Suputtamongkol, Y., White, N.J. and Pitt, T.L. (1993). Relationships among *Pseudomonas pseudomallei* isolates from patients with recurrent melioidosis. *Journal of Clinical Microbiology.* 31 (6): 1592-1596.

Deuschle, U., Kammerer, W., Genz, R. and Bujard, H. (1986). Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structure. *EMBO Journal.* 5: 2987-2994.

Dharakul, T., Songsivilai, S., Viriyachitra, S., Luangwedchakarn, V., Tassaneetritap, B. and Chaowagul, W. (1996). Detection of *Burkholderia pseudomallei* DNA in patients with septicemic melioidosis. *Journal of Clinical Microbiology.* 34(3): 609-614.

Dharakul, T. Tassaneetrithep, B., Trakulosomboon, S. and Songsivilai, S. (1999). Phylogenetic analysis of Ara(+) and Ara(-) *Burkholderia pseudomallei* isolates and development of a multiplex PCR procedure for rapid discrimination between the two biotypes. *Journal of Clinical Microbiology.* 37(6): 1906-1912.

Dharakul, T., Vejbaesya, S., Chaowagul, W., Luangtrakool, P., Stephens, H.A.F. and Songsivilai, S. (1998). HLA-DR and –DR associations with melioidosis. *Human Immunology.* 59: 580-586.

Domenighini, M., Montecucco, C., Ripka, W.C. and Rappuoli, R. (1991). Computer modelling of the NAD binding site of ADP-ribosylating toxins: active site structure and mechanisms of NAD binding. *Molecular Microbiology.* 5: 23-31.

Doolittle, R.F. (1986). *Of URFs and ORFs: a primer on how to analyse derived amino acid sequences.* University Science Books, Mill Valley. USA.

Edmond, K., Currie, B., Brewster, D. and Kilburn, C. (1998). Pediatric melioidosis in tropical Australia. *The Pediatric Infectious Disease Journal.* 17(1): 77-80.

Egan, A.M. and Gordon, D.L. (1996). *Burkholderia pseudomallei* activates complement and is ingested nut not killed by polymorphonuclear leukocytes. *Infection and Immunity.* 64(12): 4952-4959.

Elleman, T.C., Hoyne, P.A., Stewart, D.J., McKern, N.M. and Peterson, J.E. (1986). Expression of pili from *Bacteroides nodosus* in *Pseudomonas aeruginosa. Journal of Bacteriology.* 168: 574-580.

Embi, N., Suhaimi, A., Mohamed, R. and Ismail, G. (1992). Prevalence of antibodies to *Pseudomonas pseudomallei* exotoxin and whole cell antigens in military personnel in Sabah and Sarawak, Malaysia. *Microbiology and Immunology.* 36(8): 899-904.

Fernandez, L.A. and Berenguer, J. (2000). Secretion and assembly of regular surface structures in Gram-negative bacteria. *FEMS Microbiology Reviews.* 24: 21-44.

Fetherston, J.D., Bearden, S.W. and Perry, R.D. (1996). *YbtA*, an AraC-type regulator of the *Yersinia pestis* pesticin/yersiniabactin receptor. *Molecular Microbiology*. 22(2): 315-25.

Finlay, B.B. and Falkow, S. (1989). Common Themes in Microbial Pathogenicity. *Microbiological Reviews*. 53 (2): 210-230.

Forest, K.T. and Tainer, J.A. (1997). Type-4 pilus structure: outside to inside and top to bottom- a minireview. *Gene*. 192: 165-169.

Frazer, A.C. and Curtiss III, R. (1975). Production and utility of bacterial minicells. *Current Topics in Microbiology and Immunology*. 69: 1-4.

Friedland, J.S., Suputtamongkol, Y., Remick, D.G., Chaowagul, W., Strieter, R.M., Kunkel, S.L., White, N.J. and Griffin, G.E. (1992). Prolonged elevation of Interleukin-8 and Interleukin-6 concentrations in plasma and of leukocyte Interleukin-8 mRNA levels during septicemic and localized *Pseudomonas pseudomallei* infection. *Infection and Immunity*. 60 (6): 2402-2408.

Gallegos, M.T., Michan, C. and Ramos, J.L. (1993). The XylS/AraC family of regulators. *Nucleic Acids Research*. 21 (4): 807-810.

Gauthier, Y.P., Thibault, F.M., Paucod, J.C. and Vidal, D.R. (2000). Protease production by *Burkholderia pseudomallei* and virulence in mice. *Acta Tropica*. 74: 215-220.

Gefter, M.L., Becker, A. and Hurwitz, J. (1967). The enzymatic repair of DNA. I. Formation of circular λ DNA. *Proceedings of the National Academy of Sciences of the United States of America*. 58: 240-247.

Geourjon, C. and Deleage, G. (1994). SOPM: a self-optimised method for protein secondary structure prediction. *Protein Engineering*. 7: 157-164.

Gill, D.M. and Dinius, L.L. (1971). Observations on the structure of diphtheria toxin. *Journal of Biological Chemistry*. 246: 1485.

Gillam, S., Astell, C.R. and Smith, M. (1980). Site-specific mutagenesis using oligodeoxyribonucleotides: isolation of a phenotypically silent ϕX174 mutant, with a specific nucleotide deletion, at very high efficiency. *Gene*. 12: 129-137.

Gillis, M., Vanvan, T., Bardin, R., Goor, M., Hebbar, P., Willems, A., Segers, P., Kersters, K., Heulin, T. and Fernandez, M.P. (1995). Polyphasic taxonomy in the genus *Burkholderia* leading to an emended description of the genus and proposition of *Burkholderia vietnamensis* sp. nov. for N-2-fixing isolates from rice in Vietnam. *International Journal of Systematic Bacteriology*. 45(2): 274-289.

Glauser, M.P., Zanetti, G., Baumgartner, J.D. and Cohen, J. (1991). Septic shock: pathogenesis. *Lancet*. 338: 732-736.

Grange, J.M., Fox, A. and Morgan, N.L. (Editors). (1991). *Genetic Manipulation: Techniques and Applications*. Society for Applied Bacteriology. 28. Blackwell Scientific Publications.

Gray, B.L., Smith, D.H., Baldridge, J.S., Harkins, R.N., Vasil, M.L., Chen, E.Y. and Heyneker, H.L. (1984). Cloning, nucleotide sequence and expression in *Escherichia coli* of the exotoxin A structural gene of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences of the United States of America*. 81: 2645-2649.

Groisman, E.A. and Ochman, H. (1996). Pathogenicity islands: bacterial evolution in quantum leaps. *Cell*. 87: 791-794.

Groisman, E.A., Chaio, E., Lipps, C.J. and Heffron, F. (1989). *Salmonella typhimurium phoP* virulence gene is a transcriptional regulator. *Proceedings of the National Academy of Sciences of the United States of America*. 86: 7077-7081.

Haase, A., Janzen, J., Barrett, S. and Currie, B. (1997). Toxin Produced by *Burkholderia pseudomallei* Strains and Correlation with Severity of Melioidosis. *Journal of Medical Microbiology*. 46: 557-563.

Haase, A., Melder, A., Smith-Vaughan, H., Kemp, D. and Currie, B. (1995). RAPD analysis of isolates of *Burkholderia pseudomallei* from patients with recurrent melioidosis. *Epidemiology and Infection*. 115: 115-121.

Hacker, J., Blum-Oehler, G., Muhldorfer, I. And Tschape, H. (1997). Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Molecular Microbiology*. 23 (6): 1089-1097.

Halder, D., Zainal, N., Wah, C.M. and Haq, J.A. (1998). Neonatal meningitis and septicaemia caused by *Burkholderia pseudomallei*. *Annals of Tropical Pediatrics*. 18(2): 161-164.

Hanahan, D. (1985). In: *DNA Cloning Volume 1: A Practical Approach*, Glover, D.M. (ed.), IRL Press, Oxford, pp.113.

Harb, O.S. and AbuKwaik, Y. (1999). Probing the microenvironment of intracellular bacterial pathogens. *Microbes and Infection*. 1(6): 445-453.

Harley, V.S., Dance, D.A.B., Tovey, G. and Drasar, B.S. (1994). Interaction of *Pseudomonas pseudomallei* with macrophages. *Biochemical Society Transactions*. 22: 88S.

Harlow, E. and Lane, D. (1988). *Antibodies: A Laboratory Manual*. Cold Spring Harbor Books, Cold Spring Harbor, USA.

Haussler, S., Nimtz, M., Domke, T., Wray, V. and Steinmetz, I. (1998). Purification and characterization of a cytotoxic exolipid of *Burkholderia pseudomallei*. *Infection and Immunity*. 66(4): 1588-1593.

Haussler, S., Rhode, M. and Steinmetz, I. (1999). Highly resistant *Burkholderia pseudomallei* mall colony variants isolated *in vitro* and in experimental melioidosis. *Medical Microbiology and Immunology*. 188: 91-97.

Hawley, D.K. and McClure, W.R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Research*. 11: 2237-2255.

Haynes, W.C. (1957). Bergey's Manual of Determinative Bacteriology (7th edn.). Baltimore: Williams & Williams, p. 100.

Heckly, R.J. and Nigg, C. (1958). Toxins of *Pseudomonas pseudomallei. Journal of Bacteriology.* 76: 427-436.

Helling, R.B., Goodman, H.M. and Boyer, H.W. (1974). Analysis of endonuclease R.*Eco*RI fragments of DNA from lambdoid bacteriophages and other viruses by agarose gel electrophoresis. *Journal of Virology.* 14: 1235-1244.

Heng, B.H., Goh, K.T., Yap, E.H., Loh, H. and Yoe, M. (1998). Epidemiological surveillance of melioidosis in Singapore. *Annals of Medicine.* 27(4): 478-484.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E. and Holden, D.W. (1995). Simultaneous identification of bacterial virulence genes by negative selection. *Science.* 269: 400-403.

Hobbs, M., Collie, E.S.R., Free, P.D., Livingston, S.P. and Mattick, J.S. (1993). *Pil*S and *Pil*R, a two component transcriptional regulatory system controlling expression of type IV fimbriae in *Pseudomonas aeruginosa. Molecular Microbiology.* 7: 669-682.

Hopp, T.P. and Woods, K.R. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences of the United States of America.* 78: 3824-3828.

Howe, P.W., Holland, H.M., Burrow, J.C.N. and Currie, B.J. (1997). Neurological melioidosis (*Burkholderia pseudomallei*) mimicking Guillain-Barre syndrome. *Anaesthetics and Intensive Care.* 25: 166-167.

Huynh, T.V., Young, R.A. and Davis, R.W. (1985). Construction and screening cDNA libraries in λgt10 and λgt11. In: *DNA Cloning Techniques: A Practical Approach*; Glover, D. (Ed.) IRL Press. Oxford.

Iglewski, B.H., Liu, P.V. and Kabat, D. (1977). Mechanism of action of *Pseudomonas aeruginosa* exotoxin A: adenosine diphosphate-ribosylation of mammalian elongation factor 2 *in vitro* and *in vivo. Infection and Immunity.* 15: 138.

Inglis, T.J.J., Chiang, D., Lee, G.S.H. and Chor-Kiang, L. (1998). Potential Misidentification of *Burkholderia pseudomallei* by API 20NE. *Pathology.* 30: 62-64.

Inglis, T.J.J., Garrow, S.C., Henderson, M., Clair, A., Sampson, J., O'Reilly, L. and Cameron, B. (2000). *Burkholderia pseudomallei* traced to water treatment plant in Australia. *Emerging Infectious Diseases.* 6(1): 1-5.

Ismail, G., Mohamed, R., Rohana, S., Sharifah, H.S.M. and Embi, N. (1991). Antibody to *Pseudomonas pseudomallei* exotoxin in sheep exposed to natural infection. *Veterinary Microbiology.* 27: 277-282.

Ismail, G., Noor Embi, M., Omar, O., Razak., Allen., J.C. and Smith, C.J. (1987). Enzyme immunoassay for the detection of antibody to *Pseudomonas pseudomallei* exotoxin in mice. *FEMS Microbiology Letters.* 40: 27-31.

Jackson, D.A., Symons, R.H. and Berg, P. (1972). Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America.* 69:2904-2909.

Jesudason, M.V., Shanthakumari, R. and John, T.J. (1997). *Burkholderia pseudomallei* – an emerging pathogen in India. *Indian Journal of Medical Microbiology.* 15(1): 1-2.

Jones, A.L., Beveridge, T.J. and Woods, D.E. (1996). Intracellular survival of *Burkholderia pseudomallei*. *Infection and Immunity.* 64(3): 782-790.

Kanai, K., Kondo, E. and Kurata, T. (1996). Affinity and response of *Burkholderia pseudomallei* and *Burkholderia Cepacia* to insulin. *Southeast Asian Journal of Tropical Medicine and Public Health.* 3: 584-591.

Karcher, S.J. (1995). *Molecular Biology: A Project Approach.* Academic Press, Inc. London.

Karplus, P.A. and Schulz, G.E. (1985). Prediction of chain flexibility in proteins. *Naturwissens-chaften.* 72: 212-213.

Kehoe, M., Duncan, J., Foster, T., Fairweather, N. and Dougan, G. (1983). Cloning, expression, and mapping of the *Staphylococcus aureus* α-haemolysin determinant in Escherichia coli K-12. *Infection and Immunity.* 41: 1105-1111.

Kenny, D.J., Russell, P., Rogers, D., Eley, S.M. and Titball, R.W. (1999). *In vitro* susceptibilities of *Burkholderia mallei* in comparison to those of other pathogenic *Burkholderia* spp. *Antimicrobial Agents and Chemotherapy.* 43(11): 2773-2775.

Kessler, C. and Holtke, H.J. (1986). Specificity of restriction endonucleases and methylases – a review. *Gene.* 47: 1-153.

Koonpaew, S., Na Ubol, M., Sirisinha, S., White, N.J. and Chaiyaroj, C. (2000). Genome fingerprinting by pulsed-field gel electrophoresis of isolates of *Burkholderia pseudomallei* from patients with melioidosis in Thailand. *Acta Tropica.* 74: 187-191.

Korbsrisate, S., Suwanasai, N., Leelaporn, A., Ezaki, T., Kawamura, Y. and Sarasombath, S. (1999). Cloning and characterisation of a nonhemolytic phospholipase gene *from Burkholderia pseudomallei*. *Journal of Clinical Microbiology.* 37(11): 3742-3745.

Krishnaswamy, C.S. (1917). Morphia injector's septicaemia. *Indian Medical Gazette.* 52: 296-299.

Kruger, K.M. and Barbieri, J.T. (1995). The family of ADP-ribosylating exotoxins. *Clinical Microbiology Reviews.* 8: 34-37.

Kuhnert, P., Boerlin, P. and Frey, J. (2000). Target genes for virulence assessment for *Escherichia coli* isolates from water, food and the environment. *FEMS Microbiology Reviews.* 24: 107-117.

Kunakorn, M., Jayanetra, P. and Tanphaichitra, D. (1991a). Man-to-man transmission of melioidosis. *Lancet.* 337: 1290-1291.

Kunakorn, M., Petchlai, B., Khupulsup, K. and Naigowit, P. (1991b). Gold Blot for Detection of Immunoglobulin M (IgM) and IgG-Specific Antibodies for Rapid Serodiagnosis of Melioidosis. *Journal of Clinical Microbiology.* 29(9): 2065-2067.

Kurland, C.G. (1987). Strategies for efficieny and accuracy in gene expression. 1. The major codon preference: a growth optimisation strategy. *Trends in Biochemical Sciences.* 12: 126-128.

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology.* 157: 105-132.

Lath, R., Rakshekhar, V. and George, V. (1998). Brain abscess as the presenting feature of melioidosis. *British Journal of Neurosurgery.* 12(2): 170-172.

Leakey, A.K., Ulet, C. and Hirst, R.G. (1998). BALB/c and C57Bl/6 mice infected with virulent *Burkholderia pseudomallei* provide contrasting aminal models for the acute and chrnoc forms of human melioidosis. *Microbial Pathogenesis.* 24: 269-275.

Lee, M.K. and Chua, C.T. (1986). Brain abscess due to *Pseudomonas pseudomallei. Australian and New Zealand Journal of Medicine.* 16(1): 75-77.

Leelarasamee, A. (2000). Melioidosis in Southeast Asia. *Acta Tropica.* 74: 129-132.

Leelarasamee, A. and Bovornkitti, S. (1989). Melioidosis: review and update. *Reviews of Infectious Diseases.* 11: 413-452.

Leimeister-Wachter, M., Domann, E. and Chakraborty, T. (1992). The expression of virulence genes in *Listeria monocytogenes* is thermoregulated. *Journal of Bacteriology.* 174: 947-952.

Lertpatansuwan, N., Sermsri, K., Petkaseam, A., Trakulosomboon, S., Thamlikitkul, V. and Suputtamongkol, Y. (1999). Arabinose-positive *Burkholderia pseudomallei* infection in humans: case report. *Clinical Infectious Diseases.* 28(4): 927-928.

Lew, A.E. and Desmarchelier, P.M. (1993). Molecular typing of *Pseudomonas pseudomallei-* restriction fragment length polymorphisms of ribosomal RNA genes. *Journal of Clinical Microbiology.* 31(3): 533-539.

Li, X. and Hayward, A.C. (1994). Bacterial whole cell protein profiles of the rRNA group II *pseudomonads. Journal of Applied Bacteriology.* 77: 308-318.

Lim, T.K. (1997). Emerging pathogens for pneumonia in Singapore. *Annals of the Acadamy of Medicine, Singapore.* 26: 651-658.

Liu, P.V. (1957). Survey of hemolysin production among species of *Pseudomonads. Journal of Bacteriology.* 74: 718-727.

Liu, P.V. (1973). Exotoxins of *Pseudomonas aeruginosa.* I. Factors that influence the production of exotoxin A. *Journal of Infectious Diseases.* 128: 503-513.

Lobbann, P.E. and Kaiser, A.D.(1973). Enzymatic end to end joining of DNA molecules. *Journal of Molecular Biology.* 78:453-471.

Maguire, G.P., Flavell, H.D., Burrow, J.C.N. and Currie, B.J. (1998). Relapsing neurological melioidosis from the Top End of the Northern Territory. *Australian and New Zealand Journal of Medicine*.28: 219-220.

Mahan, M.J., Slauch, J.M., Hanna, P.C., Camilli, A., Tobias, J.W., Waldor, M.K. and Mekalanos, J.J. (1993). Selection for bacterial genes that are specifically induced in host tissues: the hunt for virulence factors. *Infectious Agents and Disease*. 2: 263-268.

Mahan, M.J., Tobias, J.W., Slauch, J.M., Hanna, P.C., Collier, R.J. and Mekalanos, J.J. (1995). Antibiotic-based detection of bacterial genes that are specifically induced during infection of the host. *Proceedings of the National Academy of Sciences of the United States of America*. 92: 669-673.

Mandel, M. and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*. 53:159-162.

Manzenyuk, O., Volozhantsev, N.V. and Svetoch, E.A. (1994). Identification of *Pseudomonas mallei* bacteria with the help of *Pseudomonas pseudomallei* bacteriophages. *Microbiology*. 63: 303-307.

Markova, N., Kussovski, V. and Radoucheva, T. (1998). Killing of *Pseudomonas pseudomallei* by polymorphonuclear leukocytes and peritoneal macrophages from chicken, sheep, swine and rabbits. *Zentralblatt fur Bakteriologie-International Journal of Medical Microbiology, Virology, Parasitology and Infectious Diseases*. 288: 103-110.

Marrs, C.F., Schoolnik, G., Koomey, J.M., Hardy, J., Rothbard, J. and Falcow, S. (1985). Cloning and sequencing of a *Moraxella bovis* pilin gene. *Journal of Bacteriology*. 163: 132-139.

Martin, R., Hoover, C., Grimme, S., Grogan, C., Holtke, J. and Kessler, C. (1990). A highly sensitive, nonradioactive DNA labelling and detection system. *Biotechniques*. 9: 762-768.

Masoud, H., Ho, M., Schollaarot, T. and Perry, M.B. (1997). Characterization of the Capsular Polysaccharide of *Burkholderia (Pseudomonas) pseudomallei* 304b. *Journal of Bacteriology*. 179(18): 5663-5669.

Matsuura, M., Kawahara, K., Ezaki, T. and Nakano, M. (1996). Biological activities of lipopolysaccharide of *Burkholderia* (*Pseudomonas*) *pseudomallei*. *FEMS Microbiology Letters*. 137: 79-83.

McCormick, J.B., Sexton, D.J., McMurray, J.G., Carey, E., Hayes, P. and Feldman, R.A. (1975). Human-to-human transmission of *Pseudomonas pseudomallei*. *Annals of Internal Medicine*. 83: 512-513.

Meinkoth, J. and Wahl, G. (1984). Hybridisation of nucleic acids immobilized on solid supports. *Analytical Biochemistry*. 138: 267-284.

Mertz, J.E. and Davis, R.W. (1972). Cleavage of DNA by R_I restriction endonuclease generates cohesive ends. *Proceedings of the National Academy of Sciences of the United States of America*. 69: 3370-3374.

Meyer, T.F., Billyard, E., Haas, R., Storzbach, S. and So, M. (1984). Pilus gene of *Neisseria gonorrhoeae*: chromosomal organisation and DNA sequence. *Proceedings of the National Academy of Sciences of the United States of America.* 81: 6110-6114.

Miller, J.F., Kukral, A.M. and Mekalanos, J.J. (1989). A two component regulatory system (*phoP* and *phoQ*) controls *Salmonella typhimurium* virulence. *Proceedings of the National Academy of Sciences of the United States of America.* 86: 5054-5058.

Minassian, M.A., Gage, A., Price, E. and Sefton, A.M. (1999). Impenem for the treatment of melioidosis. *International Journal of Antimicrobial Agents.* 12(3): 263-265.

Mohamed, R., Nathan, S., Embi, N., Razak, N. and Ismail, G. (1989). Inhibition of Macromolecular Synthesis in Cultured Macrophages by *Pseudomonas pseudomallei* Exotoxin. *Microbiology and Immunology.* 33(10): 811-820.

Mollaret, H.H. (1988). "L'affaire du jardin des plantes" ou comment la melioidose fit son appaition en France. *Med. Mal. Infect.* 18: 643-645.

Moore, R.A., DeShazer, D., Reckseidler, S., Weissman, A. and Woods, D.E. (1999). Efflux-mediated aminoglycoside and macrolide resistance in *Burkholderia pseudomallei. Antimicrobial Agents and Chemotherapy.* 43(3): 465-470.

Mosleh, I.M., Huber, L.A., Steinlein, P., Pasquali, C., Gunther, D. and Meyer, T.F. (1998). *Neisseria gonorrhoeae* porin modulates phagosome maturation. *Journal of Biological Chemistry.* 273(52) :35332-8.

Muller, A., Gunther, D., Dux, F., Naumann, M., Meyer, T.F. and Rudel, T. (1999). *Neisserial porin* (PorB) causes rapid calcium influx in target cells and induces apoptosis by the activation of cysteine proteases. *EMBO Journal.* 18(2): 339-352.

Murray, N.E. (1991). Special uses of lambda phage for molecular cloning. In: *Methods in Enzymology* (Miller, J.H. Ed.).Vol. 204, pp. 280-301. Academic Press, San Diego.

Naigowit, P., Kurata, T., Wangroonsub, P., Petkanjanapong, V., Kondo, E. and Kanai, K. (1993). Application of indirect immunofluorescence microscopy to colony identification of *Pseudomonas pseudomallei. Asian Pacific Journal of Allergy and Immunology.* 11(2): 149-154.

Neilson, H., Engelbrecht, J., Brunak, S. and Von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering.* 10: 1-6.

Newman, A.J., Linn, T.G. and Hayward, R.S. (1979). Evidence for cotranscription of the RNA polymerase genes *rpo*BC with a ribosome protein of *Escherichia coli. Molecular and General Genetics.* 169: 195-204.

Nigg, C., Heckly, R.J. and Colling, M. (1955). Toxin produced by *Malleomyces pseudomallei. Proceedings of the Society for Experimental Biology and Medicine.* 89: 17-20.

Norton, R., Roberts, B., Freeman, M., Wilson, M., Ashhurst-Smith, C., Lock, W., Brookes, O. and La Brooy, J. (1998). Characterization and Molecular Typing of *Burkholderia pseudomallei*: Are Disease Presentations of Melioidosis Clonally Related? *FEMS Immunology and Medical Microbiology*. 20: 37-44.

Old, R.W. and Primrose, S.B. (1989). *Principles of gene manipulation: an introduction to genetic engineering*. Fourth edn. Blackwell Scientific Publications, Oxford.

Ottemann, K.M. and Miller, J.F. (1997). Roles for motility in bacterial-host interactions. *Molecular Microbiology*. 24(6): 1109-1117.

Otterbein, C.K., Splettstoesser, W.D., Linde, H.J., Grunow, R., Wolf, H., Finke, E.J. and Neubauer, H. (1998). Development and characteriztion of a murine monoclonal antibody reactive with a 64 ka somatic antigen of *Burkholderia cepacia*. *Hybridoma*. 17(2): 143-150.

Ottow, J.C.G. (1975). Ecology, physiology and genetics of fimbriae and pili. *Annual Review of Microbiology*. 29: 79-108.

Padiglione, A., Ferris, N., Fuller, A. and Spelman, D. (1998). Brain abscesses caused by *Burkholderia pseudomallei*. *Journal of Infection*. 36(3): 335-337.

Pappenheimer, A.M. (1955). The pathogenesis of diphtheria. In: *Mechanisms of Microbial Pathogenicity*, 5th Symp. Soc. Micrbiol., Howie, J.W. and O'Hea, A.J., Eds., 40.

Paranchych, W. and Frost, L. (1988). The physiology and biochemistry of pili. *Advances in Microbial Physiology*. 29: 53-114.

Parker, J.M.R., Guo, D. and Hodges, R.S. (1982). New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*. 25: 5425-5432.

Pasloske, B.L., Finlay, B.B. and Paranchych, W. (1985). Cloning and sequencing of the *Pseudomonas aeruginosa* PAK pilin gene. *FEBS Letters*. 183: 408-412.

Patrick, S. and M.J. Larkin. (1995). *Immunological and Molecular Aspects of Bacterial Virulence*. John Wiley & Sons.

Pelicic, V., Reyrat, J. and Gicquel, B. (1998). Genetic advances for studying *Mycobacterium tuberculosis* pathogenicity. *Molecular Microbiology*. 28 (3): 413-420.

Percheron, G., Thibault, F., Paucod, J.C. and Vidal, D. (1995). *Burkholderia pseudomallei* requires Zn^{2+} for optimal exoprotease production in chemically defined media. *Applied and Environmental Microbiology*. 61(8): 3151-3153.

Perry, .B., MacLean, L., Schollaardt, T., Bryan, L.E. and Ho, M. (1995). Structural characterisation of the lipopolysaccharide O antigens of *Burkholderia pseudomallei*. *Infection and Immunity*. 63(9): 3348-3352.

Petkanjanapong, V., Naigowit, P., Kondo, E. and Kania, K. (1992). Use of endotoxin antigens in enzyme-linked immunosorbant assay for the diagnosis of *P.pseudomallei* infections (melioidosis). *Asian Pacific Journal of Allergy and Immunology*. 10: 145-150.

Pitt, T.L., Trakulosomboon, S. and Dance, D.A.B. (2000). Molecular phylogeny of *Burkholderia pseudomallei*. *Acta Tropica*. 74: 181-185.

Plum, G. and Clark-Curtiss, J.E. (1994). Induction of *Mycobacterium avium* gene expression following phagocytosis by human macrophages. *Infection and Immunity*. 62: 3135-3145.

Pongsunk, S., Thirawattanasuk, N., Piyasangthong, N. and Ekpo, P. (1999). Rapid identification of *Burkholderia pseudomallei* in blood cultures by a monoclonal antibody assay. *Journal of Clinical Microbiology*. 37(11): 3662-3667.

Proctor, R.A., Denlinger, L.C. and Bertics, P.J. (1995). Lipopolysaccharide and bacterial virulence. In: *Virulence Mechanisms of Bacterial Pathogens* (Second Edn.) American Society for Microbiology. Washington DC.

Pruksachartvuthi, S., Aswapokee, N. and Thankerngpol, K. (1990). Survival of *Pseudomonas pseudomallei* in human phagocytes. *Journal of Medical Microbiology*. 31: 109-114.

Pugsley, A.P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiology and Molecular Biology Reviews*. 57: 50-108.

Punyagupta, S., Atthasampunna, P., Jayanetra, P. and Kurathong, S. (1976). Melioidosis: A medical problem among the Thai? *Journal of the Medical Council of Thailand 1975*. 5: 419-435. In Thai.

Rao, V.K. and Progulske-Fox, A. (1993). Identification of a novel type IV pilin gene arrangement: cloning and sequence of two N-methylphenylalanine pilin genes from *Eikenella corrodens*. *Journal of General Microbiology*. 139: 651-660.

Rappuoli, R. and Pizza, M. (1991). Structure and evolutionary aspects of ADP-ribosylating toxins. In: *Sourcebook of Bacterial Protein Toxins* pp. 1-21. Academic Press Ltd.

Rattanathongkom, A., Sermswan, R.W. and Wongratanacheewin, S. (1997). Detection of *Burkholderia pseudomallei* in blood samples using polymerase chain reaction. *Molecular and Cellular Probes*. 11: 25-31.

Rimington, R.A. (1962). Melioidosis in north Queensland. *Medical Journal of Australia*. 1: 50-53.

Redfearn, M.S., Palleroni, N.J. and Stanier, R.Y. (1966). A comparative study of *Pseudomonas pseudomallei* and *Bacillus mallei*. *Journal of General Microbiology*. 43: 293-313.

Rouquette, C., Harmon, J.B. and Shafer, W.M. (1999). Induction of the *mtr*CDE-encoded efflux pump system of *Neisseria gonorrhoeae* requires MtrA, an AraC-like protein. *Molecular Microbiology*. 33(3): 651-8.

Rubin, H.L., Alexander, A.D. and Yagar, R.H. (1963). Melioidosis – a military medical problem? *Military Medicine*. 128: 538-542.

Russel, M. (1998). Macromolecular assembly and secretion across the bacterial cell envelope: type II secretion systems. *Journal of Molecular Biology.* 279: 485-499.

Russell, M.A. and Darzins, A. (1994). The pile gene product of *Pseudomonas aeruginosa,* required for pilus biogenesis, shares amino acid sequence identity with the N-termini of Type IV prepilin proteins. *Molecular Microbiology.* 13 (6): 973-985.

Salyers, A.A. and Whitt, D.D. (1994). Bacterial Pathogenesis. A Molecular Approach. ASM Press.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: a laboratory manual.* Cold Springs Harbour Laboratroy Press, Cold Springs Harbour, USA.

Sancar, A., Hack, A.M. and Rupp, W.D. (1979). Simple method for identification of plasmid-coded proteins. *Journal of Bacteriology.* 137: 692-693.

Sanford, J.P. (1990). In: Mandell, G.L., Douglas Jr., R.G. and Bennett, J.E. (Eds.) Principles and Practice of Infectious Diseases. Churchill Livingstone, New York, pp 1692-1696.

Sanford, J.P. (1995). In:Mandell, G.L., Douglas, R.G. and Bennett, J.E. (Eds.) Principles and Practice of Infectious Diseases (Fourth edn.). Churchill Livingstone, New York, pp 2003-2009.

Sanger, F., Nicklen, S. and Coulsen, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America.* 74: 5463-5467.

Seigfried, L. and Kmetova, M. (1997). Role of virulence factors in pathogenesis of *Escherichia coli* infections. *Biologica Bratislava.* 52(6): 697-705.

Sexton, M.M. and Jones, A.L. (1994). Purification and Characterisation of a Protease from *Pseudomonas pseudomallei. Canadian Journal of Microbiology.* 40: 903-910.

Sharp, R.A., Sugden, B. and Sambrook, J. (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry.* 12:3055-3063.

Shaw, C.E. and Taylor, R.K. (1990). *Vibrio cholera* O395 *tcp*A pilin gene sequence and comparison of predicted protein structural features to those of type IV pilins. *Infection and Immunity.* 58: 3042-3049.

Shea, J.E., Hensel, M., Gleeson, C. and Holden, D.W. (1996). Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium. Proceedings of the National Academy of Sciences of the United States of America.* 39: 2593-2597.

Shostak, S.(1998). Death of Life. The Legacy of Molecular Biology. MacMillan Press Ltd.

Simpson, A.J.H. and Wuthiekanun, V. (2000). Interaction of insulin with *Burkholderia pseudomallei* may be caused by a preservative. *Journal of Clinical Pathology.* 53(2): 159-160.

Simpson, A.J.H., Suputtamongkol, Y., Smith, M.D., Angus, B.J. and Rajanuwong, A., Wuthiekanun, V., Howe, P.A., Walsh, A.L., Chaowagul, W. and White, N.J. (1999). Comparison of impenem and ceftazidime as therapy for severe melioidosis. *Clinical Infectious Diseases*. 29(2): 381-387.

Singer, M. and Berg, P. (eds.) (1991). *Genes and genomes* (Third Edn.). Blackwell Scientific Publications. Oxford.

Sirisinha, S., Anuntagool, N., Intachote, P., Wuthiekanun, V., Puthucheary, S.D. and Vadivelu, J. (1998). Antigenic difference between clinical and environmental isolates of *Burkholderia pseudomallei*. *Microbiology and Immunology*. 42: 731-737.

Smith, C.J., Allen, J.C., Noor Embi, M., Othman, O., Razak, N. and Ismail, G. (1987). Human melioidosis: an emerging medical problem. *MIRCEN Journal*. 3: 343-366.

Smith, C.J., Jones, M., Dance, D.A.B., Chaowagul, W. and White, N.J. (1991). Diagnosis of melioidosis by means of an ELISA detecting antibodies to *Pseudomonas pseudomallei* exotoxin- preliminary assay evaluation. *World Journal of Microbiology and Biotechnology*. 7(1): 29-36.

Smith, H.O. and Wilcox, K.W. (1970). A restriction enzyme from *Haemophilus influenzae*. I. Purification and general properties. *Journal of Molecular Biology*. 51: 379-391.

Songsivilai, S. and Dharakul, T. (2000). Multiple replicons constitute the 6.5-megabase genome of *Burkholderia pseudomallei*. *Acta Tropica*. 74: 169-179.

Soonhammer, E.L.L., Von Heijne, G. and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. In: *Proceeding of the Sixth International Conference on Intelligent Systems for molecular Biology*, pp. 175-182. Glasgow, J., Littlejohn, T., Major, F., Lathrop, D., Sankoff, D. and Sensen, C. Menlo Park, CA: AAAI Press.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*. 98: 503-517.

Stanton, A.T. and Fletcher, W. (1921). Melioidosis, a new disease of the tropics. *Transactions of the Fourth Congress of the Far Eastern Association of Tropical Medicine*. 2: 196-198.

Stanton, A.T. and Fletcher, W. (1932). Melioidosis. Studies from the Institute for Medical Research, Federated Malay States, Vol. 21. John Bale & Sons and Danielson Ltd., London.

Steinmetz, I., Nimtz, M., Wray, V., Haussler, S., Reganzerowski, A. and Brenneke, B. (2000). Exopolysaccharides of *Burkholderia pseudomallei*. *Acta Tropica*. 74: 211-214.

Steinmetz, I., Reganzerowski, A., Brenneke, B., Haussler, S and Simpson, A. (1999). Rapid identification of *Burkholderia pseudomallei* by latex agglutination based on an exopolysaccharide-specific monoclonal antibody. *Journal of Clinical Microbiology*. 37(1): 225-228.

Steitz, J.A. (1979). Genetic signals and nucleotide sequences in messenger RNA. In: *Biological Regulation and Development. 1. Gene Expression*. Ed. Goldberger, R.F., pp. 349-399. Plenum Press, New York.

Straus, D. and Ausubel, F.M. (1990). Genomic subtraction for cloning DNA corresponding to deletion mutants. *Proceedings of the National Academy of Sciences of the United States of America.* 87: 1889-1893.

Straus, J.M., Groves, M.G., Mariappan, M. and Ellison, D.W. (1969). Melioidosis in Malaysia. Distribution of *Pseudomonas pseudomallei* in soil and surface water. *American Journal of Tropical Medicine and Hygiene.* 18: 698-702.

Strom, M.S. and Lory, S. (1993). Structure-function and biogenesis of the Type IV pili. *Annual Review of Microbiology.* 47: 565-96.

Suputtamongkol, Y., Hall, A.J., Dance, D.A.B., Chaowagul, W., Rajchanuvong, A., Smith, M.D. and White, N.J. (1994). The epidemiology of melioidosis in Ubon Ratchatani, northeast Thailand. *International Journal of Epidemiology.* 23: 1082-1090.

Tan, N.G., Ang, B. and Sethi, D.S. (1997). An Unusual Case of Sorethroat: Nasopharyngeal Melioidosis. *Singapore Medical Journal.* 38(5): 223-225.

Taylor, R.K., Miller, V.L., Furlong, D.B. and Mekalanos, J.J. (1987). Use of *phoA* gene fusions to identify a pilus colonization factor co-ordinately regulated with cholera toxin. *Proceedings of the National Academy of Sciences of the United States of America..* 84: 2833-2837.

Thamprajamchit, S., Chetchotisakd, P. and Thinkhamrop,B. (1998). Cefoperazone/Sulbactam + Co-trimoxazole vs Ceftazidime + Co-trimoxazole in the Treatment of Severe Melioidosis: A Randomized, Double-Blind, Controlled Study. *Journal of the Medical Association of Thailand.* 81 (4): 265-271.

Thomas, A.D., Norton, J.H., Forbes-Faulkner, J.C. and Woodland, G. (1981). Melioidosis in an intensive piggery. *Australian Veterinary Journal.* 57: 144-145.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research.* 22: 4673-4680.

Thummakul, T., Wilde, H. and Tantawichien, T. (1999). Melioidosis, an environmental and occupational hazard in Thailand. *Military Medicine.* 164(9): 658-662.

Timblin, C., Battey, J. and Kuehl, W.M. (1990). Application for PCR technology to subtractive cDNA cloning: identification of genes expressed specifically in murine plasmacytoma cells. *Nucleic Acids Research.* 18: 1587-1593.

Tong, S., Yang, S., Lu, Z. and He, W. (1996). Laboratory Investigation of Ecological Factors Influencing the Environmental Presence of *Burkholderia pseudomallei. Microbiology and Immunology.* 40(6): 451-453.

Trakulsomboon, S., Dance, D.A.B., Smith, M.D., White, N.J. and Pitt, T.L. (1997). Ribotype differences between clinical and environmental isolates of *Burkholderia pseudomallei. Journal of Medical Microbiology.* 46: 565-570.

Trakulsomboon, S,. Pitt, T.L. and Dance, D.A.B. (1994). Molecular typing of *Pseudomonas pseudomallei* from imported primates in Britain. *Veterinary Record.* 135: 65-66.

Tyler, S.D., Strathdee, C.A., Rozee, K.R. and Johnson, W.M. (1995). Oligonucleotide Primers Designed to Differentiate Pathogenic *Pseudomonads* on the Basis of the Sequencing of Genes Coding for 16S-23S rRNA Internal Transcribed Spacers. *Clinical and Diagnostic Laboratory Immunology*. 2(4): 448-453.

Ulett, G.C., Ketheesan, N. and Hirst, R.G. (1998). Macrophage-lymphocyte interactions mediate anti-*Burkholderia pseudomallei* activity. *FEMS Immunology and Medical Microbiology*. 21: 283-286.

Ulett, G.C., Ketheesan, N. and Hirst, R.G. (2000). Proinflammatory cytokine mRNA responses in experimental *Burkholderia pseudomallei* infection in mice. *Acta Tropica*. 74: 229-234.

Urakami, T., Itoyoshida, C., Araki, H., Kijima, T., Suzuki, K.I. and Komagata, K. (1994). Transfer of *Pseudomonas plantarii* and *Pseudomonas glumae* to *Burkholderia* as *Burkholderia* spp. and description of *Burkholderia vandii* sp. nov. *International Journal of Systematic Bacteriology*. 44(2): 235-245.

Vadivelu, J., Puthucheary, S.D., Drasar, B.S., Dance, D.A.B. and Pitt, T.L. (1998). Stability of strain genotypes of *Burkholderia pseudomallei* from patients with single and recurrent episodes of melioidosis. *Tropical Medicine and International Health*. 3 (7): 518-521.

Valdivia, R.H. and Falkow, S. (1998). In: *Methods in Microbiology*. Vol. 27. pp. 3-12. Academic Press Ltd.

Van Ness, B.G., Howard, J.B. and Bodley, J.W. (1980). ADP-ribosylation of elongation factor 2 by diphtheria toxin. NMR spectra and proposed structures of ribosyl-dipthamide and its hydrolysis products. *Journal of Biological Chemistry*. 255: 10710-10716.

Vedros, N.A., Chow, D. and Liong, E. (1988). Experimental Vaccine Against *Pseudomonas pseudomallei* Infections in Captive Cetaceans. *Diseases of Aquatic Organisms*. 5: 157-161.

Veljanov, D. and Najdenski, H. (1993). Virulence and susceptibility to phagocytosis of *Pseudomonas pseudomallei* R- and S-forms for ground squirrels. *Acta Microbiologica Bulgarica*. 30: 11-16.

Viallard, V., Poirier, I., Cornoyer, B., Haurat, J., Wiebkin, S., Ophel Keller, K. and Balandreau, J. (1998). *Burkholderia graminis* sp. nov., a rhizospheric *Burkholderia* species, and reassessment of [*Pseudomonas*] *phenazinium*, [*Pseudomonas*] *pyrrocinia* and [*Pseudomonas*] *glathei* as *Burkholderia*. *International Journal of Systematic Bacteriology*. 48(2): 549-563.

Voelker, L.L. and Dybvig, K. (1998). Characterisation of the lysogenic bacteriophage MAV1 from *Mycoplasma arthritidis*. *Journal of Bacteriology*. 180 (22): 5928-5931.

Vorachit, V., Chongtrakool, Arkomsean, S. and Boonsong, S. (2000). Antimicrobial resistance in *Burkholderia pseudomallei*. *Acta Tropica*. 74: 139-144.

Vorachit, M., Lam, K., Jayanetra, P. and Costerton, J.W. (1995). Electron microscopy study of the mode of growth of *Pseudomonas pseudomallei in vitro* and *in vivo*. *American Journal of Tropical Medicine and Hygiene*. 98: 379-391.

Vuddhakul, V., Tharavichitkul, P., Na-Ngam, N., Jitsurong, S., Kunthawa, B., Noimay, P., Noimay, P., Binla, A. and Thanlikitkul, V. (1999). Epidemiology of *Burkholderia pseudomallei* in Thailand. *American Journal of Tropical Medicine and Hygiene.* 60(3): 458-461.

Wajanarogana, S., Sonthayanon, P., Wuthiekanun, V., Panyim, S., Simpson, A.J.H. and Tungpradabkul, S. (1999). Stable Marker on Flagellin Gene Sequences Related to Arabinose Non-Assimilating Pathogenic *Burkholderia pseudomallei. Microbiology and Immunology.* 43 (11): 995-1001.

Walsh, A.L., Smith, M.D., Wuthiekanun, V., Suputtamongkol, Y., Desakorn, V., Chaowagul, W. and White, N.J. (1994). Immunofluorescence Microscopy for the Rapid Diagnosis of Melioidosis. *Journal of Clinical Pathology.* 47: 377-379.

Walsh, A.L. and Wuthiekanun, V. (1996). The laboratory diagnosis of melioidosis. *British Journal of Biomedical Sciences.* 53: 249-253.

Watts, T.H., Kay, C.M. and Paranchych, W. (1983). Spectral properties of three quaternary arrangements of *Pseudomonas* pilin. *Biochemistry.* 22: 3640-3646.

Werney, R., Kinne, J., Haydn Evans, J. and Ul Haq, A. (1997). Melioidosis in a seven year old camel, a new disease in the United Arab Emirates (UAE). *Journal of Camel Practice and Research.* 4(2): 141-143.

White, N.J. (1994). Melioidosis. *Zentralblatt fur Bakteriologie-International Journal of Medical Microbiology, Virology, Parasitology and Infectious Diseases.* 280: 439-443.

White, N.J., Chaowagul, W., Wuthiekanun, V., Dance, D.A.B., Wattanagoon, Y. and Pitakwatchara, N. (1989). Halving of mortality of severe melioidosis by Ceftazidime. *Lancet.* September 23: 697-701.

Whitmore, A. and Krishnaswamy, C.S. (1912). An account of the discovery of a hitherto undescribed infective disease occurring among the population of Rangoon. *Indian Medical Gazette.* 47: 262-267.

Wibawan, I.W.,Lammler, C., Seleim, R.S. abd Pasaribu, F.H. (1993). A haemagglutinating adhesin of group B *streptococci* isolated from cases of bovine mastitis mediates adherence to HeLa cells. *Journal of General Microbiology.* 139(9): 2173-8.

Williams, N. and Skurray, R. (1980). The conjugation system of F-like plasmids *Annual Review of Genetics.*14: 41-76.

Winstanley, C., Hales, B.A., Corkill, J.E., Gallagher, M.J. and Hart, C.A. (1998). Flagellin gene variation between clinical and environmental isolates of *Burkholderia pseudomallei* contrasts with the invariance among clinical isolates. *Journal of Medical Microbiology.* 47: 689-694.

Winstanley, C. and Hart, C.A. (2000). Presence of type III secretion genes in *Burkholderia pseudomallei* correlates with Ara(-) phenotypes. *Journal of Clinical Microbiology.* 38(2): 883-885.

Woods, M.L., Currie, B.J., Howard, D.M., Tierney, A., Watson, A., Anstey, N.M., Philpott, J., Asche, V. and Withnall, K. (1992). Neurological melioidosis: seven cases from the Northern Territory of Australia. *Clinical and Infectious Diseases.* 15: 163-169.

Woods, D.E., DeShazer, D., Moore, R.A., Brett, P.J., Burtnick, M.N., Reckseidler, S.L. and Senkiw, M.D. (1999). Current studies on the pathogenesis of melioidosis. *Microbes and Infection.* 2: 157-162.

Woods, D.E., Jones, A.L. and Hill, P.J. (1993). Interaction of Insulin with *Pseudomonas pseudomallei. Infection and Immunity.* 61(10): 4045-4050.

Wuthiekanun, V., Smith, M.D., Dance, D.A.B., Walsh, A.L., Pitt, T.L. and White, N.J. (1996). Biochemical characteristics of clinical and environmental isolates of *Burkholderia pseudomallei. Journal of Medical Microbiology.* 45: 408-412.

Wuthiekanun, V., Smith, M.D., Dance, D.A.B. and White, N.J. (1995a). Isolation of *Pseudomonas pseudomallei* from soil in north-eastern Thailand. *Transactions of the Royal Society of Tropical Medicine and Hygiene.* 89: 41-43.

Wuthiekanun, V., Smith, M.D. and White, N.J. (1995b). Survival of *Burkholderia pseudomallei* in the absence of nutrients. *Transactions of the Royal Society of Tropical Medicine and Hygiene.* 89: 491.

Yabuuchi, E., Kawamura, Y., Ezaki, T., Ikedo, M., Dejsinlert, S., Fujiwara, N., Naka, T. and Kobayashi, K. (2000). *Burkholderia uboniae* sp. nov., L-arabinose assimilating but different from *B.thailandensis* and *B.vietnamensis. Microbiology and Immunology.* 44(4): 307-317.

Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., Ezaki, T. and Arakawa, M. (1992). Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbiology and Immunology.* 36: 1251-75.

Yabuuchi, E., Kosako, Y., Yano, I., Hotta, H. and Nishiuchi, Y. (1995). Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. nov.: proposal of *Ralstonia pickettii* (Ranston, Pelleroni and Doudoroff, 1973) comb. nov.: *Ralstonia solanacearum* (Smith, 1896) comb. nov. and *Ralstonia eutropha* (Davis, 1969) comb. nov. *Microbiology and Immunology.* 39: 897-904.

Yang, H., Chaowagul, W. and Sokol, P.A. (1991). Siderophore Production by *Pseudomonas pseudomallei. Infection and Immunity.* 59(3): 776-780.

Yang, H., Kool, C. and Sokol, P.A. (1993). Ability of *Pseudomonas pseudomallei* Malleobactin to Acquire Transferrin-Bound, Lactoferrin-Bound and Cell-Derived Iron. *Infection and Immunity.* 61: 656-662.

Yu, S., Fiss, E. and Jacobs, W.R. (1998). Analysis of the exochelin locus in *Mycobacterium smegmatis*: biosynthesis genes have homology with genes of the peptide synthetase family. *Journal of Bacteriology.* 180(17): 4676-4685.

Zimmerman, S.B., Little, J.W., Oshinsky, C.K. and Gellert, M. (1967). Enzymatic joining of DNA strands: A novel reaction of diphosphopyridine nucleotide. *Proceedings of the National Academy of Sciences of the United States of America.* 57: 1841-1848.

Zoller, M.J. and Smith, M. (1983). Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods in Enzymology.* 100: 468-500.