MOLECULAR CHARACTERISATION OF A TWO-COMPONENT REGULATORY SYSTEM FROM *Burkholderia pseudomallei*

by

MAGDY ELSAYED MAHFOUZ B.Sc., M.Sc.

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 the
Public Health Laboratory Service (PHLS)
Derriford hospital, Plymouth, UK

July 2001

MOLECULAR CHARACTERISATION OF A TWO-COMPONENT REGULATORY SYSTEM FROM *Burkholderia pseudomallei*

MAGDY ELSAYED MAHFOUZ

ABSTRACT

Studies were undertaken to clone and characterise a two-component regulatory system from a clinical isolate (204) of the human and animal pathogen *Burkholderia pseudomallei*. A number of genomic libraries were constructed in *E. coli* host-vector systems and screened for the presence of a two-component system using oligonucleotide probes based on nucleotide sequence homology. Fragments of genomic DNA were cloned and sequenced and found to possess two open reading frames (ORFs) that overlap with a single nucleotide and are believed to encode a novel two-component regulatory system. A possible promoter region was identified upstream of the two ORFs, *mrgR* and *mrgS*, which read in the same direction and may represent an operon. The deduced translation of *mrgR* reveals a protein, MrgR, which possesses conserved motifs that are consistent with the phosphorylation domains and DNA-binding helix-turn-helix structure of a family of response regulatory proteins. The deduced translation of *mrgS* reveals that the MrgS protein possesses all the invariant amino acids that characterise other sensor regulatory proteins. Southern hybridisation studies showed that the *mrgRS* locus was present in 19 isolates of *B. pseudomallei* from a wide geographical derivation, but not in any closely related bacterial species, including *Burkholderia thailandensis*. The expression of the two genes was verified using antibodies developed to synthetic peptides based on sequences from the C- and N-terminal regions of MrgR and MrgS, respectively. The specificity of the antibodies was confirmed in Western blotting studies in which almost all of *mrgR* and the proximal quarter of *mrgS* were translationally fused with *malE* (MBP-MrgR and MBP-MrgS) and expressed in *E. coli* K12. The antibodies were used to probe Western blots of cellular and extracellular extracts of different isolates of *B. pseudomallei* and identified multiple bands in whole-cell lysates. The sizes of two of these bands were 24 kDa and 115 kDa, which may represent the unprocessed forms of MrgR and MrgS, respectively. It was proposed that the other bands represented either isoforms or degradation products of the full-length proteins. The recognition of all bands was abolished following pre-incubation of the antibodies with the immunising peptide but remained unaffected if an irrelevant peptide was used for this purpose. Western blot analysis demonstrated that serum antibodies from a patient with acute melioidosis recognised MBP-MrgR but not MBP-MrgS suggesting a possible role for MrgR in the disease process. The expression of *mrgR* and *mrgS* was found to be constitutive in *B. pseudomallei* that had been cultured using different combinations of temperature, pH and NaCl suggesting that the genes perform a number of biological functions. There is some evidence that at 42°C the processing of MrgR and MrgS may be altered and the possible mechanisms for this are discussed. *B. pseudomallei* grew better at 42°C and pH 5 and less well at 25°C and pH 8 and this was influenced by NaCl concentration partly reflecting the environmental distribution and intracellular nature of the pathogen. Environmental and clinical isolates of *B. pseudomallei* differed in the pH optimum for growth at 42°C. The DNA flanking the *mrgRS* locus in isolate 204 was cloned, sequenced, and seven ORFs were identified including a transcriptional regulatory gene similar to *bvgR* of *Bordetella pertussis*. Southern blot analysis using three different DNA probes revealed restriction fragment length polymorphisms (RFLPs) in the region downstream of *mrgRS*. Two distinct RFLP patterns were identified among 16 different isolates of *B. pseudomallei*. The potential effects of this variation on gene expression and protein function await further investigation.

LIST OF CONTENTS

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX II

APPENDIX IV

APPENDIX V

APPENDIX VI

APPENDIX VII

LIST OF TABLES

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX I

APPENDIX II

APPENDIX III

APPENDIX IV

APPENDIX V

APPENDIX VI

ABBREVIATIONS

A	adenine
ADP	adenine diphosphate
Ara	arabinose
ATP	adenosine triphosphate
B&D	Birnboim & Doly
bp	base pair
BSA	bovine serum albumin
C	cytosine
$CaCl_2$	calcium chloride
cfu	colony forming units
CIAP	calf intestinal alkaline phosphatase
CL3	Containment Level 3 Laboratory
cm	centimetre
CNS	central nervous system
CsCl	caesium chloride
dH_2O	distilled water
Da	Dalton
DAB	3,3'-diaminobenzidene tetrahydrochloride
DMF	dimethylformamide
°C	degree centigrade
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
ECP	extracellular products
EDTA	ethylenediamine tetra acetic acid
ELISA	enzyme-linked immunosorbent assay
Eps	exopolysaccharides
EtBr	ethidium bromide
FCA	Freund's complete adjuvant
G	guanine
g	gram
x g	acceleration due to gravity

h	hour
H_2O	water
HCl	hydrochloric acid
HEXCO	hexaminecobalt chloride
HPK	histidine protein kinase
IAA	isoamyl alcohol
IFN	interferon
Ig	immunoglobulin
IL	interleukin
IUPA	International Union of Pure and Applied Chemistry (symbols)
IPTG	isopropyl-β-D-thiogalactopyranoside
K	thousand
kb	kilobase pairs
kDa	kilodalton
kg	kilogram
L	litre
λ	lambda
LB	Luria-Bertani
LC	lethal concentration
LPS	lipopolysaccharide
M	molar
mA	milliampere
MAb	monoclonal antibody
µg	microgram
mg	milligram
µl	microlitre
µM	micromolar
µm	micrometer
MBP	maltose-binding protein
MEE	multilocus enzyme electrophoresis
MgSO4	magnesium sulphate
min	minute
ml	millilitre

mM	millimolar
mm	millimetre
MrgR	putative *B. pseudomallei* response regulator
MrgS	putative *B. pseudomallei* sensor regulator
mRNA	messenger ribonucleic acid
NCBI	National Centre for Biotechnology Information
NaCl	sodium chloride
NaOH	sodium hydroxide
ng	nanogram
NBT	nitroblue tetrazolium salt
OD	optical density
OD_{600}	optical density at 600 nm wavelength
ORF	open reading frame
p	statistical probability
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
%	percent
PFGE	pulsed field gel electrophoresis
PHLS	Public Health Laboratory Service
pfu	plaque forming units
PIR	Protein Identification Resource
pmol	picomolar
RAPD	randomly amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rRNA	ribosomal RNA
RNase	ribonuclease
rpm	revolutions per minute
RR	response regulator
RT	room temperature (21-25°C)
s	second

SDS	sodium dodecyl sulphate
SSC	salt sodium citrate
T	thymine
TBE	tris-borate-EDTA
TcF	two-component flanking protein
TE	tris-EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	tris (hydroxymethyl) aminomethane
tRNA	transfer ribonucleic acid
u	unit
U	uracil
UV	ultraviolet
V	volt
v/v	volume/volume
W	watts
w/v	weight/volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

ACKNOWLEDGEMENTS

I would like to acknowledge a number of individuals who kindly contributed to the development of this work:

Dr Martyn Gilpin for his supervision, enthusiasm, and encouragement during the conduct and writing of this thesis.

Dr Hilton Grayson for directing every step of this work and his meticulous care during the performance of this study. Also for correcting my English, invaluable constructive criticism and very helpful discussion during the research project.

Dr Jack Harris for his initial encouragement, supervision and support.

Dr David Dance for his supervision and interest in my studies.

I would like to thank my colleagues from the Molecular and Environmental Microbiology Research Group (MEMR), who have all been great Lab mates for four years. Clare Davies, my first and only friend for over 12 months, Sarah Alexander, and Laura Biggs for their continuous support and jokes.

My thanks extend to Dr Michele Kiernan, Dr Annette Wrathmell, and Dr Janet Roper for their invaluable comments and discussions and to Tean Mitchell, Marion Pill, Cynthia Palmer, Julie Soane, Tara Arrowsmith, Jo Carter, Pete Bond and Paul Russell. Also to my Egyptian friends Ahmed Abada and Abd El-Wahab Abd El-Wareth. Not forgetting the skilful technical assistance of Mrs Lynne Gay, which was greatly appreciated.

I am also grateful to all the staff of the Department of Biological Sciences, particularly Dr Colin Munn the Head of the Department, at the University of Plymouth for providing the necessary facilities.

This study would not have been possible without a grant from the Egyptian government, hence I wish to extend many thanks for their sponsorship.

I have to thank the members of staff at Tanta University, Egypt, who encouraged me to undertake this research project. Without their help I probably would have missed this great opportunity to study and to live abroad and meet new people.

And last but not least, I would like to thank my parents and my little family, Heba, Ahmed and Nahla for their patience and support during a task that inevitably proved more difficult than expected.

AUTHOR'S 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 was financed by the Egyptian government, with the aid of the Egyptian Education and Culture Bureau, London, UK.

Relevant scientific meetings and conferences were regularly attended and at which work was often presented.

PUBLICATIONS

M. E. Mahfouz, T. H. Grayson, D. A. B. Dance & M. L. Gilpin. The organisation of a two-component regulatory system from *Burkholderia pseudomallei* is remarkably similar to the *Bordetella* virulence regulon (in preparation for submission).

M. E. Mahfouz, T. H. Grayson, D. A. B. Dance & M. L. Gilpin. A restriction fragment length polymorphism in *lipA* the putative phospholipase A1 gene of *Burkholderia pseudomallei* separates isolates into two distinct groups (in preparation for submission).

M. E. Mahfouz, T. H. Grayson, D. A. B. Dance & M. L. Gilpin. The effects of pH, salt and temperature on the expression of the two-component regulatory genes *mrgR* and *mrgS* in *Burkholderia pseudomallei* (in preparation for submission).

CONFERENCES AND MEETINGS ATTENDED

2nd International Congress on Melioidosis, Bangkok, Thailand, 22-25 November, 1998.

143rd Ordinary meeting of the Society of General Microbiology, Edinburgh, Scotland, 12-16 April, 1999.

146th Ordinary meeting of the Society of General Microbiology, Warwick, England, 10-14 April, 2000. *

1st Ordinary meeting of the Peninsula Medical School, Plymouth, England, 6 June, 2001. *

UK Melioidosis Meetings

November 1997, London School of Hygiene and Tropical Medicine.

May 1998, Oxford.

December 1998, Collindale Public Health Laboratory. *

May 1999, Derriford Hospital, Plymouth.

January 2000, Imperial College, London. *

June 2000, CBC Porton Down, Salisbury.

* Indicates when work was presented by the author

Signed...Hagaly.....

Date.....23/8/01.....

CHAPTER 1

INTRODUCTION

Burkholderia pseudomallei is a saprophytic gram-negative bacillus taxonomically placed within the family Burkholderiaceae. The bacterium is an opportunistic pathogen that is capable of intracellular survival and causes melioidosis, a frequently fatal disease of humans and animals, which can be difficult to diagnose. *B. pseudomallei* is resistant to many antibiotics and prolonged multiple antibiotic treatment is often not fully effective and relapses are common (Dance, 1990). There is currently no licensed vaccine for melioidosis. Although the pathogen is mainly distributed in the soil and water of tropical regions, especially south-east Asia and northern Australia, it is able to survive and grow in a wide range of environments (Dance, 1991). The disease encompasses a broad spectrum of clinical symptoms and outcomes, including long periods of latency up to 29 years (Leelarasamee, 1998). Melioidosis is regarded as an emerging disease and there is some concern that the disease may become more widespread with the increasing growth in world trade, tourism and travel. Up to 7 million people are at risk in north-east Thailand alone (Currie *et al.*, 2000b). The attention focused on *B. pseudomallei* in recent years reflects these concerns.

Although a number of components have been recently characterised that may contribute to the virulence of *B. pseudomallei* and the pathogenesis of melioidosis, there is still a general lack of knowledge about the pathogen (Brown & Beacham, 2000). In particular, how the organism is capable of successfully fulfilling its' role as a saprophyte and a pathogen. This unique behaviour stimulated our interest in the ways in which *B. pseudomallei* interacts with its' surrounding environment. Bacteria possess highly complex systems for controlling their growth and the production of the components that are required for survival, in response to their environment. They respond to environmental change by altering gene expression via a number of sensory and regulatory pathways that can act either individually or in concert with other regulatory networks. Two-component signal transduction systems provide the most common means employed by bacteria for regulating

these adaptive responses to their environment (Stock *et al.*, 2000). In their simplest form, two-component signal systems elegantly combine the functions of a sensor, transducer and transcriptional regulator within two proteins. Without such sensory pathways bacteria would be unable to survive the onset of harsh or deleterious conditions. In order to survive and colonise a wide variety of environments, *B. pseudomallei* must be equipped with mechanisms that allow the bacterium to firstly, sense and respond to substantial changes in its' environment; secondly, suitably modify its' metabolism and growth potential; and thirdly, successfully adapt to these new circumstances. An improved understanding of the molecular basis for these adaptive responses will provide valuable information about how the bacterium survives and causes disease and assist in the formulation of new approaches for interfering with these processes. Therefore, the aims of this thesis were to:

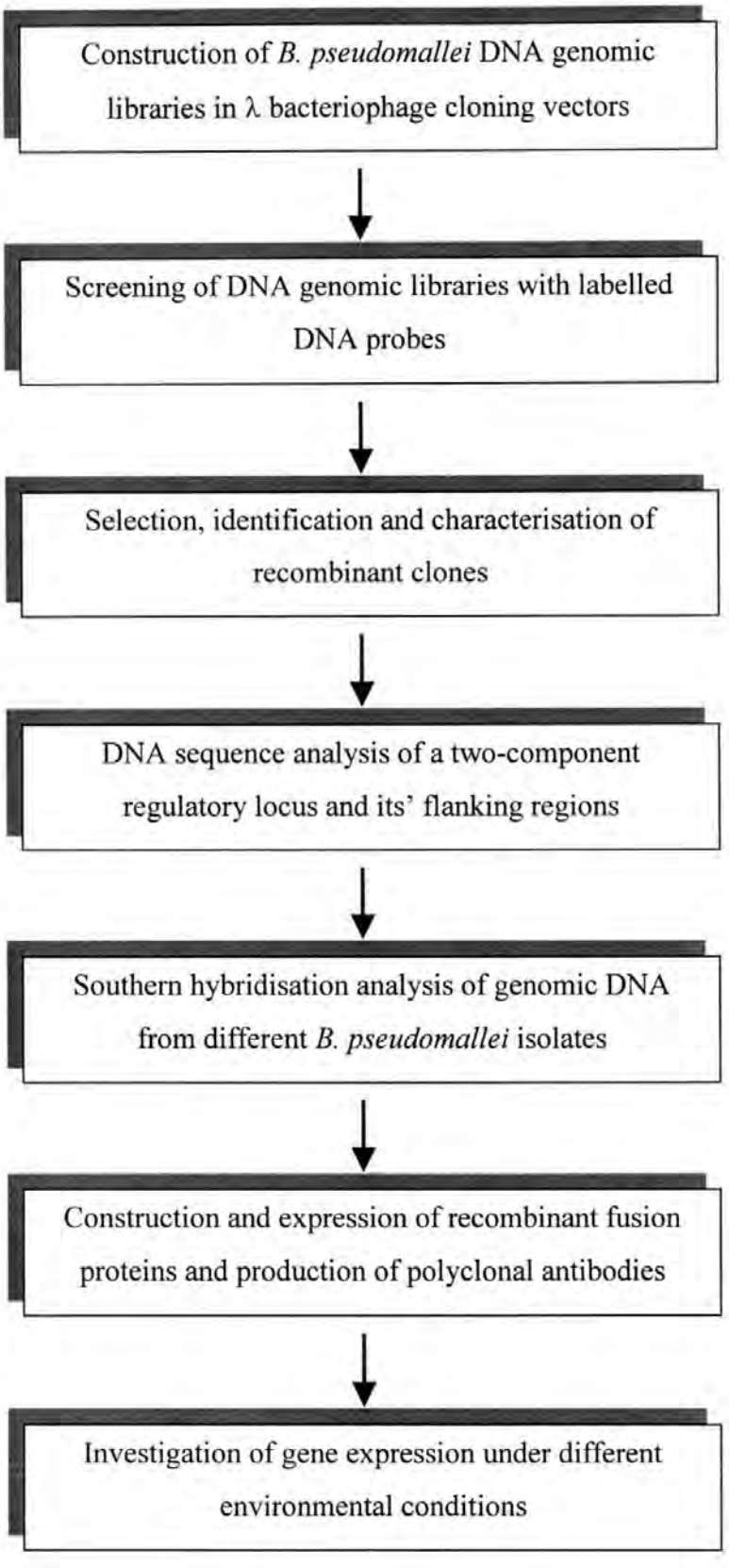
1) Identify the genes encoding a two-component regulatory system in *B. pseudomallei*,

2) Characterise the genes and the flanking regions of the genome,

3) Examine the possible conditions under which *B. pseudomallei* modifies the expression of these genes.

The direction of the study was guided by comparative examples drawn from our knowledge of other bacterial species, particularly those that share genetic or pathogenic similarities. Over the course of this research project a number of experiments were performed in order to investigate other aspects of *B. pseudomallei*, only those that concur with the major aims of the work are discussed. The major finding of this study has been the identification and molecular characterisation of a two-component signal transduction locus. It is hoped that this locus may be of diagnostic use.

The aims of the work described in this thesis were facilitated by applying molecular techniques for the study of *B. pseudomallei* according to the strategy outlined in the following flow diagram:


Construction of *B. pseudomallei* DNA genomic libraries in λ bacteriophage cloning vectors
Screening of DNA genomic libraries with labelled DNA probes
Selection, identification and characterisation of recombinant clones
DNA sequence analysis of a two-component regulatory locus and its' flanking regions
Southern hybridisation analysis of genomic DNA from different *B. pseudomallei* isolates
Construction and expression of recombinant fusion proteins and production of polyclonal antibodies
Investigation of gene expression under different environmental conditions

CHAPTER 2

LITERATURE REVIEW

2.1 *BURKHOLDERIA PSEUDOMALLEI*

2.1.1 General aspects

Burkholderia pseudomallei is a small, aerobic, non-acid-fast, non-spore-bearing Gram negative bacillus, motile by polar tuft of one to four flagella (Brindle & Cowan, 1951), the latter being a distinctive feature that distinguishes *B. pseudomallei* from the closely related non-motile species, *Burkholderia mallei*. The bacterium is oxidase positive, often shows bipolar staining, and obtains its energy for growth by respiration and not fermentation (Bokman *et al.*, 1957). Other characteristic properties include the accumulation of poly-β-hydroxybutyrate as intracellular granules; arginine dihydrolase and gelatinase activities; and resistance to aminoglycoside and polymyxin antibiotics (Levine & Wolochow, 1960; Redfearn *et al.*, 1966). Moreover, *B. pseudomallei* is known to have a nitrate reductase activity that is a requirement for anaerobic respiration and it can grow and survive under oxygen-limited conditions by switching from aerobic to anaerobic metabolism (Wongwanich *et al.*, 1996). The bacterium is the causative agent of the frequently fatal disease of humans, melioidosis (Smith *et al.*, 1987; Dance, 2000a).

B. pseudomallei is a free-living organism in the natural environment of tropical and sub-tropical regions. It is found widely in the soil and surface water of rice paddies, fields newly planted with oil palm, monsoon drains, gardens, and playgrounds in endemic areas. It is nutritionally versatile and can, for example, utilise an exceptionally wide range of organic compounds as the sole source of its carbon and energy needs (Redfearn *et al.*, 1966).

B. pseudomallei grows readily on most standard culture media, with a tendency to form microcolonies and biofilms (Vorachit *et al.*, 1993), but for initial isolations Ashdown's selective medium that contains crystal violet and gentamicin (Ashdown, 1979) is favoured. The bacterium can grow at temperatures between 18-42°C, and will tolerate 5°C for up to 190 days (Yabuuchi *et al.*, 1993), but the optimum temperature for growth *in vitro* is

between 37-42°C, reflecting the temperature of the tropical and sub-tropical climate (Dance, 2000abc). In contrast, Tong *et al.* (1996) showed that under laboratory conditions, the optimum temperature for survival appears to be between 24 and 32°C and *B. pseudomallei* was purportedly isolated at 20°C from drinking water in Bologna, Italy (Zanetti *et al.*, 2000). The evidence suggests that *B. pseudomallei* adapts well to a variety of growth conditions (Dejsirilert *et al.*, 1991). For example, it is able to grow in brain heart infusion broth of initial pH 4.5 under aerated or non-aerated conditions and survives well for up to 30 days.

Colonies of *B. pseudomallei* vary considerably in their morphology depending on the culture media and the strain of the organism, ranging from the most common rough type to smooth and sometimes even to mucoid appearance (Nicholls, 1930; Dance, 1989). Colonies often appear wrinkled, and have a strong musty, earthy odor after incubation for 2 to 3 days (Salisbury & Likos, 1970; Dance, 1990; Walsh & Wuthiekanun, 1996).

Genome analysis of *B. pseudomallei* has shown that its average gene size is 1031 base pairs, and the mol% guanine + cytosine content is 65.7% (range 55.6-74.0%). The genome is gene-rich with about 89% of the capacity used as coding sequences. It can therefore be estimated that the entire *B. pseudomallei* genome encodes about 5600 genes (Songsivilai & Dharakul, 2000).

2.1.2 Nomenclature and taxonomy

Until recently, *B. pseudomallei* was known as *Pseudomonas pseudomallei* and over the course of the last 80 years has been called *Bacillus pseudomallei*, *Bacillus whitmori*, *Pfeifferella whitmori*, *Flavobacterium pseudomallei*, *Actinobacillus pseudomallei*, *Loefferella whitmori* and *Malleomyces pseudomallei* (Howe *et al.*, 1971; Patamasucon *et al.*, 1982; Sanford, 1985; Brown & Thin, 1986).

Since the description of Haynes (1957) and until about 1992 this organism was classified as a member of the genus *Pseudomonas* (Migula, 1894) because of many

similarities in cultural, morphological and biochemical characteristics. Furthermore, this genus included human pathogens such as the opportunistic *Pseudomonas aeruginosa*, that infects wounds, burns and the lungs of cystic fibrosis sufferers and *Pseudomonas cepacia*, an opportunist pathogen in hospital patients, previously known as a phytopathogen and a cause of 'soft rot' of onions (Palleroni & Holmes, 1981). Many *Pseudomonas* species are plant pathogens, including *P. syringae* and *P. viridiflavae*, the cause of substantial post-harvest losses of fruits and vegetables, and *P. gladioli*, the cause of 'flower-rot' in gladioli. On the other hand, some strains of *P. fluorescens* and *P. putida* display a rather different interaction with plants, colonising the roots of several crops and causing significant yield increases through production of anti-microbial agents and iron scavenging compounds (Schippers *et al.*, 1987).

Palleroni & Holmes (1981) divided the genus *Pseudomonas* into five groups of species according to the results of rRNA-RNA hybridisation. Subsequently, on the basis of the 16S rRNA sequences comparisons, DNA-DNA homology values, cellular lipid and fatty acid composition, and phenotypic characteristics, the genus *Pseudomonas* has been reclassified. Yabuuchi *et al.* (1992) transferred seven species of genus *Pseudomonas* homology group II to a new genus *Burkholderia*, named after an American bacteriologist W. H. Burkholder, with the type species being *B. cepacia*. Initially, *B. cepacia*, *B. mallei*, *B. caryophylli*, *B. gladioli*, *B. pickettii*, *B. solanacearum* and *B. pseudomallei* were included in the genus. Currently more than 15 species have been placed in the genus *Burkholderia*, and these are mostly soil saprophytes or plant pathogens; the exceptions are the human pathogens *B. pseudomallei*, *B. mallei* and *B. cepacia*. Subsequently, *B. pseudomallei* and *B. mallei* were found to be genetically indistinguishable, but because of epidemiological dissimilarities, they have remained separate species.

Recently, isolates of *B. pseudomallei* have been differentiated into two biotypes which

distinguish between those that can assimilate L-arabinose (Ara^+) and those that can not (Ara^-) (Smith *et al.*, 1997). Almost all of the clinical isolates of *B. pseudomallei* are Ara^-, while both Ara^+ and Ara^- are found in the soil. The two biotypes exhibit distinct differences in their virulence in animal models of infection. Only the Ara^- isolates appear to cause disease [LD_{50} of ~10 units (cfu) in Syrian hamsters (DeShazer *et al.*, 1997a) and 10^2 cfu in mice (Smith *et al.*, 1997)], while the Ara^+ isolates are nonvirulent [LD_{50} of $>10^9$ cfu in mice]. Differences also exist between the biotypes in the production of proteinase, lipase, lecithinase and other enzymes, and in lipopolysaccharide composition. Smith *et al.* (1997) suggested that Ara^+ might be different enough from Ara^- to warrant its inclusion in a new species within the genus *Burkholderia*. Based on the difference in 16S rRNA gene sequence and other supporting evidence, Ara^- strains have currently been classified as true *B. pseudomallei* while Ara^+ strains are considered to represent a separate species termed *B. thailandensis* (Brett *et al.*, 1998).

More recently, Songsivilai & Dharakul (2000) analysed the genomes of both *B. pseudomallei* and *B. thailandensis* by pulsed field gel electrophoresis. They showed that the *B. pseudomallei* genome consisted of two large replicons indicating the presence of two chromosomes. *B. pseudomallei* has chromosomes of approximately 3563 and 2974 megabase-pairs in size, giving a total genome size of about 6.54 megabase-pairs that places it with the largest bacterial genomes. In contrast, the two chromosomes of *B. thailandensis* are about 20% smaller than *B. pseudomallei*.

2.1.3 Distribution in the environment

B. pseudomallei is endemic to south-east Asia and northern Australia, although the distribution of the bacterium elsewhere is well documented. Early surveys conducted in Malaysia (Strauss *et al.*, 1969), showed that *B. pseudomallei* was readily isolated from wet rice fields and other cleared, cultivated areas. In the urban environment of Singapore, the

bacterium was commonly found in soil/water samples from sports fields (Thin *et al.*, 1971). In Thailand, *B. pseudomallei* is more frequently found in the south of the country (Natchiangmai *et al.* 1985), from cultivated areas such as rubber plantations and rice fields, and in the soil of the north-east (Wuthiekanun *et al.*, 1995). Soil conditions in the north-east region of Thailand, the major endemic area of melioidosis, are mostly acidic in a pH ranging from 4.38 to 7.7 (Mitsuchi *et al.*, 1986). Kanai & Kondo (1994) attributed the occurrence of *B. pseudomallei* in such environments to its ability to survive under low pH (4.5). On the other hand, Dance (2000abc) considered that the uneven presentation of the bacteria could be attributed to technical factors, such as the technique used for the environmental survey, the sensitivity of different culture conditions for the isolation of environmental *B. pseudomallei* and the lack of a protocol that considers the so-called viable non-cultivable state of the organism, which is thought to be important in the environmental survival of other bacteria (Barer *et al.*, 1993). Furthermore, several other factors were found to influence the distribution of *B. pseudomallei* in soil such as temperature, rainfall, sunlight, and the physical, chemical and biological characteristics of the soil.

2.1.4 Survival

The extended survival of *B. pseudomallei* in the natural environment, especially its resistance to drying, has been demonstrated in field and experimental systems (Stanton & Fletcher, 1932; Everett & Nelson, 1975; Soy *et al.*, 1983). The organism was shown to remain viable and to retain its virulence even after 27 days of desiccation within garden soil at 27°C. It appears that *B. pseudomallei* can survive in distilled water (Wuthiekanun *et al.*, 1995) and also in moist clay soil kept in the shade at tropical room temperature, for up to 36 months (Thomas *et al.*, 1981). However, survival times were much shorter in dry, sandy soil. On the other hand, in the dry season *B. pseudomallei* may survive in a dormant state of low metabolic activity in the deeper layers of the soil with poor aeration. In rice paddies, *B.*

pseudomallei has the opportunity to survive not only in the aerobic layer of around 20 mm thickness prevailing under the surface water but also in the anaerobic layer located below the aerobic layer. Wongwanich *et al.* (1996) presumed that ammonium sulfate fertiliser will be converted to into HNO_3 by the action of nitrifying bacteria in the aerobic layer and the resulting HNO_3 will move to the anaerobic layer to serve the nitrate respiration of *B. pseudomallei*. Interestingly, Choy *et al.* (2000) reported that *B. pseudomallei* demonstrated an ability to survive for some time in 'Saffan' (alphaxalone 9 mg/ml and alphadalone 3 mg/ml solution), an injectable anaesthetic for cats and 'Savlon liquid' (cetrimide 3% and chlorhexidine 0.3% solution), an antiseptic/disinfectant solution commonly used for the preparation of a surgical site and surgical instrument decontamination.

Little is known about the behavior of *B. pseudomallei* in the natural environment. In particular, many questions remain unanswered concerning the viable non-cultivable state of the organism and its relationship with plants or other protozoans. It is possible that *B. pseudomallei* can survive in soil and water by parasitising other organisms. *B. pseudomallei* is a facultative intracellular bacterial pathogen, and other species in this category, for example, *L. pneumophila*, *Mycobacterium avium*, and *Listeria monocytogenes* have been shown to survive as endosymbionts in amoebae. Inglis *et al.* (2000) exposed *B. pseudomallei* to three different amoebic species and showed that the bacterium was able to survive within the vacuoles of free-living amoebae. This *in vitro* demonstration of the interaction between *B. pseudomallei* and free-living *Acanthamoebae* raises the possibility that a similar interaction *in vivo* might promote environmental survival of *B. pseudomallei* and subsequent human exposure. Since the parasitism of amoebae by bacteria bears a striking resemblance to the infection of human phagocytic cells, Inglis *et al.* (2000) postulated the occurrence of similar phenomena in mammalian phagocytic cells.

Other possible host-parasite relationships have been suggested. Certain strains of *B.*

cepacia can reside in the rhizosphere, the region of the soil modified as a result of the uptake and deposition of substances by a growing plant root, and in this case bacteria growing in close association with root nodules assist in the fixation of nitrogen and the uptake of minerals by the plant (Pitt *et al.*, 2000). Although a specific host for *B. pseudomallei* has not yet been described it is possible that it plays a similar role in the rice paddy field. Moreover, Bianciotto *et al.* (1996) and Ruiz-Lozano & Bonfante (1999) demonstrated the presence of a *Burkholderia* species living in the cytoplasm of *Gigaspora margarita*, a fungus that is known to establish symbiotic associations with the root tissues of more than 80% of land plants. Morphological observations show that the obligately intracellular *Burkholderia* multiplies both inside the fungal spore and in the mycelium during differentiation of the colonisation structures (Bonfante *et al.*, 1994). If this were the case with *B. pseudomallei*, it may help explain the abundance of the bacteria in certain environments, particularly cultivated tropical regions, rather than other areas.

2.1.5 Opportunistic and intracellular pathogen

Opportunistic pathogens are organisms that under most circumstances do not cause disease. In these cases virulence can be defined only with reference to the host condition. Thus, a change or inherent weakness in host resistance initiates the process of infection. Opportunistic infections are common nowadays, partly because in circumstances where specific microbial pathogens have been effectively eliminated, opportunistic infections are relatively more numerous and also because modern medical care keeps alive many people who have impaired resistance to microbial infections (Mims *et al.*, 1997). Examples of opportunistic pathogens are *P. aeruginosa,* that is essentially a free-living species sometimes present in the intestinal tract, and *B. cepacia*, an infective agent in hospital patients, but also a phytopathogen (Palleroni & Holmes, 1981). Although *B. pseudomallei* infections have been documented in otherwise healthy individuals, the organism usually behaves as an

opportunist and susceptible individuals have an underlying predisposition such as diabetes or renal disease (Chaowagul *et al.*, 1989; Woods *et al.*, 1993).

Microbial pathogens have evolved several strategies for successfully infecting host organisms. One such strategy is the colonisation of niches that allow survival and even replication in conditions that are relatively unaffected by host defence mechanisms. Intracellular niches seem to be particularly useful for the long-term survival of microbial pathogens despite ongoing host immune responses and/or antimicrobial therapy, and many diseases caused by intracellular pathogens take a chronic course (Kaufmann & Reddehase, 1989).

The clinical manifestation of *B. pseudomallei* infection and its protracted course have led investigators to suspect that the bacterium is a facultative intracellular pathogen which is capable of invading eukaryotic cells. *B. pseudomallei* has been observed to survive and multiply within phagocytes (Jones *et al.*, 1996). Wongwanich *et al.* (1996) demonstrated, by electron microscopy, the presence of *B. pseudomallei* within the phagosomes of mononuclear cells and attributed this to the ability of the bacteria to grow in an acidic environment such as occurs within phagosomes, pH 4.7 to 5.5 (Sprich, 1956). These observations indicated that *B. pseudomallei* is similar to other facultative intracellular bacteria, such as *L. monocytogenes* and *Salmonella typhi*. Macrophages and neutrophils kill ingested microorganisms by various means including a respiratory burst that leads to the production of hydrogen peroxide, superoxide radicals, and other oxygen-derived radicals, together with activation of myeloperoxidase. Lysosomes fuse with the pathogen-containing phagosome and release their acid hydrolases (glycosidases, proteases, and lipases) (Moulder, 1985). Therefore, intracellular bacteria have adopted various ways for surviving the microbicidal activity of phagocytes, including interference with reactive reactive oxygen intermediates, inhibiting phagosome-lysosome fusion (e.g. legionellae, mycobacteria),

resistance to bactericidal lysosomal enzymes (e.g. salmonellae), and evading phagosome-lysosome fusion by escaping into the cytoplasm (e.g. *Listeria*, shigellae) (Kaufmann, 1993). In order to survive inside macrophages, *Salmonella* have developed a set of at least 40 proteins that confer resistance to reactive forms of oxygen and defensins, toxic host peptides that kill bacteria (Kaufmann & Flesch, 1992). In addition, salmonellae possessing Vi capsular polysaccharide are less susceptible to killing by lysosomal contents than rough mutants that lack the long polysaccharide chains. *B. pseudomallei* also possesses a capsular polysaccharide that may contribute to serum resistance and intracellular survival. This matter is considered in more detail in section 2.4.2.5.

The mechanism by which *B. pseudomallei* are able to enter host cells is unknown. Studies *in vivo* (Dannenberg & Scott, 1958) and *in vitro* (Pruksachartvuthi *et al.*, 1990) showed that *B. pseudomallei* can invade, survive and replicate in 'professional' and 'non-professional' phagocytes including rat alveolar macrophages, human neutrophils, and cultured epithelial cell lines, such as HeLa, CHO, A459, and Vero (Jones *et al.*, 1996). An examination of the infected cells by electron microscopy confirmed the presence of intracellular bacteria located within membrane-bound vacuoles. *B. pseudomallei* was able to replicate inside human neutrophils after 16 h and in A459 epithelial cells for up to 48 h. In addition, the authors demonstrated the degeneration of the membranes of vacuoles containing the pathogen in infected cells. Furthermore, Ho *et al* (1997) observed that a small percentage of bacteria could survive the oxidative and nonoxidative bursts following phagocytosis by human neutrophils when incubated in the presence of 10% heat-inactivated rabbit or human serum. The occurrence of relapsing and late-onset infection has been attributed to the ability of *B. pseudomallei* to survive within macrophages (Sanford *et al.*, 1971; Praksachartvuthi *et al.*, 1990; Jones *et al.*, 1996).

In order to establish an infection in a susceptible host, a microbial pathogen must avoid

killing by non-specific and specific host defence mechanisms. Intracellular survival and a strong resistance to the bactericidal activity of normal human serum (Ismail *et al.*, 1988; Egan & Gordon, 1996; Pruksachartvuthi *et al.*, 1990) provides *B. pseudomallei* with this capability. The pathogen is a facultative intracellular parasite that, like mycobacteria, may remain dormant in host macrophages for many years.

2.2 MELIOIDOSIS DISEASE

2.2.1 General aspects

Melioidosis is a life-threatening disease that affects humans and animals in tropical and subtropical areas where it is known to be a major public health problem. It has been termed "the great imitator" of every infectious disease, as virtually any organ can be infected. The disease has been qualified as an emerging infection, since it has been recognised more frequently during the past two decades, both within established endemic areas and elsewhere (Dance, 2000a). The severity of the disease depends on one or more of the following: the extent to which host defences are impaired, the strain of the bacterium, and the magnitude and route of the inoculum.

2.2.2 History

A British pathologist, Captain A. Whitmore, and C. S. Krishnaswami, an assistant surgeon, at Rangoon General Hospital, Burma first identified melioidosis in 1911. They conducted a post-mortem on a 40-year-old patient suspected of having human glanders who had died of pneumonia, and the disease was first known as Whitmore's disease. The post-mortem examination revealed wide spread lesions not only in the lungs but also in the liver, spleen and kidney. It became apparent that the disease was common among the ill-nourished, neglected Rangoon people with or without morphine addiction and Whitmore reported 38 cases of the infection. Since the initial report, cases of melioidosis have been regularly reported from South East Asia, although recent interest was generated following

reports of the infection in American servicemen serving in Vietnam.

In 1921, Stanton and Fletcher coined the term melioidosis, when *B. pseudomallei* was isolated from the blood of a patient in Malaysia. The term is derived from the Greek word "melis" meaning "a distemper of donkeys (actually asses)". The disease resembles glanders, which is predominantly a pulmonary disease in asses caused by the bacterium *B. mallei* (formerly known as *Pseudomonas mallei*). The same authors in 1932 predicted that melioidosis would prove to be far more prevalent than appreciated at that time. They concluded that the wild rat was the natural reservoir of an infection that was spread via the faeces and urine to the soil and vegetation. They proposed that the infection was consequently transmitted to humans through the ingestion of contaminated food or water. Several years later, French workers in Indo-China observed that the occurrence of melioidosis often followed exposure to mud and water (Dance, 2000b), and Chambon (1955) isolated the bacterium from the mud and water of ponds and rice paddies in South Vietnam. It thus became apparent that *B. pseudomallei* was a saprophyte, and that both man and animals were likely to fall victim following exposure to the organism in soil and surface water (Dance, 2000b). Although over 300 human cases had been documented by 1957 (Smith *et al.*, 1987), melioidosis was a rare event in developed countries until 343 cases of the disease were reported among American soldiers serving in the Vietnam war forcing a significant reappraisal of the importance of the disease.

In Thailand, despite the diagnosis of an imported case in 1928 and two cases in prisoners of war in 1947, the first indigenous case was not described until 1955 (Dance, 1991). Punyagupta (1989) reviewed 686 cases of melioidosis. More recently, with increasing recognition of the disease, it has been estimated that there are from 2000 to 5000 cases of melioidosis each year in Thailand (Dharakul & Songsivilai, 1996).

2.2.3 Geographical distribution

The disease is most commonly recognised in south-east Asia and northern Australia (**Fig 2.1**). In south-east Asian countries serological evidence indicated a widespread occurrence of melioidosis in its sub-clinical asymptomatic form (Smith *et al.*, 1987). Surveys carried out in Malaysia and Thailand have revealed the presence of circulating antibodies to the causative organism in 15-30% of the normal population (Nigg 1963; Strauss *et al.*, 1969) and up to 80% of 4 year old children in north-east Thailand (Kanaphun *et al.*, 1993) whilst in northern Australia the incidence of seroconversion was reported to be 5-10% among at risk groups by Ashdown & Guard (1984).

Sporadic cases of melioidosis have been described in many other regions of the world such as the Indian subcontinent, Pacific Islands, Africa, the Caribbean and Central and South America, Iran, China, Korea, Philippines, Mexico, Indonesia, Madagascar, Turkey, Indonesia, France and UK. Although the disease may be less common elsewhere, many thousands of people may die annually from melioidosis world wide, and many millions experience minor infections and are at risk of subsequent relapse. It is estimated that about 7 million people are at risk in north-east Thailand alone (Currie *et al.*, 2000c).

It is important to note that the frequency of isolation of *B. pseudomallei* from the environment may not be directly related to the incidence of melioidosis in an area (Trakulsomboon *et al.*, 1997). Such discrepancies may be partly reflected by phenotypic differences between some clinical and environmental isolates (Anuntagool *et al.*, 1998). The disease incidence recorded by one hospital in Ubon Ratchathani in north-east Thailand is approximately 4.4 per 100 000 population per year (Suputtamongkol *et al.*, 1994). In this province melioidosis is the cause of 20% of all community-acquired septicaemias, and 40% of deaths that are attributed to septicaemia (Chaowagul *et al.*, 1989). However, it is likely that this figure is an underestimate due to problems of diagnosis. The Thai estimate

Fig 2.1: Geographical distribution of melioidosis. Shaded areas indicate known endemic regions, hatched areas and asterisks indicate sporadic cases. Adapted from Dance (1991).

compares with a smaller exposed population in the Northern Territory of Australia of around 140,000 people, but with an average annual incidence of 18 per 100,000 (Currie *et al.*, 2000c).

A relationship between the geographical distribution of *B. pseudomallei* and the prevalence of human infection has not been clearly established. Melioidosis is not evenly distributed within endemic areas, and hot spots for melioidosis were reported during the Vietnam War. Dance (2000b) attributed this to the uneven distribution of *B. pseudomallei* in soil (because of climatic, geological or ecological factors), differences in the virulence of strains of *B. pseudomallei*, regional variations in cultural practices leading to exposure (e.g. farming techniques), or to differences in the susceptibility of populations to infection.

Since the *B. pseudomallei* is nutritionally diverse and is capable of resisting a variety of environmental extremes, it is puzzling as to why a more uniform global distribution of the bacterium is not more apparent (Smith *et al.*, 1987). It is possible that the disease remains greatly under diagnosed in many areas of the tropics where sophisticated laboratory facilities are not available. Whatever the reasons it seems likely that the increasing movement of humans, animals or infected materials made easier by the huge increase in international trade, travel and tourism will provide more opportunities for melioidosis to spread from established areas to new pastures (Dance, 2000a). The observation that all recent isolations that have been reported from non-endemic areas appear to have been derived from endemic areas supports this suggestion.

2.2.4 Transmission

Transmission of *B. pseudomallei* commonly occurs in two ways: 1) by direct contact with contaminated soil and/or water through pre-existing skin cuts and abrasions (the majority of infections), and 2) inhalation or ingestion of dust particles containing the bacteria.

Direct contact with the organism through cuts and abrasions is thought to be the most frequent means of exposure. In regions where the pathogen is endemic this occurs commonly among farmers as a consequence of their occupation and among children whilst working and playing (section 2.2.5). However, it has only been proven in from 5-25% of cases (Currie *et al.*, 2000c).

Although common for some pathogens, e.g. *Bordetella pertussis* and *Mycobacterium tuberculosis*, inhalation may be a less frequent means for transmission of *B. pseudomallei*. Inhalation has been suggested as a major route of infection during the Vietnam War, where a disproportionate number of helicopter crewmen succumbed to *B. pseudomallei* infections as compared to other soldiers stationed in the same region. To explain this phenomenon, it has been proposed that the helicopter rotors acted to disturb infectious particles in the dust around landing zones, dispersed greater amounts of *B. pseudomallei* in the air, and thus facilitated the pulmonary inoculation of the crewmen (Howe *et al.*, 1971; Sanford, 1990). On the other hand, Currie *et al.* (2000a) postulated haematogenous spread to the lungs following presumptive inoculating skin injuries rather than inhalation or spread from the upper respiratory tract.

B. pseudomallei can survive at low pH (Tong *et al.*, 1996), a feature that characterises pathogens that follow a faecal-oral route e.g. enteric pathogens such as *Salmonella*, and the possibility of occasional ingestion has been raised (Howe *et al.*, 1971; Dance, 1990). It has been shown experimentally that animals can be infected via the oral or nasal mucosa or by ingestion (Whitmore, 1913; Stanton & Fletcher, 1932) although no gastrointestinal lesions were found in the infected animals. The possibility that ingestion may be important in some circumstances (Cottew, 1952) is supported by autopsy findings of the presence of infected gastro-hepatic lymph nodes in pigs (Ketterer *et al.*, 1986), and cases that are thought to be associated with contaminated drinking water (Inglis *et al.*, 1998).

Horizontal transmission of melioidosis has not yet been proved to be significant and there is no evidence for a *B. pseudomallei* carrier state in humans (Kanaphun *et al.*, 1993). Human-to-human transmission of melioidosis has been suspected in only two reports of the disease (McCormick *et al.*, 1975; Kunakorn *et al.*, 1991). However, there is evidence that the bacterium can cross the placenta to infect the foetus or pass perinatally from mother to child. In 1997, a goat with melioidosis aborted twins and died soon afterwards. *B. pseudomallei* was isolated from the mother's uterus and from the spleen of one of the aborted twins, confirming the transplacental infection. Thomas *et al.* (1988) also supported the possibility of this mode of transmission. Transmission from animal to human has been reported anecdotally (Galimand & Dodin, 1982; Choy *et al.*, 2000) although infected animals may play a role in the dissemination of the organism to new environments (Dance, 1990; Golledge *et al.*, 1992). Apart from one suspected case (Kan & Kay, 1978), arthropod-borne infection apparently does not occur naturally (Sanford, 1985), although transmission has been accomplished experimentally in guinea pigs both by the bite of the mosquito (*Aedes aegypti*) and of the rat flea (*Xenopsylla cheopis*) (Sanford, 1995). Moreover, at least two incidences of laboratory-acquired melioidosis have been reported (Green & Tuffnell, 1968).

In summary, the main route of transmission is probably via percutaneous inoculation with contaminated soil or water, although in some circumstance inhalation of aerosols and ingestion may be important, while person to person transmission is rare.

2.2.5 Occupation, age, sex, and season

In Thailand, the infection is most common in rice-farming communities, an environment where individuals are more likely to be exposed to the pathogen. In northern Thailand, 81% of all cases of melioidosis occurred in rice farmers or their families (Suputtamongkol *et al.* 1994). Moreover, during an outbreak in Australia, Merianos *et al.*

(1993) found that the risk of melioidosis was approximately 10-fold higher among individuals who were occupationally exposed to soil compared with indoor workers.

Melioidosis may present at any age (White, 1994). In a study of 423 cases of culture-positive melioidosis in Ubon Ratchatani, north-east Thailand, Suputtamongkol *et al.* (1994) observed a bimodal age distribution, with the lower peak occuring in the first 9 years of life and the highest age-specific incidence occurring at 50-59 years in women and 60-69 years in men. Men are more often affected than women or children, and male predominance has been cited in many studies (Rode & Webling, 1981; Guard *et al.*, 1984; Suputtamongkol *et al.*, 1994). The male:female ratio in the incidence of melioidosis of 1.5:1 in Thailand (Suputtamongkol *et al.*, 1994) contrasted with that of 5:1 in Singapore (Anonymous, 1995), and probably reflected work-related differences in exposure to soil and water through rice farming (Dance, 2000c). Approximately 20% of cases occur in children under 15 years of age. Recent reports from Thailand suggest a higher prevalence of the disease in children than previously recognised with a unique spectrum of symptoms, notably the presence of parotid abscesses (Dance *et al.*, 1989). Infection in the neonatal period has been reported, but rarely (Edmond *et al.*, 1998; Halder *et al.*, 1998).

Melioidosis is regarded as a disease which is predominantly associated with the tropical monsoon season and the number of cases increase as the rainy season progresses in endemic areas and farmers become busy in their fields (Congakpoon *et al.*, 1989; Sookpranee *et al.*, 1989). The risk of exposure to *B. pseudomallei*, and possible infection, are particularly high at the beginning of the wet season in the endemic areas (November-April) when ploughing and planting of seedlings are performed without protective clothing and minor skin trauma is extremely common (Chaowagul *et al.*, 1989; Leelarasamee & Bovornkitti, 1989). It is possible that *B. pseudomallei* organisms lying dormant in the deeper layers of the soil will be carried to the surface as the water table rises during the course of

the rainy season. In such moistened soil, the organisms will regenerate enhancing their potential to infect exposed humans and animals (Kanai & Kondo, 1994). Moreover, the re-activation of latent disease has been strongly associated with seasonal change and in northern Australia 85% of such cases arise during the wet season (Currie *et al.*, 2000c). The wet season is therefore an important factor in disease transmission (Golledge *et al.*, 1992).

2.2.6 Diagnosis

It seems likely that melioidosis is often under-diagnosed, especially in poor, rural areas where appropriate microbiological facilities are not available. The specific diagnosis of melioidosis requires awareness on the part of clinicians, and the existence of a laboratory capable of isolating and identifying *B. pseudomallei* (Dance, 2000a). It also demands a high degree of clinical suspicion, especially in residents or travelers from endemic areas, and particularly if the presenting features are those of fulminant respiratory failure. When overt infection occurs it may initially defy diagnosis because of its tendency to resemble other more common conditions. Leelarasamee & Bovornkitti (1989) reported that it is difficult to estimate the number of cases of melioidosis accurately since many cases are not recognised or are misdiagnosed. Among the bacterial infections it rivals syphilis as the great mimicker. The high mortality rate (approximately 95%) among acute fulminating melioidosis sufferers is mainly due to the failure to recognise the disease in the early stages of an infection (Smith *et al.*, 1987).

Smith *et al.* (1987) elucidated the major limitations that hampered the evaluation and recognition of the disease. These included: the wide array of signs and symptoms presented by acute melioidosis; the long period of latency associated with chronic melioidosis; and the unavailability of suitable diagnostic procedures for melioidosis. Consequently, diagnosis depends upon the rapid isolation and identification of the causative organism and/or the identification of a specific marker antigen, such as *B. pseudomallei* exotoxin, in the serum or

tissues of patients. The diagnostic methods for the identification of clinical and subclinical melioidosis are presented in **Fig 2.2**.

2.2.6.1 Bacterial culture and identification

Isolation of the bacterium in culture from clinical specimens is the 'gold standard' for laboratory diagnosis of melioidosis, since it is relatively simple and economical to perform (**Table 2.1**). However, because a high percentage of patients with acute septicaemia die within 24-48 h of admission, the time required for the isolation and identification of *B. pseudomallei* from blood culture, 3 days-2 weeks, is too lengthy to allow the administration of effective therapy in many cases (Dance, 1990; Sirishinha *et al.*, 2000).

The commercial API 20 NE kit, a substrate-utilisation test panel, was first used by Dance *et al.* (1989) as an easily transportable tool for the biochemical identification of *B. pseudomallei*. The kit identified 390 (97.5%) strains correctly on first testing and all but one of the remainder on second testing, although other workers have reported the misidentification of *B. pseudomallei* by this method (Inglis *et al.*, 1998). A serious problem in ascertaining the presence of *B. pseudomallei* in environmental and clinical specimens lies in isolating and differentiating the bacterium from the heterogeneous population of organisms that may be present (Dance, 1991). Mixed infection with other frequently encountered species is possible in endemic areas of melioidosis, where the numbers of *B. pseudomallei* that are present are often relatively low (Werner *et al.*, 1967). Furthermore, *B. pseudomallei* has a generation time of 45 min at 37°C which is lengthy compared with many bacterial species (Werner *et al.*, 1967) and the resultant overgrowth by faster growing bacteria can mask the pathogen's presence (Kunakorn *et al.*, 2000; Sirisinha *et al.*, 2000). *B. cepacia* is often misidentified as *B. pseudomallei* and vice versa. Both *B. pseudomallei* and *B. cepacia* are soil saprophytes, opportunistic pathogens and are similar in their genetic

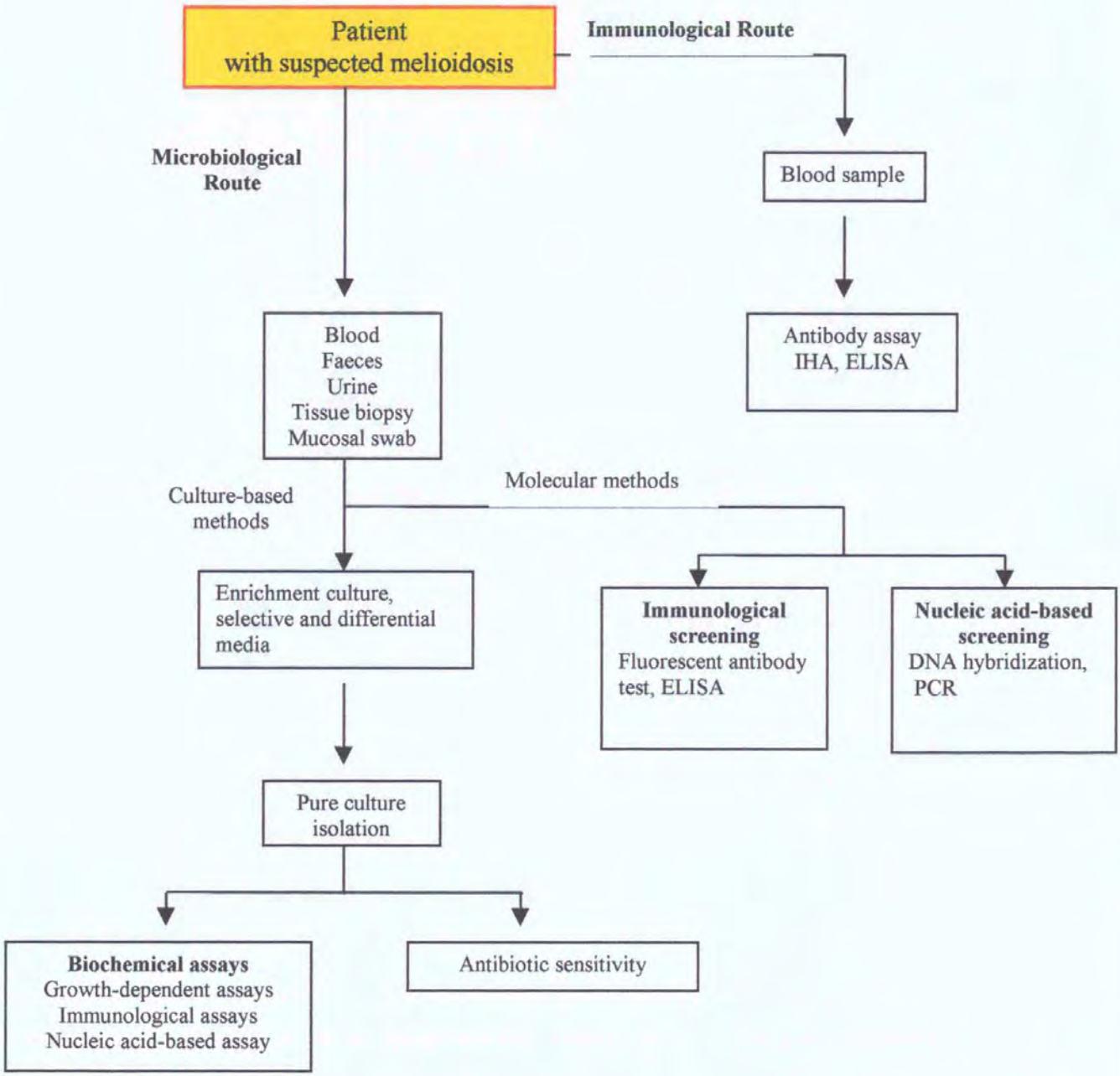


Fig 2.2: Clinical and diagnostic methods for the isolation and identification of *B. pseudomallei*. Adapted from Madigan *et al.* (1999).

Table 2.1: The results of the API 20NE test for identifying *B. pseudomallei*

Test	Results
Nitrate reduction	+
Indol (TRP)	-
Glucose fermentation	-
Arginine dihydrolase	+
Urease	-
Aesculin hydrolysis	+/-
Gelatinase	+
Galactosidase	-
Assimilation of:	
Glucose	+
Arabinose	-
Mannose	+
Mannitol	+
Maltose	-
N-acetyl-glucosamine	+
Gluconate	+
Caprate	+
Adipate	+
Malate	+
Citrate	+
Phenyl-acetate	+
Oxidase	+

makeup, colonial morphology, ability to catabolise a wide variety of organic compounds, and resistance to the aminoglycoside and polymyxin groups of antibiotics, such as gentamicin and colistin (Dance 1990). Another species that sometimes causes confusion is *Chromobacterium violaceum* a soil and water inhabitant that has recently emerged as a human pathogen, although the exact reasons for misidentification are not clear in this case (Inglis *et al.*, 1998).

Nevertheless, following isolation in pure culture, the conclusive identification of *B. pseudomallei* can be achieved by a combination of simple tests involving Gram's stain, oxidase reaction, typical growth characteristics, and antibiotic resistance (Leelarasamee, 2000).

2.2.6.2 Serological tests

A variety of serological tests for detecting host antibody responses to *B. pseudomallei* have been developed including immunofluorescence (IFA), indirect haemagglutination (IHA), complement fixation, and enzyme-linked immunosorbent assay (ELISA). These methods have limited value for routine use because of the crude nature of the antigens used (Sirisinha, 1991), occasional false-positive reactions, and the high background seropositivity of many individuals from populations within the endemic areas (Sermswan *et al.*, 1994). It is been estimated that more than 80% of children in rice-farming communities in north-east Thailand have acquired antibodies against *B. pseudomallei*, as detected by IHA, by the time they are 5 years old (Kanaphun *et al.*, 1993). Nevertheless, the IHA test, remains the most widely used serological test for the diagnosis of melioidosis, especially in non-endemic areas (Dance, 1990). A crude autoclaved culture filtrate (Ashdown, 1987) or lipopolysaccharide extracted from the organism are most commonly used as the antigen for the IHA test.

Serological methods for the detection of the soluble secreted products of *B.*

pseudomallei in blood and urine, and bacterial cells in pus, wound, sputum and throat swabs have also been developed. These include a monoclonal antibody-based assay for the quantitation of exotoxin (Ismail *et al.*, 1987), and polyclonal antibody-based ELISAs for the detection of secreted antigens (Wongratanacheewin *et al.*, 1990) and LPS (Petkanjanapong *et al.*, 1992). However, these tests have yet to be evaluated for their clinical usefulness on a large-scale (Sirisinha *et al.*, 2000). Recently, Steinmetz *et al.* (1999) developed a simple latex agglutination test for the rapid identification of *B. pseudomallei* culture isolates, using an exopolysaccharide-specific monoclonal antibody (Map 3015). A variety of *B. pseudomallei* isolates from environmental and clinical sources were tested and showed a strong and specific agglutination. In contrast, avirulent *B. thailandensis* and a variety of other bacterial species did not react.

2.2.6.3 Nucleic acid detection

There are a variety of diagnostic techniques for the detection and identification of bacteria in clinical samples based on characteristic DNA sequences. Nucleic acid-based detection systems can allow the identification of bacteria without the need for isolation in pure culture (Salyers & Whitt, 1994) or the propagation of living organisms, which is important for dangerous human pathogens such as *B. pseudomallei.* The technique is especially useful for the detection of organisms that can not easily be grown *in vitro* (Ieven, 1998) and nucleic acid-based methods have the potential to provide a rapid, specific and highly sensitive means for the detection and identification of pathogenic bacteria (Lew & Desmarchelier, 1994).

(i) Polymerase chain reaction (PCR)

PCR, a technique for the amplification of DNA sequences *in vitro*, has been widely used to assist in the diagnosis of infectious diseases, for example *Streptococcus pneumoniae* (Zhang *et al.*, 1995), *Neisseria gonorrhoeae* (Wong *et al.*, 1995). It can be highly specific,

provided that the appropriate PCR primers are chosen, thus allowing rapid identification and the speed and sensitivity of the technique make it ideal for 'high-throughput' automated screening of blood and tissue samples. Furthermore, PCR can detect a single copy of a target DNA sequence and therefore requires only small samples for analysis. Although *B. pseudomallei* survives and multiplies intracellularly (Pruksachartvuthi *et al.*, 1990), there may be substantial numbers in blood samples, therefore direct analysis of clinical specimens with PCR may aid in early diagnosis (Kunakorn & Markham, 1995).

PCR assays using primers based on conserved regions in 16S rRNA gene (Dharakul *et al.*, 1996; Brook *et al.*, 1997), 23S rRNA gene (Lew & Desmarchelier, 1994) and 16S-23S rRNA gene spacer (Kunakorn & Markham, 1995) have been reported, but this method cannot always distinguish *B. mallei* from *B. pseudomallei*. Based on the nucleotide sequence of a *B. pseudomallei*-specific DNA probe (Sermswan *et al.*, 1994), Rattanathongkom *et al.* (1997) developed a PCR assay using primers named LPS1 and LPS2. The reaction amplified a 178-bp sequence from the DNA of 100 clinical *B. pseudomallei* isolates, without amplification from the DNA of any of the 18 other bacterial species tested, including *B. cepacia* and *P. aeruginosa*. However, a study comparing these primers with primer sets based on 16S rDNA and16S-23S rRNA gene spacer for the PCR diagnosis of melioidosis using blood samples taken from septicaemic patients showed that no single set of primers could detect all strains of *B. pseudomallei* (Kunakorn *et al.*, 2000). Hence multiple sets of primers are needed for the conclusive diagnosis of *B. pseudomallei* in order to compensate for strain variation. Moreover, none of these PCR assays has been subjected to critical evaluation in the field (Sirisinha *et al.*, 2000).

(ii) Nucleic acid hybridisation

An alternative approach that is currently used for the clinical diagnosis of various bacterial diseases such as *S. typhi* (Rubin *et al.*, 1985), *Legionella* spp (Edelstein, 1986), and

Neisseria gonorrhoeae (Torres *et al.*, 1991), is nucleic acid hybridisation using a specific DNA probe. The basic principle of this method is the detection of a specific nucleic acid sequence by hybridisation to a complementary sequence of DNA or RNA, a labelled probe, followed by detection of the label (Ieven, 1998). Moseley *et al.* (1980) first applied this technology to an infectious disease, for the detection of enterotoxigenic *E. coli* by DNA-DNA hybridisation in stool samples.

A specific DNA probe (pKKU-S23L) was developed and used for the detection of *B. pseudomallei* by dot blot, and was also used to examine 60 clinical *B. pseudomallei* isolates for restriction fragment length polymorphism following *Hind*III digestion of total cellular DNA (Sermswan *et al.*, 1994). The probe showed no cross-hybridisation with total cellular DNA from closely related bacteria such as *B. cepacia* and *P. aeruginosa*. Whilst as little as 1.5 ng of genomic DNA or 4 x 10^4 bacterial cells could be detected the hybridisation technique was not sufficiently sensitive for routine diagnostic use (Lew & Desmarchelier, 1994; Sermswan *et al.*, 1994).

2.2.7 Clinical manifestations and pathology

The signs and symptoms of melioidosis can mimic various other infections, including common bacterial septicemia, malaria, tuberculosis, coccidioidomycosis, and histoplasmosis. Furthermore, when the infection manifests itself in an acute form, the disease may show gastrointestinal symptoms that can be confused with dysentery or typhoid (Stanton & Fletcher, 1932; Khaira *et al.*, 1959). Although the spectrum of clinical symptoms is broad, human melioidosis can be divided into three categories: acute, sub-acute and chronic disease.

Localised suppurative infection; acute pneumonia or an overwhelming septicaemia characterises acute forms of the disease, which is rapidly fatal. It is a severe illness, with the rapid onset (24-48 h) of ill-defined symptoms. Acute forms of the disease may follow either

a pulmonary or a septicaemic course. Acute pulmonary symptoms appear rapidly and usually involve the upper lobes of the lung. They are characterised by high fever and pulmonary distress with non-productive or productive cough, followed by the appearance of visceral abscesses and death within a few days if left untreated. The septicaemic form is characterised by the dissemination of the bacteria via the blood to various organs, and is also rapidly fatal. This form occurs predominantly in-patients with underlying diseases (Sanford, 1995), with clinical indications including malaise, meningitis, cellulitis, as well as cutaneous and subcutaneous lesions (Dance, 1990; Sanford, 1990). It is also the most common cause of fatality by community-acquired septicaemia in north-eastern Thailand and bacteraemic pneumonia in northern Australia. If improperly handled, the mortality rate of the acute septicaemic cases can be as high as 70-80% due to shock, respiratory failure, and multiple organ failure (Chaowagul *et al.*, 1989; Boonsawat *et al.*, 1990). Even when diagnosis can be accomplished quickly and antibiotic therapy is initiated soon after admission, the mortality rate is still high (Sirisinha *et al.*, 2000).

Sub-acute melioidosis is much less severe, with infrequent fatalities. It can arise from a primary infection that remains sub-acute or occur as a result of the reactivation of a previous infection. It is best described as a prolonged febrile illness characterised by multi-organ involvement, systemic abscess formation and bacteraemia. Typically exhibiting pulmonary involvement, symptoms are similar to the acute pulmonary form of disease. During the latter stages of the disease, the organism can be readily cultured from blood, pus, urine and other bodily tissues and secretions (Smith *et al.*, 1987). Most cases of melioidosis seen in non-endemic areas belong to this group (Ip *et al.*, 1995). Patients may be entirely asymptomatic (apart from pulmonary radiograph) or they may present with a disease indistinguishable from pulmonary tuberculosis, with a low-grade fever and weight loss that evolves over the course of months (Ip *et al.*, 1995; Currie *et al.*, 2000c).

Chronic illness is believed to be the most common form of the disease (Sanford, 1985). Described as a 'medical time bomb' because of the ability of asymptomatic chronic infection to flare up suddenly into a rapidly fatal form, chronic melioidosis may persist undiagnosed for decades. Diagnosis is most often achieved by post-mortem examination of infected tissues, or when clinical indications present themselves upon activation of the disease by a traumatic event (Brett *et al.*, 1997). In a serologic surveillance study of US army personnel who had been stationed in Vietnam, 1-2% had significant antibody titres to *B. pseudomallei*, suggesting acquisition of subclinical melioidosis without developing overt disease (Sanford, 1971). In non-endemic areas, diagnostic clues include a history of travel or residence in an endemic area, even in the distant past. Occasionally, patients may present with a fever of unknown origin without any obvious localising signs (Ip *et al.*, 1995).

Melioidosis may be commonly localised in the lung and most patients who have the disease in a severe form show an infection of the lungs (Smith *et al.*, 1987). However, during either acute or chronic infection the organism can be readily isolated from any organ, and lesions can form in any tissue including skin and subcutaneous tissues, bones and joints, liver, spleen, pancreas, kidney, bladder, prostate gland, genital organs, brain and meninges, parotid glands (especially in children), lymph nodes, and pericardium. Moreover, *B. pseudomallei* is similar to *K. pneumoniae* in its predilection for invading the blood stream, causing metastatic abscesses and fulminant pneumonia (Lim, 1997). The histopathology of lesions caused by *B. pseudomallei* vary from acute to chronic granulomatous inflammation, a necrotising reaction with varying numbers of neutrophils, macrophages, lymphocytes and giant cells (Li & You-wen, 1992; Wong *et al.*, 1995) that resembles the reaction caused during tuberculosis (Leelarasamee & Bovornkitti, 1989; Dance, 1990; Sirikulchayanonta & Subhadrabandhu, 1994). Likewise, examination of infected tissue reveals either necrotising suppuration or granuloma formation (Yabuuchi & Arakawa, 1993) or nodules consisting of

polymorphonuclear leukocytes (Liu, 1981). Moreover, chest radiographic changes include airspace consolidation, lung abscesses, and nodular densities, sometimes with cavitation, and, less commonly, pleural effusions (Everett & Nelson, 1975; Chaowagul *et al.*, 1989; Tan *et al.*, 1995).

Further manifestations of melioidosis are regularly described, including the so-called 'neurological melioidosis'. Although long recognised in both animals (Dannenberg & Scott, 1958) and humans (Laws & Hall, 1963), the reported occurrence of such neurological abnormalities has usually been low. In Australia, studies at Royal Darwin Hospital have documented that from 1989 to 1998, there were 232 treated melioidosis cases, 12 of which (5%) displayed neurological melioidosis (Currie *et al.*, 2000b). In Thailand, Punyagupta (1989) documented that 19/686 (3%) patients with melioidosis had CNS involvement. Symptoms comprise cranial nerve palsies, peripheral motor weakness, brain-stem encephalitis, aseptic meningitis and respiratory failure (Woods *et al.*, 1992; Lath *et al.*, 1998). In addition, although bone infection is rare, Subhadrabandhu *et al.* (1995) described 10 cases of melioidotic osteomyelitis.

In summary, melioidosis encompasses a wide clinical spectrum including acute, sub-acute and chronic symptoms. Once the individual is infected, the disease may manifest itself in many ways including local cutaneous lesions, subacute pneumonia, focal organ abscesses and fulminate septicaemia with or without multiple organ abscesses.

2.2.8 Latency and predisposing factors

Latency can be defined as the period of persistence of a pathogen in the host without concurrent damage, but with the potential for reactivation resulting in active pathogenic processes (Penn, 1992). Melioidosis can occur in individuals decades after an initially asymptomatic infection, demonstrating the pathogens persistent nature and ability to cause severe disease later in life when the host immune system is weakened (Dance, 1991). The

incubation period for acute melioidosis is between 1 and 21 days, with a mean of 9 days (Flamma *et al.*, 1969; Achana *et al.*, 1985; Wang *et al.*, 1993; Suputtamongkol *et al.*, 1994; Currie *et al.*, 2000c). It is likely that the inoculating dose, mode of infection, host risk factors and differential virulence of *B. pseudomallei* strains will influence the incubation period (Howe *et al.*, 1971; Leelarasamee & Bovornkitti, 1989). The incubation period for chronic melioidosis ranges from three weeks up to 29 years following exposure to soil or surface water (Chodimella *et al.*, 1997). The activation of latent infection is often associated with concurrent disease or injury. Indeed, the long periods of latency and frequent relapses after antibiotic treatment are characteristic features of chronic melioidosis (Dance, 1990).

The majority of patients with clinically apparent melioidosis are recognised as having underlying diseases: 76% in Malaysia (Putucheary *et al.*, 1992); 88% in Australia (Currie *et al.*, 1993); and 53% in Thailand (Suputtamongkol *et al.*, 1994). Thus it appears that melioidosis is often associated with conditions resulting in immune suppression, such as diabetes, but the disease may also occur without recognisable risk factors. During a 1 year study of patients admitted to a hospital in north-eastern Thailand for treatment of septicaemic melioidosis, 32% demonstrated pre-existing diabetes mellitus (Chaowagul *et al.*, 1989). In the Australian Northern Territory, a similar figure (36%) was observed (Currie *et al.*, 2000a), while the prevalence of diabetes mellitus among patients with melioidosis in Singapore was 64% (Lim, 1997). Similarly, it has been shown in a diabetic infant rat model of infection that such animals are far more sensitive to challenge with *B. pseudomallei* isolates than are the healthy, non-diabetic rats (Woods *et al.*, 1993). Although it is well documented that diabetics have an increased risk of contracting infectious diseases, the possible contribution of insulin to the progress of melioidosis has been investigated. Despite initial evidence showing that human insulin suppressed the growth of *B. pseudomallei* (Woods *et al.*, 1993), a recent thorough reappraisal of the interaction of insulin with 100

clinical isolates found that the growth of *B. pseudomallei* was unaffected by pure insulin (Simpson & Wuthiekanun, 2000). Nevertheless, human insulin has been shown to bind to *B. pseudomallei* cells (Kanai *et al.*, 1996).

Other specific risk factors that have been identified are alcohol and drug abuse, renal disorders, malignancy, HIV infections, connective tissue diseases, heavy occupational exposure and ethnicity (Whitmore & Krishnaswami, 1912; Whitmore, 1913; Leelarasamee & Bonornkitti, 1989; Tanphaichitra, 1989).

Human leukocyte antigen (HLA) polymorphism or HLA linked genes may substantially influence the type and intensity of the immune response that develops following *B. pseudomallei* infection and determine the course of the disease (Dharakul *et al.*, 1998). HLA molecules determine the peptide antigens that are presented to T cells and therefore selectively activate antigen-specific T lymphocyte subsets. Host genetic factors are known to be important in diseases caused by other intracellular pathogens such as tuberculosis and leprosy e.g. DR2 (Khomenko *et al.*, 1990; Mehra, 1990; Rani *et al.*, 1993). Particular HLA class II alleles were common in certain groups of melioidosis patients, especially those with severe clinical disease (DRB1*1602), as well as the relapse cases (DRBI*0701). There was no HLA association with localised melioidosis or melioidosis with pre-existing diabetes mellitus.

2.2.9 Outbreaks of melioidosis in animals

Melioidosis in animals extends the geographic distribution of the disease further into Africa (Chad, Egypt, Madagascar, and Niger) (Dance, 1991) and Europe and outbreaks of melioidosis have been reported in animals from endemic and non-endemic areas. In non-endemic areas these outbreaks have been attributed to the importation of animals from endemic areas. Fournier (1965) considered that this accounted for the introduction of melioidosis into Aruba, and also into Australia after World War II (Dance, 2000abc).

Melioidosis can develop in a wide spectrum of hosts with varying susceptibility. All species of domestic animals may occasionally be affected but the most common are goats, sheep and pigs (Choy *et al.*, 2000). However, mortality in animals such as the banded leaf-monkey, galah, horse, camel, kangaroo and dolphin have been reported (Thomas, 1981; Asche, 1991; Wernery *et al.*, 1998). Although melioidosis has occurred in dogs, cats and cattle, these animals are considered to be fairly resistant to the disease, unless immunocompromised (Choy *et al.*, 2000). Moreover, some animals, such as fowl and water buffalo, are regarded as immune to the organism (Smith *et al.*, 1987). The latter is frequently used to pull ploughs in rice paddies and is, therefore, expected to encounter the microbe on a regular basis (Brett & Woods, 2000).

Among animals, melioidosis was first described from Australia in an outbreak in sheep in 1949 in Winton, north Queensland (Cottew, 1950). Since then there have been reports of mortalities among sheep, goats, and pigs living on Caribbean islands (1957), dolphin in a Hong Kong oceanarium (1975), lambs in southern Australia, cattle in south-east Queensland (Ketterer *et al.*, 1975), and swine in the Burnett River Valley of Queensland (Ashdown & Guard, 1984). The most extensive outbreak of melioidosis in a non-endemic region occurred in France in the mid-1970s amongst the animals of a Paris zoo, which had been exposed to heavily contaminated soil. The disease subsequently spread to other zoos in Paris and throughout France resulting in a number of animal deaths and at least two human fatalities. The origin of the outbreak was traced to either infected horses imported from Iran to France, or a panda donated to France by Mao Tse Tung in 1973 (Dance, 1991).

In summary, a wide range of animals, terrestrial and aquatic are affected by melioidosis with varying severity, and outbreaks often occur following the movement of infected stock.

2.2.10 Animal models of melioidosis

Early experimental studies showed that many animal species such as hamster, guinea

pig, rabbit, monkey, rat and mouse are sensitive to *B. pseudomallei* infection (Miller *et al.*, 1948; Dannenberg & Scott, 1958; Jayanetra *et al.*, 1975). However, the recent availability of animal models for melioidosis has boosted research into pathogenesis. These models exhibit varying degrees of susceptibility to infection, ranging from diabetic infant rats, which are most susceptible, through Syrian hamsters and guinea pigs, to relatively resistant inbred and outbred stains of mice. The difference in susceptibility to *B. pseudomallei* infection has been investigated for strains of inbred mice, revealing that BALB/c mice are highly susceptible whereas C57BI/6 mice are relatively resistant (Leakey *et al.*, 1998). The course of infection in BALB/c mice was similar to acute infections in humans and the mice suffered a rapidly progressive bacteraemia, leading to death after 96 hours. In contrast, infection of C57BI/6 mice appeared to mimic chronic human melioidosis and the mice remained asymptomatic for up to 6 weeks. The reasons for these differences may be related to the greater production of the proinflammatory cytokines TNF-α, IL-1β and IL-6 in BALB/c mice, as compared with C57BL/6 mice, which may contribute to their acute susceptibility (Ulett *et al.*, 2000). In addition to enhancing the antimicrobial activity of macrophages, these proinflammatory cytokines are important mediators of endotoxic shock (Havell & Sehgal, 1991; Takasuka *et al.*, 1991). Regulation of their production is therefore critical to the outcome of infection especially in the early stages of a non-specific immune response.

2.2.11 Treatment, relapse and prevention

The treatment of melioidosis is challenging because *B. pseudomallei* is resistant *in vitro* to a variety of antimicrobial agents including penicillins, first and second generation cephalosporins and many of the aminoglycosides (Dance *et al.*, 1989; Leelarasamee & Bovornkitti, 1989; and Godfrey *et al.*, 1991). In addition, because *B. pseudomallei* can survive and multiply intracellularly, antibiotics that penetrate macrophages are thought to be a much more effective therapy (Slauch *et al.*, 1995), although it has not been proven

clinically. Hence it is important to choose the most appropriate drugs for treatment as quickly as possible in order to reduce the risk of mortality (Leelarasamee & Bovornkitti, 1989).

Except for kanamycin and ceftazidime, all the usable antibiotics are bacteriostatic and may not be effective, especially in the septicaemic form of melioidosis. Until the 1980s, the usual treatment for confirmed cases of acute melioidosis was a combination regime comprising chloramphenicol, tetracyclines and co-trimoxazole (Dance, 2000c). However, for the last decade, ceftazidime is the drug of choice for the treatment of severe melioidosis. It was shown to reduce the mortality of patients with septicaemic melioidosis from 74% to 37% when used alone (White *et al.*, 1989) and from 47% to 18.5%, when used in combination with co-trimoxazole (Sookpranee *et al.*, 1992), compared with conventional combination treatment. However, treatment can involve up to 9 months of antibiotic therapy for clearance of the infection (Leelarasamee & Bovornkitti 1989), and relapse is common.

Relapse after an apparently successful primary treatment (Chaowagul *et al.*, 1993), presumably results from the re-emergence of surviving intracellular bacteria. It has been reported to occur in 20-30% of survivors of severe melioidosis in north-east Thailand after eight weeks treatment (Dance, 1990; Currie *et al.*, 2000a). The explanation for this is the ability of *B. pseudomallei* to survive within phagocytic cells, production of a glycocalyx, and formation of microcolonies in infected tissue (Vorachit *et al.*, 1995). Vorachit *et al.* (2000) suggested that the ideal antimicrobial agents for melioidosis therapy should have a bactericidal effect, be able to penetrate phagocytic cells, and eliminate or inhibit the production of a glycocalyx.

In endemic areas, reducing the risk of exposure by wearing protective clothing or by attempting to avoid exposure to the organism are not realistic options. Therefore efforts have been directed towards prevention by the development of vaccines for effective immunisation

(Brett & Woods, 1996; Dance, 2000c).

2.2.12 Host immune responses

Host defence against infection with facultative intracellular bacteria such as *Salmonella* and *Listeria* is predominantly mediated by cellular immune mechanisms. Macrophages can kill these pathogens particularly after stimulation with gamma interferon (IFN-γ), produced by natural killer cells and T lymphocytes, which enhances antimicrobial activity (Kaufmann, 1993). Although humoral immunity may play a significant role in the resolution of many infectious diseases, this is not the case in melioidosis where high levels of circulating antibodies against the LPS provide no basis for long-term protection (Razak *et al.*, 1986; Bryan *et al.*, 1994). Ulett *et al.* (1998) demonstrated that enriched peritoneal macrophages are unable to effectively kill *B. pseudomallei* without T lymphocyte interaction and attributed the mechanisms responsible for acute and chronic clinical presentations of human melioidosis to the balance between Th1/Th2 responses. This is a similar situation for chronic infections caused by intracellular pathogens such as mycobacteria (Kaufmann, 2000) and *Leishmania major* (Heinzel *et al.*, 1993). Although *B. pseudomallei* resists killing by human macrophages it appears to be more susceptible to neutrophils (Pruksachartvuthi *et al.*, 1990).

The two major microbicidal mechanisms for phagocytes are reactive oxygen intermediates (ROI), consisting of superoxide anion, hydrogen peroxide, hydroxyl radical, and singlet oxygen (Johnston *et al.*, 1975), and reactive nitrogen intermediates (RNI), such as nitric oxide (NO) (Hibbs *et al.*, 1987). The killing mechanisms of INF-γ-activated macrophages are mediated by one or more of these oxidants, although the extent of the contributions of ROI and RNI may differ from one pathogen to another. It has been demonstrated that IFN-γ-activated macrophages inhibit the intracellular growth of *B. pseudomallei* by RNI- and ROI-dependent killing mechanisms, although RNI was more effective (Miyagi *et al.*, 1997). These findings were supported by an *in vitro* experiment

using chemically generated nitrogen- and oxygen-derived oxidants in a macrophage-free system.

Serum levels of neopterin, a stable metabolite produced by IFN-γ-stimulated macrophages, IFN-γ, IL6 and soluble interleukin-2 receptor (sIL-2R) are elevated during melioidosis, and are directly related to disease severity, with higher levels preceding a fatal outcome (Brown *et al.*, 1991; Friedland *et al.*, 1992; Supputtamongkol *et al.*, 1992). Although the early production of IFN-γ has been correlated with resistance to acute infection in a mouse model (Santanirand *et al.*, 1999) the prevailing cytokine millieu is an important factor in the outcome of acute melioidosis (Ulett *et al.*, 2000).

2.2.13 Protective antigens and potential vaccines

Vaccines are usually designed to use immune intervention to tip the balance between infection and disease. Surveys conducted in Thailand and south-east Asia have revealed that at least 15-30% of the normal population possess antibody against *B. pseudomallei* (Atthasampunna *et al.*, 1969; Strauss *et al.*, 1969). However, the high incidence of clinical disease in these areas suggests that circulating antibody is insufficient to protect the infected host against the intracellular pathogen (Razak *et al.*, 1986). Hence, there is a need to stimulate long-term cellular immunity that would provide the host with a rapid response to acute infection and also clear or contain the chronic intracellular bacteria. Preliminary research involved the use of simple preparations such as whole heat-killed bacteria (Stanton & Fletcher, 1932; Razak *et al.*, 1986), toxic filtrates (Nigg *et al.*, 1955; Razak *et al.*, 1986), sonicated cell lysates (Vedros *et al.*, 1983), and protein-polysaccharide mixture (Vedros *et al.*, 1988). Despite initially encouraging results there are as yet no licensed vaccines for melioidosis (Brett & Woods, 2000).

Different *B. pseudomallei* strains express only a single serotype of lipopolysaccharide (LPS) that is highly conserved (Pitt *et al.*, 1992), with two distinct O-polysaccharide (PS)

moieties, PS-I and PS-II (Perry *et al.*, 1995). Interestingly, the nonvirulent *B. thailandensis* (previously known as Ara^+) lacks PS-I (Brett *et al.*, 1998). Because the results of IHA testing suggest that PS is recognised in natural infections, Bryan *et al.* (1994) postulated that a vaccine that utilised the PS portion of the LPS could be potentially useful. Polyclonal and monoclonal antisera raised against tetanus toxoid-conjugated PS provided passive protection in a diabetic rat model of *B. pseudomallei* infections.

Similarly encouraging results were obtained using highly purified flagellar filaments as immunogens. Brett *et al.* (1994) developed an antibody against purified flagellin protein that effectively inhibited the motility of *B. pseudomallei*, and provided passive protection against infection. Extending this work, Brett & Woods (1996) incorporated LPS and flagellar antigens to enhance the immunological repertoire of the vaccine recipient and also to afford protection against strains that display serotypically distinct antigens of one of the two components. The use of a vaccine containing the PS portion but not the toxic lipid A component removes the toxic side effects of LPS but retains the protective advantage of the PS moiety. Consequently, a candidate vaccine, composed of *B. pseudomallei* flagellin proteins conjugated with the endotoxin derived *O* –polysaccharide (PS) antigens, was tested (Brett & Woods, 1996). Passive protection conferred by an immunoglobulin fraction has been demonstrated, however, the efficacy of the vaccine in animals has not yet been reported. Interestingly, passive immunisation experiments conducted by Steinmetz *et al.* (2000) using a single monoclonal antibody with specificity for the repeating tetrasaccharide subunit showed a significant reduction of the bacterial load in various organs in a mouse model of melioidosis. From these results they concluded that the tetrasaccharide subunit might be a promising candidate for use as a protective antigen in active immunisation experiments.

Other possibilities, including the use of killed *B. mallei* and killed *B. pseudomallei*

(Russell, personal communication) and a recombinant vaccine based on other bacteria, including *Francisella tularensis* (Kislitchkine, personal communication), are still in the experimental stages of development.

2.2.14 Molecular epidemiological studies

Information about epidemiology helps us to understand the patterns of transmission of the disease between the environment, animals and man, the extreme variation in clinical outcomes of the infected patients and the role of different bacterial strains in pathogenesis (Sermswan *et al.*, 1994; Choy *et al.*, 2000). Additionally, a typing system would be helpful in determining whether relapses are due to reinfection or to the recurrence of the original infection (Sermswan *et al.*, 1994). On the other hand, epidemiological investigations are highly dependent on the reliability and reproducibility of typing systems to differentiate between isolates (Norton *et al.*, 1998).

Molecular typing of *B. pseudomallei* has been attempted using ribotyping (Desmarchelier *et al.*, 1993; Trakulsomboon *et al.*, 1997), randomly amplified polymorphic DNA (RAPD) (Haase *et al.*, 1995; Choy *et al.*, 2000), multilocus enzyme electrophoresis (MEE) (Norton *et al.*, 1998), restriction fragment length polymorphism (RFLP) (Sermswan *et al.*, 1994), pulsed field gel electrophoresis (PFGE) (Trakulsomboon *et al.*, 1994), and multiplex PCR (Wongratanacheewin *et al.*, 2000). Generally, these methods have shown that enormous diversity exists amongst isolates from endemic and non-endemic areas.

Early serotyping of *B. pseudomallei* using the heat-stable O antigen, found limited diversity among strains and two serotypes were identified (Chambon & Fournier, 1956); one was found in Australia and southeast Asia, while the other was unique to Australian isolates (Dodin & Fournier, 1970). On the other hand, favouring the inclusion of *B. pseudomallei* in a separate species Trakulsomboon *et al.* (1997) presented evidence, by RFLP, suggesting that *B. pseudomallei* strains that utilise arabinose (later known as *B. thailandensis*)

constituted a population that was distinct from some other environmental and clinical strains.

Little is known about the strains that cause recurrent or relapsing melioidosis and their genetic relatedness with the primary strain infecting the patient. In an attempt to improve our understanding of this relationship, Desmarchelier *et al.* (1993) examined *Bam*H1 ribotype patterns among 54 isolates and concluded that isolates of similar patterns from a single patient represented a single strain. They also found that for most melioidosis patients recurrent infection was probably due to the original infecting strain, although reinfection with strains of similar types in the environment can not be ruled out. The ribotyping and PFGE studies of Vadivelu *et al.* (1998) supported these findings. Nevertheless, the work highlighted the possibility that some patients may be infected by more than one strain of *B. pseudomallei*.

Ribotyping and RAPD analysis of 10 isolates from a nontropical focus of animal and human disease over a 25 year period indicated the clonality, strain stability, and persistence of *B. pseudomallei* under circumstances where the organism was most likely introduced as a single strain (Currie *et al.*, 1994; Haase *et al.*, 1995). However, the clinical outcome of melioidosis could not be correlated with RAPD patterns (Haase *et al.*, 1995). On the other hand, Norton *et al.* (1998) compared the genetic inter-relatedness of 18 clinical isolates using RAPD and MEE in order to investigate the possible existence of an epidemiological link between cases of melioidosis that occurred at about the same time in four different locations in North Queensland in Australia. The isolates segregated into two groups that correlated with clinical presentation rather than geographical location, one group consisted mainly of patients with respiratory disease while the second consisted of isolates from patients with abscesses in the viscera rather than the lung or brain. However, these findings have not been extended to studies involving a larger group of clinical isolates.

Interestingly, a comparison of the *Bam*HI ribotypes of 350 isolates of *B. pseudomallei* recovered over 71 years from 23 countries revealed 44 distinct patterns, nineteen of which were unique to single isolates and more than three-quarters of the collection fell into just nine ribotypes (Pitt *et al.* 2000). Two ribotypes, 1 and 3, were the most common accounting for almost half the collection. Ribotype 3 was prevalent in Asian countries while ribotype 1 was more widely distributed. In addition, the clinical features of 113 melioidosis patients infected with *B. pseudomallei* strains of different ribotypes were investigated to assess whether certain strains or groups were associated with a particular clinical outcome. Disease association was suggested for 4 of the ribotypes and strains of ribotype 4 were markedly associated with a fatal outcome, while the risk of relapse with isolates of ribotype 1 was more than twice as high as for strains of other ribotypes. In addition, Wongratanacheewin *et al.* (2000) demonstrated, by multiplex PCR, that 21 type patterns were obtained from 37 *B. pseudomallei* environmental and clinical isolates from human and animals. Among human and animal isolates, 10 types were described, of which two correlated with the disseminated septicaemic melioidosis commonly reported in northeast Thailand.

2.3 HOST-PATHOGEN INTERACTIONS

Early views of pathogenicity and virulence were primarily pathogen centred although it was recognised that pathogenicity was neither invariant nor absolute (Casadevall & Pirofski, 1999). In the light of new information on the nature of the complex series of interactions between host and pathogen it becomes obvious that the outcome is as dependent on the host as it is upon the properties of the pathogen (Falkow, 1997). Any host-pathogen interaction can be reduced to two possible outcomes: those that result in damage to the host and those that result in no damage. Disease occurs when the host sustains damage that is sufficient to impair the normal functioning of cells (including necrosis, apoptosis, and malignant transformation), tissues, and organs (including granulomatous inflammation, fibrosis, and

tumor) (Sparling, 1983). The consequences of an infection will depend on a combination of factors including the virulence of bacterial pathogen, the immune status of the host and the innate resistance of the host (Dougan, 1989). However, numerous studies have demonstrated the existence of what appear to be 'universal' virulence mechanisms used by diverse bacterial species in order to fulfil the following five stages of pathogenicity. Pathogenic bacteria must be able to (i) adhere to host tissues, (ii) invade host tissues (usually), (iii) multiply in host tissues, (iv) evade host defenses, and (v) cause damage (Mitchell, 1998). Each (or all) of these stages should, somehow, involve means for avoiding or subverting the host's normal defences. The bacterial products responsible for these five biological requirements are the determinants of pathogenicity or virulence factors.

It is beyond the scope of this thesis to examine each of these aspects in detail and excellent reviews are readily available (Falkow, 1997; Finlay & Falkow, 1997; Casavell & Pirofski, 1999; Smith, 2000). Recent work has sought to improve our understanding of the molecular basis of melioidosis and this work will be considered in the light of these 'universal' aspects of bacterial pathogenicity.

2.4 MOLECULAR BASIS OF THE PATHOGENESIS OF MELIOIDOSIS

Basic understanding of the pathogenic properties of *B. pseudomallei* is the cornerstone for the control, and hence the prevention of melioidosis. A multifactorial mechanism has been proposed for the pathogenesis of melioidosis (Wongwanich *et al.*, 1996) and *B. pseudomallei* is capable of expressing an impressive array of both secreted and cell-associated antigens. However, the role of these products in the pathogenesis of disease has to date been relatively ill defined (Brett & Woods, 2000).

2.4.1 Extracellular products

Like many other bacteria, *B. pseudomallei* produces many components that are secreted into the extracellular medium, with a variety of biological properties. These extracellular

products include a siderophore, haemolysins, proteases, lipase, lecithinase (phospholipase C), and toxins.

2.4.1.1 Siderophores

Iron is an essential requirement for both *in vivo* and *in vitro* growth of nearly all microorganisms. Although the mammalian host has an abundance of iron, most is located intracellularly. The small amount of iron remaining is withheld by iron-binding or transport proteins, transferrin in the serum, and lactoferrin in mucosal secretions and specific granules of neutrophils (Spitznagal *et al.*, 1974; Crichton & Wauters, 1987). Many pathogens are able to acquire iron from the host by producing specific, water-soluble, low-molecular weight, high affinity iron binding compounds termed siderophores. They bind iron with high affinity and transport it into the bacterial cells by using specific receptor proteins and appropriate transport mechanisms (Neilands, 1993). Production of siderophores is associated with virulence in many animal pathogens (Bullen & Griffiths, 1987; Weinberg & Weinberg, 1995), e.g. *Salmonella*, *Shigella* spp., *P. aeruginosa and Staphylococcus aureus*. Moreover, the invasiveness of *E. coli* strains often correlates with the production of aerobactin, a hydroxamate-type siderophore (Jacobson *et al.*, 1988) and *B. cepacia* express at least 3 siderophore-mediated iron transport systems including pyochelin, cepabactin and azurechelin (Sokol, 1986; Meyer *et al.*, 1989). Morbidity and mortality in infected CF patients has been correlated to the production of pyochelin that may assist the dissemination of *B. cepacia* throughout the lungs (Sokol & Wood, 1988).

A siderophore, malleobactin, has been described by Yang *et al.* (1991) in 84 *B. pseudomallei* clinical isolates, grown under iron-deficient conditions. Structural and chemical analysis of the water-soluble molecule revealed that it belongs to the hydroxamate class and has no yellow-green fluorescence, which is different from the siderophores found in fluorescent Pseudomonads. Malleobactin, 1000 Da, was shown to promote iron uptake by

B. pseudomallei from both transferrin and lactoferrin with a preference for transferrin, and to reverse the growth inhibition caused by human transferrin *in vitro* (Yang *et al.*, 1993).

Very recently, Loprasert *et al.* (2000) cloned and sequenced the ferric uptake regulator gene (*fur*) of *B. pseudomallei* that encodes a 16 kDa protein that shares antigenic determinants with the Fur protein of *E. coli*. Fur protein represses the transcription of iron-regulated promoters in response to an increasing intracellular iron concentration. *Fur* mutants of *B. pseudomallei* that produced an elevated level of siderophore, were defective in their ability to repress siderophore biosynthesis, and also showed reduced ferric superoxide dismutase (FeSOD) and peroxidase activities. Because these oxidative stress response enzymes may have a role in the detoxification of reactive oxygen metabolites during the oxidative burst produced by professional phagocytes, they may assist the intracellular survival of *B. pseudomallei* (section 2.1.5). The authors postulated that *B. pseudomallei* might have a strategy to resist the toxic actions of these metabolites following enzymatic conversion by SOD, peroxidase and other oxidoreductases, as the case for *P. aeruginosa* (Hassett *et al.*, 1996) and the facultative intracellular bacteria *B. pertussis*.

2.4.1.2 Extracellular enzymes

In an attempt to study the role of excreted products in the pathogenesis of the bacterium, one hundred clinical isolates of *B. pseudomallei* from human were tested for the production of extracellular biologically active substances (Ashdown & Koehler, 1990). Over 90% of the *B. pseudomallei* isolates tested positive for lecithinase, lipase, protease and haemolysin but none was positive for elastase. In addition, four of the isolates had a particularly high level of haemolytic activity (alpha haemolysis). The haemolysin was shown to be heat labile, most active in an acid environment (pH 5.5), and cytolytic in broth culture filtrate for a variety of animal and human erythrocytes. The authors concluded that there are at least two haemolysins produced by *B. pseudomallei*, one is common and weakly

cytolytic, while the other occurs infrequently with stronger, heat-labile cytolytic activity.

It seems likely that the level of haemolytic activity produced by *B. pseudomallei* is dependent on the medium and the culture conditions. Whilst some earlier studies showed that the organism was non-haemolytic (Chambon & Fournier, 1956; Wetmore & Gochenour, 1956), other researchers have reported distinct alpha haemolysis (Crotty *et al.*, 1963; Salisbury & Likos, 1970). The haemolysins produced by many bacteria are known to contribute to their ability to cause disease, including tissue damage, intracellular survival, and virulence e.g. *E. coli* cytolytic haemolysin (Jonas *et al.*, 1993), phospholipase (Jepson & Titball, 2000), but the contribution of haemolysins to the course of *B. pseudomallei* infection remains unclear.

Sexton *et al.* (1994) characterised an extracellular protease, termed exoprotease, molecular weight 36000, which is a metal-binding enzyme requiring iron for maximal activity and belongs to the alkaline protease group of metalloproteases. Antibodies developed against an alkaline protease produced by *P. aeruginosa* cross-reacted with the *B. pseudomallei* exoprotease. The purified exoprotease was capable of digesting a variety of eukaryotic proteins including immunoglobulins. The authors detected antibodies to *B. pseudomallei* protease in the sera of melioidosis patients indicating that the protease is produced *in vivo* and that this enzyme may help to inactivate host defence mechanisms and facilitate the invasion of host tissues. Furthermore, they reported that studies in animals indicated that a *B. pseudomallei* strain deficient in protease production produced significantly less lung damage than its parent strain. Hence, they referred the ability of *B. pseudomallei* to elicit lung damage during infection to the production of the exoprotease *in vivo*. Several studies have demonstrated that purified proteases from *B. cepacia* (McKevitt *et al.*, 1989) and *P. aeruginosa* (Gray & Kreger, 1979), exert direct histopathological changes to rat lungs *in vivo* and the latter degrade a variety of host proteins and have a direct effect

on skin (Holder & Neely, 1989).

B. pseudomallei also produces another protease of 42000 Da in size (Percheron *et al.*, 1995). More recently, Gauthier *et al.* (2000) assessed the production of this protease in culture filtrates from 19 different *B. pseudomallei* strains. No correlation was observed between virulence and proteolytic activity of *B. pseudomallei*. On the contrary, the most virulent strain tested produced 17 times less protease than a less virulent strain. Therefore, they argued that the secretion products from *B. pseudomallei* are not a critical factor in the pathogenesis of acute melioidosis. This was in agreement with Brett *et al.* (1997) who showed that administration of concentrated culture supernatant, containing proteases, from *B. pseudomallei* strains of higher and lower virulence were ineffective for eliciting an acute illness in a mouse animal model of infection.

More recently, Lee & Liu, (2000) cloned and sequenced a novel protease gene, *mprA* from *B. pseudomallei*, a serine metalloprotease containing a signal sequence that could cleaved to give 47-kDa mature protein. However, no correlation between virulence and level of protease activity was found when *B. pseudomallei* was intraperitoneally injected in mice.

By using transposon mutagenesis in *B. pseudomallei*, the mechanism for the secretion of hydrolytic enzymes into the extracellular milieu (including lecithinase, lipase, and protease) has recently been elucidated. DeShazer *et al.* (1999) identified an 11.8kb chromosomal locus in the general secretory pathway of twenty-nine unique Tn5-OT182 mutants that are unable to secrete protease, lipase and lecithinase. The mutants were found to be as virulent as the wild type in the Syrian hamster model, suggesting that these secretion products have limited role in the virulence of this organism (Dharakul & Songsivilai, 1999).

Phospholipase C (PLC) is an enzyme that cleaves phosphatidylcholine, a major lung surfactant, to yield phosphorylcholine and diacylglycerol, and has been associated with cytopathology of lung tissue (Nelson *et al.*, 1994). Previous studies have implicated PLCs as

virulence factors involved in infection by pathogenic bacteria such as *L. monocytogenes* where they play a significant role in escape of the pathogen from the phagosome membrane and invasion of adjacent cells (Smith *et al.*, 1995). Recently, a nonhaemolytic phospholipase C gene (PC-PLC) has been cloned from *B. pseudomallei* (Korbsrisate *et al.*, 1999). Western blot analysis with sera from melioidosis patients indicated that they produced immunoglobulin M antibodies against this PC-PLC protein. The authors postulated that PC-PLC of *B. pseudomallei* might play a similar role to that of *L. monocytogenes.*

In summary, the precise role of proteases, haemolysins, lipases or lecithinases in the pathogenesis of *B. pseudomallei* infections remains to be elucidated.

2.4.1.3 Exotoxins

Exotoxins are proteins produced by growing bacteria that are usually secreted into the surrounding medium, but are sometimes bound to the bacterial surface and released upon lysis. They subvert, alter, or destroy host cell functions, ultimately leading to cellular dysfunction or death (Moayeri & Welch, 1998). The mechanism of exotoxin action includes assisting bacterial spread in tissues (e.g. hyaluronidase of streptococci breaks down hyaluronic acid, the ground substance of connective tissue), lysing cells either by lipase action (e.g. lecithinase of clostridia) or insertion in the membrane to form pores (e.g. α-toxin of *S. aureus*; streptolysin O of certain streptococci), or by blocking cellular protein synthesis (e.g. diphtheria and cholera toxins) (Obrig, 1994; Schaechter, 1998).

Bacterial exotoxins that inhibit eukaryotic protein synthesis do so by a variety of means. For example, ADP ribosylation of elongation factor-2 (EF-2) by diphtheria toxin and *P. aeruginosa* exotoxin A inactivates host protein synthesis (Domenighini *et al.*, 1995) and increase epithelial permeability. *Shigella* spp. shiga toxin and *E. coli* shiga-like toxin inactivate the 60S ribosomal subunits by inhibiting EF-1-dependent amino-acyl tRNA binding to the ribosome when they cleave an *N*-glycosidase bond of the adenine at position

4324 in 28S rRNA (Donohue-Rolfe *et al.*, 1988).

Investigations of *B. pseudomallei* toxins date from the 1950s. Nigg *et al.* (1955) raised the possibility that the pathogen produces exotoxic material. They observed that the inoculation of viable *B. pseudomallei* into mice or guinea pigs resulted in death within one or two days. A crude toxin preparation of bacterial filtrate, which was thermolabile, was not only lethal for mice and hamsters but also produced haemorrhagic necrotic lesions following the intradermal inoculation of normal guinea pigs. Hence, they associated both lethal toxicity and necrotic activity with the heat-labile components of the filtrate. This was substantiated by further studies (Heckly & Nigg, 1958). Studies conducted by Liu (1957) and Ismail *et al.* (1987) demonstrated that the lethal exotoxin is separate from the proteolytic enzyme. The components that possess necrotic activity are heat-stable, while the lethal toxicity of *B. pseudomallei* filtrate is heat-labile. A thermolabile toxin was purified to homogeneity and characterised as a protein of 31000 Da (Ismail *et al.*, 1987).

In a study to investigate the function of *B. pseudomallei* exotoxin at the cellular level, Mohamed *et al.* (1989) suggested that exotoxin may play an important role in the early stages of infection by exerting a direct and irreversible toxic effect on the host's macrophages. They demonstrated that both DNA and protein synthesis were inhibited in macrophages which were exposed to the exotoxin, and suggested that it does not belong to the membrane-damaging class of bacterial protein toxins such as phospholipases, haemolysins, and cytolysins. However, demonstration of the cytolethal effect required a relatively high concentration of 20 μg toxin/ml, an amount equivalent to 50-ml bacterial culture, suggesting that it is unlikely that this toxin could be the cause of the very rapid onset of septicaemic melioidosis. Nevertheless, some workers believe that exotoxins are responsible for the rapid onset of the disease and the dramatic course of septicaemic melioidosis (Haase *et al.*, 1997). Woods *et al.* (1992) postulated that 'neurological'

melioidosis (Lath *et al.*, 1998) could be attributed to an exotoxin-mediated pathology in the absence of direct infections of the central nervous system. However, Currie *et al.* (2000b) suggested that it was more likely to be due to direct invasion.

Wongwanich *et al.* (1996) tested the cytotoxic activity in fractions obtained from both culture filtrates and cell-free extracts of seven strains of *B. pseudomallei* by examining the morphological effects on cultured cells. All fractions of *B. pseudomallei* brought about complete (100%) and rapid destruction of the cultured-cells. Recently, Haase *et al.* (1997) purified the cytotoxic activity in culture filtrates and identified an exotoxin, that is active in very small amounts, had an apparent molecular weight of ~ 3000 and is most likely to be a small peptide. The fast acting nature of the toxin suggested that it may act directly on cell membranes rather than interacting with intracellular components and does not therefore represent an ADP ribosylating toxin. The authors proposed that toxin synthesis may be upregulated during the course of *B. pseudomallei* infection.

A different heat-stable extracellular toxin produced by *B. pseudomallei*, has been purified by Haubler (1998). This novel haemolytic glycolipid, with a molecular mass of 762 Da, was cytotoxic to phagocytic (HL60) and non-phagocytic (HeLa) cell lines. This activity was resistant to alkali and acid treatments, but was neutralised by albumin which suggests that it is unlikely to be of significance in the pathogenesis of *B. pseudomallei* infections (Brett & Woods 2000).

In conclusion, although a variety of toxins have been reported, the precise role of these molecules in the pathogenesis of melioidosis remains obscure.

2.4.2 Cell-associated factors

Bacterial pathogens synthesise a variety of cell-associated components that may be exposed on the surface of the organism including flagella, pili, acid phosphatase, capsule and lipopolysaccharide. These features are often associated with the progression of infection,

pathology of the disease, and the immune response of the host (Anuntagool *et al.*, 2000).

2.4.2.1 Flagella

For many pathogenic species, the contribution of flagella to virulence has not been clearly established. Nonetheless it is generally accepted that flagella make an important contribution to motility, and the expression and repression of flagellation and virulence have in some cases been shown to be linked (Moens & Vanderleyden, 1996). Moreover, the presence of flagella on the surface of pathogenic and opportunistic bacteria is thought to facilitate colonisation and dissemination from the initial site of attachment (Rozalski *et al.*, 1997). Currently, flagellar motility has been shown to be implicated in disease caused by many pathogens including *Helicobacter* (Eaton *et al.*, 1992), *Campylobacter* (Grant *et al.*, 1993), *Salmonella* (Carsiotis *et al.*, 1984), *Legionella* (Pruckler *et al.*, 1995), and *Vibrio* (Milton *et al.*, 1996).

Bacterial flagella consist of a long helical filament that is driven by a rotary motor anchored in the cell surface. The main structural protein of the flagellar filament is flagellin. Flagellin proteins from several different strains of *B. pseudomallei* have been isolated and purified to homogeneity (Brett *et al.*, 1994). One single subunit type constitutes the flagellar filaments of *B. pseudomallei* strains. This agrees with evidence suggesting that nonsheathed bacterial flagella display only a single flagellin species when disaggregated (Joys, 1988). A polyclonal antibody raised against the 43.4 kDa flagellin protein reacted with 64 of 65 *B. pseudomallei* strains that were tested. Winstanley *et al.* (1998) cloned and sequenced the flagellin genes from one environmental and three clinical isolates of *B. pseudomallei*. They compared, by PCR, the genes from four clinical and six environmental isolates and demonstrated that while the genes from the clinical isolates were either 100% identical or differed by only one nucleotide, greater genetic variation was detected among some of the environmental isolates.

It is of note that mutagenesis and animal model experiments have failed to highlight a role for the flagellum in the pathogenesis of *B. pseudomallei* infection. In an attempt to identify and characterise potential virulence determinants in *B. pseudomallei*, DeShazer *et al.* (1997) used the transposon Tn5-OT182 to identify multiple genes involved in motility, including the flagellin structural gene, *fliC*. Twenty-eight unique motility mutants containing Tn5-OT182 integrations in a variety of genes encoding proteins with homology to flagellar and chemotaxis proteins in *E. coli* and *S. typhimurium*, were identified. There was no significant difference in the virulence of the mutant strain 1026b compared to that of the wild type, in diabetic rats or Syrian hamster, suggesting that flagella and/or motility are probably not virulence determinants in these animal models of *B. pseudomallei* infection.

2.4.2.2 Pili

The first stage in the infection process is colonisation of a host mucosal surface. Many pathogenic bacteria adhere to host cells using fimbriae or pili that specifically bind to the carbohydrate moiety of glycolipids or glycoproteins. Fimbrial-associated receptor recognition permits bacteria to adhere to diverse targets ranging from inorganic substances to highly complex biomolecules. A pilus is a rod-shaped structure, which consists primarily of an ordered array of helically arranged single protein subunits called pilin. According to their morphology and function, pili are classified into conjugative (F or sex), which are rigid and have a large axial hole, and adhesive pili, which are flexible and may have a small axial hole. Adhesive pili bind to receptors on host epithelial cells, allowing pathogens to establish a niche within the host. Binding via pili may not only prevent the elimination of bacteria from a desirable niche in the host, but may also trigger the expression of a cascade of virulence factors that enable the bacterium to thrive in its new environment (Kuehn, 1997).

The occurrence of fimbriae or pili on *B. pseudomallei* had not been reported until Smith *et al.* (1987) revealed, by electron micrographs, that *B. pseudomallei* was heavily

fimbriated. They correlated the presence of pili with the haemagglutinating activity of the organism. By using experimental techniques to depilate bacterial cells, these authors demonstrated that the ability of *B. pseudomallei* to adhere to human buccal cavity epithelial cells is mediated through pili.

The production and assembly of pili involves a number of genes. Unpublished data from D. E. Woods' group in Canada suggests that *B. pseudomallei* possesses at least two types of pili. This group managed to clone and sequence genes (*fimA*) encoding a 15-kDa pilin structural protein that exhibits homology to type I pili, and the accessory genes involved in the assembly of type I pili, FimC, a chaperone protein, and FimD, an outer membrane usher protein (Reckseidler *et al.*, unpublished). Moreover, the same group cloned genes with strong homology to *pilB*, *pilC*, and *pilD* of *P. aeruginosa*. In addition, Kuehn (1997) reported that *B. pseudomallei* also contain the genes involved in the production of type IV pili that are known to mediate attachment to host epithelial cells in an increasing number of bacterial pathogens including *V. cholerae*, *P. aeruginosa*, and *N. gonorrhoeae* (Hobbs & Mattick, 1993). The pili present in *P. aeruginosa* and *B. cepacia* adhere to buccal and tracheal epithelial cells *in vitro* (Kuehn *et al.*, 1992; Wilkinson & Pitt, 1995; Toder, 1998). Electron microscopy has shown that approximately 60% of *B. cepacia* strains express peritrichous pili (Kuehn *et al.*, 1992) that facilitate attachment to the respiratory mucin from CF and non-CF individuals (Sajjan & Forstner, 1992).

2.4.2.3 Acid phosphatase

Early studies of bacterial acid phosphatases associated with the cell surface have proposed a pathogenic function (Walther *et al.*, 1984). For example, in *S. typhimurium* a regulator gene, *phoP*, controls the expression of the genes of *pag* loci including the locus for non-specific acid phosphatase (*pagN*) and other loci needed for survival within macrophages (Miller & Mekalanos, 1990). Dejsirilert *et al.* (1989) described *B. pseudomallei* acid

phosphatase and proposed that the physiological role of the molecule may be to allow the organism to obtain orthophosphate from organic phosphate sources to meet nutritional demands. Kondo *et al.* (1991a) reported the presence of heat-stable and heat-labile components of acid phosphatase that differed in pH optima, 4.2 and 5.2, respectively. The heat-stable molecule was a membrane-associated enzyme and the authors speculated that heat-stability would result from the progressive glycosylation of the molecule. Further characterisation by Kondo *et al.* (1991b) revealed the enzyme was concentrated in a glycoprotein fraction that was recognised by antibodies in serum from melioidosis patients. The *B. pseudomallei* enzyme had the highest activity among the Pseudomonad species that were tested, with a peak at around pH 5.0, similar to the pH of the phagolysosome (Egan & Gordon, 1996), and has a preference for phosphotyrosine and phosphoserine substrates (Kanai & Kondo, 1991). Hence, it was assumed that the enzyme, a glycoprotein, might be a cell-surface receptor similar to protein tyrosine phosphatase (PTP). PTP, which is considered to be a part of a two-component system involved in signal transduction, accepts and transmits messages from the environment to the intracellular millieu (Walton & Dixon, 1993). Most acid phosphatases are distributed widely in the animal, plant and microbial kingdom and are glycoproteins (Vogel & Hinnen, 1990). A similar study by Kondo *et al.* (1994) confirmed the presence of glycoproteins that are synthesised within the cell and secreted after translocation to the cell surface. They emphasised that acid phosphatase is a cell-surface enzyme and that it is a glycoprotein whose polysaccharide moiety may project outside the cell. Moreover, Kondo *et al.* (1996) established that the acid phosphatase could be readily isolated from both whole cell and supernatant fractions. Kanai *et al.* (1996) have proposed that the cell-associated form of PTP is a high affinity receptor for insulin and that this receptor-ligand interaction may be responsible for modulating the enzymatic profiles of *B. pseudomallei* isolates. The authors postulated that the PTP interacts with insulin in a

similar manner to protein tyrosine kinase which is thought to work as a regulator of cell growth, differentiation, and metabolism (Jin & Inoue, 1993).

Some bacteria have cell surface proteins that are important for adherence but do not assemble themselves into pilin-like structures (Salyers & Whitt, 1994). *P. aeruginosa* and *B. cepacia* have receptors that specifically bind tissue glycolipids in a receptor-ligand relationship (Krivan *et al.*, 1988), which raises the possibility that a similar system may operate to mediate adhesion between *B. pseudomallei* and host tissue in the early stages of infection. Kanai *et al.* (1997) demonstrated that *B. pseudomallei* binds specifically to tissue glycolipids and suggested that cell-surface acid phosphatase, with its strong affinity for glycolipids, may be a receptor molecule.

2.4.2.4 Lipopolysaccharides or endotoxins

Lipopolysaccharides (LPS) also known as endotoxins are major constituents of the outer membrane of Gram-negative bacteria that act as toxins only under special circumstances. LPS can be released from the cell surfaces of bacteria during replication, lysis, and death and cause a broad spectrum of pathophysiological effects (Rietschel *et al.*, 1994). At low concentrations, LPS elicits a series of alarm reactions including fever, activation of complement, activation of macrophages, and stimulation of B lymphocytes, while at high concentrations it produces shock and even death (Schaechter, 1998). Unlike exotoxins, endotoxins are not proteins but rather complex molecules with some exotic chemical constituents that are much more heat stable. LPS consists of a polysaccharide and a lipid A moiety that is embedded in the outer membrane (**Fig 1**, Appendix II). Lipid A is the toxic (active) portion of the molecule and because of its location in the outer membrane, it exerts its effects only when bacteria lyse. On the other hand, the polysaccharide moiety consists of a core and O side chain. Some of the diseases in which endotoxin may play an important role include typhoid fever, tularaemia, plague and brucellosis in addition to a

variety of hospital-acquired infections caused by opportunistic Gram-negative pathogens that include *E. coli*, *Proteus*, *P. aeruginosa*, *Enterobacter*, *Serratia* and *Klebsiella* (Mims *et al.*, 1995). For example, the LPS of *B. cepacia* has endotoxin-like activities including lethality and tissue necrosis in CF patients (Straus *et al.*, 1989) whereas the LPS of *H. influenzae* enhances bacterial survival in the nasopharynx and facilitates invasion across cellular barriers by disruption of ciliated epithelial cells (Johnson & Inzana, 1986).

The evaluation of the role of LPS in the virulence and pathogenicity of *B. pseudomallei* has over the years given inconsistent and controversial results (Anuntagool *et al.*, 1998). Rapaport *et al.* (1961) and Redfearn (1964) reported that *B. pseudomallei* possessed endotoxin activities but the chemical and biological characterisation was not completed until recently. The importance of LPS in the pathogenesis of melioidosis was highlighted when patients with septicaemic melioidosis displayed high levels of anti-LPS antibodies (Petkanjanapong *et al.*, 1992) and had significantly elevated levels of proinflammatory cytokines, which are known to be activated by LPS (Friedland *et al.*, 1992; Suputtamongkol *et al.*, 1992). Pitt *et al.* (1992) demonstrated that the LPS expressed by 12 clinical strains of *B. pseudomallei* which had been isolated over a period of 70 years from southeast Asia and Australia exhibited identical patterns. Similarly, Anuntagool *et al.* (1998) examined 214 isolates of *B. pseudomallei* (119 Ara$^-$ clinical, 13 Ara$^-$ soil, 70 Ara$^+$ soil, and 8 reference National Type Culture Collection strains) and identical LPS patterns in 210 of these isolates were typical for LPS from other Gram-negative bacteria. Moreover, the same 210 isolates exhibited similar immunoblot profiles using either pooled sera from patients with melioidosis or hyperimmune mouse sera.

Because the LPS of the two different *B. pseudomallei* biotypes (later considered to be two different species) appeared antigenically indistinguishable, Anuntagool *et al.* (1998) concluded that it is unlikely that LPS is related to the virulence and pathogenicity of *B.*

pseudomallei. Consistent with these studies, Iwasa *et al.* (1992) and Matsuura *et al* (1996) published reports concerning the endotoxin activities of *B. pseudomallei*. Iwasa *et al.* (1992) studied heat-treated whole cells, culture supernatants, and extracted endotoxin preparations of *B. pseudomallei* in comparison with the same preparations from *B. cepacia* and *P. aeruginosa*. Although the endotoxic effect was detected in all the samples of *B. pseudomallei*, the activity was 30 times less than that present in samples from *P. aeruginosa*. Matsuura *et al.* (1996) compared the biological properties of *B. pseudomallei* LPS with those of a reference LPS obtained from *Salmonella abortus equi*. The endotoxic activities of *B. pseudomallei* LPS, such as pyrogenic activity in rabbits, lethal toxicity in mice, and macrophage activation, are 10-100 times weaker than *Salmonella* LPS.

Moreover, as the shedding of some surface components, particularly the LPS, is known to be associated with pathogenicity and virulence in a number of Gram-negative bacteria (Rietschel *et al.*, 1996), Anuntagool *et al.* (2000) compared the shedding of LPS in cultures of Ara^+ and Ara^- isolates, using MAb-based ELISAs. Although the MAb was found to react with a 200-kDa surface antigen that was present only in the Ara^- isolates (Rugdech *et al.*, 1995; Sirisinha *et al.*, 1998), the results showed that there was no difference in the LPS quantity or shedding profiles of the two biotypes.

Studies have shown that *B. pseudomallei* can produce two structurally distinct LPS, simultaneously. Knirel *et al.* (1992) and Perry *et al.* (1995) revealed by chemical and physical methods, that one O-antigenic polysaccharide is a high-molecular weight unbranched 1,3-linked homopolymer of 2-O-acetylated 6-deoxy-β-D-manno-heptopyranosyl residues, while the other is an unbranched heteropolymer consisting of (-3)-β-D-glucopyranose-(1-3)-6-deoxy-α-L-talpopyranosyl-(1-disaccharide repeats (L-6dTal*p*: ~33% *O*-4 acetylated and *O*-2 methylated; ~66% *O*-2 acetylated).

In conclusion, *B. pseudomallei* LPS, although chemically unique, is no more potent or

unique in its biological properties than the LPS from other organisms. If anything, the LPS from *B. pseudomallei* appears to be less potent than the LPS from *E. coli*, for example, in activating macrophages.

2.4.2.5 Capsule

Exopolysaccharides (EPS), also called capsular polysaccharides, were among the first known bacterial virulence determinants in both gram-negative and gram-positive bacteria. They include structures varying from discrete capsules that are closely associated with the cell, to massive slime wall formations that are only loosely associated with the cell surface (Steinmetz *et al.*, 1995).

The presence of a slime layer around the cells of *B. pseudomallei* has been noted from the gross appearance on laboratory media (Jayanetra, 1989) and the bacterium has a biofilm producing, mucoid phenotype *in vitro* and *in vivo* (Vorachit *et al.*, 2000). The pathogenic significance of the *B. pseudomallei* capsule has been known for more than a decade. Smith *et al.* (1987) verified the presence of a capsule using light microscopy on wet and dry Indian ink preparations, and by morphological observations with an electron microscope. *B. pseudomallei* was found to be encapsulated when isolated from dead mice that had been infected with the organism. The encapsulated form could also be produced *in vitro* by growing the organism in a chemically defined medium supplemented with 1% (w/v) lactose and 0.1% (w/v) peptone.

Studies *in vitro* have determined that both capsulated and non-capsulated *B. pseudomallei* strains are equally susceptible to phagocytosis by neutrophils (PMN), and that the presence of exopolysaccharide confers some resistance to the bactericidal effects of the phagolysosomal compartment (Smith *et al.*, 1987; Pruksachartvuthi *et al.*, 1990). Ho *et al.* (1997) proposed that the capsular polysaccharide produced by *B. pseudomallei* was important in pathogenesis and may provide protective immunity to immunised hosts.

Furthermore, Vorachit *et al.* (1995) attributed chronic infections to the production of exopolysaccharide glycocalyces, as demonstrated by scanning electron microscopy of lung tissue from infected humans and animals. Using ruthenium red staining of macrophages containing ingested *B. pseudomallei*, Leelarasamee and Bovornkitti (1989) were unable to demonstrate the capsule by electron microscopy, although they described the appearance of an electron-dense layer surrounding the bacterial cells that may suggest the release during growth of one or more soluble slime-like substances. On the other hand, using the same technique, Puthucheary *et al.* (1996) revealed the existence of three morphologically distinct variants of bacterial cultures: one with a very marked and another with a less distinct electron-dense layer, termed glycocalyx, surrounding the cell wall, and a third variety devoid of such structures. They suggested that these three morphologic types represent either an expression of phase variation due to environmental factors, or different stages in the development of the extracellular material.

While Popov *et al.* (1994) demonstrated the absence of any well defined structures on the bacterial cell surface, they determined, by electron microscopy, that *B. pseudomallei* formed a protective capsule during the first few hours following infection of guinea-pigs. Without the capsule, bacterial cells were actively phagocytised by peritoneal macrophages and polymorphonuclear leukocytes and then damaged. They observed that encapsulated bacteria survived and preserved their normal ultrastructure and concluded that the capsule prevented the destruction of the bacteria by the contents of the phagolysosomes.

Steinmetz *et al.* (1995) identified and characterised a species-specific exopolysaccharide of *B. pseudomallei* and produced a monoclonal antibody (MAb) with specificity for a polysaccharide structure on the bacterial surface by immunising mice with a bacterial strain that exhibited mucoid growth characteristics (NCTC 7431). Immunoelectron microscopy showed that the MAb binds to an exopolysaccharide that had a capsule-like

appearance in the mucoid strain used for immunisation. The authors showed that all of the 12 mucoid and non-mucoid *B. pseudomallei* strains tested from geographically different tropical regions were recognised by the MAb, indicating that the expression of this exopolysaccharide is a stable character among this species. Moreover, temperature appeared to have little effect on the synthesis of the exopolysaccharide since *B. pseudomallei* strains grown at 15∘C and 37∘C reacted with the MAb. The exopolysaccharide was classed as group I polysaccharide because of its high molecular weight (>150 kDa) and synthesis below 20∘C. Importantly, specific antibodies present in the sera from a number of melioidosis patients recognised the capsular polysaccharide demonstrating that it is expressed *in vivo*. The chemical and structural nature of the polysaccharide, from a virulent clinical isolate, *B. pseudomallei* 304b, was elucidated by Masoud *et al.* (1997) and shown to have a unique structure that is not closely related to any known bacterial polysaccharide. It had a linear unbranched polymer of repeating tetrasaccharide units composed of three $_{D}$-galactose units and one 3-deoxy-$_{D}$-*manno*-octulosonic acid (KDO), containing an *O*-acetyl substituent. Similarly, Nimtz *et al.* (1997) have demonstrated that a structurally identical capsular antigen is expressed by *B. pseudomallei* NCTC 7431. Moreover, *B. pseudomallei* produces two other polysaccharides depending on the growth medium (Kawahara *et al.*, 1998).

A recent study by Ahmed *et al.* (1999) showed that capsular polysaccharide is important in facilitating the attachment of *B. pseudomallei* to respiratory epithelial cells *in vitro*, apparently without fimbriae (**Fig 2.3**). Furthermore, Steinmetz *et al.* (2000) reported that unpublished studies revealed the existence of a second exopolysaccharide in virulent *B. pseudomallei* strains. Monosaccharide analysis showed that it does not contain KDO but contains galactose and other sugar components.

The capsular polysaccharide functions in many ways to aid in the survival of bacteria in various hostile environments. In addition to a role in mediating direct interaction between

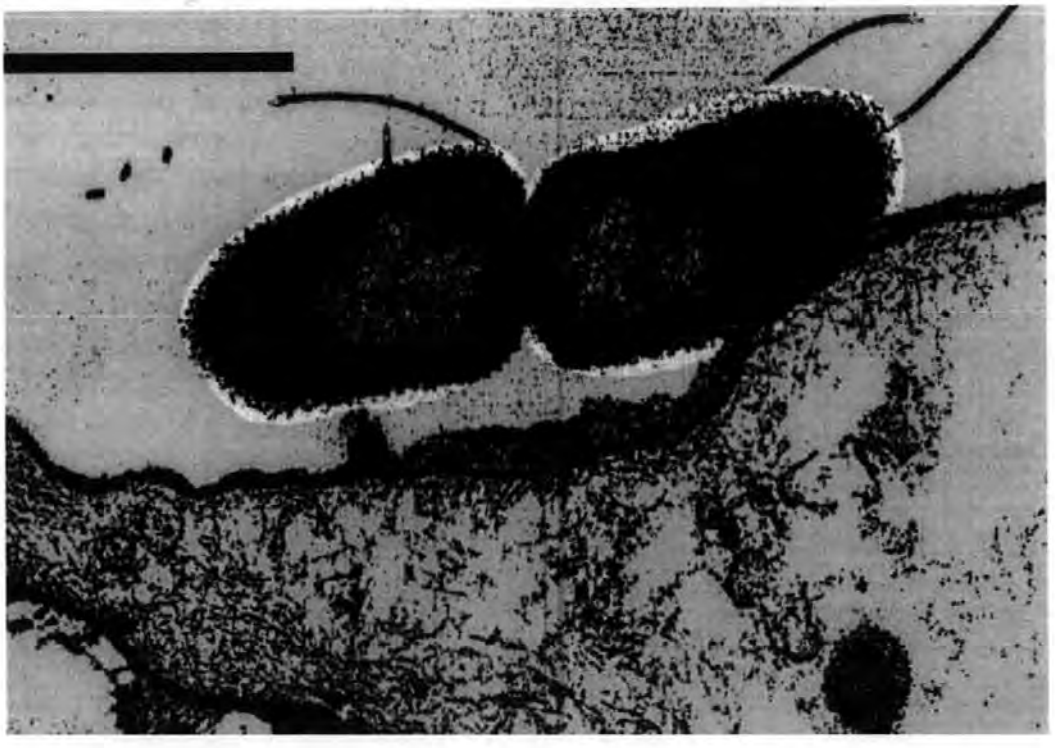


Fig 2.3: A transmission electron micrograph of *B. pseudomallei* (X 3800) attached to a pharyngeal epithelial cell, stained with ruthenium red. A thin, granular layer on the surface of the bacteria and the cell indicates a polysaccharide layer. Adapted from Ahmed *et al.* (1999).

the bacterium and its immediate environment, a number of possible functions have been suggested for polysaccharide capsules including adherence, prevention of desiccation, and resistance to both specific and non-specific host immunity (Costerton *et al.*, 1987; Roberts, 1996). In the case of *P. aeruginosa*, the production of large amount of an alginate-like polysaccharide, conferring the mucoid phenotype of most of the strains isolated from the lungs of CF patients, is associated with the virulence of this bacterium (Govan & Deretic, 1996). This event often coincides with a marked deterioration of the pulmonary function in CF. Similarly, some *B. cepacia* strains have been shown to produce capsule that correlates well with virulence (Nelson *et al.*, 1994; Wilkinson & Pitt, 1995). During saprophytic existence, capsules may protect bacteria from desiccation, concentrate minerals and nutrients, reduce contact with hydrophobic or charged molecules, and enhance attachment to surfaces. On the other hand, during infection many bacterial pathogens are able to combat the bactericidal activity of complement and phagocytes by encapsulation with structurally diverse polysaccharides and it may be only in the late phase of the infection that the host produces specific anticapsular antibodies. The regulation of capsule synthesis in *E. coli* and several other bacteria is under the control of two-component regulatory systems.

2.5 TWO-COMPONENT REGULATORY SYSTEMS

2.5.1 Introduction

In order to meet the challenges posed by a competitive, changing and often unpredictable environment, bacteria have devised sophisticated signalling systems for eliciting adaptive responses to environmental variation (Parkinson, 1993). Movement of bacteria toward favourable conditions and away from harmful ones is one of these adaptive responses, to seek optimal living conditions. Another approach is a regulated response to adaptive opportunities. Since bacteria are manifestly capable of such adaptation, they must possess the means to gather and analyse environmental information and to transduce this

information to the cellular apparatus concerned with mounting the response (Dorman & Bhriain, 1992).

Cells respond to changes in their physical and chemical environment by generating intracellular signals that lead to the modification of their structure, metabolism and movement (Alex & Simon, 1994). Stress conditions of many sorts, such as antibiotics, heavy metals, turgor pressure, or starvation, elicit changes in gene expression that lead to effective coping responses (Parkinson, 1993). In pathogens, chemical, temperature, or other conditions characteristic of the host environment elicit the production of components that are essential to ensure survival (Bliska *et al.*, 1993). Bacterial cells must have elaborate signalling networks to ensure that they make an appropriate response which, once commenced, must be ushered to a successful conclusion (Parkinson, 1993).

2.5.2 Protein phosphorylation and signal transduction

Protein phosphorylation is one of the most widely used means for regulating biological processes in prokaryotes and eukaryotes (Mizuno, 1998). Since the discovery of regulatory protein phosphorylation by Krebs, Fischer, and colleagues in the mid-1950s, protein phosphorylation-dephosphorylation has been considered to play a key role in cellular mechanisms for the detection, transmission, and integration of intra- and extra-cellular signals (Fischer & Krebs, 1989). There are two classes of enzymes that regulate signalling through the phosphorylation and dephosphorylation of proteins, namely protein kinases and protein phosphatases. Protein phosphorylation is ideally suited for signal transduction and the shuttle of information from the cell surface to the interior, since protein kinases have an in-built mechanism for signal amplification, and protein phosphatases provide a means of reversing the signal when the stimulus is removed (Hunter, 1995).

Proteins can be phosphorylated on acidic (aspartate, glutamate), basic (arginine, histidine, lysine), hydroxy (serine, threonine, tyrosine), or cysteine residues (Bourret *et al.*,

1990). The enzymes that catalyse protein phosphorylation, protein kinases, can be classified according to the amino acid species they phosphorylate: namely, protein serine/threonine kinases (PSKs); protein tyrosine kinases (PTKs); protein histidine kinases (PHKs); and also a small subfamily of PSKs can phosphorylate both serine/threonine and tyrosine residues. Eukaryotes almost exclusively target the hydroxyl amino acids serine, threonine, and tyrosine for signal transduction purposes, while bacteria favour the use of histidine and the carboxyl amino acids as phosphoacceptors (Kennelly & Potts, 1996).

In bacteria, it has been well documented that signals are processed via a system that involves phosphorylation, which entails the formation of high-energy phosphoamino acids, phosphohistidine and phosphoaspartate (Alex & Simon, 1994). Moreover, both pathogenic and non-pathogenic bacteria have evolved signal transduction systems to control the co-ordinate expression of virulence determinants and to develop adaptive responses to their environment.

2.5.3 Two-component regulatory proteins

2.5.3.1 Brief history

All living cells must sense changes in their environment and respond appropriately in order to adapt successfully to their environment. To meet this need, bacteria commonly employ a sophisticated signal transduction strategy that involves at least two protein components of signal transduction. They are therefore referred to as the "two-component regulatory system" and their sole purpose is to allow for regulation (Stock *et al.*, 1989, 1990, 2000; Parkinson, 1993; Hoch & Sailhavy, 1995; Appleby *et al.*, 1996). This family of proteins allows bacteria to sense and respond to numerous changes in their environment and it was originally defined on the basis of protein sequence similarities between pairs of "sensor" and "regulator" proteins. Generally, the genes encoding the sensor and effector pair are located near each other, often in the same operon.

The first analyses of two-component system pathways were carried out in bacterial systems in 1986 when the predicted amino acid sequences of two regulatory proteins in *Bradyrhizobium* sp. were seen to be virtually identical to the NR_{II} and NR_{I} proteins of *Klebsiella pneumoniae* (Nixon *et al.*, 1986). In addition, the regulators of nitrogen metabolism (NtrB/NtrC), and the chemotactic regulators (CheA/CheY) were identified (Ninfa & Magasanik, 1986; Hess *et al.*, 1988; Keener & Kustu, 1988). Soon, the dual component nature of these systems was described. Currently, the two-component signal transduction systems constitute the most common class of bacterial signal transduction proteins and control a broad range of bacterial responses. Apparently, the operation of the individual systems vary widely, but their primary activation processes are probably all based on the same chemical mechanism of Mg^{2+}-dependant protein phosphorylation (Volz, 1995) and a common phosphorylation reaction mechanism exists in most, if not all, two-component regulatory systems. This is further emphasised by the finding that various protein kinases are capable of correctly phosphorylating noncognate regulator proteins *in vitro*, although to a much lesser extent than cognate regulators (Bourret *et al.*, 1991).

In its simplest form, the two-component signal transducing system consists of a stimulus-dependent autophosphorylating sensor kinase and its substrate, termed the response or effector regulator that mediates an adaptive response by a change in gene expression (Mizuno, 1998).

2.5.3.2 Bacterial sensor histidine kinases

The histidine protein kinase superfamily, also called sensor kinases, includes molecules that monitor certain environmental parameters, and modulate accordingly the functions of the response regulator. Most sensor proteins are located in the cytoplasmic membrane with the transmitter portion of the molecule projecting into the cell. Many, but not all, have two transmembrane domains that anchor the protein to the cell membrane, separated by a loop

through the periplasmic space between the inner membrane and cell wall. The periplasmic domains of these proteins are structurally unrelated and have diverse receptor functions reflecting the variety of chemical and physical stimuli they detect (Stock *et al.*, 2000).

The N-terminal portion of the histidine protein kinase functions as an input domain, detecting environmental stimuli directly or interacting with an upstream receptor, while at the C-terminal end of the protein is the transmitter module. In a typical sensory kinase, this catalytic domain is generally 240 amino acids long and contains several blocks of residues that are conserved among histidine protein kinases. One of these is the H box, typically located near the N-terminus of the module, and includes a histidine residue at which the proteins autophosphorylate at position N-3 of the imidazole ring through the formation of a phosphoamidate, using the γ–phosphoryl group of ATP. Mutations in this region frequently have a deleterious effect on kinase activity. Four additional blocks of conserved amino acids constitute the N (asparagine), G1, F (phenylalanine) and G2 boxes. The G1 and G2 boxes are glycine-rich sequences (DXGXG and GXGXG) that bear a resemblance to the nucleotide binding motifs of other proteins. On the other hand, the function of the N and F boxes are unknown (**Fig 2.4a**) (Parkinson, 1993; Alex & Simon, 1994; Appleby *et al.*, 1996; Mizuno *et al.*, 1996). However, not all kinases contain all five conserved motifs, for example, CheA lacks the conserved H box and is phosphorylated at a histidine in a non-conserved amino-terminal domain. Once autophosphorylated, the histidine kinase serves as a relatively high-energy molecule that acts as a phospho-donor for its cognate response regulator (Egger *et al.*, 1997; Mizuno, 1998).

While histidine kinases were thought to be present only in prokaryotes, they have recently been identified in eukaryotic systems including *Saccharomyces cerevisiae* (yeast), *Schizosaccharomyces pombe*, *Arabidopsis thaliana* (plant), *Neurospora crassa* and *Dictyostelium discoideum* (Alex *et al.*, 1996; Wurgler-Murphy & Saito, 1997). Similar to

bacterial hybrid histidine kinases, the majority of the eukaryotic histidine kinases which have been characterised to date include both a conserved histidine kinase domain and a response regulator type domain in the same protein (Egger *et al.*, 1997). The prevalence of histidine kinase/response regulators in organisms from all three kingdoms (eubacteria, archaebacteria, and eukaryota) attests to the early evolution of this signal transduction mechanism (Loomis *et al.*, 1997). Moreover, while histidine kinases have not yet been isolated from vertebrate tissues, there is evidence in the nucleotide sequence data banks indicating that genes encoding members of this family are present in the human genome and are likely to play a significant role in signal transduction (Loomis *et al.*, 1997).

2.5.3.3 Response regulators

The substrate for the histidine kinase is the response regulator, which is phosphorylated at a specific aspartic acid residue. A typical response regulator consists of two domains: a conserved N-terminal receiver domain, followed by a variable C-terminal signal output domain that has an effector function. The overall organisation of each of the response regulators is tailored to a specific effector function, while their receiver domains are highly conserved, implying a ubiquitous mechanism of activation by Mg^{2+}-dependent protein phosphorylation (Pao & Saier 1995; Mizuno *et al.*, 1996). The lower level of sequence conservation within the C-terminal region of response regulators perhaps reflects different mechanisms of transcriptional regulation (Bourret *et al.*, 1991).

Response regulators are characterised by, and share, a common receiver domain of about 125 amino acids in length in the N-terminal region, which contains the conserved motifs that define membership in the response regulator family (**Fig 2.4b**). The conserved residues include an invariant aspartate (D1) residue which serves as the site of phosphorylation and is located around the centre of the receiver domain, a pair of aspartate residues (DD) preceding D1 at the N-terminal end, a threonine (T), and a lysine residue (K)

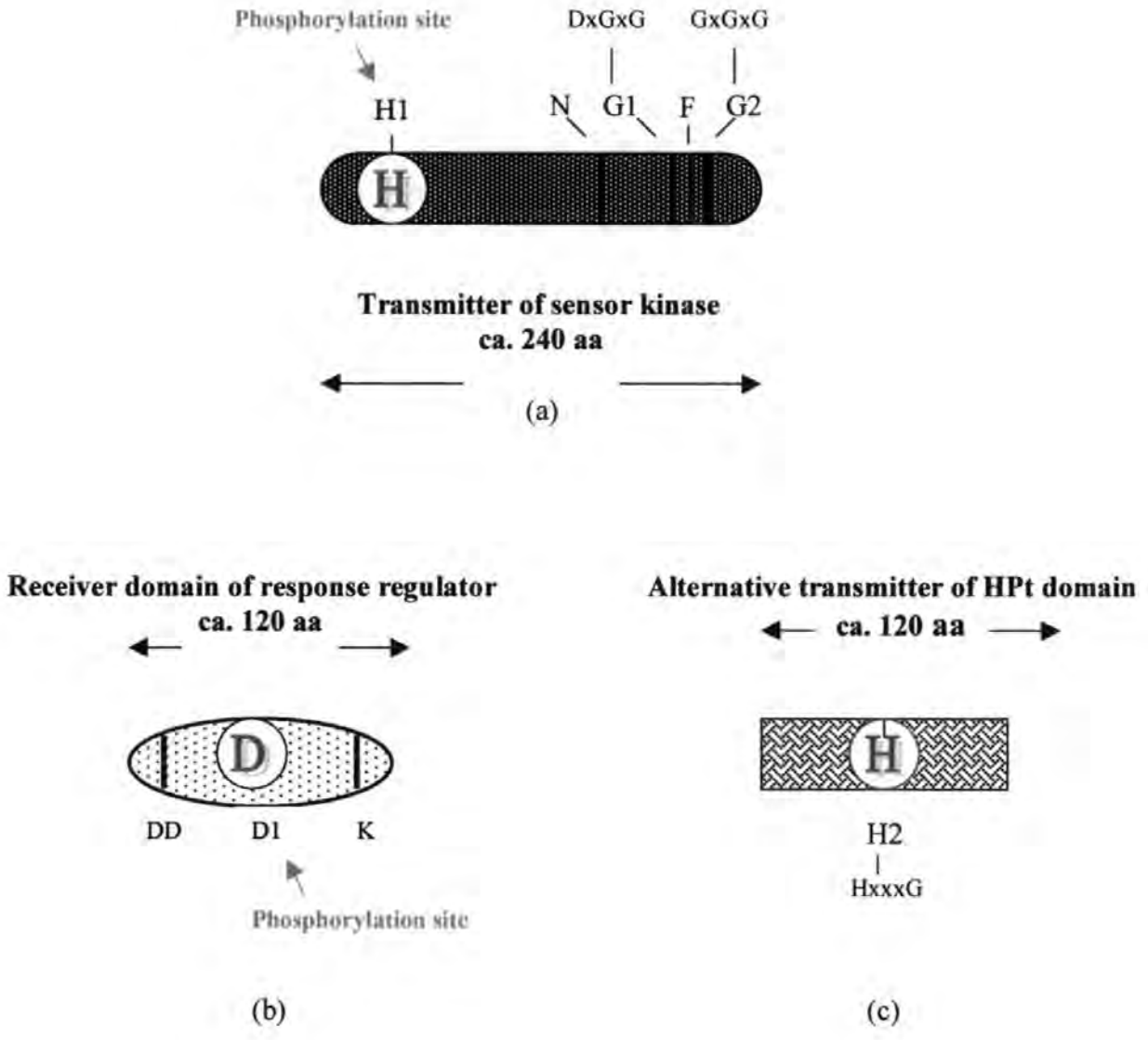


Fig 2.4: Structural features of (a) the transmitter, (b) receiver and (c) HPt domains, with conserved signature motifs. Conserved residues are as in the text. Adapted from Parkinson (1993); Mizuno *et al.* (1996); Mizuno (1998).

at the C-terminal end of the receiver domain, which all contribute to the acidic pocket for the phosphorylation site (Stock *et al.*, 1989; Egger *et al.*, 1997). Pao & Saier (1995) analysed the sequences of forty-nine response regulators and a multiple alignment of the receiver domains of these sequences showed that only one residue is fully conserved in all of these domains. This was the aspartyl residue, D1, the site of phosphorylation.

Receiver-containing proteins can catalyze the transfer of the phosphoryl group from the histidine of the sensor kinase to a conserved aspartate residue. Alternatively, they can phosphorylate themselves using a variety of small donor molecules (but not ATP) such as acetyl phosphate, imidazole phosphate, or phosphoramidate as alternate phosphodonors (Feng *et al.*, 1992; Lukat *et al.*, 1992), demonstrating that the response regulator can catalyse phosphoryl transfer independently of the histidine kinase. Additionally, receivers can catalyse the hydrolytic loss of their phosphoryl groups (Hess *et al.*, 1988), with half-lives ranging from a few seconds to many minutes (**Fig 2**, Appendix II).

The phosphorylation state of the receiver domain modulates the activity of a unique C-terminal output domain, commonly a transcriptional regulator, to elicit an adaptive response to the stimulus (Appleby *et al.*, 1996). Accordingly, the response regulator mediates changes in gene expression or cell behavior in response to environmental stimuli (Mizuno, 1997). The aspartate residue D1 is modified through the formation of an acyl phosphate on acquisition of the phosphate from the phospho-histidine of its cognate transmitter. The phosphorylated receiver is relatively unstable, and is readily converted to its nonphosphorylated form (the *in vitro* half-lives of phospho-CheY and phospho-OmpR are a few seconds and several hours, respectively) (Bourret *et al.*, 1991). Furthermore, the structure of receiver modules is well understood and the pattern of conserved residues implies that all response regulators will have a similar structure. X-ray crystallography studies of the *S. typhimurium* (Stock *et al.*, 1989) and *E. coli* (Volz & Matsumura, 1991)

CheY and NarL (Baikalov *et al.*, 1996) proteins proved a barrel-like arrangement of five sets of alternating β strands and α helices, in the sequence starting with β_1 and ending with α_E. (**Fig 2.5**).

Most response regulators are positive transcriptional regulators, containing DNA recognition sequences in their carboxyl-terminal domains and have an effector domain linked to the N-terminal domain. It is apparent that the N-terminus in its dephosphorylated state inhibits the activity of the effector domain(s) and phosphorylation acts to relieve this inhibition. Phosphorylation induces a conformational change by movement of the α_E helix that swings the attached C-terminal domain into a different orientation with respect to the regulatory domain (Stock *et al.*, 1995). The crystal structure of NarL (**Fig 2.6**) provided the first structural determination for the C-terminal domain of the FixJ superfamily (Baikalov *et al.*, 1996). It showed that the C-terminal DNA-binding domain, with 62 amino acids, is a compact bundle of 4 α helices, of which the middle 2 form a helix-turn-helix motif (H-T-H) closely related to other H-T-H DNA-binding proteins (**Fig 2.7**). The 2 domains are connected by an α helix of 10 amino acids. In the nonphosphorylated form of NarL, the C-terminal domain is turned against the receiver domain in a manner that would preclude DNA binding. The NarL structure was considered as a model for other members of the FixJ or LuxR family of bacterial transcriptional activators. On the other hand, due to the instability of the phosphorylated form of CheY, the crystal structure represented the nonphosphorylated state of the protein. Therefore the nature of the precise conformational changes which occur upon phosphorylation remain elusive (Egger *et al.*, 1997).

Based on extensive sequence homology in the less conserved C-terminal domain, bacterial response regulators can be separated into five families. These are represented by: (i) CheY family (including RegA, SpoOF, and Xcc1), with ca. 130 amino acids, has only the N-terminal receiver domain, and possesses no output or DNA binding domains; (ii) OmpR

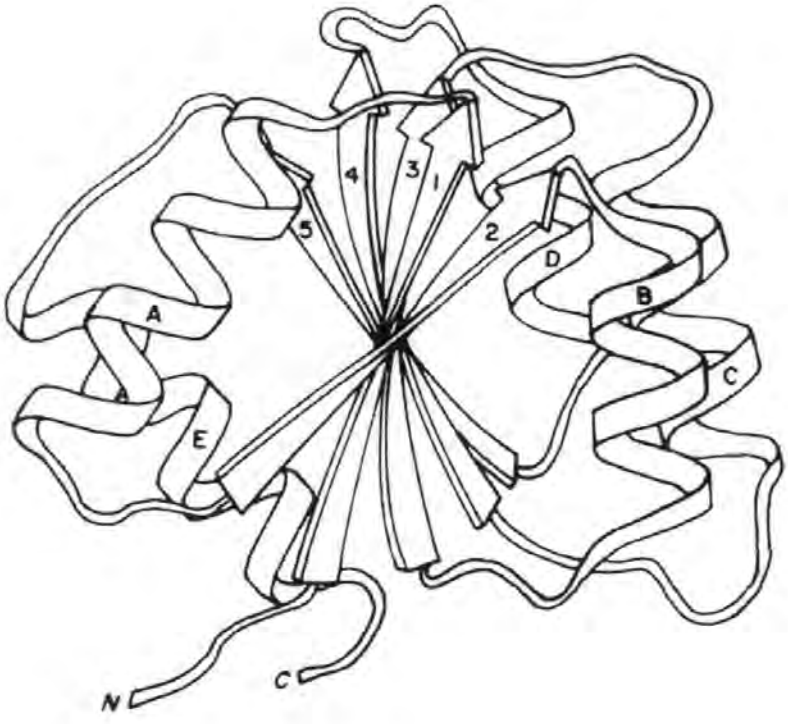


Fig 2.5: Tertiary structure of the *S. typhimurium* CheY response regulator. The protein is composed of a five-stranded parallel β-sheet (1-5) surrounded by five α-helices (A-E). Adapted from Stock *et al.* (1989).

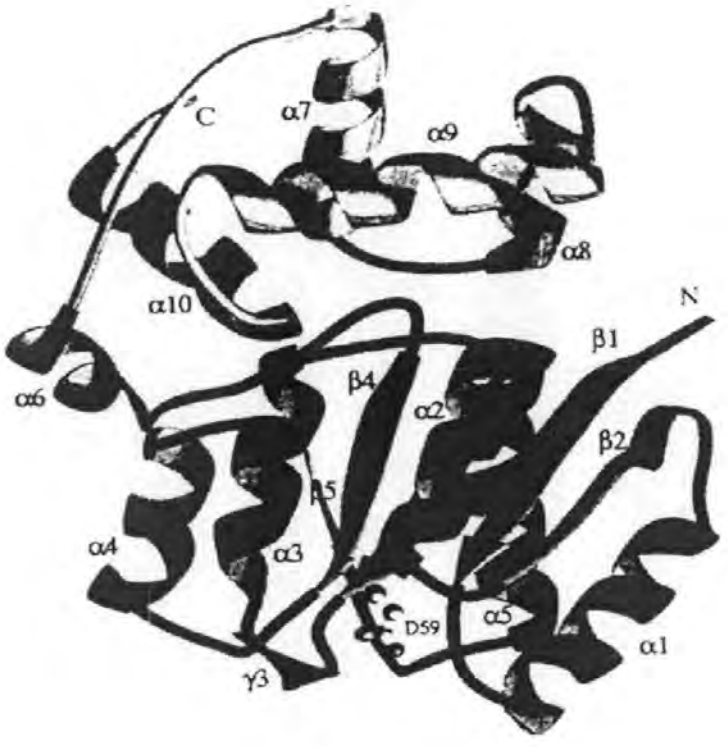


Fig 2.6: Secondary structure of the *E. coli* NarL response regulator. The N-terminal domain has five α/β motifs, while the C-terminal domain consists of the rest of the molecule with α8 and α9 forming a Helix-turn-helix motif (H-T-H).

Fig 2.7: The orientation of the helix-turn-helix motif (in blue) of the *E. coli* bacteriophage 434 repressor in the major groove of the DNA double helix. The 'N' and 'C' indicate the N and C terminals of the motif, respectively. (Adapted from Brown, 1999)

family (including ArcA, BaeR, BasR, CopR, CreB, CpxR, KdpE, NisR, PetR, PhoB, PhoP, ResD, SparR, SphR, TctD, ToxR, and VirG), has a single C-terminal receiver domain and ca. 230 amino acids in length; (iii) NtrC family (including AlgB, DctD, HydG, and PilR), has 2 effector domains C-terminal to the receiver domain, and ca. 460 amino acids in length; (iv) FixJ or LuxR family (including DegU, EvgA, FimZ, NarL, RcsB, and UhpA), has a single C-terminal domain and ca. 220 amino acids in all; (v) Unclassified (including BvgS, CheB, PleD, RcsC, and VirA), exhibit a wide variation in structure and length (Baikalov *et al.*, 1996; Egger *et al*, 1997).

2.5.3.4 Hybrid kinases

While most bacterial histidine-aspartate (His-Asp) phosphorelay systems involve direct transfer from the phosphorylated histidine sensor kinase to the conserved aspartyl residue of the response regulator, more sophisticated signal transduction proteins of bacteria contain domains that resemble both the histidine kinases and the response regulators in one primary sequence. This type of signal transducer is referred to as the "hybrid sensory kinase" (Alex & Simon, 1994; Mizuno *et al.*, 1996; Egger *et al.*, 1997). *E. coli* alone has at least 5 hybrid kinases (ArcB, BarA, EvgS, RcsC, and TorS) (Mizuno, 1997), while 16 hybrid sensor kinases were found in a photosynthetic cyanobacterium, *Synechocystis sp.* Others have been recognised in a wide variety of other bacteria (BvgS, LemA, RpfC, RteA and VirA). The response regulator domain of these proteins is thought to function as a pseudo-substrate inhibitor of the kinase domain. Upon detection of the stimulus, the kinase domain undergoes autophosphorylation, and then transfers this phosphate to the regulator domain, thus relieving inhibition. Phosphorylation of the true substrate can then take place. These ideas are largely based on work on ArcB, which controls the anaerobic repression of several operons in *E. coli* by modulating its downstream response regulator, ArcA (Alex & Simon, 1994).

2.5.3.5 Histidine-containing phosphotransfer

The histidine-containing phosphotransfer (HPt) domain was discovered very recently, but now is known to play an important role in some, but not all, His-Asp phosphotransfer signalling systems (Appleby *et al.*, 1996). In a given sequence of a signal transducer, the HPt domain is the most difficult to recognise, because its amino acid sequence is quite variable, and does not resemble that of the authentic histidine site of the transmitter (Ishige *et al.*, 1994). However, in a typical HPt domain an invariant and phosphorylated histidine residue (H2) is followed by a short characteristic stretch of amino acids (**Fig 2.4c**). Based on sequence criterion, a number of bacterial signal transducers were predicted to contain HPt domains such as ArcB and BarA of *E. coli* and BvgS of *Bordetella pertussis*. Unlike authentic transmitters, the HPt domain does not exhibit any catalytic function and as yet there is no evidence for intramolecular phosphotransfer within the sensor protein histidine kinase. The HPt domain thus appears to serve solely as a passive intermediate molecule, or substrate, in His-Asp phosphotransfer by acquiring/transferring a phosphoryl group from/to another signalling domain (Mizuno, 1998).

2.5.3.6 Cross-talk

Cross-talk is the term given to non-specific interactions between non-partner sensor and regulator proteins. It occurs between parallel signal transduction pathways as an information-processing network that integrates signals from a wide range of sensory inputs to co-ordinate the activities of an equally diverse array of output response strategies (Stock *et al.*, 1989). Despite the highly conserved nature of sensor and regulator proteins, cross-talk is minimal (Parkinson, 1995) but can occur, although phosphotransfer rates between family members are at least two orders of magnitude slower (Egger *et al.*, 1997). This suggests that the regulatory systems possessed by some species have the potential to transduce signals from a diverse series of inputs or alternative phosphate donors and to collectively influence

several regulatory responses (Gervais & Drapeau, 1992). Biochemical evidence is now available for interactions between members of the two-component family, in particular EnvZ/OmpR, NtrB/NtrC, CheA/CheY and CheA/CheB (Ninfa *et al.*, 1989; Stock *et al.*, 1989). One other striking example of integrated signalling pathways in *B. subtilis*, is between the pathways controlling phosphate utilisation (PhoR/PhoP), aerobic and anaerobic respiration (ResE/ResD), and sporulation (KinA-B/Spo0A) (Msadek, 1999). Consequently, once the cell commits to sporulation, respiration and phosphate utilisation are down-regulated. Such a networking control system seems to assist the bacteria by increasing the flexibility of their environmental responses.

2.5.3.7 Targets for antimicrobial therapy and vaccines

The prevalence of two-component systems in a wide variety of bacterial species has stimulated interest and some limited success in designing new classes of broad-spectrum antimicrobials that can block these signalling pathways (Miller *et al.*, 1989; DiRita & Mekalanos 1989; Deretic *et al.*, 1991). The most attractive reason for targeting two-component systems is that they are used by pathogenic bacteria to control the expression of virulence genes required for infectivity such as *Salmonella typhimurium* PhoP/PhoQ and *B. pertussis* BvgA/BvgS (Stock *et al.*, 2000). General histidine kinase or response regulator inhibitors could potentially be candidates for broad-spectrum antibiotics. Bacterial histidine kinase inhibitors and inhibitors of aspartyl response regulator DNA-binding have been identified (Roychoudhury *et al.*, 1993). Natural inhibitors of two-component systems such as unsaturated fatty acids were shown to be noncompetitive inhibitors of the autophosphorylation activity of KinA (Strauch *et al.*, 1992), CheA (Roychoudhury *et al.*, 1993), and NtrB (Barrett & Hoch, 1998). Furthermore, a number of compounds that not only inhibit two-component functions but also affect bacterial growth, are currently under investigation. These include hydrophobic tyramines (Barrett *et al.*, 1998), salicylanilides

(Machielag *et al.*, 1998), and benzoxazines (Barrett & Hoch, 1998).

The possibility that strains of pathogenic bacteria with mutations in a specific two-component system may be good candidates for vaccine use has been tested. The rationale for this approach is that some two-component systems may regulate gene expression only during certain stages of infection. Mutants deficient for such systems might still survive within the host long enough to elicit an immune response before becoming growth-limited due to the mutation. Mutants of *Salmonella typhi* were created by deletion of the *phoP/phoQ* two-component system that regulates survival of this organism within the macrophage. When tested in human volunteers, the mutants induced greater intestinal and humoral immune responses than did four doses of the approved oral typhoid vaccine. This work establishes a precedent for studying the use of two-component mutants as vaccines against other important pathogens (Miller *et al.*, 1993; DiRita, 1999).

2.5.4 Circuitry of signalling systems

The functionally critical step in two-component signal transduction is the interaction between the histidine kinase and the response regulator (Stock *et al.*, 1995). Bacterial signalling proteins are constructed from modular components: input sensing domains; output effector domains; and transmitter and receiver domains for promoting protein-protein communication (**Fig 2.8**). Assembling these elements in various configuration constructs signalling circuits (Parkinson & Kofoid, 1992).

The circuitry of signalling and the phosphorylation events, His-Asp phosphotransfer, involved have been demonstrated in several systems. The molecular communication between the sensors and their cognate regulators imply phosphotransfer reactions (phosphorylation and dephosphorylation) (Mizuno *et al.*, 1996). In order to trigger a reaction, the histidine kinase sensing domain directly bind ligands or detect other physical stimuli. Alternatively, more complex schemes involve indirect detection of signals through

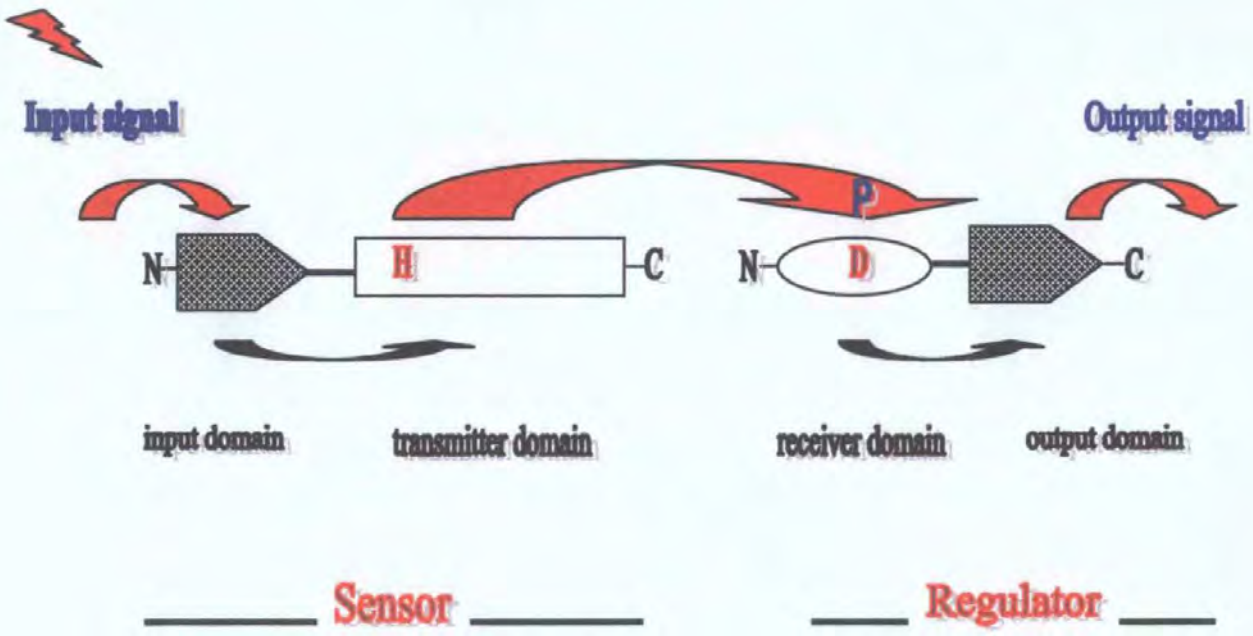


Fig 2.8: A model for two-component signal transduction proteins, showing the information flows through noncovalent controls exerted by one domain upon another (black arrows) and by phosphorylation (coloured arrows) involving histidine (H) and aspartate (D) residues. (Adapted from Parkinson & Kofoid, 1992)

interaction with other protein components (Stock *et al.*, 2000). The general mode of phosphorelay involves the following steps:

1- Autophosphorylation of the sensor histidine kinase at a conserved histidine.

2- The histidine kinase serves as a phospho-donor for the response regulator, which receives the phosphate onto a conserved aspartate within the phospho-accepting domain.

3- A phosphorylation induced conformational change in the response regulator that permits a regulatory target to be engaged, which in turn allows the regulation of downstream effectors, commonly by mediating changes in gene expression.

4- Down-regulation of the pathway by dephosphorylation via phosphatases (Egger *et al.*, 1997).

The rate of phosphotransfer depends on both the rate of autophosphorylation and the specific protein-protein interactions between the sensor histidine kinase and the response regulator (Egger *et al.*, 1997). Following phosphotransfer, the sensor histidine kinase and response regulator dissociates, which has been demonstrated in the case of CheA and CheY (Schuster *et al.*, 1993). In the absence of phosphorylation, the C-terminal domain of response regulators is not usually active. It has been proposed that the N-terminal domain exerts an inhibitory function, and when phosphorylated the conformation of the C-terminal domain is altered in a way that permits activation (i.e. DNA binding).

2.5.5 Signals/ stimuli

A wide variety of signalling tasks are handled by two-component regulatory systems in order to mediate the response to a particular stimulus or set of stimuli. One response to changing conditions is to move to a more "favourable" locale. Another possible course of action is to adapt the cell to the new environment, either by changing enzyme activity or by altering the expression of specific genes or groups of genes, either as a temporary adjustment or to establish a new long-term state e.g., the sporulation response to starvation.

Signal transduction pathways account for the control of stimulus-response systems including host detection and invasion leading to symbiosis or pathogenesis; metabolic adaptation to changes in chemical concentrations (carbon, nitrogen, electron acceptor, and phosphate); physiological responses to changes in medium osmolarity; chemotaxis; stress-induced differentiation, such as sporulation and fruiting-body formation; stress leading to antibiotic synthesis, capsule synthesis, gene transfer, heterocyst formation, or sporulation (Parkinson & Kofoid, 1992). In addition, nutrient and toxin levels, acidity, temperature, humidity, viscosity, or light and many other conditions can change rapidly and unexpectedly. For example, the nitrogen system monitors the ratio between two metabolites, glutamine and α-ketoglutarate, and in this way, determines the availability of nitrogen. The chemotaxis system monitors the level of several small-molecule attractants and repellents, the molecular identities of which are also clear. By contrast, EnvZ and OmpR mediate the response to osmolarity, a property of the environment altered by varying the concentration of several small molecules, independent of the identity of these molecules (Van Alphen & Lugtenberg, 1977).

The adaptive responses in bacteria range from rapid transit changes in motility to long-term global reorganisations of gene expression and cell morphology (Stock *et al.*, 1989) and the response time varies widely, from milliseconds as in the case of the chemotaxis response, to hours which are required for initiating sporulation. Underlying this flexibility is a series of protein domains that apparently have been modified and rearranged in evolution to accommodate the regulatory requirements of a wide variety of systems (**Table** 1, Appendix II).

2.5.6 *E. coli* and two-component systems

The recently completed genome sequence of *E. coli* has allowed the compilation of a complete list of genes encoding the two-component signal transduction proteins (**Table 2,**

Appendix II). Mizuno (1997) identified at least 62 open reading frames as putative two component signal transducers in this single species. Among these there were 32 response regulators, 23 orthodox sensory kinases and 5 hybrid sensory kinases, while the rest had atypical domains lacking the phosphorylated histidine residue. The hybrid sensory kinases are RcsC, TorS, ArcB, EvgS and BarA. RcsC contains both the transmitter and receiver domains, while the other four contain HPt as well (Mizuno, 1997). Moreover, fourteen ORFs exhibit extensive similarity to members of the OmpR subfamily not only in the receiver domains, but also in the output domain. Seven other ORFs appear to be members of the NarL subfamily and four are members of the NtrC subfamily. Since OmpR, NarL, and NtrC response regulators are known to function as specific DNA-binding transcriptional regulators, it can thus be assumed that all these 25 members are DNA-binding transcriptional regulators. The response regulators are implicated in gene regulation in response to environmental parameters, yet each of them is postulated to regulate (activate or repress) a specific subset of genes (Mizuno, 1998). In only 17 cases have the physiological consequences of the adaptive responses induced by the sensor/regulators been previously, albeit incompletely, documented. The chromosomal positions of the coding sequences specifying each of these putative signal transducers are scattered evenly over the *E. coli* genome (**Fig 3,** Appendix II). In most instances, a cognate pair of sensor/regulator is located next to each other, and most likely, in the same transcriptional unit (or operon). The order of these pairs of genes (5'-sensor/3'- regulator or 5'-regulator/3'-sensor) and the transcriptional direction relative to the chromosome (direct or complementary) appears to be random (Mizuno, 1997).

2.5.7 Two-component systems in other bacteria

The homologous of two-component regulatory systems are found in both gram-negative and gram-positive bacteria, including the genera *Agrobacterium*, *Bradyrhizobium*,

Bacillus, *Bordetella*, *Enterobacter*, *Klebsiella*, *Myxcococcus*, *Pseudomonas*, *Rhizobium*, *Staphylococcus*, and *Salmonella* (Stock *et al.*, 1989). However, the number of two-component proteins differs greatly in different bacteria. Recently, the entire genomic sequences have been completed for 10 prokaryotic microorganisms, including *Archaea*. A brief inspection of these databases revealed that *Haemophilus influenzae* contains 4 sensors-5 regulators; *Helicobacter pylori*, 4 sensors-7 regulators; *Synechocystis* sp., 32 sensors-38 regulators; *Mycoplasma gentitalium*, none; *Methanococcus jannashii* (archaebacterium), none; and *Methanobacterium thermoautotrophicum* (archaebacterium), 15 sensors-9 regulators. *Bacillus subtilis* appears to possess a similar number of sensor-regulator pairs as *E. coli*. Many prokaryotic microorganisms are considered to be exposed to a much greater variety of environments than many eukaryotes and this may reflect the greater number of His-Asp phosphotransfer signalling systems to be found in these bacteria (Mizuno, 1998). The complete genome sequence of *P. aeruginosa* contains the highest number of putative genes encoding two-component regulatory systems of all bacterial genomes sequenced to date, 64 response regulators and 63 histidine kinases (Rodriguez *et al.*, 2000). The authors attributed this high number of two-component systems to the relatively large size of the *P. aeruginosa* genome (6264 Kb), and to the nature of the organism as an opportunistic pathogen that needs to adapt rapidly to a wide variation in environmental conditions. Recently, a two-component regulatory system, termed *irlRS* for invasion related locus, has been identified in *B. pseudomallei* (Jones *et al.*, 1997). The system was detected by using Tn5-OT182 mutagenesis to generate mutants deficient in the ability to invade an established pulmonary cell line, a human type II pneumocyte cell line. The predicted proteins shared considerable homology with other two-component systems involved in the regulation of heavy metal (such as Cd^{2+} and Zn^{2+}) resistance in other organisms.

2.5.8 The Rcs (regulation of capsule synthesis) regulatory system

The rcs system, found in a number of enteric bacteria, is a typical example of a histidine protein kinase/response regulatory system that acts as a mechanism for survival under difficult environmental conditions.

2.5.8.1 The *E. coli* Rcs proteins and capsule synthesis

Although Rcs proteins are now known to exist in a number of gram-negative bacteria, they were first identified in *E. coli*. Consequently, they are the best characterised of all regulators and so Rcs-like proteins from other organisms are generally identified on the basis of homology with the *E. coli* protein.

Strains of *E. coli* as well as other enterobacteria can be provoked to produce a mucoid slime in response to external stimuli or mutations in key regulatory loci (Kelley & Georgopoulos, 1997). Bacterial capsular polysaccharides are classified into two groups by chemical and physical criteria. In general, group I polysaccharides contain uronic acid as the acidic component, have high molecular mass and are co-expressed with specific O polysaccharides. In contrast, group II polysaccharides contain a large variety of acidic components and have a relative low molecular mass (Jann & Jann, 1990; Virlogeux *et al.*, 1996).

There is considerable information about the regulation of group I polysaccharide biosynthesis (Coplin *et al.*, 1990; Stout, 1994), and the regulatory strategies used for colanic acid biosynthesis, an extracellular slime produced by *E. coli* K-12 under appropriate conditions, usually serves as the reference model for capsular synthesis (Virlogeux *et al.*, 1996). *E. coli* strains produce a variety of cell surface polysaccharides. Some of these polysaccharides are serotype specific, such as the lipopolysaccharide (LPS) O side chains (O antigen) and the capsular (K antigen) polysaccharides (CPS). There are approximately 160 different O antigens in *E. coli*, and more than 70 biochemically distinct K antigens are

recognised (Jann & Jann, 1987). The K antigens are classified into two groups (designated I and II) by chemical, physical, and genetic criteria. Other cell surface polysaccharides, such as colanic acid (M antigen) and enterobacterial common antigens, are not serotype specific and are found in many different strains of *E. coli* and in other enteric bacteria. The O and K antigens are thought to act as virulence determinants in pathogenic *E. coli* by providing resistance to complement-mediated serum killing and phagocytosis (Jann & Jann, 1987).

In contrast to K antigen, the function(s) of colanic acid is unclear, since this polymer is generally produced only at low growth temperatures or on nitrogen-limited, carbon-rich media (Houng *et al.*, 1992). Recent experiments suggest that colanic acid capsular production may represent an adaptive response that provides cells with the means to survive in unfavourable conditions prevailing outside the host such as desiccation and osmotic shock (Brill *et al.*, 1988; Stout, 1994).

The general biosynthetic pathway for colanic acid capsular polysaccharide begins with the synthesis of the component sugars as nucleotide sugar intermediates, which is followed by assembly, transport, and modification of the polymer (Markovitz, 1977). The genes thought to be involved in the latter steps of assembly, transport and modification are called the *cps* genes (Stout, 1996). In *E. coli* K-12 the *cps* genes are located near the *his* and the *rfb* (encoding protein O antigen LPS synthesis) gene clusters located at 45 min on the chromosome (Boulnois & Roberts, 1990). The *cps* gene cluster is composed of 5 complementation groups (*cps*A-F genes) that encode polymerisation, transport and modifications functions. Transcriptional activation of the *cps* genes depends on a two-component regulator encoded by *rcsC* and *rcsB* (47 min), which are separated by 196 bp and transcribed in opposite directions. They act in concert with the product of the *rcsA* gene (43 min), an unstable and limiting regulatory protein that is a substrate for the Lon protease, a major ATP-dependent protease (Gottesman *et al.*, 1985; Gottesman & Stout, 1991).

However, studies using gene disruption and complementation revealed an *rcsA*- independent pathway for capsule synthesis and the requirement for RcsA can be partially bypassed by overproducing RcsB (Brill *et al.*, 1988). Therefore, RcsA is considered as an auxiliary factor, which may interact with RcsB to form a heterodimer required for increased transcription of the colanic acid synthesis (*cps*) genes (**Fig 2.9**). The RcsAB dimer was thought to act as a class I activator (Ishihama, 1993) and helps in attracting the RNA polymerase or in stabilising the transcriptional complex (Wehland & Bernhard, 2000). RcsC has a negative regulatory effect that was described through mutations that increase the level of polysaccharide by more than 10-fold (Gottesman *et al.*, 1985). Recessive mutation in *rcsC*, such as *rcsC137* a mutant which lacks autophosphorylation activity and, therefore, is locked in a conformation that leads to constitutive activation of RcsB, results in overexpression of capsular polysaccharide (Stout & Gottesman, 1990). On the other hand, RcsB was genetically defined by the isolation of mutations that block the expression of *cps* genes in mutant hosts (Gottesman *et al.*, 1985). Moreover, another positive regulator RcsF, a putative membrane protein that has no significant sequence similarity with any known sensor proteins, has been proposed to be a kinase that may exerts its action on colanic acid synthesis through activation of RcsB (Gervais & Drapeau, 1992). RcsB appears to be essential for *cps* expression (Brill *et al.*, 1988), and RcsA cannot activate *cps* expression in the absence of RcsB (Brill *et al.*, 1988; Torres-Cabassa & Gottesman, 1987). From the above, RcsB has three major roles, namely, interaction with RcsC, interaction with RcsA, and interaction with the *cps* promoter (Gupte *et al.*, 1997).

2.5.8.2 Additional possible roles for RcsB and RcsC proteins

(i) RcsB

It is possible that different regulatory networks, using common factors such as RcsB, have been adopted for the control of synthesis of different capsules. The environmental

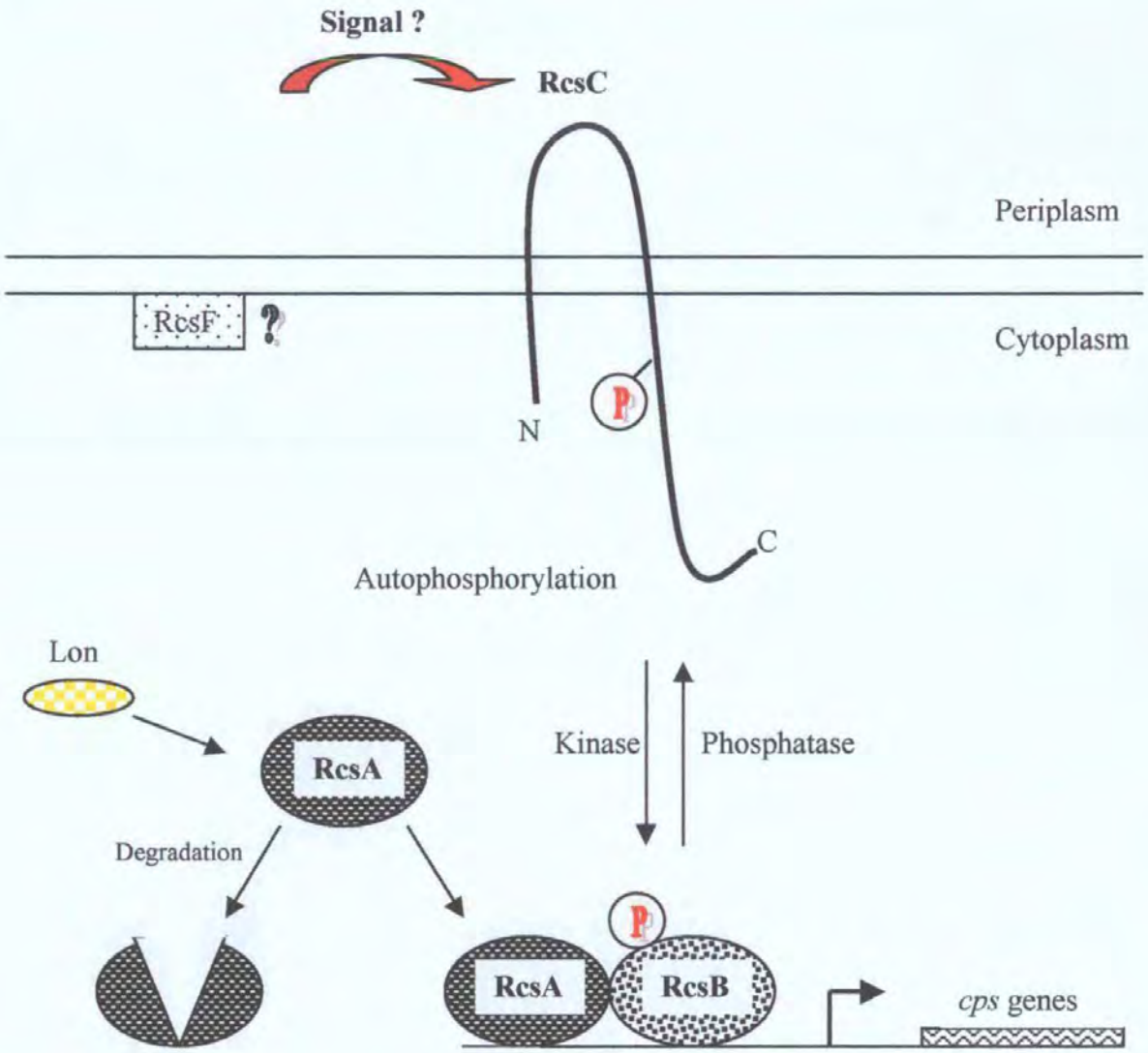


Fig 2.9: Model of the RcsB-RcsC two-component signalling system and capsule regulation in *E. coli*. (Adapted from Stout & Gottesman, 1990 and Kelly & Georgopoulos, 1997)

circumstances that dictate the need for one type most likely make the production of the other unnecessary (Russo & Singh, 1993). The RcsB protein required for the regulation of colanic acid capsule is present in the cytoplasm and can be used for regulation of different kinds of capsules without the energy waste involved in the synthesis of new and different regulators for a different capsule.

Generally, most two-component systems have single kinase acting on one regulator (e.g. NtrB$\rightarrow$NtrC or EnvZ$\rightarrow$OmpR). However, different numbers of these elements may be linked together allowing one kinase to initiate multiple responses (distributive circuit) or multiple kinases to initiate a common response (integrative circuit) (Bourret *et al.*, 1991). The ability to switch or upregulate expression under specific circumstances (e.g. at specific phases during infection) would almost certainly provide a selective advantage for pathogens (Gottesman *et al.*, 1985).

In *E. coli* 04/K54/H5, an extraintestinal pathogenic isolate, its group II capsular polysaccharide is an important virulence determinant and confers serum resistance. Russo & Singh (1993) demonstrated that RcsA is a negative regulator of that capsule, and its action appears to be mediated through RcsB, as is the case in *E. coli* K-12. Likewise, in *E. coli* and *Klebsiella pneumoniae,* RcsB affects the production of both group I and II capsular polysaccharides, which is not surprising considering that the function of these capsules appears to be divergent, with respect to serum sensitivity (McCallum *et al.*, 1989). The environmental circumstances that dictate the need for one type most likely make the production of the other unnecessary. Jayaratne *et al.* (1993) examined the function of the *rcs* regulatory system in expression of the K30 capsular polysaccharide, which is produced at 37∘C and has been implicated as a virulence factor (Orskov *et al.*, 1971). They indicated that the *rcs* system is not essential for expression of low levels of the group I capsular polysaccharide in lon+*E. coli* K30, although it appeared to be more important for achieving

high levels of expression.

On the other hand, *rcsB* can act as a positive regulator of *Klebsiella* K2 capsule production. Wacharotayankun *et al.* (1992) demonstrated, by Northern hybridisation, that *rcsB* enhanced a long strand of mRNA from *Klebsiella cps* genes indicating that *rcsB* regulates the expression of K2 capsular polysaccharide at the transcriptional level.

Gervais and Drapeau, (1992) and Gervais *et al.* (1992) suggested that RcsB has a role in the regulation of *ftsZ*, a gene encoding a septum protein that is critical for cell division, and proposed that RcsB increases the transcription of *ftsZ* thus accounting for the restoration of colony formation of *ftsZ* mutant cells. They identified multicopy clones carrying *rcsB* among those that restored the ability of *ftsZ* mutants to grow on medium without salt; this effect is independent of RcsA. Increasing the amount of RcsB led to an increase in the expression of an *ftsAZ-lacZ* fusion. Gervais *et al.* (1992) demonstrated the capacity of *rcsB* to autoregulate its own synthesis, particularly in the presence of NaCl.

The 5' untranslated sequence of *rcsB* contains several regulatory signals. Deleting this region or part of it results in a gene that is not sufficiently expressed to promote colanic acid synthesis and *ftsZ* expression. The regulatory signals present are the σ^{54} promoter which is located immediately upstream of the rcsB initiation codon, a consensus LexA-binding site sequence containing a presumptive σ^{70} promoter, and another uncharacterised sequence that appears to be essential for its expression.

(ii) RcsC

RcsC was identified during a search for genes controlling the expression of the *tol* genes, which encode proteins involved in the maintenance of cell envelope integrity in *E. coli*. Mutations in any of the *tolQRA* genes render the cell sensitive to drugs and bile salts and induce the release of periplasmic enzymes. A *rcsC* mutant strain, *rcsC338*, was transformed with plasmids containing operon fusions to the *tolQ*, *tolR* or *tolA* genes. The

results showed that in the absence of an external signal, RcsC would not phosphorylate its transmitter protein. Under such conditions, capsule synthesis is not activated while *tolQAR* genes are expressed to maintain cell-envelop integrity. On the other hand, in the presence of a signal indicating some alteration to the cell surface, RcsC switches its control by activating capsule synthesis as a stress response to protect the cell (Clavel *et al.*, 1996).

2.5.8.3 The Rcs system in other bacterial species

The *rcs* regulatory network has been shown to be widespread among bacterial species, and may serve as a common regulatory system for the expression of structurally different polysaccharides in bacteria that occupy diverse ecological niches (Jayaratne *et al.*, 1993). Homologs of Rcs proteins have been identified in *Erwinia* spp, *K. pneumoniae*, and *S. typhi*. In *Erwinia stewartii*, a primary determinant of pathogenicity is the production of extracellular polysaccharide (EPS). *Erwinia* spp use RcsA to regulate expression of both the species-specific extracellular polysaccharide and a leavan polymer (Toress-Cabassa *et al.*, 1987; Coleman *et al.*, 1990). The gene promotes the transcription of at least two operons (*cps*) involved in EPS synthesis. Possibly, both RcsB and RcsC, regulate the production of group II capsular polysaccharide.

There have been few *in vivo* studies on the possible role, if any, of RcsB and RcsC in virulence. Bereswill & Geider (1997) showed that overexpression of the *rcsB* gene in *Erwinia amylovora* from high-copy-number plasmids stimulated the synthesis of the acidic exopolysaccharide amylovoran. In addition, inactivation of *rcsB* by site-directed mutagenesis created mutants that were deficient in amylovoran synthesis and avirulent on host plants.

The capsular polysaccharide of *S. typhi*, a linear homopolymer of α-1,4 2-deoxy-2-*N*-acetylgalactosamine uronic acid, is produced by all strains of *S. typhi*, and *Salmonella paratyphi* as well as from some strains of *Salmonella dublin* and *Citrobacter freundii*.

Production of Vi antigen is partially controlled by the two-component regulatory system RcsB-RcsC and by the positive regulator TviA, which interacts with RcsB to promote optimal transcription of the genes involved in Vi antigen synthesis (Virlogeux *et al.*, 1996). However, there are some specific features in the mechanisms controlling Vi antigen synthesis in *S. typhi* that are distinct from those involved in colanic acid synthesis in *E. coli*: (i) RcsA is not involved in Vi antigen synthesis; (ii) TviA is the auxiliary protein interacting with RcsB to promote transcription of the *viaB* locus; and (iii) TviA is not subjected to Lon-dependent degradation (Virlogeux *et al.*, 1996).

Furthermore, the production of *S. typhi* invasion-promoting Sip proteins, secreted proteins required for invasion and entry into epithelial cells, and flagellin was reported to be modulated by the RcsB-RcsC regulatory system in response to changes in environmental osmolarity. This regulation occurs at both the transcriptional and post-translational levels (Arricau *et al.*, 1998). The RcsB-RcsC modulation of expression in response to changes in the osmolarity of the growth medium was attributed to the adaptive response of *S. typhi* to the changing environments encountered during the different stages of pathogenesis.

In summary, two-component regulatory systems are widespread amongst bacterial species and are essential for regulating responses to changing environments. Therefore, in many instances these systems make a vital contribution to the pathogenesis of disease.

CHAPTER 3

GENERAL MATERIALS AND METHODS

3.1 REAGENTS AND MEDIA

Details of the formulation of reagents and composition of media used during this study are given in Appendix III. When necessary, reagents and media were sterilised by autoclaving (121°C, 15 min). Heat-sensitive components, such as antibiotics, were filtered using a syringe and membrane filter of 0.2 or 0.45 μm mesh (Sartorius Ltd).

Unless stated otherwise, analytical grade chemicals were obtained from Sigma, or BDH, while enzymes for the restriction and modification of DNA were obtained from Gibco BRL, Boehringer Mannheim, Promega Ltd, Stratagene or New England Biolabs. All PCR reactions were performed with reagents obtained from Roche (Lewes, UK).

3.2 BACTERIOLOGY

3.2.1 Bacterial strains

The *B. pseudomallei* strains that were used in this work are described in **Table 3.1**, while other bacterial species are listed in **Table 3.2**. The *E. coli* strains shown in **Table 3.3** were used for general cloning procedures, the propagation of the plasmids that were constructed during this study and for the manipulation of cloned genes.

3.2.2 Laboratory culture and storage of bacteria and phage

All strains of *B. pseudomallei* were grown on LB medium at 37°C statically, unless otherwise stated, within an authorised Containment Level 3 Laboratory (CL3) either at the PHLS, Derriford Hospital, Plymouth by Dr D.A.B Dance, or at 401A Davy Building, University of Plymouth by Dr M.L. Gilpin. The identity of *B. pseudomallei* was confirmed by the API 20NE biochemical test.

E. coli K12 strains were routinely incubated aerobically on LB medium at 37°C and broth cultures were shaken at 200 strokes/min, overnight. When required, antibiotics were added to the autoclaved growth medium after cooling to less than 50°C (**Table 1**, Appendix III).

Table 3.1: Isolates of *B. pseudomallei* that were used in this study

Strain	Source	Sample	Geographical origin	Date
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
204	Human	Blood	Thailand	-
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	-
576	Human	Blood	Thailand	-
E8	Environment	Soil	Thailand	-
E25	Environment	Soil	Thailand	-
Zhan 1	Human	Blood	Malaysia	
Hainan 1	Human		China	
Hainan 2	Human		China	
Hainan 3	Human		China	
Hainan 4	Human		China	
Hainan 55	Human		China	
Hainan 106	Human		China	
Hainan 706	Human		China	
Jie 187	Human		China	

Table 3.2: Bacterial strains that were used for DNA hybridisation studies

Strain	Source
B. thailandensis E27	T. Pitt
B. thailandensis E82	T. Pitt
B. thailandensis E254	T. Pitt
B. thailandensis E255	T. Pitt
B. thailandensis E256	T. Pitt
B. thailandensis E260	T. Pitt
B. cepacia I non-epidemic	T. Pitt
B. cepacia II non-epidemic	T. Pitt
B. cepacia IIIa epidemic	T. Pitt
B. cepacia IIIa non-epidemic	T. Pitt
B. cepacia IIIb non-epidemic	T. Pitt
B. cepacia IIIc epidemic	T. Pitt
B. cepacia IV non-epidemic	T. Pitt
B. cocovenenans LMG 11626	T. Pitt
B. plantarii LMG 10908	T. Pitt
B. vietnamiensis LMG 6998	T. Pitt
B. vietnamiensis LMG 6999	T. Pitt
B. vietnamiensis LMG 10620	T. Pitt
B. vandii LMG 10620	T. Pitt
P. aeruginosa	University of Plymouth culture collection

Table 3.3: Strains of *E. coil* K12 that were used in this study

Strain	Genotype	Source
SY327 λpir	F⁻*ara*D D(*Lac pro*) argE (Am) *rec*A56 Rif^R nalA	N. Fairweather
DH5α	F⁻, *end*A1, *hsd*R17 (r-k m+k), *sup*E44, *thi*, *rec*A1, *gyr*A96, *rel*A1, Δ*lac*ZM15	M. L. Gilpin
LE392	F⁻ *hsd*R574 (r_K-, m_K+) *sup*E44 *sup*F58 lacY1 or Δ(*lac*ZY)6 *gal*K2 *gal*T22 *met*B1 *trp*R55	Promega
XL1-Blue	*sup*E44 *hsd*R *thi*-1 *thr*-1 *leu*B6 *lac*Y1 *ton*A21	M. L. Gilpin
Y1088	*ΔlacU*169 *sup*E *sup*F *hsd*R⁻ *hsd*M$^+$ *met*B rtpR *ton*A21 *pro*C□Tn5 (pMC9). PMC9=pBR322-*lac*^{Q}	Promega
XLOLR	Δ (*mcr*A) 183 Δ (*mcr*CB-*hsd*SMR-*mrr*) 173 *end*A1 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* [F' *pro*AB *lac*I^q ZΔM15 Tn10(Tetr)] Su⁻	Stratagene

For long term storage of bacterial strains, single colonies were removed using a sterile inoculating loop and grown overnight at 37°C in the appropriate broth and 0.5 ml aliquots were mixed with 0.5 ml of sterile glycerol, cryoprotectant, and stored at -20°C or at -80°C. For short-term storage, cultures on agar plates were sealed with parafilm and stored at 4°C for up to eight weeks. To initiate growth after storage at -20°C or -80°C, 50-100 µl of stock culture was inoculated into 10 ml LB broth or streaked onto agar plates, and incubated at 37°C prior to inoculating media containing the appropriate antibiotics.

3.2.3 Sterility and viability checks of *B. pseudomallei* culture products

B. pseudomallei is a Category 3 pathogen, therefore all products that were derived from viable cultures that were to be used outside the CL3 facility were rigorously checked to confirm sterility. Initially, two independent samples were processed and their entire contents were tested for viability by inoculating 0.1 ml aliquots onto LB agar and into LB broth, followed by incubation at 37°C for 3 days. If no growth was obtained, then a 10% volume of all subsequent preparations were checked by inoculating broth and agar plates with 0.1 ml aliquots and incubating at 37°C for 3 days prior to release.

3.3 PLASMIDS AND PHAGE

The plasmids and phage shown in **Table 3.4** were used for constructing gene libraries (λGEM-11, λZAP), subcloning for DNA sequencing (pUC18, $pBSK^+$) or gene expression (pMAL-C2) and for constructing a suicide vector for mutagenesis studies (pCVD442). Phage stocks were stored with a few drops of chloroform at 4°C.

3.4 DNA MANIPULATION

3.4.1 Preparation of chromosomal DNA

Prior to the availability of the CL3 facility at the University of Plymouth, *B. pseudomallei* chromosomal DNA was the generous gift of Dr T. Pitt, Central PHLS, London.

Table 3.4: Plasmid and phage vectors that were used in cloning studies

Vector	Description	Source
pUC18	ColE1 ori, *lac*Z, Amp^r, polylinker *Eco*RI-*Hind*III	Lab collection
pBlueSK$^+$	ColE1 ori, *lac*Z', Amp^r	Stratagene
pCVD442	Amp^r, *pir*-dependent R6K replicon, *sacB*	N. Fairweather
pMAL-c2	*lac*Z, Amp^r, *lac*I^q, *mal*E	New England Biolabs
Lambda GEM-11	Purified bacteriophage λ left and right arms, stuffer fragment removed by *Bam*HI digest	Promega
Lambda Zap Express	ZAP express predigested Gigapack cloning kit-BamHI/CIAP-treated	Stratagene

3.4.1.1 Minipreps

For *B. pseudomallei*, 10 ml of an overnight culture was centrifuged (3,000 x g, 20 min, 4°C) and the cell pellet was washed in 10 ml of STE (Appendix III). After re-centrifugation, as above, the cell pellet was resuspended in 5 ml of solution A (Appendix III) for 5 min at RT, then mixed with 1 ml of solution B (Appendix III). After 5 min at room temperature (RT), 1.2 ml of solution C (Appendix III) was added and mixed then 0.7 ml of solution D (Appendix III) was added and the mixture was placed in a water bath at 80°C for 2 h with occasional mixing. Finally, the mixture was checked for sterility (section 3.2.3) while the remaining lysate was stored at 4°C or at -70°C awaiting the outcome (3 days). For further purification of the DNA, the lysate was treated using the Puregene DNA isolation kit as detailed in section 3.4.1.2.

3.4.1.2 Puregene DNA isolation kit

For genomic DNA extraction from bacterial species other than *B. pseudomallei*, the Puregene DNA isolation kit was used (Gentra Systems, Minneapolis, USA). The protocol used was supplied by the manufacturer for the extraction of genomic DNA from 0.5 ml overnight cultures of gram-negative bacteria. Cultures were centrifuged (13,000 x g, 2 min) and the supernatant removed. The cell pellet was resuspended in 300 µl of cell lysis solution then incubated at 80°C for 5 min. RNase solution (1.5 µl) was mixed with the cell lysate and the samples were incubated at 37°C for 30 min before cooling to RT. Protein precipitation solution (100 µl) was added then the samples were vortexed (20 s) and centrifuged (13,000 x g, 3 min, RT) to pellet the precipitated protein. The supernatant, containing the DNA, was poured into a clean tube containing 300 µl isopropanol and the samples were gently mixed by inverting the tubes 50 times. The tubes were centrifuged (13,000 x g, 1 min, RT), the supernatant was removed and the pelleted DNA was air dried for 15 min. The DNA was rehydrated overnight at RT following the addition of 50 µl of TE buffer.

3.4.2 Preparation of plasmid DNA

3.4.2.1 Small scale preparation

The small scale preparation of plasmid DNA from bacterial cultures was performed as described by Birnboim & Doly (1979). Aliquots (1.5 ml) of overnight cultures were centrifuged (13,000 x g, 30 s, RT). The pelleted cells were resuspended in 100 µl of B & D I solution (Appendix III) and left at RT for 5 min. The bacterial cells were lysed by adding 200 µl of freshly prepared B & D solution II (Appendix III); the contents were gently mixed and incubated on ice for 10 min. After the addition of 150 µl of ice cold B & D solution III (Appendix III), the solution was mixed, placed on ice for 10 min to precipitate cell debris, denatured proteins, chromosomal DNA and SDS. The tube was centrifuged (13,000 x g, 5 min, RT) and the supernatant was transferred to a fresh microfuge tube. The plasmid DNA was precipitated by mixing with 1 ml of cold 100% ethanol, then incubating at -70°C for 15 min. After centrifugation (13,000 x g, 5 min), the plasmid DNA pellet was resuspended in 100 µl of B & D solution IV (Appendix III), reprecipitated by adding 250 µl of 100% cold ethanol and centrifuged as above. The pellet was washed with 70% ethanol, centrifuged (13,000 x g, 5 min, RT) and the supernatant was discarded. Finally, the precipitated plasmid DNA was air dried then resuspended in 10 µl of TE buffer. Alternatively, the addition of B & D IV was omitted and the plasmid DNA pellet was instead resuspended in 50 µl of RNase solution, incubated at 37°C for 1 h, and the plasmid DNA purified by phenol/chloroform extraction (section 3.4.4.2).

3.4.2.2 Large scale preparation

The extraction of plasmid DNA on a large scale was performed by a modification of the above method. Bacterial cultures (200-500 ml) were harvested by centrifugation (6,000 x g, 15 min, 4°C) using a MSE, Europa 24M (UK). The pellet was resuspended in 10 ml of B & D solution I for 5 min RT, then 1 ml of lysozyme solution (10 mg/ml) was added with

mixing and incubated for 5 min. Freshly prepared B & D solution II (20 ml) was added, mixed by gently inverting the centrifuge bottle and the suspension was placed on ice for 10 min. After the addition of 15 ml of ice-cold B & D solution III, the mixture was placed on ice for 10 min and centrifuged (9,000 x g, 10 min, 4°C). The supernatant was carefully transferred to a fresh tube, warmed to room temperature and mixed with 0.6 volumes of isopropanol for 20 min at RT. Plasmid DNA was pelleted by centrifugation (5,000 x g, 30 min, 20°C), washed in 70% ethanol, then centrifuged as above, and the pellet was air dried and resuspended in 2.5 ml of TE buffer.

3.4.2.3 Minipreps

Rapid extraction of small quantities of pure plasmid DNA from bacterial cultures was performed using the Miniprep plasmid DNA extraction kit (Bio-Rad, UK). Aliquots (1.5 ml) of overnight cultures were centrifuged (13,000 x g, 30 s, RT) and resuspended in 200 µl of cell resuspension solution. Cell lysis solution (250 µl) was added, gently mixed, and 250 µl of neutralization solution was added with mixing. Samples were centrifuged (13,000 x g, 5 min, RT) and the plasmid DNA contained in the supernatant was added to 200 µl of diatomaceous matrix and purified using a spin filter. The matrix was washed twice (13,000 x g, 30 s, RT) in wash buffer and the DNA was eluted into a clean tube with 100 µl of 'Analar' H_2O or TE prewarmed to 50°C (13,000 x g, 1 min, RT).

3.4.2.4 Maxipreps

Rapid extraction of large quantities of plasmid DNA from bacterial cultures was performed using the Maxiprep plasmid DNA isolation kit from Promega, UK. Overnight cultures (100-500 ml) were centrifuged (6,000 x g, 10 min, 4°C) and the pellet was resuspended in 15 ml cell resuspension solution. Cell lysis solution (15 ml) was added, the solution was mixed gently by inverting the tube several times and then centrifuged (11,000 x g, 15 min, RT). The supernatant was filtered through microcloth, and transferred to a fresh

centrifuge tube. To precipitate the plasmid DNA, a 0.5 volume of isopropanol was added, mixed gently and the solution was centrifuged (14,000 x g, 15 min, RT). The supernatant was discarded, the DNA pellet was resuspended in 2 ml TE buffer, and 10 ml of DNA purification resin was mixed into the solution. The DNA/resin mix was transferred into a Maxicolumn, and washed with column wash solution (25 ml) and 80% ethanol (5 ml) using a vacuum manifold. After washing, the Maxicolumn was inserted into a 50 ml polypropylene tube (Falcon, UK) and centrifuged (1500 x g, 5 min, RT), then placed into a fresh 50 ml tube. Preheated (65°C) TE buffer was added (1.5 ml) and left for 1 min before the plasmid DNA was eluted by centrifugation (1500 x g, 5 min, RT). Finally, in order to remove residual resin the DNA solution was passed through a 0.2 μm membrane filter using a syringe and the resulting plasmid DNA was stored at 4°C.

3.4.3 DNA precipitation

3.4.3.1 Ethanol precipitation

To a solution containing DNA, a one tenth volume of 3 M sodium acetate (pH 6.0) was added and mixed followed by three volumes of cold ethanol. After thorough mixing the solution was stored at -20°C overnight or at -70°C for 20 minutes to precipitate DNA. The DNA was pelleted (13,000 x g, 5 minutes, 4°C), the supernatant was carefully removed and the pellet was rinsed with 70% (v/v) ethanol. After centrifugation, the pellet was dried then resuspended in the appropriate amount of TE buffer or 'Analar' water. The concentration and purity of the DNA were determined as described in section 3.4.6.1.

3.4.3.2 Isopropanol precipitation

DNA was precipitated by adding 0.6 volumes of *iso*-propanol (2-propanol) to a DNA solution, incubated for 1 h at RT, then centrifuged (10,000 x g, 15 min, RT). The supernatant was removed and the DNA pellet washed with 70% ethanol and processed as for ethanol precipitation (section 3.4.3.1).

3.4.4 DNA purification

3.4.4.1 Caesium chloride density gradient ultracentrifugation

The method used is described by Sambrook *et al.* (1989). CsCl (2.5 g) was added to a solution of DNA in TE buffer (2.5 ml). A volume (250 µl) of ethidium bromide solution (10 mg/ml in TE buffer) was added and after thorough mixing the preparation was left in the dark at RT for 18 h. This allowed time for ethidium bromide-protein complexes to form and these were removed by centrifugation (13,000 x g, 3 min). The clarified supernatant was transferred and sealed by heat into 3.3 ml Quick-Seal centrifuge tubes (Beckman, USA) and centrifuged at 80,000 rpm for 24 h at 20°C using a fixed angle TLA-100.3 rotor in a TL100 benchtop ultracentrifuge (Beckman, USA). DNA molecules are separated from impurities according to their densities into two discrete bands. The top band contained chromosomal DNA and nicked circular plasmid DNA, whilst the lower band represented the closed circular plasmid DNA. The red pellet at the bottom of the tube contained ethidium bromide/RNA complex and the proteins lay on the surface of the gradient. After the run, DNA bands were visualized with a long-wave UV lamp. A needle was placed in the top of the tube to allow air in, and the plasmid band was removed by side puncture of the tube with a needle and syringe that inserted just below the lower plasmid DNA band, measured and transferred to a sterile plastic 5 ml bijoux. Ethidium bromide was removed from purified DNA solutions by organic phase extraction. An equal volume of isopropanol saturated with water and CsCl was added and the two phases were shaken vigorously. The phases were allowed to separate for 3 min and the upper phase consisting of ethidium bromide and isopropanol was discarded. Fresh isopropanol mixture was added again and the extraction was repeated until all traces of the pink colour of ethidium bromide disappeared. At this stage, plasmid DNA contained in the clear aqueous phase was precipitated by ethanol (section 3.4.3.1) and the CsCl was removed by successive washes in 70% ethanol. The

plasmid DNA pellet was air dried then resuspended in either 200 µl TE buffer for high copy number plasmids or 100 µl TE buffer for low copy number plasmids.

3.4.4.2 Phenol chloroform extraction

Initially, the method for the purification of nucleic acids from cellular extracts involved a first step of emulsifying the DNA solution with an equal volume of liquefied phenol (equilibrated with 0.1 M Tris-HCl, pH 8.0). This organic solvent precipitates proteins but leaves the DNA in aqueous solution. The mixture was then centrifuged (13,000 x g, 2 min, RT) to give an upper aqueous phase containing DNA. The aqueous phase was removed carefully, leaving the interface behind, and transferred to a new microfuge tube. To this an equal volume of phenol/chloroform (Appendix III) was added and thoroughly mixed by vortexing. The mixture was centrifuged (13,000 x g, 30 s, RT) and the upper aqueous layer, containing DNA, was carefully removed and transferred into a clean tube. An equal volume of water-saturated chloroform plus IAA (24 parts chloroform: 1 part isoamyl alcohol) was added, mixed and centrifuged as above and the upper aqueous layer was removed and transferred to a new tube. The chloroform/IAA procedure was repeated. The DNA was then precipitated with ethanol, pelleted (13,000 x g, 5 min, RT), washed in 70% ethanol and dried. The dried pellet was resuspended in TE buffer and stored at 4°C.

3.4.4.3 Rapid purification with Prep-A-Gene kit

The Prep-A-Gene silica-based DNA purification kit (Bio-Rad, UK) provides a fast and effective purification and concentration of DNA. Purification was performed using the protocol supplied by the manufacturer for the recovery of DNA from solutions (protocol 3.2). The amount of Prep-A-Gene matrix that was required was estimated using the capacity figure of 0.2 µg supercoiled DNA/µl Prep-A-Gene matrix. A quantity of DNA purification binding buffer equal to three times the combined volumes of Prep-A-Gene matrix suspension and DNA containing solution was added to the sample followed by the addition

of the predetermined amount of Prep-A-Gene matrix. The mixture was incubated at RT for 10 min with frequent agitation then centrifuged (13,000 x g, 30 s, RT). The supernatant was removed and the pellet containing DNA was rinsed by resuspension in a volume of binding buffer equivalent to 25 X the volume of the added matrix and vortexed. The mixture was centrifuged as above and the Prep-A-Gene pellet was washed twice with 25 X matrix volume of wash buffer. After the second wash and centrifugation, all traces of wash buffer were removed and tubes were left with lids open at either 55°C for 1 min or at RT for 5 min to allow residual alcohol to evaporate. Finally, the DNA was eluted by resuspending the matrix pellet in at least one pellet volume of elution buffer and incubating at 50°C for 5 min. The samples were centifuged as above and the DNA-containing supernatant was carefully removed to a clean tube and stored at 4°C.

3.4.5 Restriction endonuclease digestion of DNA

Class II restriction endonucleases are used extensively in molecular cloning. These enzymes recognise and cleave the DNA molecule at a specific sequence. The amount of enzyme required for complete digestion depends on the amount of DNA to be cut, the number of times the enzyme could be expected to cut the DNA sample, and the specific activity of the enzyme. One unit of enzyme is generally defined as the amount that is required to completely digest 1 μg of bacteriophage λ DNA in 1 h at 37°C.

3.4.5.1 Single enzyme digestion

Typically, in a microfuge tube, 0.2-1 μg DNA was combined with 1 μl (8-20 unit) of enzyme, 2 μl of recommended restriction endonuclease buffer supplied by the manufacturer at 10 x concentration and the reaction volume was made up to 20 μl final volume with analytical grade H_2O (Sigma, Poole, UK). The components were mixed, centrifuged (13,000 x g, 3 s) and unless otherwise stated, the tubes were incubated at 37°C for 3-18 h. Digests were analysed by electrophoresis on a 0.8% (w/v) agarose gel in comparison with DNA

molecular size markers. In some cases, enzymes were inactivated by a 10 min incubation at 68°C. Heat-stable enzymes were removed by phenol/chloroform extraction (section 3.4.4.2) or Prep-A-Gene extraction (section 3.4.4.3). Larger amounts of DNA were digested by scaling up the reaction mixture, ensuring that the enzyme volume was always kept below 10% of the total volume to prevent non-specific cleavage of DNA due to the high concentration of glycerol that is present in stock enzyme solutions.

3.4.5.2 Double enzyme digestion

For double enzyme digests where both enzymes require the same buffer, or a compatible buffer, the reactions were performed as for single enzyme digestion but using half the volume of each enzyme. If the enzymes required different buffers, the DNA was digested in two sequential single digests.

3.4.6 Electrophoresis of nucleic acids

3.4.6.1 Agarose gel electrophoresis and DNA quantification

Agarose gel electrophoresis is the standard method for separating DNA molecules according to their size, and was carried out using gel electrophoresis apparatus supplied by Amersham Pharmacia Biotech (Milton Keynes, UK). Agarose gels were prepared in TBE buffer at a concentration that varied between 0.6-1% (w/v) agarose depending on the size of fragments to be resolved (**Table 2**, Appendix III). Agarose powder was added to 1 X TBE buffer, heated in a microwave oven until molten, cooled to 60°C before ethidium bromide was added to a final concentration of 0.5 μg/ml and mixed by swirling. Molten agarose was then poured into UV-transparent gel cast (Pharmacia) provided with a toothcomb. Once solidified, the gel was submerged in 1 X TBE buffer in the electrophoresis tank, and the well-forming comb was carefully removed. DNA samples were mixed with 1/6 volume of 6 X gel loading buffer (Appendix III) then loaded into the wells. DNA ladder was run alongside in a separate lane to provide molecular size markers. Three molecular markers

were used, namely, the 1 Kb ladder (Gibco BRL, Paisely, UK), the 100 bp DNA ladder (Gibco BRL, UK) and the λ-*Eco*RI ladder (Sigma, Poole, UK). Samples were electrophoresed at 120V until resolved. DNA bands were visualised using long wavelength ultra violet (UV) illumination and photographed using either a Polaroid CU5 camera system (Eastman Kodak, New York, USA) or a gel documentation system (UVI TEC, UK).

For visual quantification of DNA fragments on agarose gels, a set of standard amounts of undigested λ phage DNA markers were run in parallel with the samples of unknown concentration. Because the amount of fluorescence is proportional to the total mass of DNA, the quantity of DNA in the sample can be estimated by comparing the fluorescence yield of the sample with that of a series of standards (Sambrook et al., 1989). This approach allows a rough estimation of the amount of DNA in a given sample and is capable of detecting as little as ~5 ng of double-stranded DNA. Usually a set of four known amounts (12, 60, 125 and 250 ng) of double stranded λ phage DNA (Sigma, Poole, UK) were used.

Alternatively, for pure samples, without substantial amounts of contaminants such as proteins or phenol, spectrophotometric measurements for the determination of DNA concentration was used. The method is based on the different absorbance spectra of nucleic acids and proteins, since DNA and RNA both strongly absorb UV light with a maximum absorbance at 260 nm. Samples of DNA were appropriately diluted, placed in quartz cuvettes and the readings were taken at wavelength 260 nm (Pharmacia PU 8720 spectrophotometer) that allows the calculation of the concentration of nucleic acid. An optical density (OD) of 1 corresponds to approximately 50 μg/ml for double stranded DNA, 40 μg/ml for single-stranded DNA and RNA, and 20 μg/ml for oligonucleotides (Sambrook et al., 1989). When DNA concentrations are low (<250 ng/ml) spectrophotometric measurements may not be accurate.

3.4.6.2 Isolation of DNA fragments from agarose gels

The recovery of DNA molecules from agarose gels was carried out using the Prep-A-Gene DNA purification system (section 3.4.4.3). Following electrophoresis the agarose gel was examined under UV light and the gel slice containing the desired band was excised using a sterile razor blade, then transferred to a pre-weighed sterile microfuge tube. The tube was weighed again and the weight of the gel slice was calculated. The amount of Prep-A-Gene matrix that was required was estimated using the capacity figure of 0.2 μg supercoiled DNA/μl Prep-A-Gene matrix. A volume of purification binding buffer equivalent to three times the combined volumes of gel and matrix was added and the tube heated in a water bath at 55°C for 5 min. The Prep-A-Gene matrix was fully resuspended then added to the dissolved gel slice and mixed by inversion. The purification proceeded as described in section 3.4.4.3.

3.5 CLONING INTO PLASMID VECTORS

Vectors were digested with the appropriate restriction endonuclease (section 3.4.5) and phosphatase treated if required (section 3.5.1). Restriction endonuclease activity was removed by either phenol/chloroform extraction (section 3.4.4.2) or by rapid purification with the Prep-A-Gene kit (section 3.4.4.3).

3.5.1 Alkaline phosphatase treatment of linearised vector

Alkaline phosphatase is capable of removing 5'-terminal phosphate groups from nucleic acid molecules that are required for DNA ligation. Hence phosphatase treatment of vector DNA will prevent self-ligation and simplify cloning because phosphorylated DNA molecules can still be ligated into the vector. Linearised vector DNA (1 μg) was dephosphorylated using 2 units of calf intestinal alkaline phosphatase (CIAP) (Boehringer, Mannheim) in the presence of CIAP reaction buffer in a total volume of 30 μl at 37°C for 1

h. A further 2 units of enzyme were added after 30 min. Because CIAP can interfere with the efficiencies of subsequent ligation and transformation, the alkaline phosphatase was inactivated by incubating the reaction for 10 min at 68°C and removed completely by phenol/chloroform extraction (3.4.4.2) or with the Prep-A-Gene kit (3.4.4.3).

3.5.2 Ligation of blunt-ended DNA restriction fragments

All blunt ends are suitable for ligation to one another. Usually a fragment to vector ratio of 4:1 was used, but often a range of ratios was used to ensure efficient ligation of the insert into the vector. The insert and vector were mixed with of 1 µl of T4 DNA ligase (Gibco BRL 1U/µl) and ligation buffer (Gibco BRL), and diluted with 'Analar' H_2O to a final volume of 30 µl. The condensing agent hexamine cobalt chloride (Appendix III) was added to a concentration of 1 µM, if necessary, to increase molecular crowding and improve ligation efficiency. The mixture was incubated at 16°C for 18 h. Once completed, the successful ligation was confirmed by agarose electrophoresis in comparison with both the digested vector and DNA markers. Ligation mixture was either used directly to transform *E. coli* or frozen at -20°C until required.

3.5.3 Ligation of cohesive-ended DNA restriction fragments

Ligation of cohesive DNA molecules was performed as for blunt ended fragments but with reduced quantities of DNA and T4 ligase. Generally, the amount of enzyme needed in blunt end ligation is found to be 10 times greater than that needed to achieve sticky-end ligation (Perbal, 1988).

3.5.4 Transformation of recombinant DNA into *E. coli* K12

3.5.4.1 Calcium chloride transformation

The usual method that was used for introducing recombinant plasmids into *E. coli* was by transformation of the host following treatment with calcium chloride (or other cations) to induce competency. A single colony of the appropriate host strain of *E. coli* was cultured in

10 ml of LB broth (section 3.2.2). A 1.5 ml aliquot of the bacterial culture was centrifuged (13,000 x g, 30 s, RT), the cells were resuspended in 200 µl of ice-cold 75 mM $CaCl_2$, briefly vortexed, and placed on ice for 30 min. The cells were pelleted (13,000 x g, 30 s, RT) and then gently resuspended in 100 µl of fresh ice-cold $CaCl_2$ (75 mM) using a micropipette. Between 50-100 ng of plasmid DNA in no more than 20 µl volume was added, mixed gently and left on ice for 10 min. The mixture was then placed at RT for 10 min, ice for 30 min, heat shocked for exactly 90 s in a water bath at 42°C, and immediately replaced on ice for a further 2 min. A volume of 0.9 ml of LB broth was added and the transformation mixture incubated at 37°C for 90 min. After the incubation period, different volumes of the transformed cell suspension was spread onto LB agar plates containing the appropriate selective antibiotic. If desired, 0.2% glucose was added to the LB agar to repress expression of the *lac* promoter that is present in certain plasmids. For volumes of less than 50 µl of transformed resuspension, sterile LB broth (100 µl) was added to the plate first, to facilitate spreading. Plates were then incubated at 37°C for overnight or up to 36 h. The plasmid DNA contained in the resulting transformants was extracted (section 3.4.2) and analysed (sections 3.4.5; 3.4.6) in order to identify the desired clone.

3.5.4.2 Transformation using frozen competent cells

Several colonies were picked from a freshly streaked LB agar plate and inoculated into 5 ml of LB broth. After vortexing, the suspension was used to inoculate 200 ml of LB broth and grown in a water bath at 37°C with shaking (275 rpm) to A_{600} ~ 0.6, usually for 2-3 h. The culture was then chilled on ice for 20 min, and centrifuged (1000 x g, 15 min, 4°C). The supernatant was decanted and the tubes were inverted on a paper towel for 5 min to allow the last traces of media to drain away. The cell pellet was gently resuspended in 60 ml of transformation buffer (RF1 medium, Appendix III) by vortexing and pipetting, then incubated on ice for 20 min. The cells were pelleted as described above, resuspended in 16

ml of RF2 medium (Appendix III), and then placed on ice for 15 min. Aliquots of 200 µl were distributed into sterile 1.5 ml microfuge tubes, flash frozen in liquid nitrogen, and stored at -70°C.

A tube of frozen competent cells was removed from -70°C and thawed on ice. Plasmid DNA (50 ng) was added in a volume <10 µl, and the tube was mixed gently and immediately chilled on ice for 30 min. The transformation mixture was treated as described above (section 3.5.4.1).

3.5.5 Blue/white screening for recombinant clones

Insertional inactivation of the functional *lacZ* gene that is present in certain plasmid vectors including pUC18 and λ bacteriophages (such as λGT-11) provides a convenient means of identifying recombinant clones. Expression of the functional *lacZ* gene yields a blue colour in the presence of the chromogenic substrate X-Gal.

Transformants were spread onto LB agar with the appropriate selective antibiotic added and containing a 4 ml LB agar overlay supplemented with that contained 45 µl of X-Gal (20 mg/ml) and 45 µl of IPTG (24 mg/ml) solutions. The plates were incubated at 37°C for 24 h, then examined for the presence of blue (lac^+, non recombinant), and white (lac^-, recombinant) colonies.

3.6 CLONING OF A PCR PRODUCT INTO A PLASMID VECTOR

3.6.1 Blunting and kinasing reaction

The amplified DNA fragments were purified using the Prep-A-Gene kit (section 3.4.4.3). A volume of 1-16 µl containing 20-1000 ng DNA was added to a tube containing 1 µl Klenow fragment of DNA polymerase, 2 µl of 10X blunting/kinasing buffer, 1 µl of polynucleotide kinase, and sterile dH_2O to a final volume of 20 µl. The contents were mixed gently, centrifuged (13,000 x g, 5 s, RT) and incubated at 37°C for 30 min. The blunted-

kinased amplicon was subsequently purified with the Prep-A-Gene kit (section 3.4.4.3).

3.6.2 Ligation of DNA into pUC18

The pUC18 plasmid DNA (**Fig 3.1**) was purchased as a 'Ready to go' formula (Sureclone Kit, Pharmacia). The purified blunted-kinased amplicon (5 µl) was mixed with 1 µl (50 ng) of *Sma*I pre-digested and dephosphorylated pUC18, 10 µl of 2X ligation buffer, 1 µl of 0.1 M DTT solution, 1 µl of T4 DNA ligase, and sterile dH_2O to a final volume of 20 µl. After gentle mixing and centrifuging (13,000 x g, 5 s, RT), the mixture was incubated overnight at 16°C for overnight.

3.6.3 Selection and identification of recombinant clones

E. coli DH5α was transformed using the ligated DNA (section 3.6.2) and recombinants were selected by blue/white screening as described in section (3.5.5). Recombinant plasmid DNA was isolated from *E. coli* using the Miniprep plasmid DNA extraction kit (section 3.4.2.3), digested with *Eco*RI and *Pst*I (section 3.4.5.2) to release inserted DNA fragments, and checked by agarose gel electrophoresis (section 3.4.6.1). *Eco*RI and *Pst*I restriction sites are located either side of the *Sma*I site within the polylinker of pUC18.

3.7 RESTRICTION MAPPING

It is often essential to obtain a restriction map of cloned fragments before additional manipulations can be carried out. The cloned DNA is usually cut with a variety of restriction enzymes that cleave the vector at specific sites and the sizes and numbers of fragments generated by each enzyme are determined. By performing a series of single and multiple digests (sections 3.4.5.1 and 3.4.5.2) with a range of enzymes, a complete restriction map of the cloned DNA can be deduced (Nicholl, 1998).

3.8 SEQUENCING OF DNA

Sequencing grade DNA was prepared from clones using the caesium chloride density

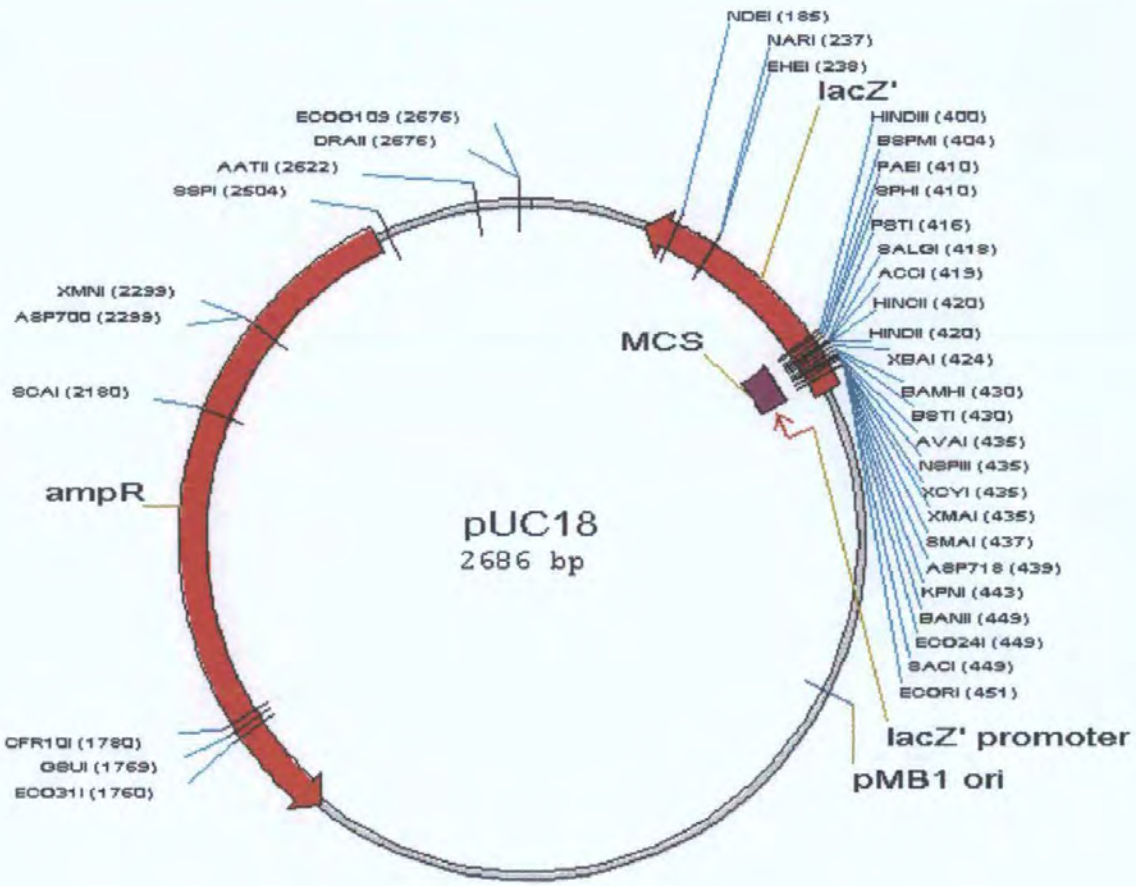


Fig 3.1: Schematic presentation of the pUC18 cloning vector.

gradient ultracentrifugation (section 3.4.4.1) or using the Maxiprep plasmid DNA isolation kit (section 3.4.2.4). Dye terminator or dye primer DNA sequencing was performed by MWG-Biotech Ltd (Milton Keynes, UK) using M13 universal and reverse primers or custom synthesised primers when needed.

3.9 COMPUTER SOFTWARE AND SEQUENCE ANALYSIS

In order to rapidly examine nucleotide sequences for the presence of a variety of structural and functional features including restriction endonuclease recognition sites, secondary structures in single stranded DNA and RNA molecules, and the presence of open reading frames, a variety of computer analysis software packages were used. Only one strand of DNA known as the sense strand, contains coded information and directs RNA synthesis, and the other complementary strand is known as an antisense strand. Messenger RNA is made from 5' to the 3' end and is read in successive groups of three bases, codons, each of which codes for an amino acid. A common way of showing that a piece of DNA could encode a protein is to demonstrate the presence of an open reading frame, or ORF, in the nucleotide sequence. An open reading frame comprises an initiation codon followed by a uninterrupted sequence of codons in the same reading frame terminated by a stop codon (TAA, TAG, TGA). It is well established that most genes are longer than 50 codons, the average length is 317 codons for *E. coli* (Brown, 1999). To identify possible ORFs, translations were carried out in all possible reading frames, i.e. six reading frames, three in one direction and three in the reverse direction of the complementary strand.

Computer analyses of the nucleotide and deduced amino acid sequences were performed using DNASIS sequence analysis software package (Hitachi, UK). Gene promoter predictions were aided by prokaryotic sequence prediction program (Harley & Reynolds, 1987), and only those with a high probability score (>0.8) were selected. Searches for nucleotide and amino acid sequence similarities were done using the FASTA and

BLAST (Basic Local Alignment Search Tool) programs (Pearson & Lipman, 1988; Altschul *et al.*, 1990) in the EMBL, NCBI, GenBank and SWISS-PROT databases, which were available through the computing facilities provided by the Department of Biological Sciences, University of Plymouth, England. Protein sequences were analysed using the PHDTOPOLOGY at PredictProtein Server, EMBL, Germany, while the Prosite program was used for motif search. Multiple alignment was performed with NSA and CLUSTAL alignment programs (Higgins & Sharp, 1988) at IBCP, Lyon France. Hydrophobicity was analysed by PEPWINDOW (Kyte & Doolittle, 1982). Helix-turn-helix motifs were predicted using HELIXTURNHELIX (Dodd & Egan, 1990). Transmembrane regions were predicted using the TopPRED2 (Von Heijne, 1992) and DAS (Cserzo *et al.*, 1997) programs at IBCP, Lyon France.

3.10 SOUTHERN HYBRIDISATION

Molecular hybridisation is the formation of double-stranded nucleic acid molecules by sequence-specific base pairing of complementary single strands. Southern hybridisation provides a means for detecting DNA fragments that are complementary to a labelled DNA probe (Southern, 1975). The fragmented DNA is firstly separated according to size by agarose gel electrophoresis then transferred to a membrane support for ease of handling.

3.10.1 Capillary transfer of DNA

Genomic DNA (3 μg) was obtained as described in section 3.4.1 from the appropriate bacterial strains. The DNA was digested to completion with *Eco*RI (section 3.4.5.1) and separated by agarose gel electrophoresis (3.4.6.1) along with bacteriophage λ DNA digested with *Hind*III and labelled with DIG-dUTP (Roche, Lewes, UK). The gel was electrophoresed until the bromophenol blue marker dye reached the bottom of the gel. Following electrophoresis, the agarose gel was trimmed to remove any excess, depurinated by washing in 0.25 M HCl for 10 min, and denaturated by soaking the gel in denaturation

solution (Appendix III) for 30 min with gentle agitation on a rotary shaker. The gel was then neutralized by immersion in 200 ml of neutralization solution (Appendix III) for 30 min at RT with constant shaking.

A Southern blot capillary transfer apparatus was used (Scotlab, UK). The reservoir was filled with 10X SSC transfer buffer. The agarose gel was laid on the wick using cling film to cover the remaining exposed wick and prevent transfer buffer from "short circuiting" the system. Positively charged nylon membrane (Amersham, UK) cut to fit the gel was pre-wetted in distilled water and positioned carefully over the gel. Three pieces of pre-wetted 3MM filter paper cut to the same size as the gel were laid on top. Air bubbles were removed as each layer was added by gently rolling a glass rod over the surface. A stack of paper towels, which serves to draw the transfer buffer through the gel by capillary action, was laid on top of the 3MM sheets followed by the blotter lid. Following overnight transfer of the DNA the nylon membrane was removed, marked for orientation and washed in 5X SSC for 5 min at RT. The filter was dried at RT and either stored at 4°C or processed for hybridisation.

3.10.2 Labelling of DNA fragments with digoxigenin-dUTP (DIG)

A commercial non-radioactive DNA labelling and detection kit (Boehringer-Mannheium, Germany) was used according to the manufacturer instructions. The method relies upon the incorporation of digoxigenin-labelled deoxyuridine triphosphate (DIG-dUTP) into a DNA probe using Klenow fragment of DNA polymerase and random primers, which generates a range of digoxigenin-labelled DNA fragments from the template DNA or using Taq polymerase and specific primers in the polymerase chain reaction (PCR).

3.10.2.1 Randomly primed labelling

The randomly primed DNA labelling reaction is fast (1 h) and results in the incorporation of one digoxigenin molecule about every 20-25 nucleotides, depending on

base content, in the newly synthesized DNA probe.

DNA template (1 μg) was denatured in a boiling water bath for 10 min then immediately chilled on ice. A solution containing 10 X hexanucleotide primer (2 μl), 10 X dNTP labelling mixture (3 μl) (dATP, dGTP, dCTP and Digoxigenin-11-dUTP) and 1 μl of Klenow enzyme (1 U/μl) was mixed with the denatured DNA template and 'Analar' water to a final volume 20 μl. The reaction tube was incubated at 37°C for 1 h or preferably overnight. The Klenow fragment was inactivated with 2 μl 0.2 M EDTA pH 8.0 and the DNA precipitated with ethanol (3.4.3.1), dried and resuspended in 30 μl of H_2O. Alternatively, the DNA probe was purified using the Prep-A-Gene kit (section 3.4.4.3). Finally, the labelled probe was stored at -20°C until required.

3.10.2.2 PCR labelling

The PCR-labelling reactions were carried out using standard reaction conditions (**Table 3**, Appendix III) in a 50 μl volume containing 5 nmol DIG-dUTP. Reactions mixtures contained in thin-walled 0.5 ml tubes were overlaid with 50 μl mineral oil (Sigma, Poole, UK) and prepared on ice. After denaturation of the template DNA (96°C, 10 min) the amplification procedure was carried out as described in program 1 (section 3.13.2). Incorporation of DIG-dUTP into the amplicon was confirmed by agarose gel (1%) electrophoresis (3.4.6.1). The presence of DIG-dUTP in a PCR amplicon retards the migration of the product during electrophoresis when compared with non-labelled amplicon.

3.10.2.3 Purification of DIG-Labelled DNA

All DIG-labelled DNA probes were purified using the Prep-A-Gene kit (Bio Rad, UK) as described in section 3.4.4.3.

3.10.3 Hybridisation of DIG-labelled probes to nucleic acids

Prior to hybridisation, nylon membranes were exposed to UV radiation (254 nm) for 2 min to permanently bind the DNA to the nylon filters. All of the following steps were

carried out in sealed glass tubes of 20 cm length and 8 cm diameter using a hybridisation oven (Techne, USA). Hybridisation was performed at high stringency (68°C) in order to minimise non-specific binding of the probe to the membrane. Filters were prehybridised for 2 h at 68°C in prehybridisation solution (Appendix III). DIG-labelled probe was denatured by boiling for 10 min, chilled on ice, and 5 µl (100-500 ng) of denatured probe was added to 20 ml of hybridisation solution and incubated with the filter overnight at 68°C. Following hybridisation, membranes were washed two times in wash buffer 1 (Appendix III) for 5 min at RT and two times in wash buffer 2 (Appendix III) for 15 min at 68°C. The filters were air dried at RT.

3.10.4 Detection of DNA-DNA hybrids

All the following steps were performed at room temperature. Nylon filters were washed briefly in detection buffer 1 (Appendix III) and incubated for 30 min in 100 ml blocking solution (Appendix III) then probed with 10 µl of anti-digoxigenin antibody conjugated to alkaline phosphatase in 20 ml of detection buffer 1 for 30 min. Unbound antibody was removed by washing 2 x 15 min with 100 ml of buffer 1. Filters were equilibrated with 20 ml of buffer 3 (Appendix III) for 5 min, then incubated in the dark with 10 ml of substrate solution (45 µl NBT-solution and 35 µl X-phosphate-solution in 10 ml of buffer 3). Following colour development, the reaction was stopped by washing filters with 50 ml of detection buffer 4 (Appendix III) for 5 min.

3.11 CONSTRUCTION OF A *B. PSEUDOMALLEI* GENOMIC LIBRARY IN BACTERIOPHAGE λGEM-11

Genes were cloned from the genome of *B. pseudomallei*, following the construction of genomic libraries. The library should represent the entire genome of an organism as a set of overlapping cloned fragments, produced in a random manner, and maintained in a stable

form with no misrepresentation of sequences. Most genomic libraries are constructed using a method called shotgun cloning where genomic DNA digests are ligated to a suitable cloning vector that has been cut with the same enzyme or an enzyme that generates compatible ends. These recombinant DNA molecules can then be propagated in a suitable host organism.

The number of clones required depends on the size of the genome and the vector used and the size of a library can be calculated on the basis of the probability of a particular sequence being represented in the library. The number of clones (N) required for a specific genomic library can be estimated using the formula of Clarke & Carbon (1976): $N = \ln (1 - P) / \ln (1- a/b)$ where P is the desired probability of a particular sequence being represented (typically 0.95 to 0.99), *a* is the average size of the DNA fragments to be cloned and *b* is the size of the genome, expressed in the same units as *a*.

The size of the *B. pseudomallei* genome has been recently determined (Songsivilai & Dharakul, 2000). *B. pseudomallei* has two chromosomes of approximately 3.56 and 2.97 Mb in size, giving a total genome size of about 6.54 million base-pairs. Replacement bacteriophage λ vectors such as λGEM-11 (**Fig 3.2**) in which about 20 kb of the phage genome (stuffer fragment) can be substituted with cloned DNA. A genomic library was constructed by ligating the products of a partial *Sau*3A restriction digest of *B. pseudomallei* chromosomal DNA with purified λ GEM-11 left and right arms (Promega, UK). Assuming average of 15 kb inserts, a collection of at least 1400 recombinant molecules must be examined to find a particular recombinant. However, since recombinants are created randomly and there is a chance that a particular insert will not be included in the putative genomic representatives, it is desirable to examine double that number.

3.11.1 *Sau3A* partial digestion of chromosomal DNA

In order to generate *B. pseudomallei* DNA fragments with a suitable range of sizes for constructing a genomic library in λGEM-11 constant amounts of DNA were digested with

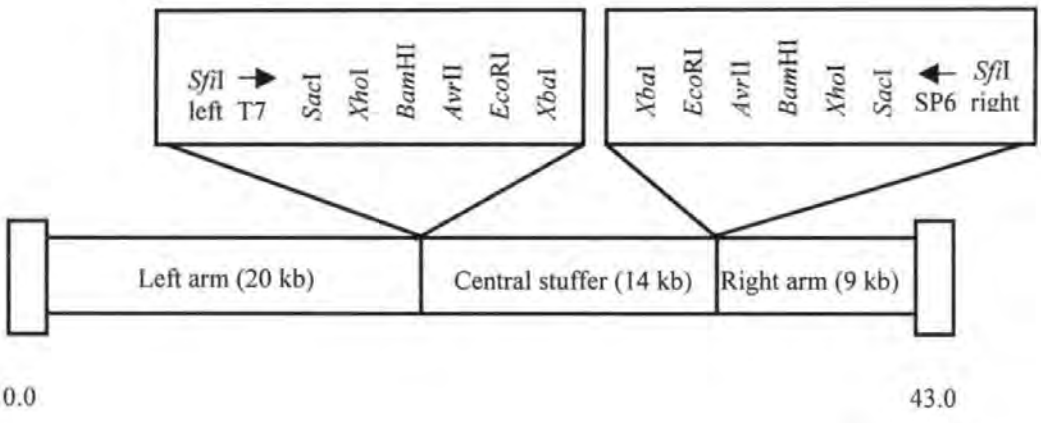


Fig 3.2: Schematic presentation of the λGEM-11 cloning vector (Stratagene).

varying amounts of *Sau*3A enzyme for a fixed incubation period. Genomic DNA (18 μg) was mixed with 10 X restriction enzyme buffer (15 μl) and 'Analar' water to a final volume of 150 μl. Nine tubes were prepared. Then 30 μl of this mixture was dispensed into tube 1 and 15 μl into each of the remaining tubes. *Sau*3A enzyme (5 U) was added to tube 1, mixed thoroughly, a 15 μl volume transferred to tube 2, mixed well and the twofold serial dilution was continued through to tube 8. Tube 9 was the undigested control. All of the tubes were incubated at 37°C for 30 min. After incubation, the reactions were stopped by adding 3 μl 0.1 M EDTA pH 7.5 to each tube or by heat inactivation at 65°C for 10 min. The *Sau*3A digested DNA (2 μl) was analysed by agarose gel electrophoresis (section 3.4.6.1) to determine which tube contained the desired fragment size distribution. The reaction that had a distribution of DNA fragments within the range of approximately 9-23 kb were selected for the construction of a λGEM-11 library. The remaining DNA (13 μl) was purified using phenol/chloroform (section 3.4.4.2).

3.11.2 Ligation of chromosomal fragments with λ arms

The preferred substrate for *in vitro* packaging is a concatamer, an end to end chain in which recombinant molecules are flanked on both sides by a *cos* site, and to favour concatamer formation, the total DNA concentration should exceed 100 μg/ml. For a ligation reaction, a molar ratio of 2:1 vector to insert is a good compromise. The concentration of the *Sau*3A digested *B. pseudomallei* DNA was determined by measuring the absorbance at 260 nm (section 3.4.6.1). Approximately 1 μg of this DNA was mixed with 2 μg of cleaved, dephosphorylated λ GEM-11 *Bam*HI vector arms (Promega) in the presence of 2 units of T4 ligase, 1 μl of 10 X ligase buffer and 'Analar' H_2O to a final volume of 10 μl. The ligation mixture was incubated for 2 h at RT, then overnight at 4°C.

3.11.3 *In vitro* packaging of ligated DNA

In vitro packaging of recombinant phage particles was carried out using packaging extracts (Promega), that contain all of the structural proteins necessary for the assembly of lambda DNA into infective phage particles.

The packaging extracts (50 µl) were thawed on ice. The ligation mixture (10 µl) was added to the extract, mixed by gently tapping the bottom of the tube several times, and incubated at RT for 3 h. Following incubation, 445 µl of SM phage buffer (Appendix III) and 25 µl of chloroform were added to the tube, mixed by inversion, and left to stand for 2 min to allow the chloroform to settle to the bottom of the tube; chloroform prevents bacterial growth. The phage genomic library was stored at 4°C until used; the phage particles normally remain infective for years.

3.11.4 Titration of packaged phage DNA

The day before titration of packaged phage DNA, a single colony of *E. coli* LE392 was picked from a freshly streaked NZCYM agar (Appendix III) plate, and used to inoculate a 10 ml NZCYM broth (Appendix III) supplemented with 0.2% maltose. The presence of maltose acts to induce the expression of *malB* that encodes a phage λ receptor and thereby promotes efficient adsorption of the phage to the *E. coli* host (Sambrook et al., 1989). The culture was incubated in a shaking water bath at 37°C, 250 rpm, until OD_{600} reached 0.6-0.8.

To investigate the bacteriophage λ titre, i.e. the number of plaque forming units per ml, a ten-fold dilution series of the packaging mix from 10^{-1} to 10^{-6} was made in SM buffer. In duplicate sterile screw cap glass bottles, 100 µl of each dilution was added to 200 µl of the prepared *E. coli* LE392 culture, mixed gently and incubated at 37°C for 20 min to allow the adsorption of the phage. Soft molten NZCYM agarose (Appendix III) (3 ml) was added, mixed by gentle swirling, and then poured immediately to overlay pre-warmed NZCYM agar supplemented with 0.2% maltose by tilting the plates back and forth. Bubbles, if any,

were quickly removed by poking with a sharp, sterile needle. When the agarose had set, the plates were inverted and incubated at 37°C overnight. The titre of the phage particles was determined by counting the number of plaques or zones of lysis on the bacterial cell lawn, for each dilution. The number of plaque-forming units (pfu/ml) was then established, according to the formula:

Titre (pfu/ml) = number of plaques (pfu) x dilution factor x 1000 / volume plated (μl)

3.11.5 Purification of recombinant bacteriophage

3.11.5.1 Picking bacteriophage λ plaques

Recombinant phage particles were selected from lysis plates using a sterile Pasteur pipette to punch out a plug of agar containing the desired plaque. The plug was then ejected into a sterile microfuge tube containing 50 μl of SM buffer and 5 μl chloroform. This suspension was vortexed briefly, left at RT for 1-2 h to allow the bacteriophage particles to diffuse out of the agar and was stored at 4°C. In addition, a number of non-recombinant phage plaques were picked and processed in this manner to serve as negative controls in further experiments.

3.11.5.2 Preparation of plate lysate stocks from single plaques

Purified phage was isolated from recombinant and non-recombinant phage by the plate lysate method. For maximum yield, the phage titre should be adjusted to give semi-confluent lysis after approximately 12 hours of incubation. Using the appropriate phage dilution for semi-confluent lysis determined in section (3.11.4), multiple lysis plates were prepared as described (section 3.11.4). Following complete (confluent) lysis of the bacterial lawn each plate was overlaid with 2 ml of SM buffer and incubated at RT for 2 h with frequent agitation to allow the bacteriophage particles to elute. With the help of a sterile bent glass rod (spreader), the top agarose was scraped and transferred to sterile 50 ml centrifuge tube. The tube was gently shaken at RT for 30 min, then centrifuged (9,000 x g, 10 min, 4°C). The

cleared supernatant was carefully pooled, transferred to a fresh tube and 0.3% (v/v) chloroform was added. The lysate was stored at 4°C for up to 6 months.

3.11.6 Isolation of phage λ DNA from recombinant clones

3.11.6.1 Small scale preparation using the Wizard system

The Wizard λ prep DNA purification kit (Promega, Madison, USA) was used according to the protocol supplied by the manufacturer. To a sterile polypropylene tube containing a 10 ml volume of bacteriophage lysate (section 3.11.5.2), 40 µl of resuspended nuclease solution was added and the mixture was incubated at 37°C for 15 min. Phage precipitant (4 ml) was added with mixing and the tube was left at RT for 30 min. After centrifugation (10,000 x g, 10 min), the supernatant was discarded and the pellet was resuspended in 500 µl of phage buffer and transferred to a fresh 1.5 ml microfuge tube. Insoluble particles were removed by centrifugation (10,000 x g, 10 s) and the supernatant transferred to a new microfuge tube. Thoroughly mixed purification resin (1 ml) was added to the supernatant and mixed by inverting the tube several times. The phage DNA was then purified using Wizard minicolumns attached to a vacuum manifold. The resin/phage DNA was washed with 2 ml of 80% isopropanol and dried briefly. The minicolumn was then removed, placed into a fresh sterile 1.5 ml microfuge tube and centrifuged (10,000 x g, 2 min, RT) to remove any residual isopropanol. Phage DNA was eluted from the minicolumn by adding 100 µl of pre-heated (80°C) TE buffer and immediately centrifuging (10,000 x g, 20 s, RT). A second elution step was performed, to obtain maximal recovery with 50 µl of preheated TE buffer and the purified phage DNA was stored at 4°C.

3.11.6.2 Small scale preparation using the Bio-Rad system

To 0.8 ml of a bacteriophage plate lysate, prepared as described in section 3.11.5.2, 1 µl of freshly prepared DNase (10 mg/ml) was added and the mixture was incubated for 30 min at RT. To a sterile polypropylene tube containing 400 µl of DNA binding buffer, the lysate

was added, mixed by inversion and left at RT for 5 min, to lyse the phage particles. A volume of 30 µl of Prep-A-Gene diatomaceous matrix was added, and the tube was vortexed, left at RT for 5 min and the phage DNA contained in the mixture was purified using a spin filter. The matrix was washed twice in wash buffer, centrifuged (13,000 x g, 30 s, RT) and the phage DNA was eluted into a clean tube with 100 µl of 'Analar' H_2O or TE prewarmed to 50°C (13,000 x g, 1 min).

3.11.6.3 Large scale preparation of phage λ from liquid cultures

This method affords a high titre stock of λ bacteriophage for subsequent use. To ensure success it is necessary to strictly control the bacterial cell density and the ratio of phage to bacterial cells, also called the multiplicity of exposure (MOE). After infection of a small bacterial culture with λ bacteriophage at low MOE, a large volume of medium is inoculated with the infected culture. Initially the concentration of bacteriophage is low and uninfected cells continue to multiply rapidly for several hours until all of the bacterial cells are infected by phage.

A 20 ml volume of NZCYM broth, supplemented with 0.2% maltose, was inoculated with a single colony of *E. coli* LE392, and grown at 37°C with shaking (300 rpm) until OD_{600} ~ 0.6 (equivalent to ~ 4.8 X 10^8 cells/ml). The MOE was calculated based on the known bacteriophage titre (section 3.11.4). Different MOEs were chosen; 0.1, 1, 10, and 100, where MOE 1 represents a 1:1 ratio of phage: bacterial cells. Phage propagation medium (Appendix III), 500 ml volumes in 2 L flasks, was warmed to 37°C prior to the addition of different ratios of bacteriophage: bacterial cells. The flasks were incubated at 37°C with shaking for about 12 h. By that time successful lysis had occurred in one or more of the flasks which could be seen as clear or slightly translucent in appearance, with strings of cell debris around the bottom of the flask. By using a range of MOEs, at least one of the flasks showed signs of complete lysis and could be processed further. The contents of the

flask(s) were harvested and 0.25ml of chloroform per 500 ml culture was added to lyse any remaining bacterial cells. The lysate was left for 30 min at 37°C, in a shaking water bath to ensure thorough mixing.

3.11.6.4 Polyethylene glycol precipitation of bacteriophage λ

The phage lysate prepared in section 3.11.6.3 was cooled to RT and DNase I and RNase A were added to a final concentration of 1 μg/ml each. The lysate was centrifuged (10,000 x g, 10 min) to remove cell debris. The supernatant was decanted into a 1 L sterile measuring cylinder and NaCl added (0.5 M final concentration) and 10% (w/v) polyethylene glycol (PEG 6000). The top of the measuring cylinder was covered with parafilm and the contents were shaken gently at RT until the PEG had dissolved. High salt concentration promotes the dissociation of the bacteriophage particles from the bacterial debris. The mixture was left to stand at 4°C overnight, to allow the bacteriophage particles to precipitate. Precipitated phage was recovered by centrifugation (8,000 x g, 30 min, 4°C) and the phage pellet was resuspended in 30 ml SM buffer and kept at 4°C. The suspension was transferred into a 40 ml Beckman Quickseal tube, heat sealed, and centrifuged (18,000 x g, 16 h, 20°C). The bacteriophage pellet was resuspended in 0.5 ml SM buffer and stored at 4°C.

3.11.6.5 Extraction of λ DNA from high titre phage stocks

To a 100 μl volume of a high titre λ stock in a 1.5 ml microfuge tube, 10 μl of 2M Tris-HCl-0.2M Na_2EDTA (pH 8.5) solution was added and mixed thoroughly. To this an equal volume of formamide (110 μl) was added, and the mixture was left to stand at RT for 1 h before adding an equal volume of double distilled water (220 μl). The phage DNA was precipitated with ethanol (section 3.4.3.1), resuspended in 100 μl of TE buffer and purified by phenol/chloroform extraction (section 3.4.4.2).

3.12 CONSTRUCTION OF A *B. PSEUDOMALLEI* GENOMIC DNA LIBRARY IN λ ZAP EXPRESS

The λ ZAP express vector (38.9 kb) is a phagemid that contains within it a region flanking the multiple cloning site that can be excised *in vivo* and recovered as a recombinant plasmid. A diagramatic representation of the λ ZAP Express cloning vector is shown in **Fig 3.3**. The number of recombinants required for complete coverage of the *B. pseudomallei* genome in the λZAP Express vector, assuming 10 kb inserts, was estimated (section 3.11) to be a minimum of 2000 plaques.

3.12.1 Digestion of *B. pseudomallei* genomic DNA

The restriction enzyme *BclI* was chosen for the digestion of *B. pseudomallei* DNA. The procedure was performed as in section 3.4.5.1, but for optimal enzyme activity the reaction was incubated at 50°C overnight.

3.12.2 Ligation of chromosomal fragments with λ ZAP arms

Lambda ZAP express vector pre-digested with *Bam*HI and CIAP treated was supplied by Stratagene (Amsterdam, Netherlands). The vector accommodates inserts ranging in size up to 12 kb; insertion into the polylinker inactivates the *lacZ* gene carried by the vector and recombinants give clear rather than blue plaques on IPTG-X-gal agar. A ligation reaction between the vector and insert DNA was performed according to the manufacturer's manual. A volume of 1 µl of the ZAP express vector (1 µg/µl) was mixed with 1 µl of *BclI* digested *B. pseudomallei* genomic DNA (0.4 µg) in the presence of 2 µl T4 DNA ligase (2 U) and 0.5 µl 10 X ligase buffer. The mixture was centrifuged briefly then incubated at 4°C overnight. The sample was heated (68°C, 10 min) to inactivate the ligase. Finally, the ligation mixture was purified as described in section 3.4.4.3, and the recombinant DNA was ready for packaging into phage particles.

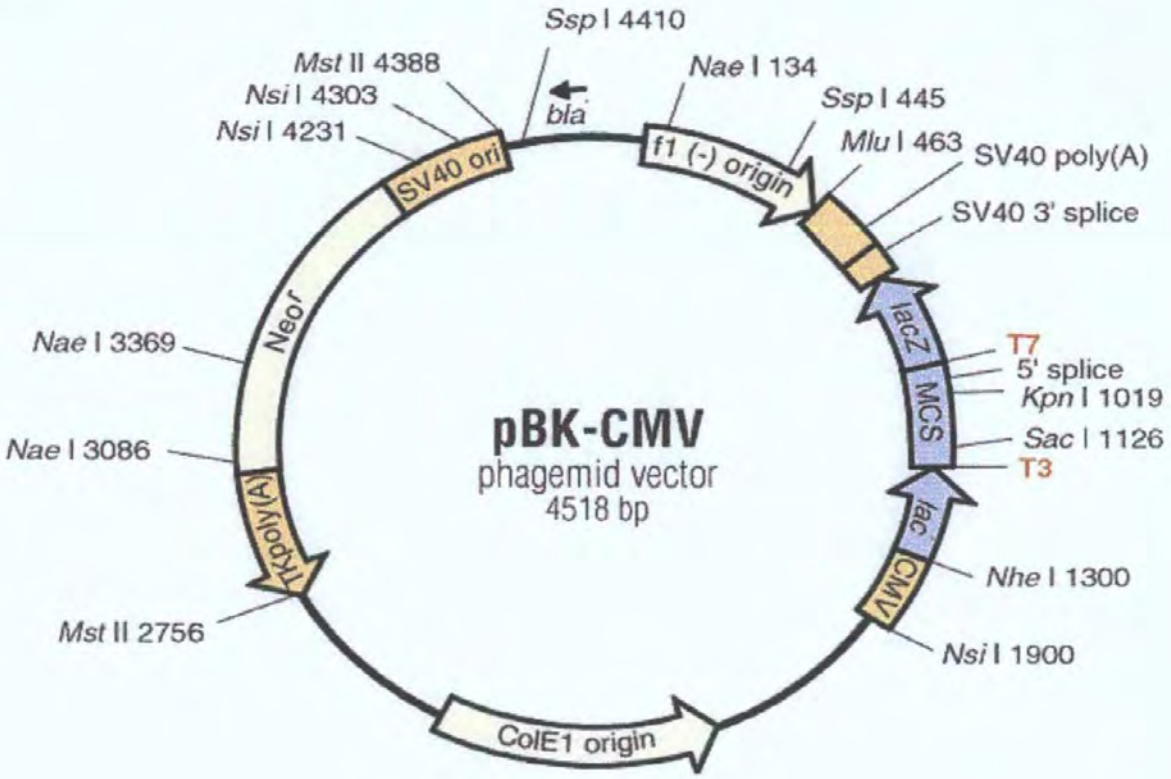


Fig 3.3: Physical map of the pBK-CMV cloning vector (Stratagene)

3.12.3 Preparation of *E. coli* XL1-Blue

E. coli XL1-Blue was cultured in 10 ml NZCYM broth supplemented with 0.2% maltose, at 37°C for 6 h (OD_{600} ~ 0.8-1). The culture was centrifuged (1,000 x g, 10 min) and the cell pellet was resuspended and diluted to an OD_{600} of 0.5 using 10 mM $MgSO_4$.

3.12.4 Packaging λ ZAP Express using Gigapack III Gold

Packaging extracts were thawed on ice. Ligated DNA (4 µl) (section 3.12.2) was immediately added, mixed with a pipette tip and the tube was centrifuged briefly. Following incubation at RT for 2 h, 500 µl of SM buffer was added and mixed, then 20 µl of chloroform was gently mixed in and the tube was centrifuged briefly to sediment the debris. The supernatant containing the phage was either used immediately or stored at 4°C until required.

3.12.5 Checking recombination efficiency using blue-white colour selection

Recombinant λZAP phage were recovered on NZCYM agar plates as follows. Firstly 1 µl of the packaged reaction was mixed with 200 µl of *E. coli* XL1-Blue culture prepared as described in section 3.12.3. The mixture was incubated at 37°C for 15 min, to allow the phage to adsorb then 3 ml of soft NZCYM agarose containing 15 µl of 0.5 M IPTG and 50 µl of X-gal (250 mg/ml) was added, mixed vigorously and overlayed on pre-warmed (37°C) NZCYM agar. The agar was allowed to set (10 min), then the plates were inverted and incubated at 37°C overnight. The plates were examined visually for the presence of recombinant phage plaques (clear) and non-recombinant phage plaques (blue colour). The recombination efficiency of the method was determined by comparing the relative numbers of blue and clear plaques.

3.12.6 Titration of the λZAP Express genomic library

E. coli XL1-Blue was cultured in a 10 ml volume of LB broth supplemented with 0.2%

maltose and 10 mM $MgSO_4$ at 30°C overnight with shaking. A 100 µl volume of this culture was used to inoculate a fresh 10 ml LB broth, supplemented with 0.2% maltose, that was incubated for about 3 h at 37°C with shaking until OD_{600} ~ 1.0. The culture was then diluted to OD_{600} ~ 0.5 with fresh LB +maltose broth. Ten fold serial dilutions of phage in SM buffer were prepared (10^{-2} to 10^{-6}) and 1 µl of each dilution was added to separate tubes containing 200 µl of the prepared *E. coli* XL1-Blue culture. The tubes were placed at 37°C for 15 min, then 3 ml of soft NZCYM agarose (50°C) was added to each tube, mixed vigorously and overlaid immediately on NZCYM agar. The agar was allowed to set, and the plates were inverted and incubated at 37°C overnight. The number of plaques was counted and the phage titre was calculated according to the dilution factor.

3.12.7 *In vivo* excision of the pBK-CMV plasmid vector

The λZAP express vector is designed to allow simple and efficient *in vivo* excision and recircularization of a phagemid contained within the λZAP vector to form a plasmid, pBK-CMV containing the cloned insert. The *in vivo* excision requires the use of a helper phage and compatible strain of *E. coli*, XLOLR.

Phage plaques were selectively picked from lysis plates and transferred to sterile microfuge tubes containing 500 µl of SM buffer and 20 µl of chloroform. The tube was vortexed, to release phage particles from the agar plug, and stored at 4°C overnight. To excise pBK-CMV, 200 µl of XL1-Blue culture (OD_{600} 1.0- section 3.12.3) were combined with 250 µl of the λ ZAP phage stock and 1 µl ExAssistant helper phage. After incubation at 37°C for 15 min, 3 ml of NZCYM broth were added and the incubation was continued for 3 h at 37°C with shaking. The tube was heated at 70°C for 20 min to inactivate the parent λ phage and kill host bacteria, then centrifuged (2,000 x g, 15 min). The supernatant contained the pBK-CMV phagemid packaged as single stranded filamentous phage particles. To

recover the excised phagemids, 10 µl of the supernatant was added to 200 µl of a fresh culture of *E. coli* XLOLR cells (OD_{600} ~1.0) and incubated at 37°C for 15 min. To the infected cells, 300 µl of NZCYM broth was added and the mixture was incubated at 37°C for 45 min. Finally, 200 µl of this culture was spread onto LB agar supplemented with kanamycin (50 µg/ml) and incubated at 37°C overnight. The resultant colonies, containing the pBK-CMV plasmid, were screened for the presence of recombinant insert DNA (section 3.4.2.3).

3.13 AMPLIFICATION OF DNA BY POLYMERASE CHAIN REACTION (PCR)

3.13.1 Primer design

A variety of PCR Primers were designed using DNASIS sequence analysis software (Hitachi, UK) and the most suitable were chosen from among them by using Amplify Software (Engels, 1993). In some cases primers were designed to include restriction sites to simplify cloning of the PCR products. A number of factors should be considered when designing PCR primers in order to improve the chances of a favourable outcome. Ideally, primers should be approximately 10 to 30 base pairs (bp), with a random base distribution, a GC content of 50-60%, minimal self-complementarity, and minimal homology with any other region of the DNA template. Oligonucleotide primers were synthesised by Sigma-Genosys (Cambridgeshire, UK). Once synthesised, primers were resuspended in TE buffer (Sigma, Poole, UK) to give an overall concentration of 20 pmol/µl, and stored at -20°C.

3.13.2 Preparation of PCR reactions

The use of PCR to amplify a DNA fragment using specific DNA oligonucleotides was performed using a Perkin Elmer TC1 DNA thermal cycler (Roche, Lewes, UK). Precautions include using a set of micropipettes dedicated only to PCR and using filter tips, to avoid DNA contamination and cross-contamination. All pipettes, tubes and racks were exposed to

UV light for 20 min prior to set up, in a cabinet using filtered air. The reaction components were assembled on ice in sterile 0.5 ml thin walled Eppendorf tubes and mixed by vortexing. Each 50 µl reaction mixture was overlaid with 50 µl of mineral oil (Sigma, Poole, UK). Additionally, each set of reactions incorporated a negative control with the DNA template replaced with HPLC water. The reaction mixture is as shown in **Table 3**, Appendix III. Program 1 was used for the amplification of fragments less than 1 kb in size, while program 2 was used to amplify fragments longer than 1 kb (**Table 3.5**). Unless otherwise stated, all PCR reactions were carried out for 35 cycles of amplification following an initial 2 min denaturation at 96°C.

Table 3.5: General cycling conditions for PCR

Step	Temperature	Program 1 Time (s)	Program 2 Time (s)
Denaturation	96°C	30	60
Annealing	60-65°C	30	60
Elongation	72°C	90	120

After amplification, samples were stored at 4°C until the products were analysed by gel electrophoresis as described in section 3.4.6.1.

3.14 SDS-PAGE AND WESTERN BLOTTING

3.14.1 Preparation of bacterial cells

B. pseudomallei cultures were grown at 25°C, 37°C and 42°C for 48 h, then centrifuged (4,000 x g, 15 min, RT). The cell pellet was resuspended in 50 µl 2X SDS-PAGE buffer without bromophenol-blue (Appendix III) and boiled, to destroy proteolytic activity, aid in cell breakage, protein solubilization and SDS binding. The suspension containing the soluble whole-cell proteins was then stored at -20° C until required.

3.14.2 Determination of protein concentration

A microassay procedure based on the method of Bradford (1976) was used to determine the concentration of proteins in bacterial cell lysates using the Bio-Rad Protein Assay Kit (Bio-Rad, UK). A series of standard protein solutions were made using bovine serum albumin (BSA) dissolved in water in the range 0-20 μg/ml. Bacterial cell lysates were diluted 100 times with modified sample buffer (that lacks bromophenol-blue). A volume of 40 μl of each of the diluted bacterial cell lysate and BSA standards was mixed with 160 μl of dye reagent concentrate in separate wells of a microtitre plate. After a period of 5 min to one hour, the OD of the samples was read at 595 nm. The BSA readings were used to construct a standard curve. Unknown protein concentrations of samples were automatically determined by OPTI Analyzer microplate reader (Molecular Devices, USA) and the results were expressed as μg protein/ml.

3.14.3 SDS-polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a widely used analytical technique for the separation and characterization of complex mixtures of proteins and the estimation of the relative molecular size of a protein. SDS-PAGE was carried out using the Laemmli discontinuous buffer system (Laemmli, 1970). Two types of electrophoretic apparatus used in this study; the Mini-Protean II system (Bio-Rad, UK) and the SE600 vertical system (Amersham Pharmacia Biotech, UK). The apparatus was assembled according to the manufacturer's instructions. The glass plates were cleaned with detergent and washed with ethanol prior to use. For routine work, a 12% acrylamide resolving gel with a 5% acrylamide stacking gel (Appendix III) was prepared. Polymerisation of acrylamide gels was achieved using an ammonium persulphate (APS)/TEMED catalyst. Following polymerisation the well-forming comb was carefully removed and the sample wells were rinsed with dH_2O and running buffer (Appendix III)

before the gel assembly was clamped into the electrophoresis chamber, which was filled with running buffer. Protein samples were made up using either 2 X stock sample buffer or modified sample buffer (Appendix III). Samples were loaded into each well along with either standard molecular weight markers (Sigma, Poole, UK), or pre-stained molecular weight markers (Novex, USA) and electrophoresis was carried out at constant voltage. Minigels were run for 2-3 h at 120 V, without cooling, while large gels were run for 7-8 h at 80 V, with cooling. When the marker dye had reached the bottom of the gel, the electrophoresis apparatus was switched off, disassembled and the gels were either stained for total protein or processed for Western blotting.

3.14.4 Staining of total proteins

The most common method for visualising protein bands in SDS-PAGE gels is staining with Coomassie blue R-250. Gels were stained for 1 h at RT in Coomassie stain (Appendix III) in a covered plastic tray, and destained in destaining solution overnight, with gentle agitation. The wet gel was photographed under white light.

3.14.5 Western blotting

Proteins that had been separated by SDS-PAGE were electrophoretically transferred to an immobilizing membrane, consisting of nitrocellulose to facilitate immunological detection (Towbin *et al.*, 1979). SDS-PAGE gels were equilibrated in transfer buffer (Appendix III) for 10 min. A gel sandwich was then assembled, consisting of a sponge, two sheets of 3MM-Whatman paper, the SDS-PAGE gel, nitrocellulose membrane (pore size 0.2 μm Schleicher & Schuell), and a further two sheets of 3MM paper and a sponge. Air bubbles were removed by rolling the layers with a glass rod and the whole sandwich was placed into the blotting apparatus (Bio-Rad Transblot System, Bio-Rad, UK) with the nitrocellulose filter towards the positive electrode. The tank was filled with transfer buffer and electrotransfer was carried out using a constant current of 30 mA for 12-16 h. Following

transfer, the apparatus was switched off and disassembled. The efficiency of transfer was visually determined by the migration of protein molecular weight markers into the membrane. The nitrocellulose membrane was washed in PBS for 5 min and placed in 1% casein blocking solution (Appendix III) at RT for 2 h with gentle shaking, to block the free protein binding sites on the membrane. The nitrocellulose membrane was probed with specific primary antibody diluted in blocking solution and incubated overnight at 4°C. All steps were carried out at RT on an orbital shaker. Any unbound antibody was removed by rinsing the membrane in blocking solution 3 times, each for 10 min. The bound antibody was detected with a secondary antibody (usually, peroxidase-conjugated swine anti-rabbit immunoglobulin, P217, DAKO, UK) at the appropriate concentration for 2 h. The membrane was then washed three times in PBS and the bound antibody was visualised by immersing the blot in development solution (Appendix III). Following development the blot was rinsed thoroughly with PBS, dried between sheets of 3MM-Whatman paper and photographed.

CHAPTER 4

CLONING AND SEQUENCE ANALYSIS OF A TWO-COMPONENT REGULATORY SYSTEM

4.1 INTRODUCTION

Most, if not all, bacteria possess sophisticated signal transduction systems that involve at least two protein components i.e. two-component regulatory systems. This family of proteins allows bacteria to sense and respond to numerous changes in their environment and it was originally defined on the basis of protein sequence similarities between pairs of "sensor" and "regulator" proteins. Generally, the genes encoding the sensor and effector pair are located near each other, often in the same operon. The features of two-component systems have been thoroughly reviewed in Chapter 2. Currently, over 100 examples of these protein systems have been recognised, and members of the two-component systems have been cloned and characterised in more than 50 different bacterial species (Parkinson & Kofoid, 1992). The widespread prevalence of two-component systems implies that they provide an essential means for mediating a wide variety of adaptive responses by bacterial cells to environmental stimuli (Parkinson, 1993), including the establishment and maintenance of infectious states in host organisms (Egger *et al.*, 1997). This chapter identifies and characterises, on the basis of sequence homology, the genes encoding a two-component regulatory system from *B. pseudomallei*.

B. pseudomallei is able to survive and grow in a wide range of environmental conditions, varying from the rice paddies of tropical regions to temperate regions of Australia and also within various mammalian hosts (reviewed in sections 2.1.3-2.1.5). The ability to adapt and respond to such varied circumstances implies a requirement for signal transduction systems that control the expression of the genes that are needed to meet these challenges. Although two-component regulatory systems are commonly found in many bacterial pathogens, they have only recently been described in *B. pseudomallei* by Jones *et al.* (1997) who reported the

identification of the *irl* genes, *irlS* similar to sensor kinases and *irlR* similar to response regulators. On the basis of sequence homology it was proposed that the *irl* genes may be involved in the resistance of *B. pseudomallei* to heavy-metals.

The primary aim of this work was to identify and characterise the genes encoding an additional two-component regulatory system from *B. pseudomallei*. The first step in this process was the development of a strategy for cloning the genes in the absence of a readily selectable phenotype. Generally, the lack of an obvious phenotype can be circumvented by screening genomic libraries with DNA probes or with specific antibodies. Because specific antibodies were unavailable an oligonucleotide probe was developed using PCR and primers based on the highly conserved sequences of regulators from other bacteria. Hence, the approach may be summarised as follows:

1- Development of an oligonucleotide probe using PCR.

2- Construction and screening of a genomic library using λ GEM-11 bacteriophage.

3- Screening of a genomic library using λGT-11 bacteriophage.

4- Construction and screening of a genomic library using λ ZAP, on the basis of the information obtained from 1-3 above.

4.2 MATERIALS AND METHODS

4.2.1 Construction and screening of a λGEM-11 genomic library

4.2.1.1 Library construction

The construction of a *B. pseudomallei* strain 204 genomic library in λGEM-11 bacteriophage was performed as outlined in section 3.11. *B. pseudomallei* DNA was obtained as described (section 3.4.1) and a partial *Sau*3A digest (section 3.11.1) was used in the preparation of the library, which enabled insertion into the λGEM-11 *Bam*HI site. A size-selection

procedure was followed to isolate DNA fragments sizes suitable for insertion into the vector arms, between approximately 9 and 23 kb. The construct was then transduced into *E. coli* LE392 strain (**Table 3.3**) and plated on NZCYM-agarose (Appendix III) containing 0.2% maltose.

4.2.1.2 Development of a DNA probe

As a part of a broad search for genes in *B. pseudomallei* that might encode products of potential interest for further study, the routine sequencing of a fragment of *B. pseudomallei* DNA revealed a striking similarity with the response regulatory genes from *E. coli* and *S. typhi*. Oligonucleotide PCR primers were designed on the basis of this sequence data in order to amplify a 375 bp internal fragment of the response regulatory gene from *B. pseudomallei*. The nucleotide sequences of the forward and reverse primers are written in 5' to 3' orientation in **Table 4.1** and the primers were designed as described in section 3.13.1. In order to reduce the amplification of non-specific PCR products, various concentrations of the components of the PCR reactions were tested including primers concentrations (10, 20 and 100 pmol/50 µl), DNA template (10, 20 and 50 ng/50 µl), and primer annealing temperatures (45, 50, 60 and 65°C).

Table 4.1: PCR primers for the amplification of a regulatory gene from *B. pseudomallei*

Forward primer (FF)	GATTTCACGATGCATCAGGCGAA
Reverse primer (RR)	TTCTGGATCGCCGCGATGTCCGTG

4.2.1.3 Immobilisation of λ plaques on nitrocellulose filters

An appropriate number of plaques, 4000, was plated out from *B. pseudomallei* λGEM-11 genomic library. A 100 µl of the bacteriophage dilution corresponds to 10^3 plaques was used to infect 100 µl of *E. coli* LE392 as described in section 3.11.4. The plates were cooled at 4°C for

1 h, to prevent transfer of the soft agarose layer. Gridded positively-charged nitrocellulose filters (Schleicher & Schuell) were numbered, marked for the orientation and placed gently onto the surface of the top agarose so that it came into direct contact with the plaques without trapping any air bubbles. Each filter became a mirror image of the pattern of recombinants on the plate. Plates were marked asymmetrically with filter marks. Transfer was allowed to proceed for 30 min at 4°C, after which the nitrocellulose filters were peeled away from the plate. Filters were then processed to denature the DNA in the samples by immersing, plaques side up, in the denaturing solution (Appendix III) for 3 min, then into neutralising solution (Appendix III) for another 3 min. Filters were rinsed in 2 x SSC solution and placed (plaque) DNA side up, on paper towels to dry for 30 min at RT. After complete drying, the filters were sandwiched between two sheets of 3MM-Whatman paper and baked for 2 h at 80°C in a vacuum oven (Townson & Mercer Ltd, Runcorn, England) to fix the DNA to the filters. The filters were immediately processed for hybridisation.

4.2.1.4 Hybridisation of a labelled DNA probe to nitrocellulose filters

The DNA probe was labelled either by random priming (section 3.10.2.1) or PCR (section 3.10.2.2). Following probe labelling, hybridisation of DIG-labelled probes to nucleic acids was performed as described in section 3.10.3.

4.2.1.5 Detection of labelled DNA hybrids

All the following steps were performed at room temperature. Nitrocellulose filters were washed briefly in detection buffer 1 (Appendix III) and incubated for 30 min in 30 ml blocking solution (Appendix III) then probed with 10 µl of anti-digoxigenin antibody conjugated to alkaline phosphatase in 20 ml of detection buffer 1 for 30 min. Unbound antibody was removed by washing 2 x 15 min with 100 ml of buffer 1. Filters were then equilibrated with 20 ml of

buffer 3 (Appendix III) for 5 min. Each filter was placed in an individual sterile Petri dish (50 mm) and then incubated in the dark with 10 ml of substrate solution (45 µl NBT-solution and 35 µl X-phosphate-solution in 10 ml of buffer 3). Following colour development, the reaction was stopped by washing filters with 50 ml of detection buffer 4 (Appendix III) for 5 min.

4.2.1.6 Selection of positive clones

The DNA of recombinant plaques to which the labelled probe has hybridised appears as intense dark spots on the filters, relative to a weak background. To pick positive plaques, the positive spots were circled with a (red) marker and oriented with respect to the original plate. Once filter and plate were aligned, positive plaques along with several negative controls were picked using a sterile Pasteur pipette and placed into a sterile microfuge tube containing 50 µl of SM phage buffer and two drops of chloroform as described in section 3.11.5.1.

4.2.1.7 Rescreening of recombinant clones

In order to purify a single homogeneous λ clone, stocks of positive phage plaques were diluted and plated as described (section 3.11.4), but at a lower plaque density to enable the isolation of a single recombinant. Nitrocellulose filters were prepared and treated as described in sections 3.12.1-3.12.3. For each putative positive, a single, pure plaque was picked.

4.2.1.8 Isolation of recombinant phage DNA

A stock of each positive phage was made using the plate lysate procedure (section 3.11.5.2) and the titre was established (section 3.11.4). λ phage DNA from recombinant (positive) clones was extracted using Wizard Lambda Preps (section 3.11.6.1).

4.2.2 Screening of a λGT-11 genomic library

λGT-11 (**Fig 4.1**) is an insertion vector that has a single recognition site for *Eco*RI, which

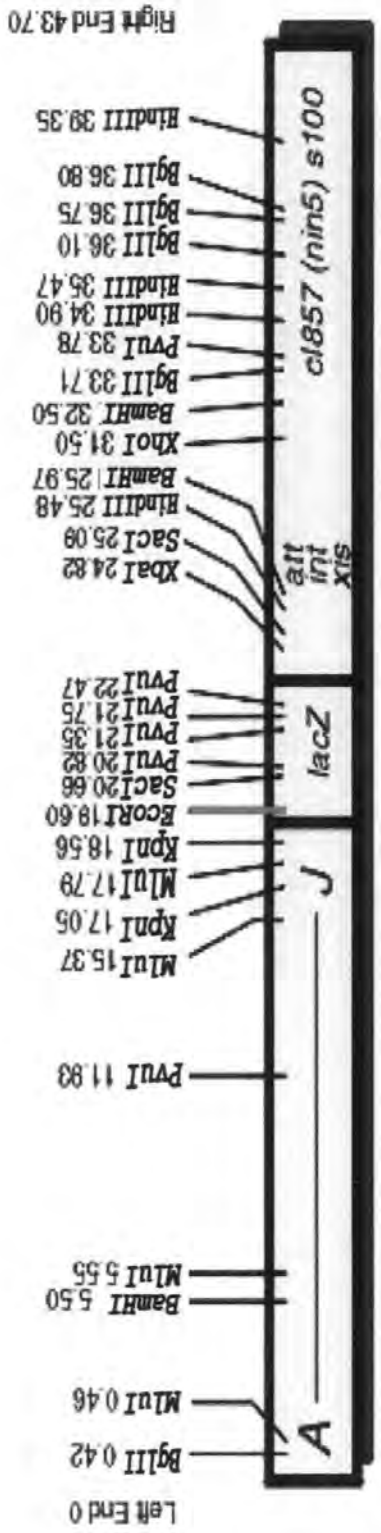


Fig 4.1: Schematic presentation of the λGT-11 cloning vector (Stratagene).

enables DNA fragment to be inserted into the λ genome. This generates left and right arms of 32.7 and 10.6 kb respectively, which can accept insert DNA fragments up to approximately 7.6 kb in length. The single *Eco*RI site for cloning is located at the carboxy-terminal end of the *β*-galactosidase (*lacZ*) gene present in the vector molecule. Insertion of foreign fragments into this site disrupts the function of *lacZ* and alters the phenotype of the phage. When plated using a host strain that is deficient in the expression of *lacZ* (such as *E. coli* Y1088), it is possible to check the efficiency of recombination using blue/white screening (section 3.5.5).

B. pseudomallei strain 204 genomic library was previously constructed in λGT-11 by C. Davies, University of Plymouth. *B. pseudomallei* genomic DNA obtained as outlined in section 3.4.1 was completely digested with *Eco*RI (section 3.4.5.1), then ligated to λGT-11 DNA (section 3.5.3) that had been pre-cut with *Eco*RI and pre-treated with alkaline phosphatase (Promega, UK). Recombinant DNA was then packaged into phage particles (sections 3.11.3) and the titre (3×10^7 pfu/ml) was determined (section 3.11.4). The library was stored at 4°C. The number of clones that represents *B. pseudomallei* genomic library in λGT-11, assuming 5 kb inserts, was estimated to be at least 4000 plaques (section 3.11). The library was plated out as described for λGEM-11 but using *E. coli* Y1088 as a bacterial host strain (**Table 3.3**).

Aliquots of the packaged phage containing $1\text{-}2 \times 10^3$ in a volume <50 μl were mixed with 200 μl of *E. coli* strain Y1088 grown to $A_{600} \sim 0.6$ in NZCYM broth supplemented with 0.2% lactose at 37°C. The mixture was plated as described in section 3.11.4. An appropriate number of plates (8), containing about 8000 plaques that fully represented the library, were overlaid with gridded nitrocellulose filters (0.45 μm, Schleicher & Schuell) and screened using a DNA probe (section 4.2.1). Positives were rescreened (section 4.2.1.7) and λ DNA was isolated using standard minipreps (section 3.11.6.2).

4.2.3 Construction and screening of a λZAP express genomic library

A *B. pseudomallei* strain 204 genomic library was constructed in λZAP Express vector as described in section 3.12, and plaque hybridisation screening was conducted as outlined in section 4.2.1.4 using a DIG-dUTP labelled DNA probe. The number of plaques to be screened was estimated (section 3.12) and approximately 5000 plaques were screened in order to ensure full coverage of the genome.

4.3 RESULTS

4.3.1 Construction of a *Sau*3A partially digested *B. pseudomallei* genomic library

A library of *B. pseudomallei* chromosomal DNA fragments was constructed in λGEM-11 as described in section 3.11. Genomic DNA from *B. pseudomallei* strain 204 was partially digested with *Sau*3A (section 3.11.1) to create a series of overlapping fragments for ligation into λGEM-11 arms. The enzyme recognises a tetra-nucleotide DNA sequence (GATC) that would be expected to occur on average once every 4^4 (256) bp in a random DNA sequence and the fragments produced can be directly ligated into *Bam*HI-cut vector, after agarose gel size fractionation and selection of the appropriate size, 9-23 kb. Test digests were performed with serial dilutions of *Sau*3A incubated with 2 μg of *B. pseudomallei* DNA for 30 min. Following agarose gel electrophoresis (**Fig 4.2**), it was found that 0.078 unit of *Sau*3A per μg of DNA (Lane 8) produced the greatest proportion of fragments in the size range of 9-23 kb. After ligation (section 3.11.2), the recombinant phage were packaged (section 3.11.3) and the titre of the library was determined (section 3.11.4).

A 100 μl volume of a 10^{-3} dilution of the λGEM-11 library yielded approximately 1100 plaques i.e. ~ 5.5 x 10^7 pfu/ml. Assuming that the average size of the *Sau*3A fragments was 15

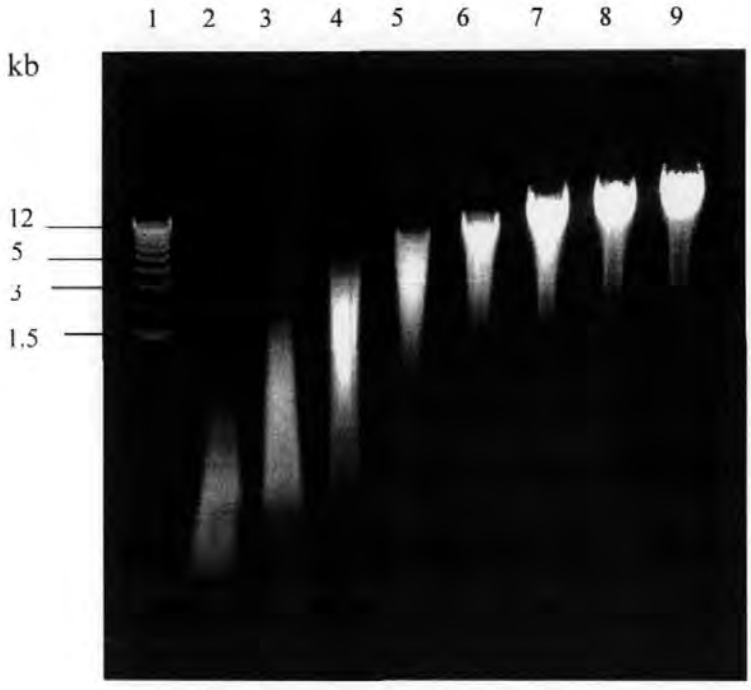


Fig 4.2: Agarose gel electrophoresis of serial *Sau*3A digests of *B. pseudomallei* 204 genomic DNA. The sizes of DNA fragments were estimated from 1 kb ladder (Lane 1).

kb, then the library stock (10 ml) contained a total of 15 x 5.5 x 10^8 = 8.25 x 10^9 kb of *B. pseudomallei* genomic DNA i.e. approximately 1.3 x 10^6 times the size of the *B. pseudomallei* genome providing ample scope for screening.

4.3.2 Screening with a DNA probe

4.3.2.1 PCR amplification of a response regulator gene

PCR primers were developed on the basis of nucleotide sequence homology with the response regulatory genes of *E. coli* and *S. typhi* (section 4.2.1.2). For PCR reactions, a concentration of 10 pmol of each PCR primer and 10 ng of *B. pseudomallei* genomic DNA template were used in a 50 µl reaction volume. The best results were obtained when the temperatures that were used for primer annealing were 45°C for the first cycle, 50°C for the second and third cycles, and 60°C for the remaining thirty-three cycles. Apart from these modifications the PCR reactions were carried out as described in section 3.13.2. The PCR products that were obtained were in the range of 300 to 500 bp in size and were isolated from agarose gels (section 3.4.6.1), blunt kinased (section 3.6.1) and ligated to *Sma*I blunt-ended-dephosphorylated plasmid pUC18 that was available in a 'ready to go' formula (section 3.6.2). The ligated molecules were transformed into *E. coli* DH5α by standard procedures (section 3.5.4.1). A number of recombinants were recovered and found to possess plasmid DNA containing inserts of the appropriate size, following *Eco*RI and *Pst*I double digestion (sections 3.4.5.2 & 3.6.3). Seven clones were selected for double stranded sequencing of the inserted DNA (section 3.11). One clone, named pRS7, was found to contain a DNA insert of 375 bp and following translation of the nucleotide sequence a search of the protein databases revealed that this portion of the deduced amino acid sequences was highly homologous to a variety of prokaryotic response regulator proteins.

4.3.2.2 Preparation of a DNA probe

Plasmid pRS7 was digested with *Eco*RI and *Pst*I (section 3.4.5.2) and the 375 bp DNA insert was isolated from an agarose gel (section 3.4.6.2), purified (section 3.4.4.3), labelled by random priming (3.10.2.1) and purified (section 3.10.2.3). The probe was designated RS7.

4.3.3 Isolation of positive clones

Screening of approximately 4,000 recombinant phage plaques from a λGEM-11 genomic library using the oligonucleotide probe RS7 hybridised at high stringency resulted in the identification of five positive clones, designated MRG1-MRG5, an average of one positive per 800 plaques screened (**Fig 4.3**). The five clones were picked from the appropriate plates (section 3.11.5.1) and rescreened with the RS7 probe. All but one, MRG5, were subsequently shown to be true positives (**Fig 4.4**).

4.3.4 PCR screening

PCR was used to confirm the presence of the correct DNA insert in clones MRG1-4. Internal PCR primers were designed to amplify a 234 bp fragment of the RS7 sequence (**Table 4.2**). The reactions were performed as described in section 3.13.2, using 20 ng of template DNA and following agarose gel electrophoresis (section 3.4.6.1) clones MRG1-4 were each found to possess a 234 bp amplicon.

Table 4.2: Internal primer sequence for RS7 probe

Forward primer (rs7F1)	CCGATCATCGTGTTCACGATGCTG
Reverse primer (rs7R1)	TTCCTTCGGCGTCAACGCGTTGAA

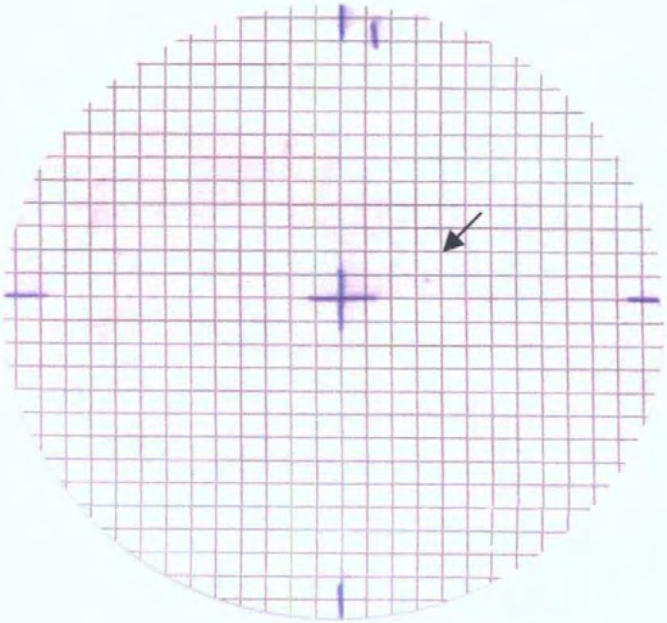


Fig 4.3: A nitrocellulose filter showing a positive plaque from the *Sau*3A λGEM-11 genomic library.

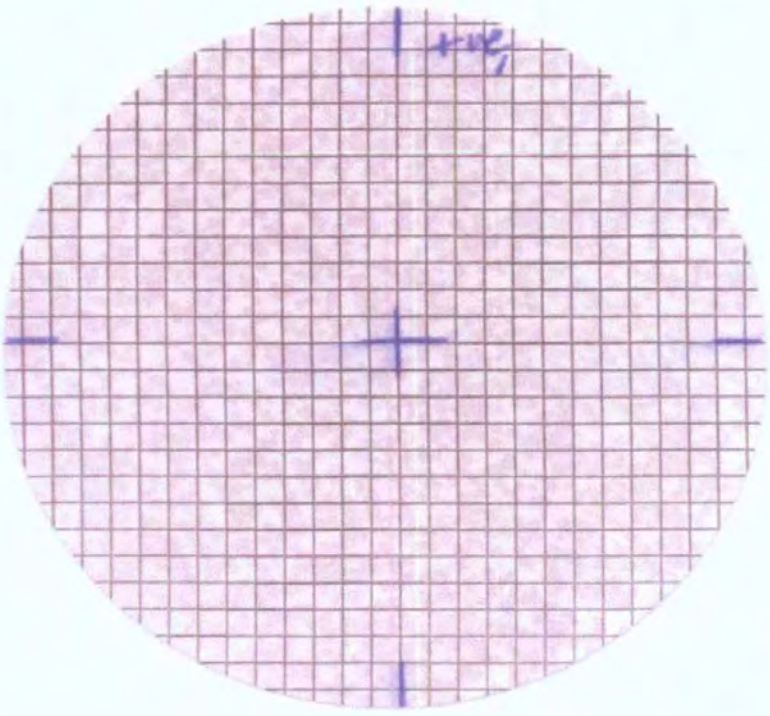


Fig 4.4: A nitrocellulose filter showing the results of rescreening a positive clone.

4.3.5 Restriction enzyme digestion of recombinant phage DNA

Restriction endonuclease *Xho*I was used to digest phage DNA from clones MRG1-4 in order to determine the size of the DNA inserts. *Xho*I cuts just outside the *Bam*HI cloning site; hence digestion with *Xho*I should release the cloned fragment and reveal any internal sites. Approximately 0.3 μg of recombinant λ DNA was digested as described in section 3.4.5.1, and checked on 0.8% agarose gel (section 3.4.6). Analysis of the restriction digests (**Fig 4.5**) showed that the four positive clones, MRG1-4 contained inserts of 11.4, 9.3, 8, and 12.1 kb, respectively. However, as so little phage DNA was actually isolated after several trials, the λ DNA was used as template for further PCR as discussed below (section 4.3.11) and it was decided to screen another *B. pseudomallei* gene bank, which had been constructed using the λGT-11 vector.

4.3.6 Screening of a λGT-11 genomic library

A library of *B. pseudomallei* strain 204 genomic DNA that had been constructed in λGT-11 (section 4.2.3) was screened with the labelled DNA probe, RS7. The library had previously been used to successfully screen for a putative *B. pseudomallei* type IV pilin gene (C. Davies; unpublished data). Screening of 8000 plaques with RS7 resulted in the isolation of 2 positive clones designated MRG6 and MRG7, which on further investigation were found to be identical. One positive clone, MRG6, was selected for further analysis. Approximately 0.3 μg of MRG6 DNA was digested with *Eco*RI as described in section 3.4.5.1, and found to contain two *Eco*RI inserts with lengths of 2.2 and 4.3 kb.

4.3.7 Subcloning of MRG6 DNA inserts into pUC18

In order to determine which of the two *Eco*RI fragments inserted into MRG6 possessed the RS7 sequence, both inserts were subcloned into pUC18. Approximately 1 μg of MRG6 DNA

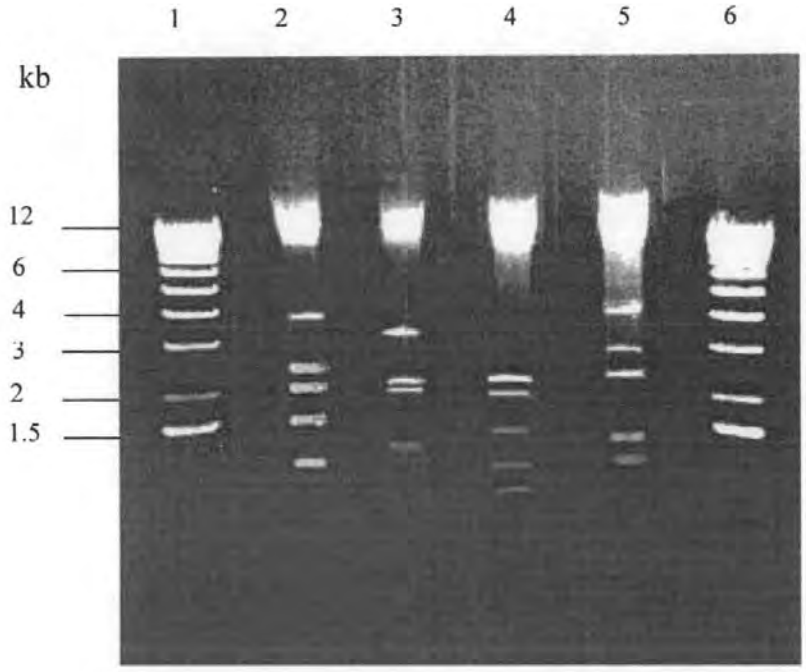


Fig 4.5: Restriction analysis of recombinant λGEM-11 phage DNA digested with *Xho*I. The sizes of fragments were estimated from 1 kb DNA ladder (Lane 1 & 6).

Track	Phage	No. of DNA fragments	Approximate size of DNA fragments (kb)	Total size (kb)
1	MRG1	4	1.1, 1.6, 2.2, 2.5, 4.0	11.4
2	MRG2	5	1.4, 2.1, 2.3, 3.5	9.3
3	MRG3	3	1.0, 1.2, 1.5, 2.0, 2.3	8
4	MRG4	5	1.2, 1.4, 2.5, 3.0, 4.0	12.1

was digested with *Eco*RI (section 3.4.5.1) and purified with the Prep-A-Gene kit (section 3.4.4.3). Approximately 0.5 μg of pUC18 DNA was digested with *Eco*RI (section 3.4.5.1), treated with alkaline phosphatase (section 3.5.1) and then ligated to the purified insert DNA. The ligated DNA was transformed into *E. coli* strain DH5α (section 3.5.4) and recombinant clones were recovered and screened as described (section 3.5.5). Plasmid DNA was extracted from *E. coli* clones (section 3.4.2.3) and examined for the presence of the appropriately sized DNA inserts using *Eco*RI digestion (section 3.4.5.1) and agarose gel electrophoresis (section 3.4.6.1). Two subclones were isolated and the plasmids carrying the 2.2 kb and the 4.3 kb fragments were designated pMRG1 and pMRG2, respectively.

4.3.8 Southern blot analysis of subclones

Southern hybridisation was carried out to check pMRG1 and pMRG2 for the presence of the RS7 sequence and to confirm that the cloned DNA fragment originated from *B. pseudomallei* DNA. Purified plasmid DNA of pMRG1 and pMRG2 and *B. pseudomallei* genomic DNA were digested to completion with *Eco*RI (section 3.4.5.1), separated by agarose gel electrophoresis (section 4.6.1) and probed with RS7 as described in section 3.10.

After enzymatic detection, RS7 was shown to hybridise with the 4.3 kb *Eco*RI insert from pMRG2 but not the 2.2 kb insert from pMRG1 (**Fig 4.6**). Furthermore, a 4.3 kb band was observed in the *B. pseudomallei* genomic DNA digest confirming the origin of DNA. Subsequently, pMRG2 was used as a source of plasmid DNA for nucleotide sequencing and other manipulations. Although pMRG1 did not possess the appropriate DNA insert the plasmid was eventually sequenced and found to encode sequences that were closely similar to glutamine synthetase (*glnA*) and the nitrogen regulator (*ntrB*) from many other bacterial species (Appendix IV).

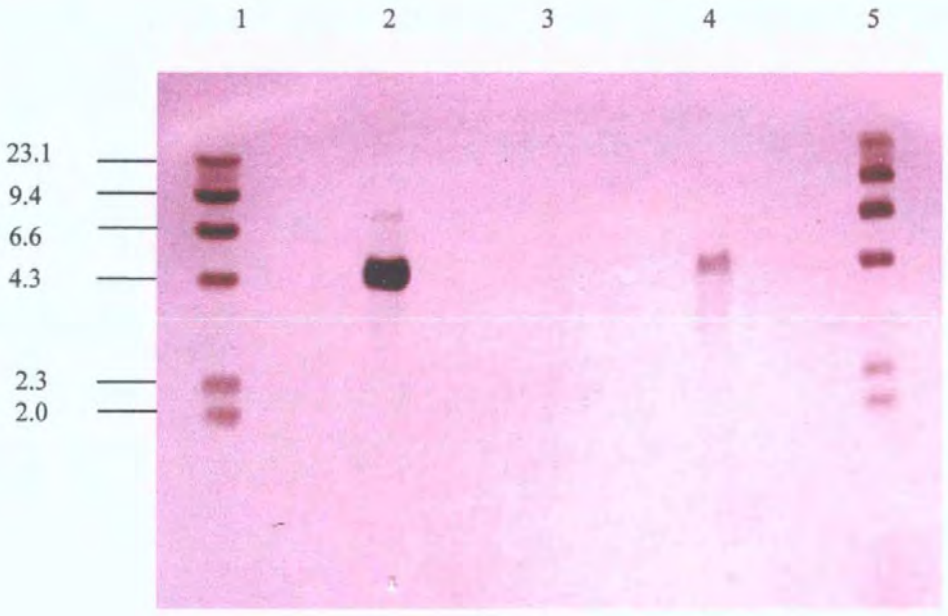


Fig 4.6: The detection of RS7 DIG-labelled probe in pMRG1, pMRG2 and *B. pseudomallei* 204 genomic DNA by Southern hybridisation. Lanes 1 & 5: λ-*Hin*dIII ladder, Lane 2: *Eco*RI digested pMRG2, Lane 3: *Eco*RI digested pMRG1 and lane 4: *Eco*RI digested *B. pseudomallei* genomic DNA.

4.3.9 Restriction mapping of pMRG2

Plasmid pMRG2 was cleaved with the following enzymes in a series of single digests (section 3.4.5.1): *Eco*RI, *Eco*RV, *Pvu*II, *Hind*III, *Mlu*I, *Bcl*I, and *Pst*I. Following the digestions, the resultant DNA fragments were separated by agarose gel electrophoresis (section 3.4.6.1) and the restriction sites were mapped. The restriction analysis of plasmid pMRG2 is summarised in **Fig 4.7**, while the deduced restriction map is presented in **Fig 4.8**.

4.3.10 Nucleotide sequencing of pMRG2

Nucleotide sequencing of pMRG2 showed that it contained an *Eco*RI fragment consisting of 4324 bp of *B. pseudomallei* strain 204 DNA. The sequence was compared with the nucleotide sequences of other bacteria stored in the GenBank database. A preliminary check suggested that the fragment contained at least one complete gene sequence and another incomplete gene sequence downstream. This is explained more fully in section 4.4. On this basis it was deemed necessary to identify and clone the DNA encoding the remainder of the second gene and also the flanking region.

4.3.11 Cloning of downstream sequences by PCR amplification

Based on the sequence data obtained in section 4.3.10, two sets of PCR primers (**Table 4.3**) were developed to attempt to amplify the downstream sequence from the DNA purified from the λGEM-11 library clones MRG1-4. The primers were designed to amplify fragments between the 3' end of the known sequence obtained above and the 5' sequence of λGEM-11 arms in both possible orientations (**Fig 4.9**). PCR reactions were carried out as described in section 3.13.2, and primers GEMF and 4.3 R (1st orientation) amplified a 532 bp DNA fragment from MRG4, which was subcloned into pUC18 as described in section 3.6, and sequenced after purification (section 3.8) of recombinant plasmid pMRG3. DNA sequence analysis revealed

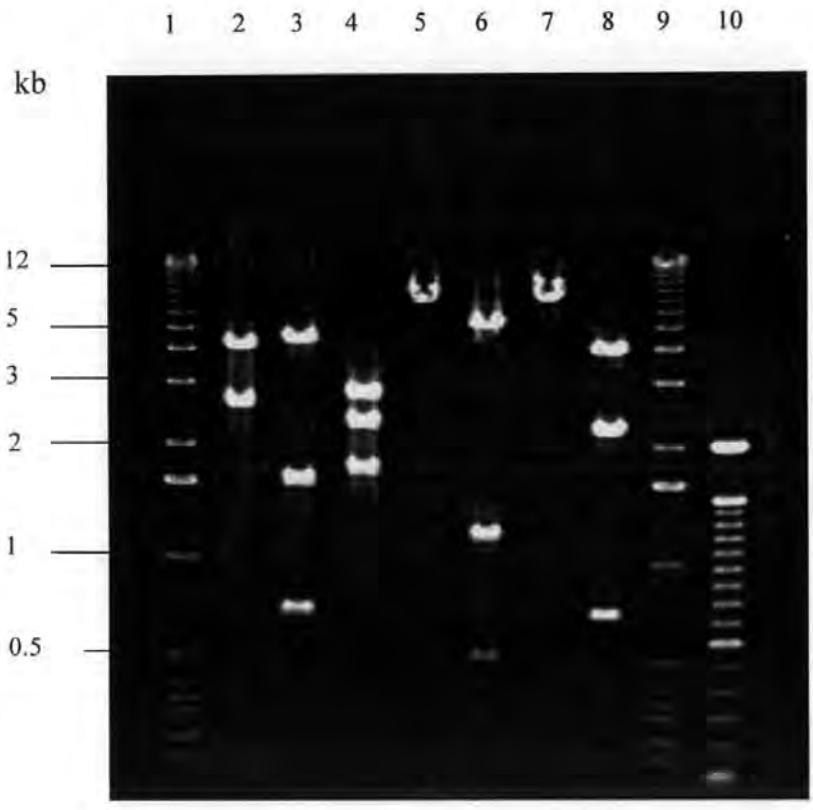


Fig 4.7: Agarose gel electrophoresis of restriction enzyme digests of pMRG2. The sizes of fragments were estimated from 1 kb (Lane 1 & 9) and 100 bp (Lane 10) DNA ladders.

Track	Enzyme	No. of DNA fragments	No of cleavage Sites on vector insert		Approximate size of DNA fragments (kb)	Total size (kb)
2	*Eco*RI	2	1	2	2.68, 4.32	7.00
3	*Eco*RV	3	0	3	0.74, 1.66, 4.6	7.00
4	*Pvu*II	3	2	1	1.8, 2.4, 2.8	7.00
5	*Hind*III	1	1	0	7.00	7.00
6	*Mlu*I	3	3	0	0.5, 1.2, 5.3	7.00
7	*Bcl*I	1	0	1	7.00	7.00
8	*Pst*I	3	1	2	0.7, 1.4, 4.9	7.00

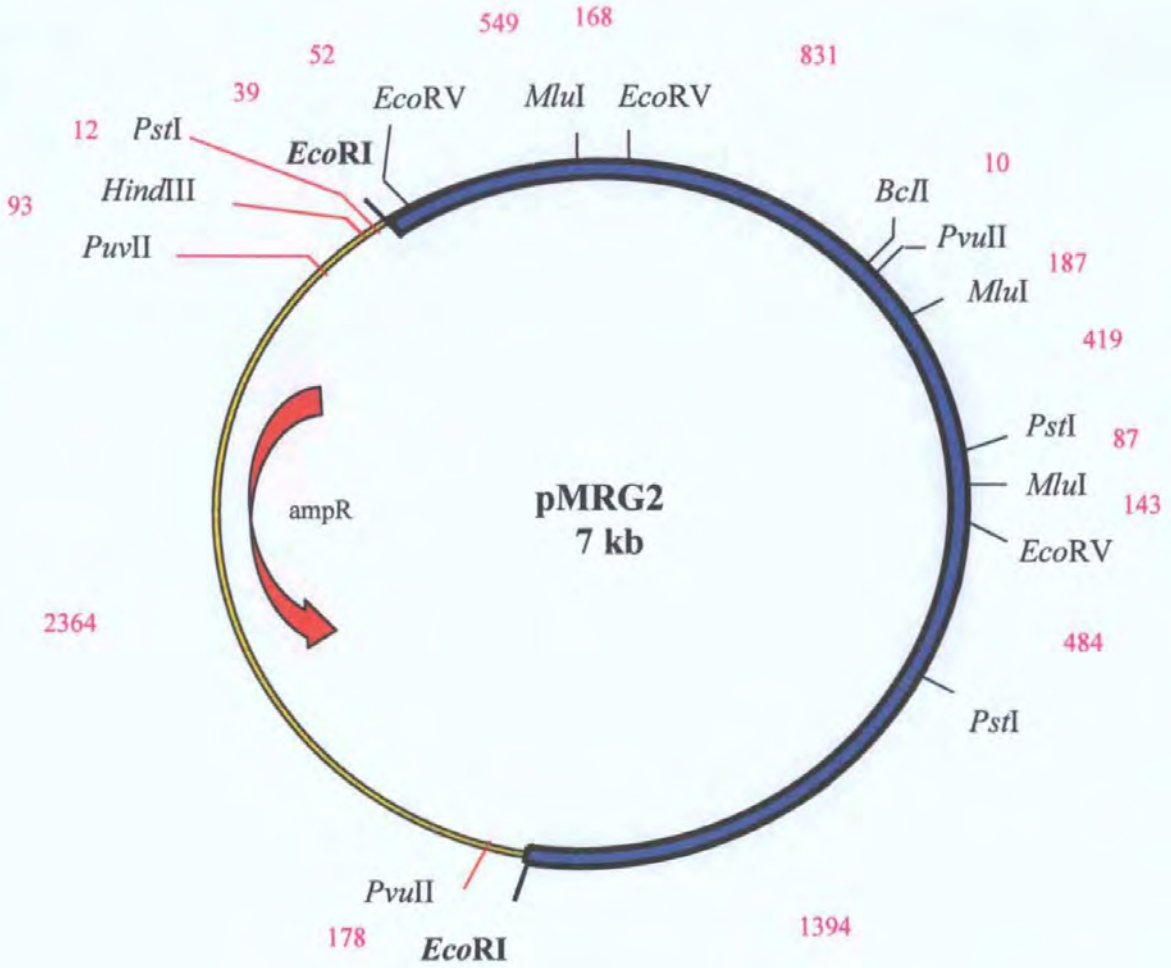
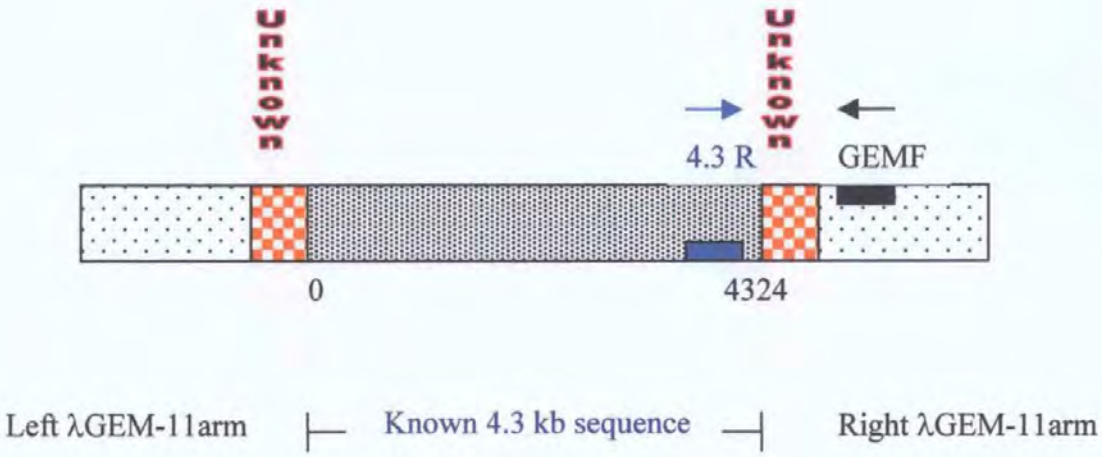


Fig 4.8: Restriction map of pMRG2 deduced from single enzyme digests. The 4.3 kb insert of *B. pseudomallei* DNA is represented by the blue colour. ampR is the gene encoding ampicillin resistance, while numbers represent the number of base pairs of DNA between adjacent restriction sites.

Putative 1st orientation



Putative 2nd orientation

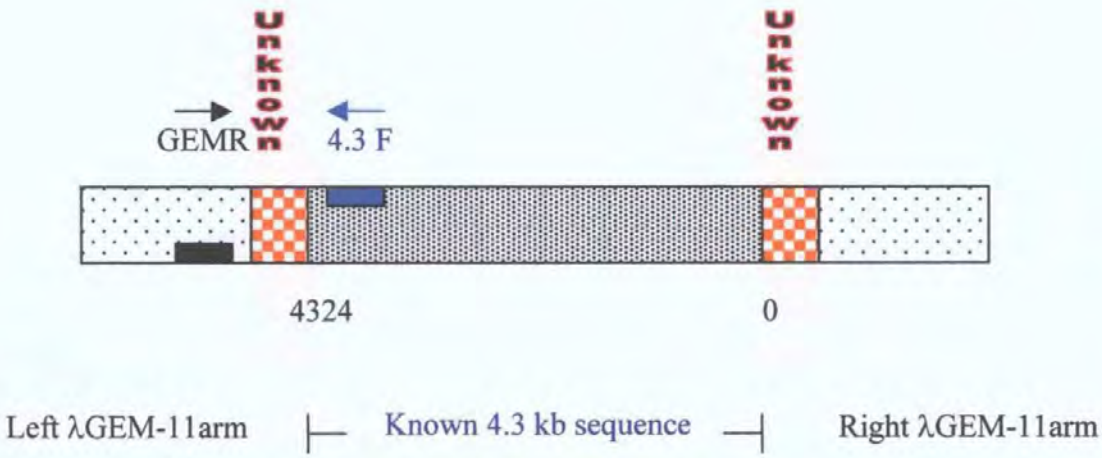


Fig 4.9: Strategy for cloning the sequence downstream from the 4.3 kb *Eco*RI fragment (cloned from a λGT-11 library) based on specific primers and λGEM-11 arms primers. Arrows show the direction of amplification.

that plasmid pMRG3 contains inserted DNA of 532 bp with a 176 bp overlap that exactly matched the sequence of the 3' end of pMRG2 insert DNA.

Table 4.3: Specific primer sequences for 3' pMRG2 insert and 5' λGEM-11 arms

Forward primer (GEMF)	TCTGGGTCGACTCTAGGCCTAAAT
Reverse primer (4.3 R)	ATGAAAGGGCGCATCGATCTGGAA
Forward primer (4.3 F)	AAGCAGCCACTGGTAGTCCTTTAG
Reverse primer (GEMR)	GTCTTGCAGACAAACTGCGCAACT

4.3.12 Construction of a λZAP genomic library

Based on the sequence data obtained from both pMRG2 and pMRG3, another genomic DNA library was constructed using the phage vector λZAP Express. The sequence was scanned for suitable restriction enzyme cleavage sites and *Bcl*I was chosen. Chromosomal DNA (3 μg) from *B. pseudomallei* 204 (section 3.4.1) was completely digested with *Bcl*I (section 3.4.5.1) and the library was constructed by ligating the products of this digest with *Bam*HI and CIAP treated λZAP Express vector as explained in section 3.12.

A 1 μl volume of the 500 μl packaged library DNA, when mixed with 200 μl of *E. coli* strain XL1-Blue (section 3.12.5), yielded 1380 recombinants and 127 nonrecombinants. This represents a total yield of approximately 7×10^5 recombinant pfu in the library. Assuming that the average size of *Bcl*I fragment inserted into λZAP Express was 4 kb, then the library contained a total of $4 \times 7 \times 10^5 = 2.8 \times 10^6$ kb of *B. pseudomallei* genomic DNA, approximately 4.3×10^2 times the size of the *B. pseudomallei* genome.

4.3.13 Probe preparation and labelling

PCR primers that were developed (section 3.13.1) on the basis of sequence data for pMRG2 and pMRG3 (**Table 4.4**) were used to amplify an 829 bp fragment from *B. pseudomallei* genomic DNA. DIG-dUTP was incorporated into the PCR reaction mixture as described in section 3.10.2.2, and the labelled probe, RS8, was used to screen the λZAP library.

Table 4.4: Primers developed on the basis of the sequences obtained from pMRG2 and pMRG3

Forward primer (MR2F)	GCTTGATGCAGATCGTCAACGATG
Reverse primer (MR3R)	GCGAGGCTATACATGCTGATTTCC

4.3.14 Screening of a λZAP genomic library

The screening of the λZAP Express library was essentially performed as for the λGEM-11 library (section 4.2.1.2). Screening of approximately 5000 plaques, with RS8 (section 4.3.14) resulted in the isolation of three putative positive clones. The clones were designated MRG8 to MRG10. These plaques were picked as described in section 3.11.5.1, and rescreening of the three putative positive plaques confirmed that MRG8, MRG9 and MRG10 were all true positives.

4.3.15 *In vivo* excision of a recombinant pBK-CMV plasmid vector

λZAP Express is designed to allow *in vivo* excision and recircularisation of any cloned insert contained within the lambda vector. The plasmid DNA was excised from the three putative positive clones (section 3.12.7) in the form of the kanamycin-resistant pBK-CMV phagemid vector containing the inserted DNA. A 200 volume µl of XL1-Blue (OD_{600} ~ 1.0) were infected with 1 µl ExAssistant helper phage and 250 µl of phage stock for each of the

MRG8, MRG9 and MRG10 clones. After the rescue process detailed in section 3.12.7, several colonies were picked from LB-kanamycin plates and the plasmid DNA was extracted and screened for the presence of recombinant insert DNA (section 3.4.2.3). Following extraction, selected plasmids were digested with *Eco*RI, *Pst*I and *Sac*I in a series of single digests as described in section 3.4.5.1. The restriction analysis was resolved by agarose gel electrophoresis and showed that the plasmid DNA from each of the three positive clones (MRG8-10) were indistinguishable, having the same restriction fragment pattern. Subsequently, the plasmid pMRG4 from one of the clones was selected for further restriction mapping and DNA sequencing.

Plasmid pMRG4 was cleaved with the following enzymes in a series of single digests: *Mlu*I, *Sac*I, *Sma*I, and *Hind*III, which were resolved by agarose gel electrophoresis (3.4.6.1). Restriction analysis of pMRG4 showed that it contains 8.4 kb of inserted DNA (**Fig 4.10**). Furthermore, DNA sequence information derived from pMRG4 supported the restriction analysis and showed that it carries 8404 bp of inserted DNA. This sequence overlaps the full length of inserted DNA in pMRG3 and 2706 bp of pMRG2 (**Fig 4.11**), including the 3' end of the second putative gene sequence in addition to a further 4.3 kb of flanking sequence information. The physical map of pMRG4 is shown in **Fig 4.12**.

In summary, it was necessary to construct and screen three *B. pseudomallei* genomic libraries in λGEM-11, λGT-11 and λZAP Express vectors, in order to isolate three overlapping genomic clones; pMRG2, pMRG3 and pMRG4 spanning approximately 10 kb of the *B. pseudomallei* chromosome (**Fig 4.13**).

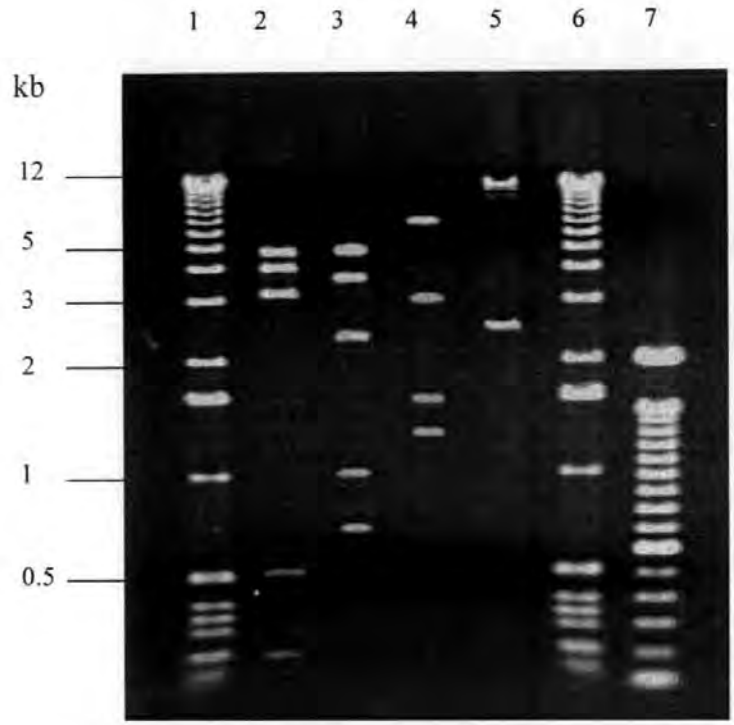


Fig 4.10: Agarose gel electrophoresis of restriction enzyme digests of pMRG4. The sizes of fragments were estimated from 1 kb (Lane 1 & 6) and 100 bp (Lane 7) DNA ladders.

Track	Enzyme	No of DNA fragments	No of cleavage sites on vector	insert	Approximate size of DNA fragments (kb)	Total size (kb)
2	*Mlu*I	5	1	4	0.163, 0.506, 3.3, 4.049, 4.887	12.90
3	*Pst*I	5	1	4	0.714, 1.079, 2.246, 3.79, 5,093	12.90
4	*Sma*I	4	1	3	1.296, 1.561, 2.989, 7,074	12.90
5	*Hind*III	2	1	1	2.46, 10.44	12.90

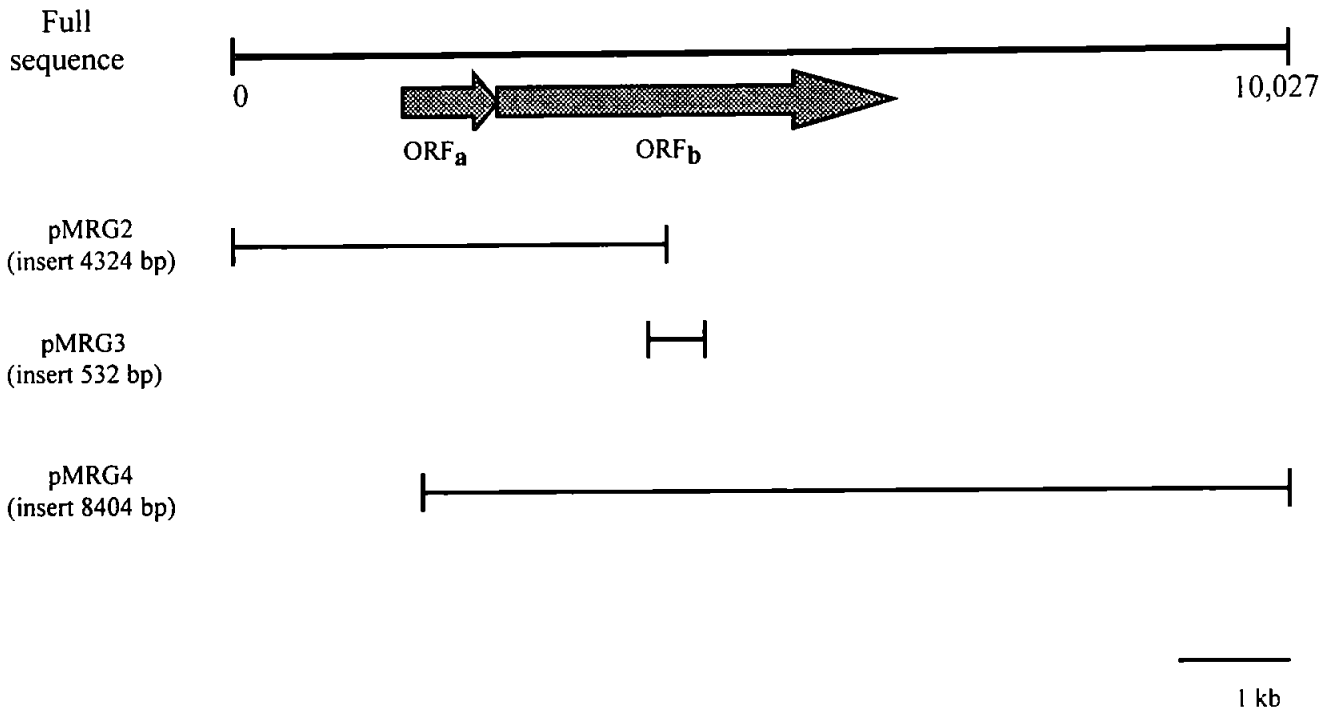


Fig 4.11: Schematic presentation of the total sequence overlaps between different clones of *B. pseudomallei* in pMRG2, pMRG3 and pMRG4. Arrows represent open reading frames (ORF) and reading direction, while numbers indicate the size of each genomic DNA fragments.

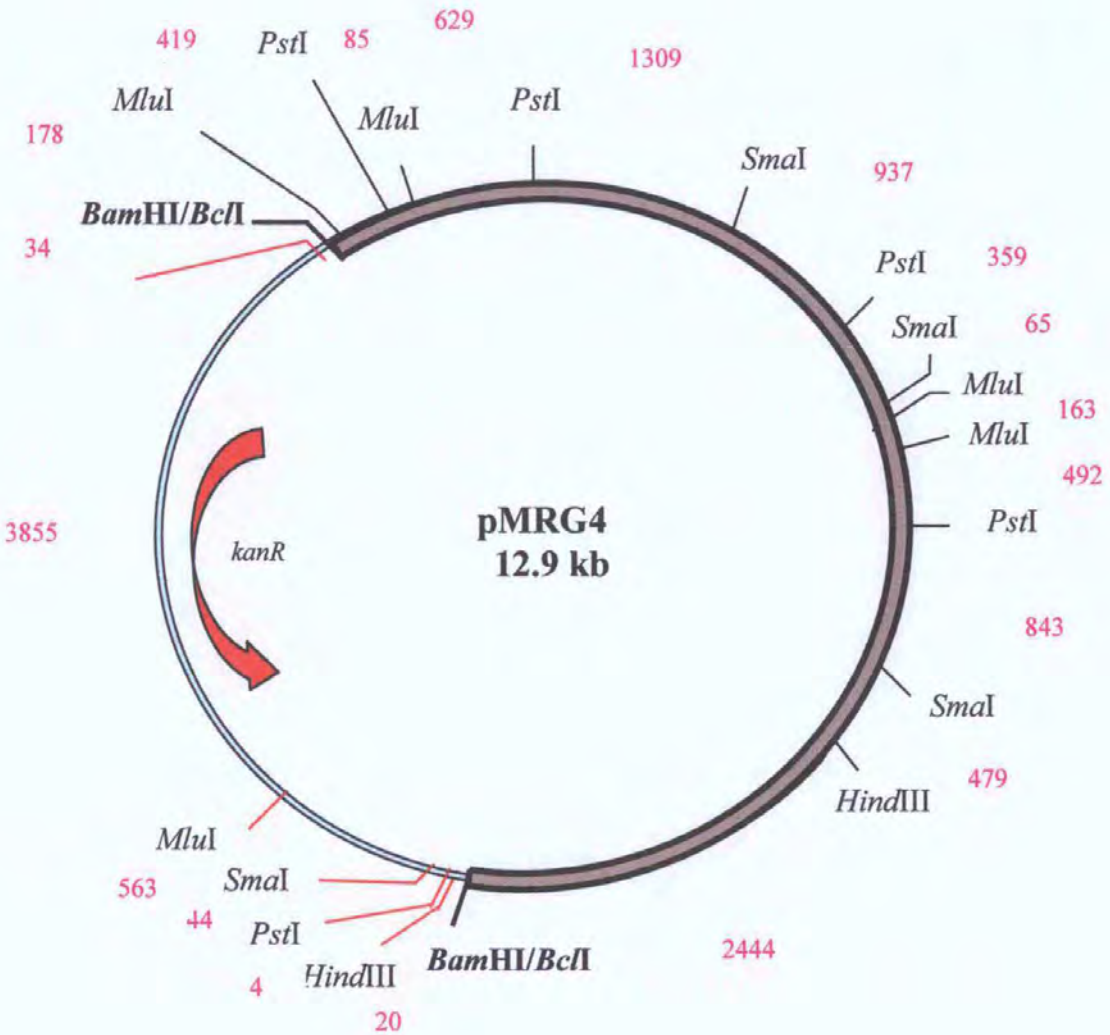


Fig 4.12: Restriction map of pMRG4 deduced from single enzyme digests. The 8.4 kb insert of *B. pseudomallei* DNA is represented by the grey colour. *kanR* is the gene encoding kanamycin resistance, while numbers represent the number of base pairs of DNA between adjacent restriction sites.

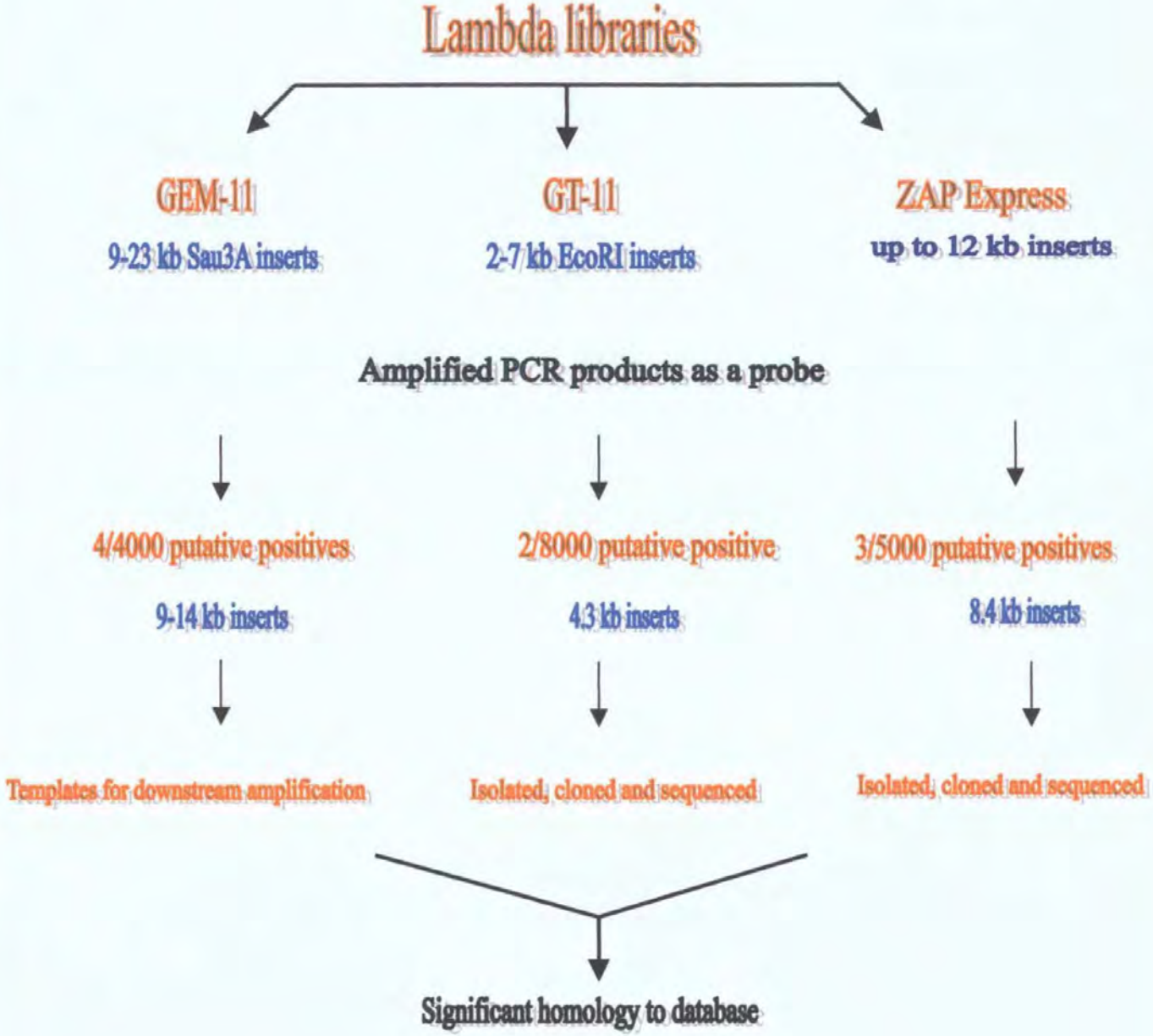


Fig 4.13: A summary of the results obtained from screening lambda libraries of *B. pseudomallei* genomic DNA with DNA probes.

4.4 SEQUENCE ANALYSIS OF A TWO COMPONENT REGULATORY SYSTEM

4.4.1 Nucleotide sequence analysis

The raw sequence data obtained from pMRG2, pMRG3 and pMRG4 was scanned for the presence of a variety of features including restriction endonuclease cleavage sites, open reading frames, initiation codons, -10, -35 and +1 regions, Shine-Dalgarno sequences (Shine & Dalgarno, 1974), transcriptional termination loops and stop codons.

4.4.1.1 Restriction endonuclease cleavage sites

A complete restriction map of the sequenced 4.3 kb *Eco*RI fragment was assembled. The predicted restriction sites were found to exactly match the observed endonuclease cleavage patterns (**Table 1**, Appendix IV).

4.4.1.2 Open reading frames

The presence of open reading frames (ORFs) was determined according to the methods outlined in section 3.9. This chapter will present the results for two major ORFs, while the flanking sequences will be considered in the next chapter.

The first ORF, ORFa, consists of 666 nucleotides (**Fig 4.14**). The non-coding region upstream of ORFa was checked for the presence of characteristic promoter sequences. A potential Shine-Dalgarno ribosome binding site (GGAG) was found 5 nucleotides upstream of the ATG translation start. Preceding this translational start site, a possible promoter structure with an RNA polymerase recognition site (-35 region) and a Pribnow box (-10 region), where the RNA polymerase begins to unwind the DNA for eventual transcription, was found. The -35 region, TTGCAT, shows extensive homology with the *E. coli* consensus sequence of promoter regions, while the -10 region, TATgca, showed a limited homology with three of the six conserved bases (small letters denote nucleotides which are not identical). The distance

separating the putative -10 and -35 regions is 25 bases, which is more than the average separation between these two regions (17 bases). This may have some bearing on the relative levels of expression of the gene. The start point for transcription of bacterial genes (+1) is generally located about 10 bp downstream from the centre of the -10 region, and in 90% of cases it is a purine (A or G). At 9 bp downstream from the -10 region, a putative transcriptional start (A) is located. A second less likely predicted promoter sequence is located downstream from the first promoter region. Both promoters have the same -35 sequence (TTGCAT), while the second predicted promoter region has a limited similarity in -10 region (cAtgag) to that of *E. coli*. The distance between the second putative -10 and -35 regions is 8 bases, considerably less than the average separation distance. ORFa is apparently terminated by the termination signal TGA, however, downstream of the stop codon no putative transcriptional terminators were predicted. Analysis of the DNA base composition showed that the individual base content for ORFa was A, 17.56%, C, 29.57%, G, 34.38% and T, 18.46% and codons with C or G were preferentially used (**Table 2**, Appendix IV). The sequence was GC rich; mol% G + C = 64%.

The second putative reading frame, ORFb, consists of 3237 nucleotides (**Fig 4.14**), and seems to be transcribed in the same direction as ORFa. The initiation codon of ORFb overlaps by a single nucleotide with the termination codon of ORFa. Eight nucleotides upstream from the ATG start codon and in the distal portion of ORFa, there is a purine-rich sequence that possibly represents a potential Shine-Dalgarno ribosome binding site (AAGGA) for the high level expression of the gene. ORFb is apparently terminated by the presence of a stop codon TGA. Downstream of the stop codon in the non-coding region, an inverted repeat capable of forming a stem-loop structure, is found, with a stem of 10 bases composed predominantly of G's and C's but lacking a T-rich region at its 3'-end. A free energy (ΔG) of -26 kcal, was estimated for the

Fig 4.14: Nucleotide sequence of two-open reading frames (ORF) overlapping by a single nucleotide. The numbering of the sequence starts at the first nucleotide of the ORF**a** initiation codon. The putative -10, -35 regions, transcriptional start site (+1), potential ribosomal binding sites (RBS) and initiation codons are highlighted and underlined, while putative stop codons are in red bold Italics. The two ORF overlap is marked by an arrow. The *Eco*RI site shown represents the end of genomic insert in pMRG2, the *Sau*3A site is the end of genomic insert pMRG3, while the *Bcl*I site is the start of genomic insert in pMRG4. Two convergent arrows in the non-coding region downstream from the ORF**b** indicate a putative transcriptional termination region.

-35

TTCACGGCCGATTCGCGGCATGGCGAGGCGGGGCGCACGAGCGTCGGCGCCG**TTGCAT**TCGAGC

-10 +1 -35

GAATCGCGCTCGCAAGGC**ATATGC**AGGAACGC**A**GGCAATCGTTATCGTT**TTGCAT**GGAACGCGC

-10 +1 RBS

ATGAGCGTGCGCCGG**A**CCTCGTTCGTGCGGCACGGAGGCGCTTTCCGCTTTTTC**GGAG**CCGCC

10	20	30	40	50	60
ATGTCCAACG	TTGCCCTGCA	TACAAGAAAA	GTCGTCGTGG	CCGACGATCA	TCCGATCGTC
70	80	90	100	110	120
CTGCGTGCGG	TGACGGATTA	CGTCAATTCG	CTGCCGGGTT	TTCACGTCGT	GGCGTCGGTT
130	140	150	160	170	180
TCGTCGGGGG	ACGCGCTGTT	GTCCGCGATG	CGGGAGCAGG	AGGTCAATCT	CGTCGTGACC
190	200	210	220	230	240
GATTTCACGA	TGCATCAGGC	GAACGACGAC	AAGGACGGCT	TGCGCTTGAT	CTCGCACCTG
250	260	270	280	290	300
ATGAGGGCGT	ACGAGCGCAC	GCCGATCATC	GTGTTCACGA	TGCTGACCAA	TAGCGGCGTG
BclI 310	320	330	340	350	360
ATCAGCCAGC	TGTGCCGCAT	GGGCGTGGCG	GGGCTCGTCG	GCAAGGAGGA	GGAGATCGCC
370	380	390	400	410	420
GAGCTCGGGC	GCGTATGCGT	GAGCGTGGCG	CGCGGCGTCA	GCCAATCGTT	GTCGCCCGGC
430	440	450	460	470	480
ATGGCTCACC	GGCTTGCCGC	CGTCGGCAGC	ATCAGGCCGG	GCGAGGCGGC	ATTCAACGCG
490	500	510	520	530	540
TTGACGCCGA	AGGAACTGGA	GGTCGTGCGG	CTGTTTACGG	GCGGCATGTC	GCTCACGGAC
550	560	570	580	590	600
ATCGCGCGCA	CGCTGAATCG	CTCGCTCGGG	ACCGTGTCGA	CGCAGAAGCG	CTCGGCCATG
610	620	630	640	650	RBS 660
CGCAAGCTGC	ACGTGGACAC	CAACGTCGAT	CTCATCAACT	GCGCCCGCGA	GC**AAGGA**CTG

Fig 4.14 continued

670 680 690 700 710 720

CTCTGATGCA AGGACTTCTG CAAGAGCTCG ACGGATCGCC GCTGAGGAAG TTCTATTCGC

730 740 750 760 770 780

TCGAGTCGAA TCTGAAGCGC GAGCGGCGGG TCTTCACGAT CGTCATCGTG CTGCTCGTCT

790 800 810 820 830 840

GCGCGGCCCT CAGCATCGCG GCCATGACCG TCACCGGCTT GTTCCAGACC GCTTTCCGGC

850 860 870 880 890 900

AGGAGGAGCA ATCCGCGCGC ATCCACGAAA AGGAAGTGGT CGACGTGTTC CTGCAGCGCC

910 920 930 940 950 960

GCATGATGTT GACGACGGCA AGCCTCGTGC TGCAACTGCG GATGAACGGC GCGCCTTCGG

970 980 990 1000 1010 1020

CGCTGAACGT GCCGGCGCCG AACGCGTGCA CGCCGATGGC CCACAATGTG CGCGACGATG

1030 1040 1050 1060 1070 1080

CGATCCTGCG CGAGAGCTGC GATTACACGG TGCAGTTGCT GGCCAACTCG GGGCAGACGC

1090 1100 1110 1120 1130 1140

CGAGCGTCGA AATGGTGACG GCCGACGGTT CGGTCGGCTA TGGATATCTG TTTCCGACGG

1150 1160 1170 1180 1190 1200

GCGACCTGAG CGCGCTACGC TCCAGCACGC CGTCCGAACT CGTGTCGGCC GTGCTCGAGC

1210 1220 1230 1240 1250 1260

GCTACGGCAA GCGCGGCCTG GACCCGCTGG AAGCCGCGCG CAAGAAGCGG ATTCTCTGGT

1270 1280 1290 1300 1310 1320

TCGCGGTGGG CCGCGGCGGG CGCGGCGAGG AGCTGCATAT GATCGGCGCG TCAGTCGTGT

1330 1340 1350 1360 1370 1380

TCAAGGACGA GCGGCTCTAC GCGCTCGTCT TGACGAGCGT GGATCTTCAC AGCCTCGTTT

1390 1400 1410 1420 1430 1440

CGCCGATCGA GCGCGCCGGC CGCGTGCAGC AGCCGGTCGT CGTGGATTCC GAGGGCGTGC

1450 1460 1470 1480 1490 1500

CGCTCGTGAA CGCGGACGAC GCGGAAACGG TCCGGAAGGT CGACGGGCGG CTCGCCGGAC

1510 1520 1530 1540 1550 1560

AACAGGATGG CCTGTATCAC TGGATTCCCG GCTTCGGGTG GGCCCTACGC CGTCCCGCGC

1570 1580 1590 1600 1610 1620

CGTTTTCCGG TTTCGGGCAC ATGACGTATC TGCTTCCGCT CGATCTGCAG TTGCGCTCGA

1630 1640 1650 1660 1670 1680

TGCGCTACGA GTTGAGCCTC GTCGGCGGCG CGACGCTCGT GCTGATCGTG TTGCTGTTCG

1690 1700 1710 1720 1730 1740

TCGCGTTCCG GTACTGGAAT TACCGGTTCT TGACGCGCAT CTACGAGGAA GCGTCGCGCG

1750 1760 1770 1780 1790 1800

Fig 4.14 continued

```
CGCTCGAGAG CGAAATGCTC AACCATCTGC TGGTTCATGC GACGCCGGTC GGGTTGTGCA
      1810       1820       1830       1840       1850       1860
TCGTGCGGCG GGCGACACTG GAGATCGTCG TCGCCAACCC GATCGCGCGC ACGATGCTCG
      1870       1880       1890       1900       1910       1920
GCTTGCGGCT GTCGGACCGG CACCTGCCGC AGGAATTGCT GAGCGCGTTC GAATCGTCGC
      1930       1940       1950       1960       1970       1980
TGGCCGAGCA GGACACCCAG TCCGACGACG CGCGCATTTT CCAGTTCCCG TTCACGCTGT
      1990       2000       2010       2020       2030       2040
CGCGCGCCGG GCATGCGGCG GTCCATATCG AAATCACGTA CGCGCCCGCG ATGCTGAACG
      2050       2060       2070       2080       2090       2100
CGCGGGAGGT GTTCTTTTGC GCGATCACGG ACATGACGGC GCACCACCAG GCGGAGATCC
      2110       2120       2130       2140       2150       2160
TGCTGCGCGA GGCGAAGCTG ACGAGCGACG CGGCGGCCAA GGCGAAGGTG GCGTTCTTCG
      2170       2180       2190       2200       2210       2220
CATCGATGAG CCATGAAATC CGCACGCCGC TGTCGTCGCT CGTGGGCAAC ATCGAGCTGA
      2230       2240       2250       2260       2270       2280
TCGCGCGCGG GCCGCTCGCG CCCGAGCAGC AGGCGCGCGT GAAGGCGATG GAGACGTCGG
      2290       2300       2310       2320       2330       2340
CGCGCGGCTT GATGCAGATC GTCAACGATG TGCTCGATTT CTCGAAGATC GACGTGGGCG
      2350       2360       2370       2380       2390       2400
AGCTGAGCCT CATGGAGGAG TGGTCGAACA TCGCCGAGCT GCTCGACCGG CTCGCGCTCT
      2410       2420       2430       2440       2450       2460
CGCACGCGCC GCTCGCGACG CAGCAGGGTT TGAAGTTCTA CATGGTGTTC GATCGCAGCC
      2470       2480       2490       2500       2510       2520
TGCCCGCGCG GCTCTACTTC GATCCGATCC GGGTCTCGCA GATCGTGAAC AATCTGCTGA
      2530       2540       2550       2560       2570       2580
GCAACGCGCT GAAGTTCACG CCGTCCGGCA AGATCGTGCT GCGCGCCGGC TGGCGTGCCG
      2590       2600       2610       2620       2630       2640
GCGCGCTCGA AATCAGCGTG ACGGATTCCG GCATCGGCAT CCCCGATGAC CTGAAGCACC
      2650       2660       2670       2680       2690       2700
GCCTCTTCCT GCCCTTCACG CAGGGCGACA GCAACCGGCT GCGGCAGGCA CGCGGCACCG
      2710       2720       2730       2740       2750       2760
GCCTCGGATT GTCGATCTGC GCGCGTCTGT GCGAGCTGAT GAAAGGGCGC ATCGATCTGG
      2770       2780       2790       2800       2810       2820
AAAGCACCGT CGGCGTGGGA ACCCGGATCG CGGTGACACT GCCGCTCGGC GTGTCGGAGG
      2830       2840       2850       2860       2870       2880
CCGATTCGAG CGATGCGTAC TGGACGCTTC CGTATCGGCG CGTGGCCGTG CTCGGTCGCG
```

Fig 4.14 continued

2890 2900 2910 2920 2930 2940

CACAGGAAAA TCTCGAGTGG CTGGCCAACC TGTTCGACCC GGGCGTCACG GCCGTGACGG

2950 2960 2970 2980 2990 3000

CTTTCTCGCG GCCGGCCGAG CCGATCGATG CGCACGCGCA CGATTTCCTG ATGGTCACCG

ECORI 3010 3020 3030 3040 3050 3060

ACGAATTCTC GCCGGCCGAG GTGCTGCCGT GGTGGAGGCG GCCGGACTCG ATCGTGTGGG

3070 3080 3090 3100 3110 3120

TCGGGCAGGC CGGCCCGCTC GTGCCGAGAC GGCGCGACGA CGGCGGAGTG GAAATCAGCA

3130 3140 3150 3160 3170 3180

TGTATAGCCT CGCGGGGCTG AAATCCGCGA CTCACATGCT CGCGGCCGGC CGCACGGCGC

3190 3200 3210 3220 3230 3240

TCGCCGAAGC GGGGCACGAG CCGCCGGGAG CCGAGGCGGG AATGACGGTG CTGATCGCCG

3250 3260 3270 ***Sau3A*** 3280 3290 3300

AGGACAATCT GCTCAACCGC AGCCTGCTGC TCGATCAGCT GACGACGCTG GGCGTGCGGG

3310 3320 3330 3340 3350 3360

TCATCGAGGC GAAGAACGGC GAGGAGGCGC TCGCGTTGCT GTTGAAGGAG CCGGTGGACG

3370 3380 3390 3400 3410 3420

TCGTGATGAC CGACATCGAC ATGCCGATGA TGGACGGTTT CCAGTTGCTC GCCGAGATGA

3430 3440 3450 3460 3470 3480

GGCGGCTCGG CATGACGATG CCGGTGTACG CGGTGAGTGC GAGCGCGCGG CCGGAAGATG

3490 3500 3510 3520 3530 3540

TGGCGGAAGG CCGGGCGCGC GGCTTTACCG ACTATCTCGC GAAGCCGGTT TCGCTCGAGC

3550 3560 3570 3580 3590 3600

GGCTCGAGAC GGTGGTACGC GCATGTTGCA GCGCGCCGGC GGGCGCGCGC GCCGACGAAG

3610 3620 3630 3640 3650 3660

ACGCGCAGGA CGAACTGCCG GGCCTACCCG ACGTGCCGCC CGCCTATGCG AGCGCGTTCG

3670 3680 3690 3700 3710 3720

TCGCGCAGGC CGGCAGCGAA ATCGCGGAAT TCGACGCGAT CCTGCGCGAA CGCGCCTGCC

3730 3740 3750 3760 3770 3780

GAAACTGCGG CGGTGGCTGC ACGGCGTATC GGGCGGCATC GCGGTCCTCG GGCCTTCCGC

3790 3800 3810 3820 3830 3840

ACTGCATGAG CAATGCCAGG AGCTTCGAGC CTACGCGCGC GAATCCGGCG AATGGAATCG

3850 3860 3870 3880 3890 3900

CGAAATCGAA CTGCAGGCGC TGGCCATTCG GAACGCGCTG GAGCGAATGG TCGCGGCGCT

3902

GACGAGCGCGTGATCGGCGCGGGCGCGGCCCTGCCGGAGCGGGCCCGCGCCCGCCATGGAC

GGCGCGGCGCGGGGGTGTGCGCGGCGTCGCACGTTCCGGCGCGGATTCGAAGATTTGTGATTC

stem-loop structure (**Fig 4.15**). This might act as a strong transcriptional terminator for the putative gene.

The absence of secondary structure around the stop codon of ORFa suggests that transcription does not terminate at that point. It is known that secondary structure around the termination site greatly influences the efficiency of transcriptional termination. In addition, the lack of an obvious promoter region for ORFb suggests that transcription may be under the control of the promoter region proposed to precede ORFa. Together these features suggest that both ORFs may be transcribed as a polycistronic unit under the control of a single promoter.

Analysis of the DNA base composition showed that the individual base content was A, 16.31%, C, 31.32%, G, 35.24% and T, 17.12%. Hence, the overall codon usage (**Table 3**, Appendix IV) was characterised by a strong preference for C or G in the third positions, which are the positions most free to drift toward the overall genomic mol% G + C content. The mol% G + C was 66.5%.

4.4.2 Deduced amino acid sequence analysis

Translation of ORFa revealed that it encodes a protein of 221 amino acids (**Fig 4.16**) with a predicted molecular mass of 23,860 kDa and pI of 7.82. The secondary structure prediction indicated that 47% of the residues should form α-helices, 12% should be in the β sheet configuration and 41% forming loops (**Fig 4.17**). The amino acid composition as deduced from the translated nucleotides is shown in **Table 4.5**. The translated protein was found to be rich in the nonpolar hydrophobic residues valine (12.2%), leucine (11.3%) and alanine (9.1%).

The predicted amino acid sequence was used to conduct a BLAST search of the amino acid sequences that are deposited in public databases. Protein sequence analysis revealed strong homology throughout the protein to the two-component response regulatory proteins of many

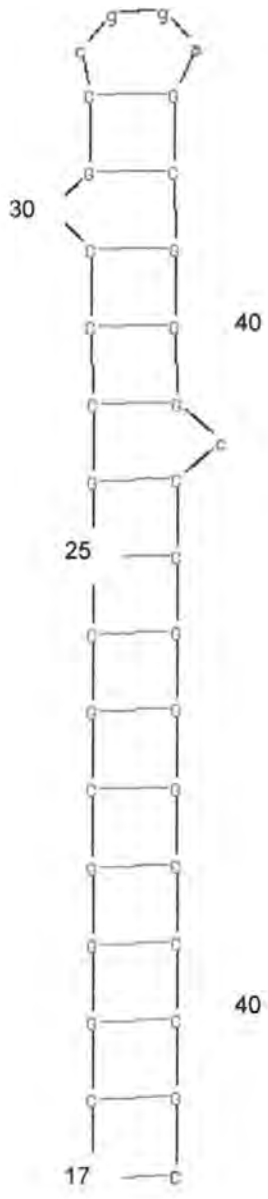


Fig 4.15: Predicted secondary structure of transcription termination formed by the inverted repeat sequence at the 3' end of the *mrg* locus; stem-loop structure. Numbers indicate position relative to the *mrgS* stop codon.

MSNVALHTRKVVVA**DD**HPIVLRAVTDYVNSLPGFHVVASVSSGDALLSAMREQEVNLVVT
DFTMHQANDDKDGLRLISHLMRAYERTPIIVFTMLTNSGVISQLCRMGVAGLVGKEEEIA
ELGRVCVSVARGVSQSLSPGMAHRLAAVGSIRPGEAAFNALTPKELEVVRLFTGG**MSLTD**
IARTLNRSLGTVSTQKRSAMR**K**LHVDTNVDLINCAREQGLL*

Fig 4.16: The deduced amino acid sequence of ORFa (MrgR protein). Conserved amino acids in other response regulator proteins are in bold. Putative Helix-Turn-Helix DNA binding motif is in red colour. Conserved residues in response regulatory proteins including D, T and K are highlighted and underlined.

Table 4.5: Amino acid residue composition of MrgR (%)

A: 9.1	C: 1.4	D: 5.0	E: 5.0	F: 2.3
G: 7.2	H: 3.2	I: 4.1	K: 2.7	L: 11.3
M: 4.1	N: 4.1	P: 2.7	Q: 2.7	R: 7.7
S: 8.1	T: 6.3	V: 12.2	W: 0.0	Y: 0.9

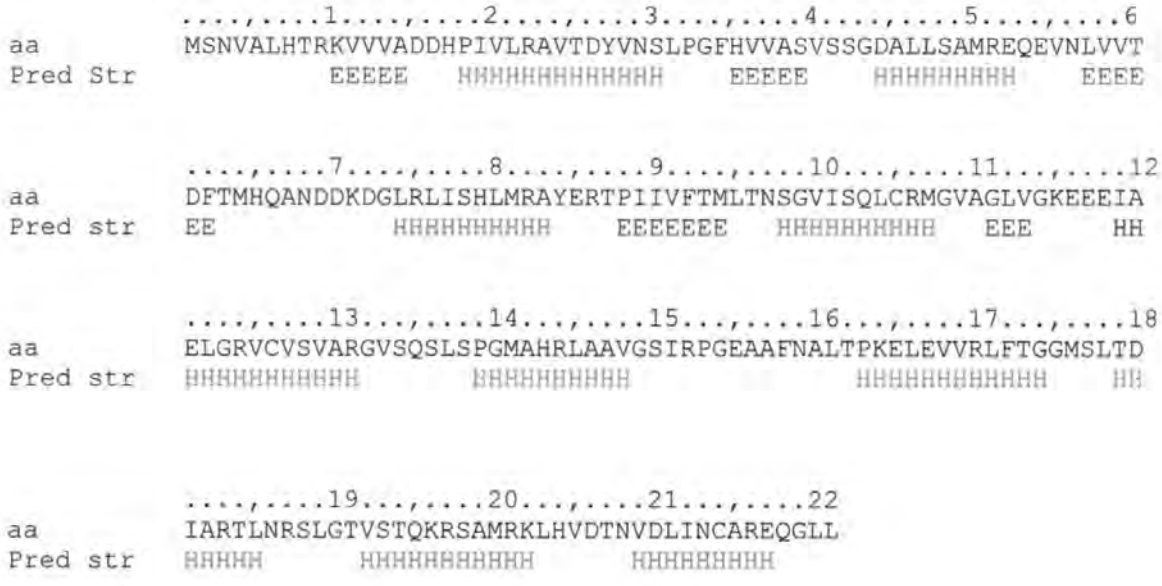


Fig 4.17: The predicted secondary structure of MrgR. H represents helix, while E is extended sheet.

different bacterial species. On multiple alignment (**Fig 1**, Appendix IV), the highest scores obtained were with response regulator of capsular synthesis (RcsB) in *Erwinia amylovora*, *E. coli*, *Proteus mirabilis* and *S. typhi* exhibiting 68.9, 68.5, 68.4 and 68.1% homology. The most significant similarities are summarised in **Table 4.6**. On the basis of this similarity the protein was designated MrgR (for melioidosis agent regulatory gene regulator). The C-terminal domain of the protein contained a typical helix-turn-helix motif (position 156 to 177) common to response regulators of the FixJ or LuxR family.

Comparison of the MrgR protein sequence with other response regulatory proteins revealed that it possesses the conserved residues characterised to this family of proteins. A conserved aspartic acid residue (D) present in all aligned proteins (position 61) is thought to be the phosphorylation site in many of the regulatory proteins that is important for regulatory activation. In addition, several other conserved motifs were identified in MrgR. These include a pair of aspartate residues (DD) at positions 15 and 16 at the N-terminal end, a threonine (T) at position 191, and a lysine residue (K) at position 202 at the C-terminal end. All of these residues may contribute to the acidic pocket for the phosphorylation site.

Translation of ORFb shows that it encodes a protein of 1078 amino acids (**Fig 4.18**) with a predicted molecular mass of 118,227 KDa and a theoretical pI of 6.04. The secondary structure prediction indicated that 44% of the residues should form α-helices, 14% should be in the β sheet configuration and 41% forming loops (**Fig 4.19**). The amino acid composition as deduced from the translated nucleotides is shown in **Table 4.7**. The translated protein was found to be rich in hydrophobic nonpolar residues of leucine (12.3%) and alanine (11.6%) as well as the hydrophilic positive amino acid arginine but with lesser amounts of cysteine (1.3%), tryptophan (1.3%) and histidine (1.9%).

Table 4.6: Homology of MrgR with the response regulatory proteins from other bacteria

Bacteria	Protein	Function	Similarity %	Identity %
Erwinia amylovora	RcsB	Capsule synthesis	68.9	31.1
Escherichia coli	RcsB	Capsule synthesis	68.5	29.4
Proteus mirabilis	RcsB	Capsule synthesis	68.4	29.4
Salmonella typhi	RcsB	Capsule synthesis	68.1	29.7
Erwinia amylovora	HrpY	Type III secretion system	65.8	23.9
Ralstonia solanacearum	VrsD	Virulence genes expression	64	23
Bordetella pertussis	BvgA	Putative virulence factor	63	28
Rhodococcus erythropolis	BpdT	?	61.4	26.6
Pseudomonas syringae	GacA	Global activator	61.4	23.5
Pseudomonas syringae	FixJ	Nitrogen fixation	61.4	23.5
Pseudomonas viridiflava	RepB	Global activator	61.4	23.5
Vibrio cholerae	GacA	Global virulence regulator	61	23.5
Pseudomonas stutzeri	NarL	Nitrate reductase	60.5	26
Pseudomonas aureofaciens	GacA	Global activator	60.4	23.1
Paracoccus denitrificans	MoxX	Methanol utilization control	60	25
Salmonella typhimurium	UhpA	Hexose phosphate uptake	59.6	25
Escherichia coli	NarP	Nitrate/nitrite regulator	59	24.8

MQGLLQELDGSPLRKFYSLESNLKRERRVFTIVIVLLVCAALSIAAMTVTGLFQTAFRQE
EQSARIHEKEVVDVFLQRRMMLTTASLVLQLRMNGAPSALNVPAPNACTPMAHNVRDDAI
LRESCDYTVQLLANSGQTPSVEMVTADGSVGYGYLFPTGDLSALRSSTPSELVSAVLERY
GKRGLDPLEAARKKRILWFAVGRGGRGEELHMIGASVVFKDERLYALVLTSVDLHSLVSP
IERAGRVQQPVVVDSEGVPLVNADDAETVRKVDGRLAGQQDGLYHWIPGFGWALRRPAPF
SGFGHMTYLLPLDLQLRSMRYELSLVGGATLVLIVLLFVAFRYWNYRFLTRIYEEASRAL
ESEMLNHLLVHATPVGLCIVRRATLEIVVANPIARTMLGLRLSDRHLPQELLSAFESSLA
EQDTQSDDARIFQFPFTLSRAGHAAVHIEITYAPAMLNAREVFFCAITDMTAHHQAEILL
REAKLTSDAAAKAKVAFFASMSHEIRTPLSSLVGNIELIARGPLAPEQQARVKAMETSAR
GLMQIVNDVLDFSKIDVGELSLMEEWSNIAELLDRLALSHAPLATQQGLKFYMVFDRSLP
ARLYFDPIRVSQIVNNLLSNALKFTPSGKIVLRAGWRAGALEISVTDSGIGIPDDLKHRL
FLPFTQGDSNRLRQARGTGLGLSICARLCELMKGRIDLESTVGVGTRIAVTLPLGVSEAD
SSDAYWTLPYRRVAVLGRAQENLEWLANLFDPGVTAVTAFSRPAEPIDAHAHDFLMVTDE
FSPAEVLPWWRRPDSIVWVGQAGPLVPRRRDDGGVEISMYSLAGLKSATHMLAAGRTALA
EAGHEPPGAEAGMTVLIAEDNLLNRSLLLDQLTTLGVRVIEAKNGEEALALLLKEPVDVV
MTDIDMPMMDGFQLLAEMRRLGMTMPVYAVSASARPEDVAEGRARGFTDYLAKPVSLERL
ETVVRACCSAPAGARADEDAQDELPGLPDVPPAYASAFVAQAGSEIAEFDAILRERACRN
CGGGCTAYRAASRSSGLPHCMSNARSFEPTRANPANGIAKSNCRRWPFGTRWSEWSRR*

Fig 4.18: The deduced amino acid sequence of ORFb (MrgS protein). Two potential membrane-spanning hydrophobic stretches located at the N-terminal region of the protein are in red colour. Conserved residues in HPKs including, the H residue corresponding to putative autophosphorylation site, N, G_1 (DTGVG), F, and G_2 (GTGLG) boxes are highlighted and underlined.

Table 4.7: Amino acid residue composition of MrgS (%)

A: 11.6	C: 1.3	D: 5.1	E: 6.4	F: 3.6
G: 6.9	H: 1.9	I: 4.0	K: 2.1	L: 12.3
M: 3.0	N: 2.6	P: 5.2	Q: 3.1	R: 8.3
S: 7.0	T: 4.8	V: 7.5	W: 1.3	Y: 2.0

Fig 4.19: The predicted secondary structure of MrgS. H represents helix, while E is extended sheet.

```
          ....,....1....,....2....,....3....,....4....,....5....,....6
aa        MQGLLQELDGSPLRKFYSLESNLKRERRVFTIVIVLLVCAALSIAAMTVTGLFQTAFRQE
Pred_str   HHHHHH        E HHHHHHHHH  EEEEEEEHHHHHHHHHHHHHHHHHHHHHHH

          ....,....7....,....8....,....9....,....10...,....11...,....12
aa        EQSARIHEKEVVDVFLQRRMMLTTASLVLQLRMNGAPSALNVPAPNACTPMAHNVRDDAI
Pred_str  HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH                      HHHH

          ....,....13...,....14...,....15...,....16...,....17...,....18
aa        LRESCDYTVQLLANSGQTPSVEMVTADGSVGYGYLFPTGDLSALRSSTPSELVSAVLERY
Pred_str  HHHHHHHHHHHHHH     EEEEEEE     EEEEE             HHHHHHHHHH

          ....,....19...,....20...,....21...,....22...,....23...,....24
aa        GKRGLDPLEAARKKRILWFAVGRGGRGEELHMIGASVVFKDERLYALVLTSVDLHSLVSP
Pred_str  HH    HHHHHHHHHHHHHH         EEEE  EEEEE    EEEE EEE    E

          ....,....25...,....26...,....27...,....28...,....29...,....30
aa        IERAGRVQQPVVVDSEGVPLVNADDAETVRKVDGRLAGQQDGLYHWIPGFGWALRRPAPF
Pred_str           EEEEE     EE   HHHHHHHHHHHH      EE

          ....,....31...,....32...,....33...,....34...,....35...,....36
aa        SGFGHMTYLLPLDLQLRSMRYELSLVGGATLVLIVLLFVAFRYWNYRFLTRIYEEASRAL
Pred_str     HHHHE         HHHHEEEE   HHHHHHHHHHHHHHHHHHHHHHHHHHHHHH

          ....,....37...,....38...,....39...,....40...,....41...,....42
aa        ESEMLNHLLVHATPVGLCIVRRATLEIVVANPIARTMLGLRLSDRHLPQELLSAFESSLA
Pred_str  HHHHH   HHHEEHHHHHHHHHHHHHHHHH  HHHHH          HHHHHHHHHHHHH

          ....,....43...,....44...,....45...,....46...,....47...,....48
aa        EQDTQSDDARIFQFPFTLSRAGHAAVHIEITYAPAMLNAREVFFCAITDMTAHHQAEILL
Pred_str  H       EEEEEEEEEEEE    E EHHHH  HHHHHHHHHHHHHHHHHHHHHHHHHH

          ....,....49...,....50...,....51...,....52...,....53...,....54
aa        REAKLTSDAAAKAKVAFFASMSHEIRTPLSSLVGNIELIARGPLAPEQQARVKAMETSAR
Pred_str  HHHHHHHHHHHHHHHHHHHHHHHH      HHHHHHHHHH    HHHHHHHHHHHHHH

          ....,....55...,....56...,....57...,....58...,....59...,....60
aa        GLMQIVNDVLDFSKIDVGELSLMEEWSNIAELLDRLALSHAPLATQQGLKFYMVFDRSLP
Pred_str  HHHHHHHHHHHHH     EEE    HHHHHHHHHHHHHHHHHHH    EEEE

          ....,....61...,....62...,....63...,....64...,....65...,....66
aa        ARLYFDPIRVSQIVNNLLSNALKFTPSGKIVLRAGWRAGALEISVTDSGIGIPDDLKHRL
Pred_str          HHHHHHHHHHHHHH     EEEEEE      EEEEEE       HHHHHHH
```

Fig 4.19 continued

```
          ....,....67...,....68...,....69...,....70...,....71...,....72
aa        FLPFTQGDSNRLRQARGTGLGLSICARLCELMKGRIDLESTVGVGTRIAVTLPLGVSEAD
Pred_str  HHHH               HHHHHHHHHHHH   EEEE      EEEEEEE

          ....,....73...,....74...,....75...,....76...,....77...,....78
aa        SSDAYWTLPYRRVAVLGRAQENLEWLANLFDPGVTAVTAFSRPAEPIDAHAHDFLMVTDE
Pred_str           HHHHHHHHHHHHHHHHHHHH    EEEEEE            EEEEE

          ....,....79...,....80...,....81...,....82...,....83...,....84
aa        FSPAEVLPWWRRPDSIVWVGQAGPLVPRRRDDGGVEISMYSLAGLKSATHMLAAGRTALA
Pred_str                EEEEE             EEEEE     HHHHHHHHHHHHHHHH

          ....,....85...,....86...,....87...,....88...,....89...,....90
aa        EAGHEPPGAEAGMTVLIAEDNLLNRSLLLDQLTTLGVRVIEAKNGEEALALLLKEPVDVV
Pred_str  H          EEEEEE   HHHHHHHHHHHHH   EEEE   HHHHHHHHHH    EE

          ....,....91...,....92...,....93...,....94...,....95...,....96
aa        MTDIDMPMMDGFQLLAEMRRLGMTMPVYAVSASARPEDVAEGRARGFTDYLAKPVSLERL
Pred_str  EEE        HHHHHHHHHH     EEEEE    HHHHHHHHHH      E    HHHH

          ....,....97...,....98...,....99...,....100..,....101..,....102
aa        ETVVRACCSAPAGARADEDAQDELPGLPDVPPAYASAFVAQAGSEIAEFDAILRERACRN
Pred_str  HHHHHHHHHHH                          EEE       HHHHHHHHHH

          ....,....103..,....104..,....105..,....106..,....107..,....108
aa        CGGGCTAYRAASRSSGLPHCMSNARSFEPTRANPANGIAKSNCRRWPFGTRWSEWSRR
Pred_str      EEEEEEE                            HHHHHH
```

Comparison of the protein sequence with other proteins in the database showed that it shares homology with a variety of two-component histidine kinase environmental sensors. The similarity of the protein to other sensor regulators is summarised in **Table 4.8**. Sequence alignments (**Fig 2**, Appendix 2) showed that the protein, distributed over its entire length, exhibits the highest score, over 57% similarity with 22% identity, to RcsC of *E. coli*. RcsC is known to be the sensor regulator of capsule synthesis in a number of bacterial species including *Erwinia amylovora*, *Proteus mirabilis* and *S. typhi*. It may perform its function by sensing environmental stimuli, hence the protein was designated MrgS (for melioidosis agent regulatory gene sensor).

Examination of the hydropathy profile of the N-terminal putative sensory domain of MrgS protein (**Fig 4.20, Table 4.9**) revealed two stretches of hydrophobic amino acids of sufficient length to span the cell membrane, 25 and 27 amino acid residues, at positions 29-53 and 286-312 respectively. These putative transmembrane sequences are characteristic of the N-terminal region of sensor proteins. Furthermore, the C-terminal portion of MrgS protein contains the five blocks of conserved functional subdomains (Stock *et al.*, 1995) characteristic of the histidine protein kinase family. A conserved histidine that is the putative site of phosphorylation at position 502, and the N, G_1 (DTGVG), F, and G_2 (GTGLG) boxes at positions 619, 646-650, 660 and 676-680 respectively. These conserved motifs presumably form a nucleotide-binding surface within the active site. Additionally, one or more of the several conserved aspartic acid residues (D) that are present in many sensor proteins including, MrgS, at sites 859, 897, 902, 909 and 948 may correspond to the extra response regulator domain that is a characteristic of hybrid histidine kinases.

Table 4.8: Homology of MrgS with the sensor regulatory proteins from other bacteria

Bacteria	Protein	Function	Similarity %	Identity %
Escherichia coli	RcsC	Capsule synthesis	56.7	22.3
Vibrio harveyi	LuxQ	Sensor kinase	52.3	16
Salmonella typhi	SpiR	Virulence regulator	49.4	20
Escherichia coli	TorR	Sensor kinase	49.2	18.8
Vibrio cholerae	VieS	Sensor kinase	48.3	18
Rhodobacter spheroides	DmsO/TmaO	Sensor kinase	46.5	18.3
Klebsiella pneumoniae	KvgS	Virulence sensor kinase	46	15.7
Pseudomonas syringae	GacS	Sensor kinase	45.8	13.5
Pseudomonas syringae	LemA	Sensor kinase	45	13
Pseudomonas tolassi	RtpA	Extracellular protease	44.6	12.5
Erwinia carotovora	RpfA	Sensor kinase	44.3	14
Bordetella bronchiseptica	BvgS	Virulence sensor kinase	44.2	15.7
Escherichia coli	ArcB	Aerobic respiration control	44.2	16.3
Pseudomonas fluorescens	ApdA	Antifungal antibiotic	44	13.3
Pseudomonas tolassi	-	Multiple function regulator	44	12
Escherichia coli	EvgS	Virulence sensor kinase	43.9	18.2
Pseudomonas aeruginosa	LemA	Sensor kinase	43.7	12.3
Escherichia coli	BarA	Sensor kinase	43.3	12
Escherichia coli	PhoR	Alkaline phosphate sensor	37.6	15.5
Proteus mirabilis	RcsC	Capsule synthesis	34.3	15.2
Streptomyces coelicolor	-	Sensor kinase	32.5	13

Table 4.9: The predicted cellular location of MrgS domains

Amino acid residues	Predicted area
1-28	Inside region
29-53	Membrane helix
54-287	Outside region
286-312	Membrane helix
313-1078	Inside region

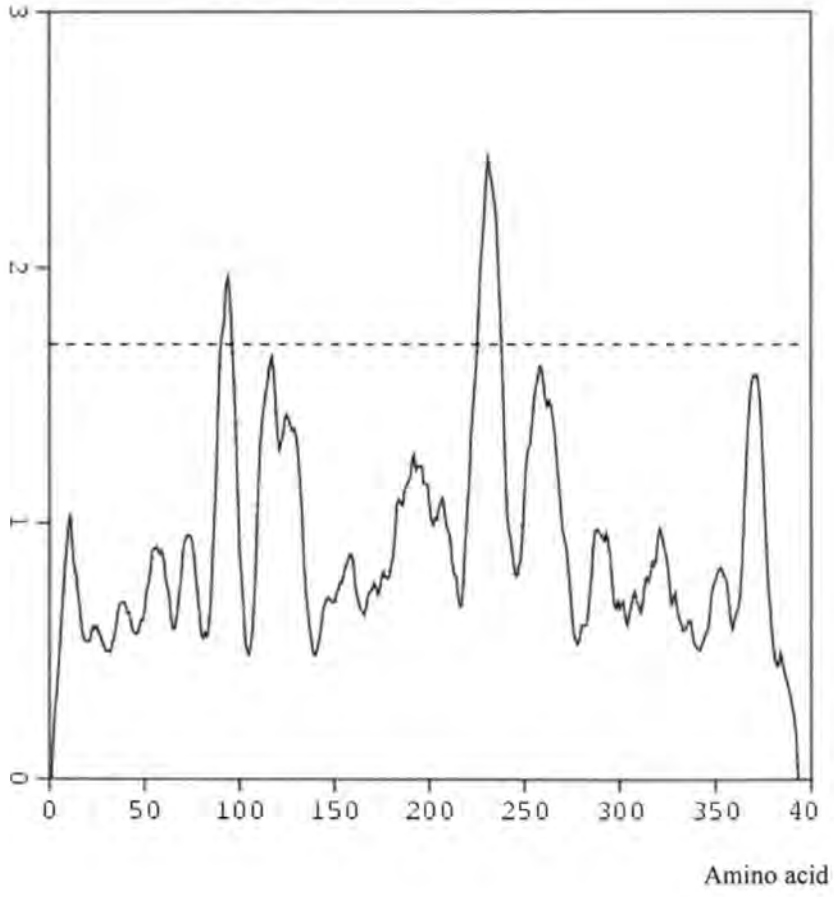


Fig 4.20: A hydropathy prediction showing proposed transmembrane regions in MrgS. Values above the horizontal dashed axis indicate hydrophobic character and those below the axis indicate hydrophilic character.

4.5 DISCUSSION

This chapter describes the cloning and characterisation of novel two-component regulatory genes from *B. pseudomallei* strain 204, which necessitated the construction and screening of three gene libraries using λGEM-11, λZAP and λGT-11. The λGEM-11 library had been used to successfully screen for the *B. pseudomallei phoP* gene, encoding a response regulator involved in two-component signal transduction (Sangiambut *et al.*, 2000).

B. pseudomallei is widely distributed in the tropical environment and is an opportunistic pathogen of humans and animals. It demonstrates a remarkable ability to survive under a wide range of environmental conditions both *in vivo* and *in vitro*, although little is known about the genetic determinants that control the expression of the genes that provide an adaptive response to environmental changes. The capacity to adapt to changing environmental conditions increases a cell's chances of survival. In bacteria, the sensory devices gathering the environmental information often involve two-component regulatory system. Furthermore, many factors that enhance the capability to interact with host cells and to cause disease are under the control of two-component regulatory systems.

Various strategies have been employed by other researchers attempting to identify the genes encoding bacterial two-component systems, mostly based on the conserved amino acid motifs that are found among the family members. These include DNA hybridisation with degenerate oligonucleotide probes corresponding to the conserved amino acid sequence of the histidine protein kinase (HPK) G2 box (Osbourn *et al.*, 1990), PCR with degenerate primers derived from the conserved N and G2 boxes (Bayles, 1993; Parkinson, 1995; Stock *et al.*, 1995), and complementation of HPK-deficient *E. coli* mutants (Aiba *et al.*, 1993; Utusumi *et al.*, 1994). In the present study, PCR primers were designed on the basis of nucleotide sequence

homology with the response regulatory genes of other bacteria and consequently a putative two-component regulatory system from *B. pseudomallei* was identified and characterised. HPK and response regulator (RR) genes are frequently organised as an operon, and this is also true for the nucleotide sequences characterised in this work.

A PCR product of 234 bp was used as a hybridisation probe to screen genomic libraries, and a 4.3 kb DNA fragment was cloned and sequenced from *Eco*RI genomic digests. Computer-aided analysis of the nucleotide sequence identified one complete ORF, ORFa, and immediately downstream, another partial OFR reading in the same direction, ORFb, that extended beyond the sequence of the 4.3 kb region. PCR amplification from DNA extracted from λGEM-11 clones using specific primers complementary to the nucleotide sequences of partial ORFb and λGEM-11 arms, permitted the downstream sequencing of a further 268 bp of ORFb but not the entire coding region. However, screening a λZAP Express library resulted in the identification of an 8.4 kb DNA fragment that was subsequently found to possess the full length of ORFb, including the termination codon.

The putative genes encoded by ORFa and ORFb were designated *mrgR* and *mrgS* for melioidosis agent regulatory gene regulator and sensor, respectively. The mol% G + C of *mrgR* (64%) and *mrgS* (66.5%) are similar to the overall content of the *B. pseudomallei* chromosome (65.7%) (Songsivilai & Dharakul, 2000). In *E. coli*, the promoter sequences for many genes have been compiled and analysed (Harley & Reynolds, 1987). Some of the highly conserved bases that have been identified are 5'-TTGACA-3' around the -35 region and 5'-TATAAT-3' around the -10 region, with a usual spacing ranging from 15 to 21 nucleotides between the two conserved regions (Hawley & McClure, 1983). However, the spacing between the -10 region and the transcriptional start site can vary from 7 to 41 nucleotides (Shiuan et al., 1994). Through

sequence comparison, two putative promoter regions are located upstream from the SD sequence associated with ORFa. It is not known which of these promoters is required for transcription of *mrgR*, however, the process may be enhanced by the presence of more than one promoter region. Nevertheless, further experimentation is required in order to decisively determine the promoter for these genes. The initiation codon of ORFb is preceded by a putative SD sequence that is located within the final 14 bases of ORFa. However, there was no evidence of a well-defined promoter region for ORFb or a transcriptional termination sequence for ORFa. Moreover, the stop codon of the ORFa overlaps with the start codon of ORFb by a single nucleotide, TGATG. It may be possible that both genes are transcribed as a polycistronic unit under the control of the promoter upstream of ORFa. In bacteria, genes with related functions are often located together in a group known as an operon. This group of genes has a single promoter site, and is transcribed into a single polycistronic mRNA molecule. It has been suggested that if the termination of translation of a gene is very close to or even overlaps the translation-initiation site of the next gene, then expression of these genes can be coupled at the translational level (Das & Yanofsky, 1984). DNA sequencing of *E. coli* operons has identified several overlaps for *trp* (Oppenheim & Yanofsky, 1980), *his* (Barnes & Tuley, 1983), *gal* (McKenney *et al.*, 1981), and *ompR/envZ* (Forst & Inouye, 1988) and this also occurs in the *irl* operon of *B. pseudomallei* (Jones *et al.*, 1997). The single base overlap between ORFa and ORFb suggests that translational termination of *mrgR* could immediately initiate translation of *mrgS*. Because bacterial genes with related functions often are grouped together in operons, it can be speculated that both *mrgR* and *mrgS* are involved in the regulation of a specific function. Furthermore, on the basis of strong homology with the sequences of genes from other bacteria, it is most likely that the function is that of a two-component regulatory system.

The translated sequences of ORFa and ORFb encode proteins with M_r of 23,860 and 118,267, respectively. Searches for similarity among protein sequences can identify well-conserved motifs as well as families of related proteins in which homology extends over one or several domains of proteins that possess similar functions (Pabo & Sauer, 1984; 1992). The deduced amino acid sequence of ORFa was substantially similar to numerous response regulator proteins. The sequence was most similar to RcsB, the regulator of capsular synthesis of gram-negative bacteria such as *Erwinia amylovora*, *E. coli*, *Proteus mirabilis*, and *S. typhi*. On the other hand, the protein translated from ORFb was most similar to bacterial sensor regulators especially RcsC of *E. coli*. In all known rcs systems, the *rcsB* gene is located adjacent to the *rcsC* gene and is transcribed in the opposite direction. Furthermore, an intergenic region exists between *rcsB* and *rcsC* in *E. coli* (196 bp) (Stout & Gottesman, 1990) and also in *S. typhi* (102 bp) (Virlogeux *et al.*, 1996). In contrast, *mrgR* and *mrgS* seem to be physically linked and are transcribed in the same direction.

Response regulators are often transcription factors whose affinity for DNA is modulated by phosphorylation. X-ray crystallography studies of the *S. typhimurium* (Stock *et al.*, 1989) and *E. coli* (Volz & Matsumura, 1991) CheY and NarL (Baikalov *et al.*, 1996) proteins confirmed the structural conservation in the N-terminal domain. They revealed a barrel-like arrangement of five sets of alternating β strands and α helices, in the sequence starting with β_1 and ending with α_E. The β strands align to form a hydrophobic inner core with the helical segment wrapped around the outside of the molecule. There is a correlation between structure and function, in that the residues that have an important role in phosphorylation are located in an acidic pocket at one end of the barrel. These include a pair of aspartates near the N-terminus (Bourret *et al.*, 1990; Brissette *et al.*, 1991), a lysine near the C-terminus (Lukat *et al.*, 1991;

Stewart, 1993), and a centrally located aspartate that is the site of phosphorylation, positioned in the loop between β_3 and α_C. (Stock *et al.*, 1989; Volz & Matsumura 1991). These and other characteristic sequences are present in over 90 known receiver-containing proteins, suggesting that all receiver modules could be α/β barrels like CheY (Stock *et al.*, 1989; Parkinson, 1993).

The site of phosphorylation in many of regulatory proteins is an aspartic acid residue residing 60-80 amino acids from the N-terminus (Volz, 1993). The corresponding aspartates in NRI, PhoB (Makino *et al.*, 1994), and VirG (Jin *et al.*, 1990) have also been shown to be phosphorylated. In CheY, the critical role of the three conserved aspartates was shown experimentally. Substitution of Asp-12, Asp-13, or Asp-57 with asparagine, glutamate, or lysine suggests that residue 12 must be an acidic amino acid and residues 13 and 57 must both be aspartates in order to maintain CheY function *in vivo* (Bourret *et al.*, 1990). MrgR possesses several candidate aspartate residues that could possibly be involved in phosphorylation at positions 61, 68, 69 and 71. However, phosphorylation at a specific residue requires the presence of an acidic pocket in the secondary structure of the protein next to the phosphorylation site. The acidic pocket is formed by a β-sheet with two further aspartate residues at its C-terminus end. This structure is located at the N-terminus of the protein in question. Such a putative β-sheet, VVVADD, is found within the first 20 amino acid residues of MrgR. Multiple alignment of MrgR with CheY (**Fig 4.21**), the protein involved in flagellar control in *E. coli*, and that is known to be phosphorylated at aspartate 57, confirmed the presence of this phosphorylation motif.

Response regulators have been classified into five families according to sequence homologies in their DNA-binding domains (Volz, 1993; Stock *et al.*, 1995). On the basis of sequence similarity, the MrgR protein should be assigned to the FixJ or LuxR family, which

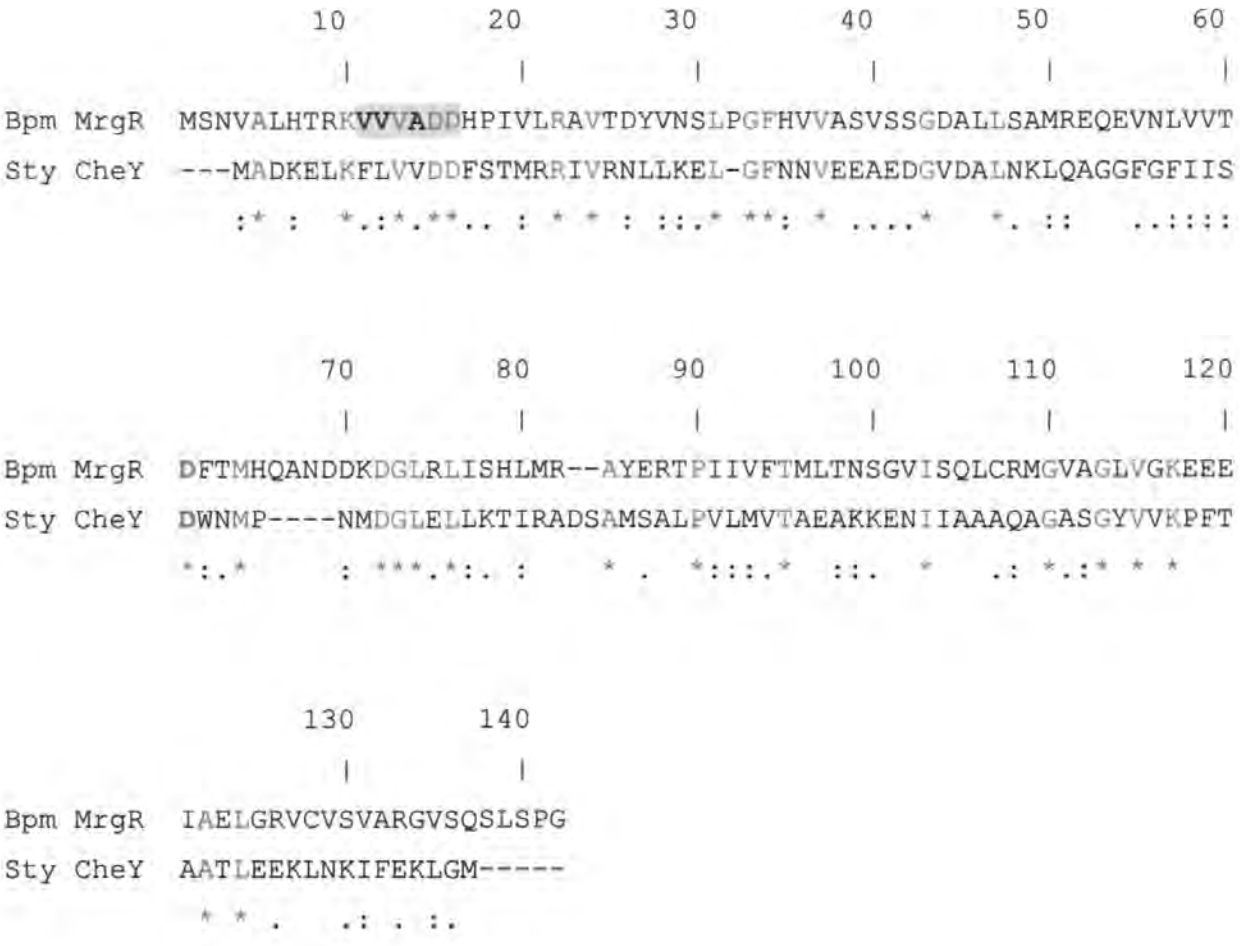


Fig 4.21: Alignment of the N-terminus end of MrgR and *S. typhimurium* CheY response regulator. The proposed acidic pocket of MrgR is highlighted and the putative phosphorylated residue, aspartate 'D', is in red bold. Stars indicate identical residues.

includes NarL of *E. coli*. Proteins from this family are cytoplasmically located and are thought to bind to specific promoter sequences located upstream of the regulated genes. Consistent with this, MrgR is predicted to be a cytoplasmic protein, based on hydrophobicity profiles. Furthermore, sequence analysis of MrgR predicts that the C-terminus of this protein possesses a helix-turn-helix (H-T-H) DNA binding domain, which further supports a role for MrgR in transcriptional regulation. Proteins that bind to DNA usually have an important role in the regulation of gene expression and the control of DNA replication. The H-T-H domain projects away from the rest of the protein thereby allowing one of the helices, the recognition helix, to bind in the major groove of the DNA, while the other helix acts to stabilise the binding structure. On the basis of what is known of response regulator proteins, MrgR is likely to regulate a variety of genes that remain to be identified.

According to sequence similarity, MrgS seems to belong to a large family of sensor histidine protein kinases (HPKs). Sensor kinases are generally integral membrane proteins that respond to specific environmental signals. Most sensor proteins are located in the cytoplasmic membrane with the transmitter portion of the molecule projecting into the cytoplasm. More than one hundred HPKs have now been sequenced from many species of bacteria and the organisation of sequence motifs has been well established, with only a few specific residues being conserved in almost all members of the family. The N-terminal portion of the HPK functions as an input domain, detecting environmental stimuli directly or interacting with an upstream receptor, while at the C-terminal end of the protein is the transmitter module. In a typical sensor kinase, this catalytic domain is generally 240 amino acids long and contains several blocks of residues that are conserved among HPKs. Amino acid replacements at various sites within these conserved regions curtail or eliminate autokinase activity (Oosawa *et al.*,

1988; Liu, 1990; Yang & Inouye, 1991). Structural predictions for the N-terminal half of the MrgS protein suggest the presence of two hydrophobic domains separated by a 234-amino-acid loop that probably extends into the periplasm, perhaps to sense a specific signal. The C-terminal half of MrgS appears to be cytoplasmic and contains the conserved motifs that are generally considered to be specific to most HPKs. These include the autophosphorylation site (H box) at which a conserved histidine (H^{502}) is phosphorylated following signal activation. In addition, four blocks of conserved amino acids are also present, constituting the N^{619} (asparagine), $G1^{646-650}$, F^{660} (phenylalanine) and $G2^{676-680}$ boxes. The G1 and G2 boxes are glycine-rich sequences (DXGXG and GXGXG) that bear a resemblance to the nucleotide binding motifs of other proteins and are thought to be the site of nucleotide binding, kinase and phosphatase activities. Hence, the putative topology (**Fig 4.22**), together with the conserved motifs strongly suggests that MrgS is involved in a signal transduction pathway. Comparisons of the amino acid sequences adjacent to the invariant histidine residue permit the separation of HPKs into four distinct subfamilies (Forst & Roberts, 1994). MrgS was found to belong to the group I subfamily that is characterised by the consensus sequence HEIRTPL (**Table 4.10**).

Several aspartic acid residues, which are highly conserved in most of the aligned HPKs, were identified close to the C-terminal end of the MrgS protein. The presence of one of these residues may suggest that MrgS belongs to a separate group of more sophisticated signal transduction proteins, the hybrid kinases, that contain domains resembling both the histidine kinases and the response regulators in one primary sequence. The response regulator domain of these proteins is thought to function as a pseudosubstrate inhibitor of the kinase domain. Following an appropriate stimulus, the kinase domain undergoes autophosphorylation, and then transfers this phosphate to the regulator domain, thus relieving inhibition. Phosphorylation of

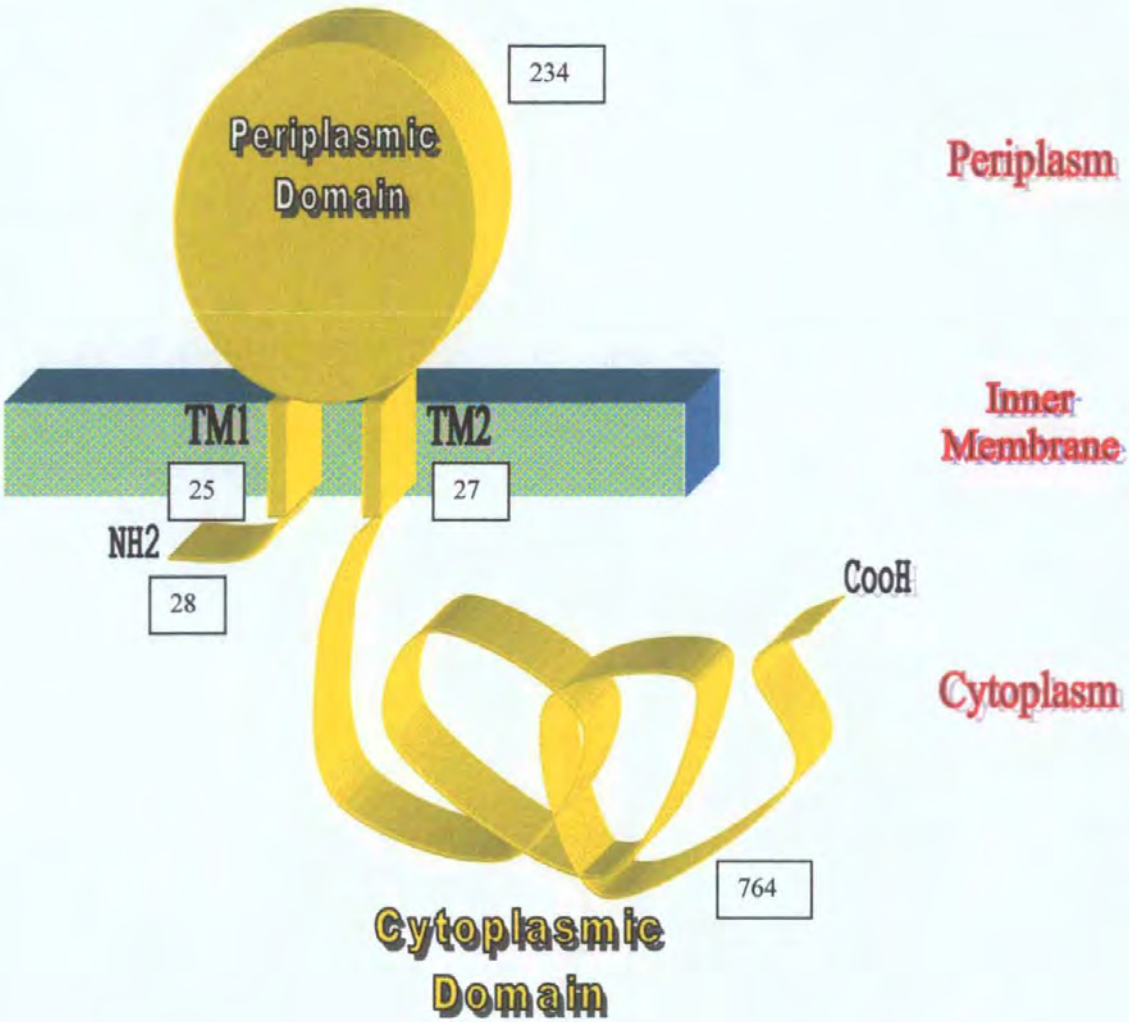


Fig 4.22: Schematic representation of the predicted topology of the MrgS protein. TM indicates a transmembrane region and numbers represent the length of each protein fragment.

Subfamily	Representative member	Sequence 1 2 3 4 5 6 7
Group I	EnvZ	H - ⊖ - ○ - ⊕ - ● - P - L
Group II	VirA	H - ⊖ - ○ - N - ● - P – L
Group III	CheA	H - ▲ - ○ - ⊕ - ● - ● - ●
Group IV	DegS	H - ⊖ - ● - ● - ● - ● - ●

Table 4.10: Consensus sequence comparison of region I of the histidine kinase sensor molecules. Capital letters indicate invariant positions: H, His; P, Pro; L, Leu; N, Asn. Symbols indicate positions that contain conserved residues as follows: ⊖ represents Asp or Glu; ○ represents Ile, Leu, Phe, or Met; ⊕ represents Lys or Arg; ● represents positions that are variable. The individual members of the subfamilies were derived from Stock *et al.*, 1989; Parkinson, 1993. Group I subfamily: EnvZ, KdpD, RcsC, CpxA, ArcB, BarA, PhoR, PleC, LemA, KinA, BvgS, VanS, NtrB, CreC. Group II subfamily: VirA, FixL, PgtB, and DctB. Group III subfamily: CheA and FrzE. Group IV subfamily: UhpB, DegS, and NarX. MrgS belongs to the group I histidine kinase sensors.

the true substrate can then take place. These ideas are largely based on work on ArcB, which controls the anaerobic repression of several operons in *E. coli* by modulating its downstream response regulator, ArcA (Alex & Simon, 1994).

The similarities between MrgR-MrgS and the RcsB-RcsC signal transduction system in other bacterial species suggests that the *B. pseudomallei* proteins may perform a similar function. The *rcs* genes that are involved in the regulation of colanic acid capsule synthesis have been well studied in *E. coli*. Two pathways that regulate the synthesis of colanic acid capsular polysaccharide have been recognised. Lon and RcsA compose one regulatory pathway, where Lon is an ATP-dependent protease that degrades several proteins, including RcsA, which is a positive regulator of capsule synthesis (Torres-Cabassa & Gottesman, 1987). RcsA is, therefore, limiting for capsule expression. The RcsB-RcsC two-component regulatory proteins provide the second pathway. RcsC is a membrane kinase that receives a stimulatory signal, such as temperature shock (Allen *et al.*, 1987), desiccation (Ophir & Gutnick, 1994), osmotic shock (Sledjeski & Gottesman, 1996), or perturbations in the structure of the outer cell membrane (Clavel *et al.*, 1996), and is converted to an active protein kinase (Gottesman & Stout, 1991). The activated RcsC protein may activate and deactivate RcsB through cycles of phosphorylation and dephosphorylation. The phosphorylated and activated RcsB protein may, in turn, stimulate transcription of the target *cps* genes (Stout & Gottesman, 1990; Gottesman & Stout, 1991, Stout, 1994). However, phosphorylated RcsB activates the transcription of *cps* genes to a maximum level only when RcsA is present (Stout *et al.*, 1991). Therefore, both pathways require RcsB for the high-level expression of *cps* genes in *E. coli* (Gottesman, 1985; Gupte *et al.*, 1997; Ebel & Trempy, 1999).

In addition, it has been proposed that the RcsB and RcsC proteins of *E. coli* perform

functions other than regulating capsular synthesis. Experiments by Gervais and Drapeau (1992) and Gervais *et al.* (1992) implied that RcsB has a role in the regulation of *ftsZ*, a gene encoding a septum protein that is considered to be essential for cell division, and proposed that RcsB increased the transcription of *ftsZ* thus accounting for the restoration of colony formation in *ftsZ* mutant cells. Consequently, Gervais *et al.* (1992) raised the possibility that *rcsB* is a global regulon involved in a variety of cellular functions, which would explain its occurrence in organisms that do not synthesise colanic acid. On the other hand, RcsC was believed to play a role in the maintenance of cell envelope integrity in *E. coli.* Homologs of Rcs proteins have been identified in *Erwinia* spp (Bernhard *et al.*, 1990; Poetter & Coplin, 1991), *K. pneumoniae* (Allen *et al.*, 1987; McCallum & Whitfield, 1991), and *S. typhi* (Houng *et al.*, 1992). Because these bacteria are normally found in very different environments, they probably do not respond to the same environmental signals; however, they do use the same molecular pathway(s) to stimulate expression of their respective capsule biosynthesis genes (Stout, 1996). In *S. typhi*, production of Vi antigen, the capsular polysaccharide that correlates with virulence, is partially controlled by the two-component regulatory system RcsB-RcsC and by the positive regulator TviA, which interacts with RcsB to promote the optimal transcription of the genes involved in Vi antigen synthesis (Virlogeux *et al.*, 1996).

CHAPTER 5

MOLECULAR CHARACTERISATION OF THE *mrgRS* LOCUS AND FLANKING REGIONS IN DIFFERENT ISOLATES OF *B. PSEUDOMALLEI* AND OTHER *BURKHOLDERIA* SPECIES

5.1 INTRODUCTION

B. pseudomallei has a wide geographical distribution and host range (sections 2.2.3, 2.2.9) and it is therefore of interest to examine the *mrgRS* locus for the presence of variation among isolates from different sources. Studies of genetic variations at specific loci may yield valuable information about the pathogenicity and epidemiology of bacterial pathogens (Watt & Dean, 2000). For example, an examination of the genome sequences of *Mycobacterium tuberculosis* and *Mycobacterium bovis* has revealed that although only very small differences exist in the gene structure of these pathogens this variation is sufficient to determine host specificity (Behr *et al.*, 1999). These small differences in genetic composition must be reflected in functional differences in gene expression either at the level of transcription, translation or post-translation.

The purpose of this chapter was to compare the *mrgRS* locus and flanking regions of 19 isolates of *B. pseudomallei* from a wide range of sources, including Asia (Thailand, Vietnam, Singapore, Malaysia, China), Africa (Burkina Faso, Madagascar), Europe (France), and Australia. A selection of clinical and environmental isolates was included. In addition, in attempting to understand the function of the *mrgRS* locus, DNA sequence analysis of the flanking regions was performed and the presence of open reading frames and potential genes was determined.

5.2 MATERIALS AND METHODS

The materials and methods for the work performed in this study have been previously described in the following sections: genomic DNA isolation (section 3.4.1), restriction enzyme digestion (section 3.4.5), agarose gel electrophoresis (section 3.4.6.1), nucleotide and protein sequence analysis (section 3.9), Southern blotting (section 3.10), labelling DNA probes (sections 3.10.1, 3.10.2), DNA sequencing (section 3.8), and PCR (section 3.13).

5.3 RESULTS

5.3.1 Nucleotide sequencing of the regions flanking the *mrgRS* locus

Sequence information for the *mrgRS* flanking regions was obtained from pMRG2 (section 4.3.10) and pMRG4 (section 4.3.15). One plasmid, pMRG2, contains the entire *mrgR* gene and ~2.3 kb of the *mrgS* gene in addition to the sequence of approximately 1.3 kb upstream from the *mrgR* initiation codon. The other plasmid, pMRG4 contains the remaining 0.9 kb distal portion of *mrgS* as well as ~ 4.3 kb of sequence downstream from the *mrgS* stop codon.

5.3.2 Sequence analysis of the *mrgRS* flanking regions

The assembled nucleotide sequence of the regions flanking the *mrgRS* locus was analysed according to the general rules for identifying open reading frames (ORFs) and prokaryotic promoters described in sections 3.9 and 4.4.1.2. The analysis included the identification of restriction endonuclease cleavage sites, ORFs, initiation codons, -35, -10 and +1 regions, Shine-Dalgarno sequences, transcriptional termination loops and stop codons. Consequently, a single ORF was found upstream of the *mrgRS* locus while 6 ORFs were present downstream of the *mrgRS* locus. A schematic presentation of the ORFs and the direction of transcription is shown in **Fig 5.1**, and the characteristics of the ORFs are summarised in **Table 5.1**. The amino acid sequences deduced for each ORF are shown in **Fig 5.2**, and summarised in **Table 5.2**, while a comparison between amino acid composition of each putative protein is summarised in **Table 5.3**. Deduced amino acids were evaluated, and only those of major interest will be discussed.

The single ORF upstream of *mrgRS*, ORFc, ends 297 nucleotides upstream from the initiation codon of *mrgR*. Thirty-two nucleotides downstream of the proposed TAG stop codon in the non-coding region, an inverted repeat capable of forming a stem-loop structure transcriptional termination, is found, with a stem of 7 bases. A free energy (ΔG) of -9 kcal,

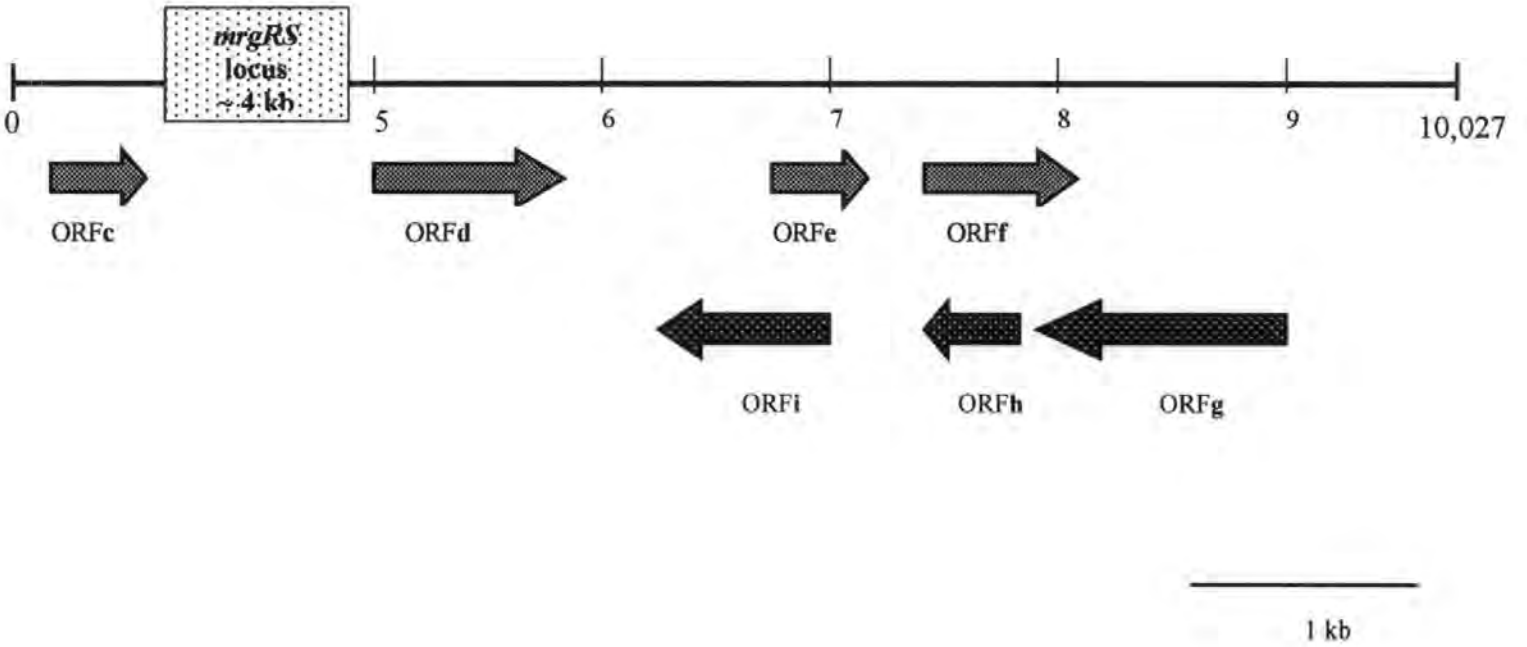


Fig 5.1: Schematic presentation of the flanking regions of the *mrgRS* locus of *B. pseudomallei*. Arrows represent open reading frames (ORF) and direction of transcription.

TcF1

MREISSHADSRSIRRADTFSLWIGSVLGNDVARLSRRSDGGLILMEHTTSIGWAARSHARSE
HACDRISQWILVLMPIVFVAFCVLLISLWYIDILTGWLIAAFALLIEAMLIGIGFFVFVMVR
ISHRRRREREAKTALCRANQPDDADVPFPVGFDGVKTVHCMFGSGQVGGFRVSCYNTNCPLM
RGATETERLPGQAERA

TcF2

MEAGITRRRGAFFAREHEHGLSLDEADLERAFEAGEFVMDYQPIVSIRSGAVTGAEATIRWD
HPEWGALPQRVVHAAADRLGAASRIAGHAFGEACRQLARWKQQGAGVSALSMRLSGAQLGAE
SVFERLALSIEAFDISSSQFTLEVPEMSAPEESLSLLDRLKRLRQKGYGIVLGDFGAHHTAM
STLMVLPVTGVKFGESFTERLPGSPTAEAILSSVSRLAHDLGFTLTVSGVENGRQFELLRRF
RDIELQGAFLFEPMRAEVWHERVNPRALSQLNLREPPF

TcF3

VSRVRAAEWARWRDNRRAVDRGARHVSLLHASRGRRFQAPPCAAMCRPPCASARRDARGMGA
RTRYLAFPHETLSRRCGASSMRRRLDFLSKRSDMAQTFIAVKTILPGDLFFPADAASALCRD
RAREAPAGSGGLTDREPTVLRHSAGIRAGRFFIGSA

TcF4

VSTSSASPWTTYSLPTAGTADERSGPNITIPPVARPGRASGASRCGFAAKIASNSNCGAAAA
YRSRPESMRPRVRLKKPAIIRSVCRLAAARRAEQATRYGEMKRPARRKHVMPGGWRYFPQPP
AWPARGLVSGRATSQGSYGTLRANARCRGSSEGAKARLPRLALVQIIGCQGKRRQISERVLP
AGFIREHLESGMICPFQFELLGVIQRSGLGVIDMLVGEVIDDFSYLADEVLRWEKIGFVNIL
DSVHRG

TcF5

MATELDPPDSFDPFDSIDSYDLFDSLDERQTNFTLSCAMDWGQNIKNPTTDILGKLIIGYLG
KSTKTKNNRFSIVWGPAVYVDPLSGYAANVTAIFKNTSNDKDYRLAFSGTNPNSQFADIVIE
DNDVKNLTPFSKYVSDCPPDALIANGTAKALDVTLQTLPSNSSASLTDFLKNLPKESTLTVT
GHSLGGVLASTFALYFSQSSGYPFKNVFCHSFAAATAGNVAFAAYSDGSLTGKMKRLWHRYD
VIPSSWNLSTMNRIKYIYEPNLFPTEDFIRKIRKIVYYFSNKHIYYAQPGSLDHSQQFELEG
AYHSGFKVFTDEAGWQHTFGYLAPLSLTADDLNEGQPWQPSLGTLA

TcF6

MTCFLRAGLFISPYRVACSARRAAARRQTLRMIAGFFNLTRGRIDSGRLLYAAAAPQFELLA
IFAANPHRLAPDARPGRATGGIVIFGPDLSSAVPAVGREYVVQGEAELVDTRAYLGEDGFLG
RPALVGQTGLEFDFEFLESLICNHFVTLSNF

TcF7

MREREIARASAHAARIAPSRRARRAAHGGTRRGLKPPPARRMQKRHMPRSAIHGAPIVAPPR
PLRRPHPGHARRPFTRRPRMRLRMRSNAHGRACATRRRPARRAIATNRESARRGLPRSLEFD
CTDRGCATESKFLEMSAQPARRARSTMQARARTTRGRVSNNCDFRAKTSEAMSGRGFAMSRP
VGPRRAISAGERPRAGRDRVCRRRPHRAARATSNAPCRRACAHGHASASDARSAGHRIAQDW
RAKA

Fig 5.2: Amino acid sequences (TcF1-TcF7) derived from the translated nucleotide sequences of ORFc-ORFi, respectively. Properties of each putative protein are indicated in text.

Table 5.1: Computer predictions of open reading frames and associated characteristics. For promoters, capital letters denote nucleotides that are identical to *E. coli* consensus sequences

ORF	Size	Start	Stop	RBS	Promoter		Promoter probability score	mol% G+C
					-35	-10		
C	609	ATG	TGA	GAAAGG	Not obvious			59%
D	861	ATG	TGA	GGGAGAG	Not obvious			65%
E	483	GTG	TGA	GAAAGG	TTGAtt	TacAgT	0.97	71%
F	765	GTG	TAA	GAGA	TTGAgA	aATgAT	0.89 & 0.83	59%
G	1071	ATG	TGA	AAGGAA	TTGAtc	TATAtT	0.9	51%
H	351	ATG	TGA	GGAAAA	Not obvious			61%
I	759	ATG	TGA	GAAAG	TTGctg	cATAcT	0.92	72%

Table 5.2: Properties of proteins deduced from ORFc-ORFi in the *mrgRS* flanking regions.

ORF	Protein designation	Mol wt (Da)	pI	Major similarity to other proteins	Other features
C	TcF1	22,556	8.90	PssL *R. leguminosarum*	Transmembrane domain
D	TcF2	31,527	5.89	BvgR *B. pertussis*	Two conserved domains
E	TcF3	17,700	11.70	FimZ *S. typhi*	Lacks HTH motif
F	TcF4	27,665	10.93	AlgP *P. aeruginosa*	Similar to eukaryotic histones
G	TcF5	39,493	5.23	Serine lipases	G-X-S-X-G lipase motif
H	TcF6	16,856	8.60	PscD *P. aeruginosa*	Similar to type III export protein
I	TcF7	28,251	12.28	-	Signal peptide sequences and high arginine content

Table 5.3: A comparison of the amino acid compositions of the deduced proteins TcF1-TcF7

Residue	Composition (%)						
	TCF1	TCF2	TCF3	TCF4	TCF5	TCF6	TCF7
A Ala	9.41	11.89	16.88	11.81	7.87	14.19	18.25
C Cys	2.79	0.35	3.12	2.36	0.84	1.94	2.78
D Asp	4.95	3.50	5.62	2.36	7.58	3.87	2.38
E Glu	4.46	9.44	2.50	4.72	2.81	5.16	3.17
F Phe	5.45	5.94	5.00	2.76	6.46	8.39	1.98
G Gly	7.92	8.74	7.50	9.84	6.18	9.68	6.35
H His	2.97	3.15	2.50	1.18	1.97	1.29	4.37
I Ile	7.92	3.50	2.50	5.91	4.78	4.52	2.78
K Lys	0.99	1.40	1.25	3.15	5.92	0.00	1.98
L Leu	8.91	11.19	6.88	6.69	9.27	11.61	2.38
M Met	3.47	2.45	2.50	1.97	1.12	1.29	3.57
N Asn	1.98	1.05	0.62	1.97	6.18	2.58	1.98
P Pro	2.97	4.20	5.62	7.09	5.62	5.16	8.33
Q Gln	1.98	3.85	1.25	3.15	3.09	2.58	1.59
R Arg	10.40	9.44	18.75	12.06	2.25	10.32	24.60
S Ser	7.92	8.04	8.12	9.06	9.83	4.52	7.14
T Thr	4.46	3.85	3.75	3.54	7.58	4.52	4.37
V Val	7.43	5.94	3.75	5.91	4.21	5.81	1.59
W Trp	2.48	1.40	1.25	1.57	1.69	0.00	0.40
Y Tyr	0.99	0.70	0.62	2.36	5.06	2.58	0.00

was estimated for the stem-loop structure (**Fig 5.3**). The deduced translation of ORFc codes for a protein of 202 amino acid residues, with a calculated molecular weight of 22,556 and pI 8.9. The predicted protein was designated TcF1. Amino acid sequence analysis showed that TcF1 has some similarity to a polysaccharide export protein, PssL, from *Rhizobium leguminosarum*. Protein multiple alignment revealed 50 % similarity with 16% identity over a span of ~180 amino acids (**Fig 5.4**). Analysis of the amino acid sequence of TcF1 by DAS transmembrane prediction (section 3.9) suggests that it is a membrane protein. A strong hydrophobic domain of 55 residues (70-124) may contain a transmembrane domain allowing the C-terminal portion of the molecule to project into the cytoplasm (**Fig 5.5**).

To elucidate the possible function of TcF1, sequence information from the *B. pseudomallei* genome sequencing project database at the Sanger Centre, was utilised. The Sanger Centre is funded to sequence the genome of the *B. pseudomallei* clinical isolate, K96243. The nucleotide sequence of ORFc was used as a template to search this database and extract nucleotide sequences upstream of the ORFc sequence. An assembled fragment of 149,217 bp in length, contig 98, was found to contain an identical sequence to ORFc (99% identity) and an additional 24.2 kb of sequence information from the region upstream of ORFc. Immediately upstream of ORFc, within the additional 24.2 kb sequence, an ORF was identified, ORFx, that is preceded by a Shine-Dalgarno binding site (GAGGAGG) and a strong promoter sequence with a predicted probability of 0.97. ORFx ends 389 nucleotides upstream of the initiation codon of ORFc. The deduced amino acid sequence of ORFx was found to encode a protein, TcFx, of 1057 amino acids with molecular weight 97,905 and pI 3.68. Hydropathy analysis revealed that TcFx has two strongly hydrophobic regions that are capable of spanning the cell membrane (**Fig 1**, Appendix V). Blast searches indicated that TcFx is similar to outer membrane proteins from a number of bacterial species including *S. typhi*, *E. coli* and *Y. enterocolitica* (**Table 1**, Appendix V). Protein multiple alignment

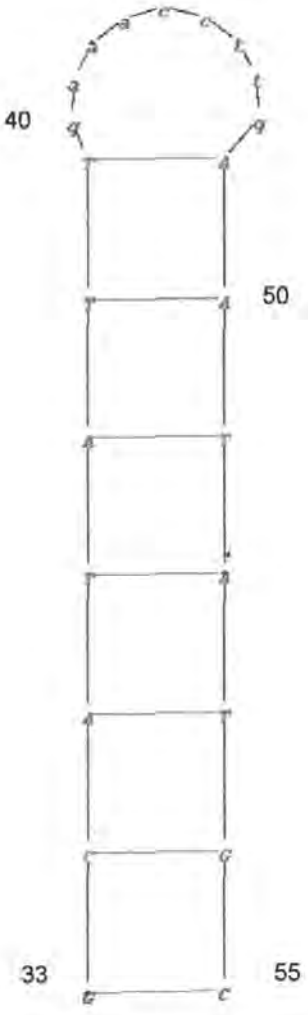


Fig 5.3: Predicted secondary structure of transcription termination formed by the inverted repeat sequence at the 3' end of ORFc; stem-loop structure. Numbers indicate distance from the proposed stop codon.

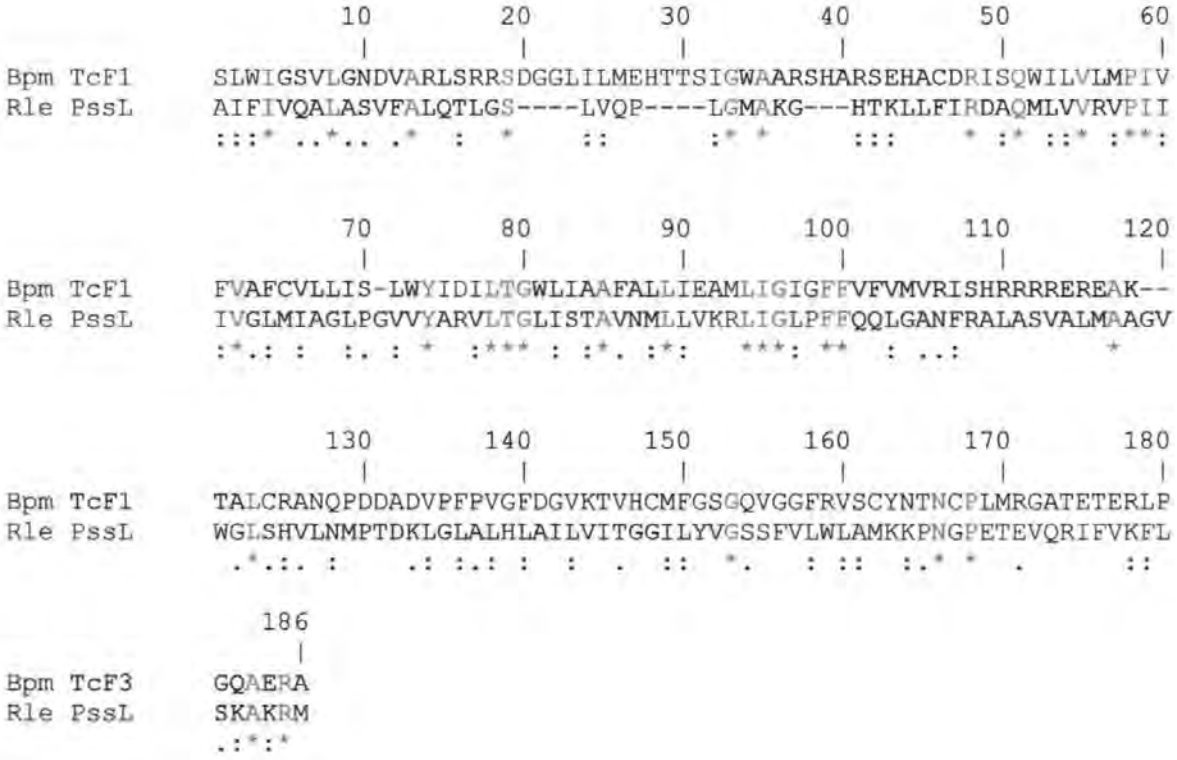


Fig 5.4: Alignment of TcF1 with PssL protein over a span of 186 amino acids. Stars indicate identical residues.

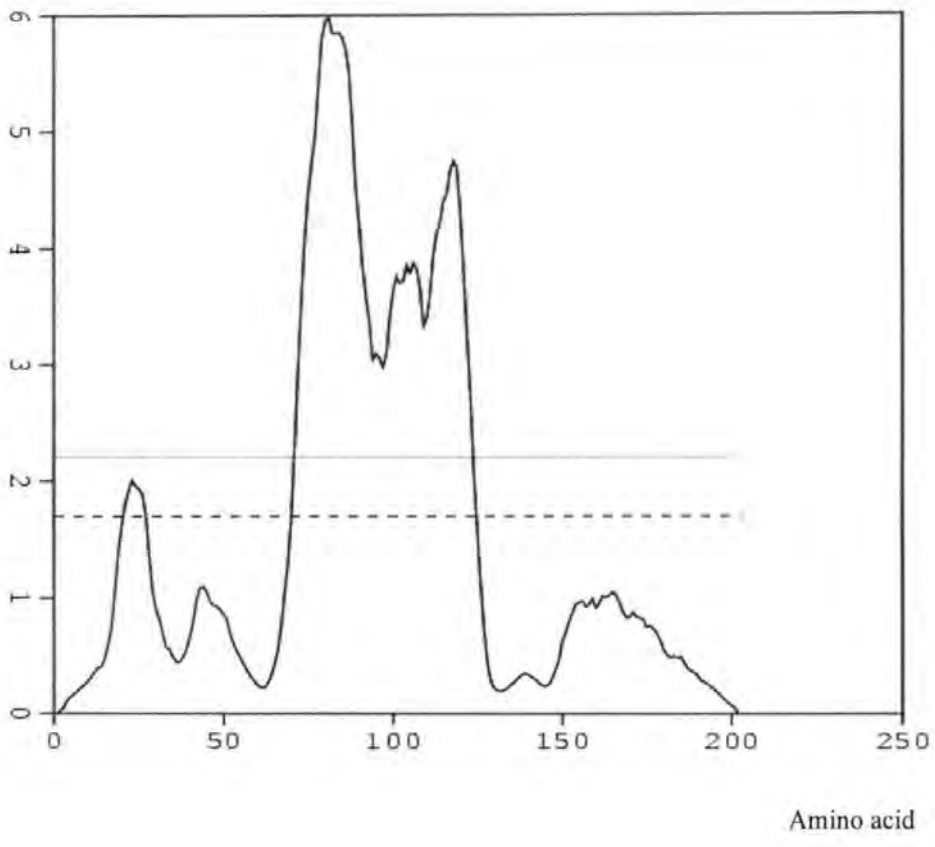


Fig 5.5: A hydropathy prediction showing possible transmembrane regions of TcF1. Values above the horizontal dashed axis indicate hydrophobic character.

showed that TcFx shares 60% and 63% similarity and 23.6% and 21% identity with surface proteins of the plant pathogen *Xylella fastidiosa* and *S. typhi*, respectively. This increases to 91% similarity and 60% identity over a span of 80 amino acids toward the C-terminal end of the proteins (**Fig 2**, Appendix V). Using the Prosite program a protein kinase C phosphorylation site, SER, was identified within this region. *In vivo*, protein kinase C exhibits a preference for the phosphorylation of serine or threonine residues found close to a C-terminal basic residue (Woodget *et al.*, 1986).

Two hundred and twenty bases downstream from the *mrgRS* locus, on the same DNA strand, an open reading frame, ORFd, was identified and its sequence was predicted to be transcribed in the same direction as *mrgRS*. ORFd encodes a protein of 286 amino acids, designated TcF2, with a predicted molecular weight of 31,527 and pI 5.89. A search of the GenBank database with the BLAST program showed that TcF2 displays homology to at least 50 proteins found in a variety of bacteria (**Table 2,** Appendix V). Although the precise functions of many of these proteins are unknown, their open reading frames are within or are tightly linked to operons with well-defined functions. For example, *bvgR* is required for regulation of all known *bvg*-repressed genes in *Bordetella pertussis*, the causative agent of whooping cough, and is located immediately downstream from the *bvgAS* locus encoding a two-component regulatory system that upregulates virulence gene expression in *B. pertussis*. Additionally, FixL of *E. coli* is a sensor regulator that is involved in the control of nitrogen fixation and is a member of the family of two-component regulatory proteins. Using the Boxshade computer program, sequence alignment of TcF2 with a number of these similar proteins (**Fig 5.6**), showed that TcF2 possesses most of the conserved residues that are present in two major domains. Domain I consists of 55 amino acid residues with 23 highly conserved residues and domain II consists of approximately 90 amino acid residues with 39 highly conserved residues. Several members of this family of proteins have either domain I

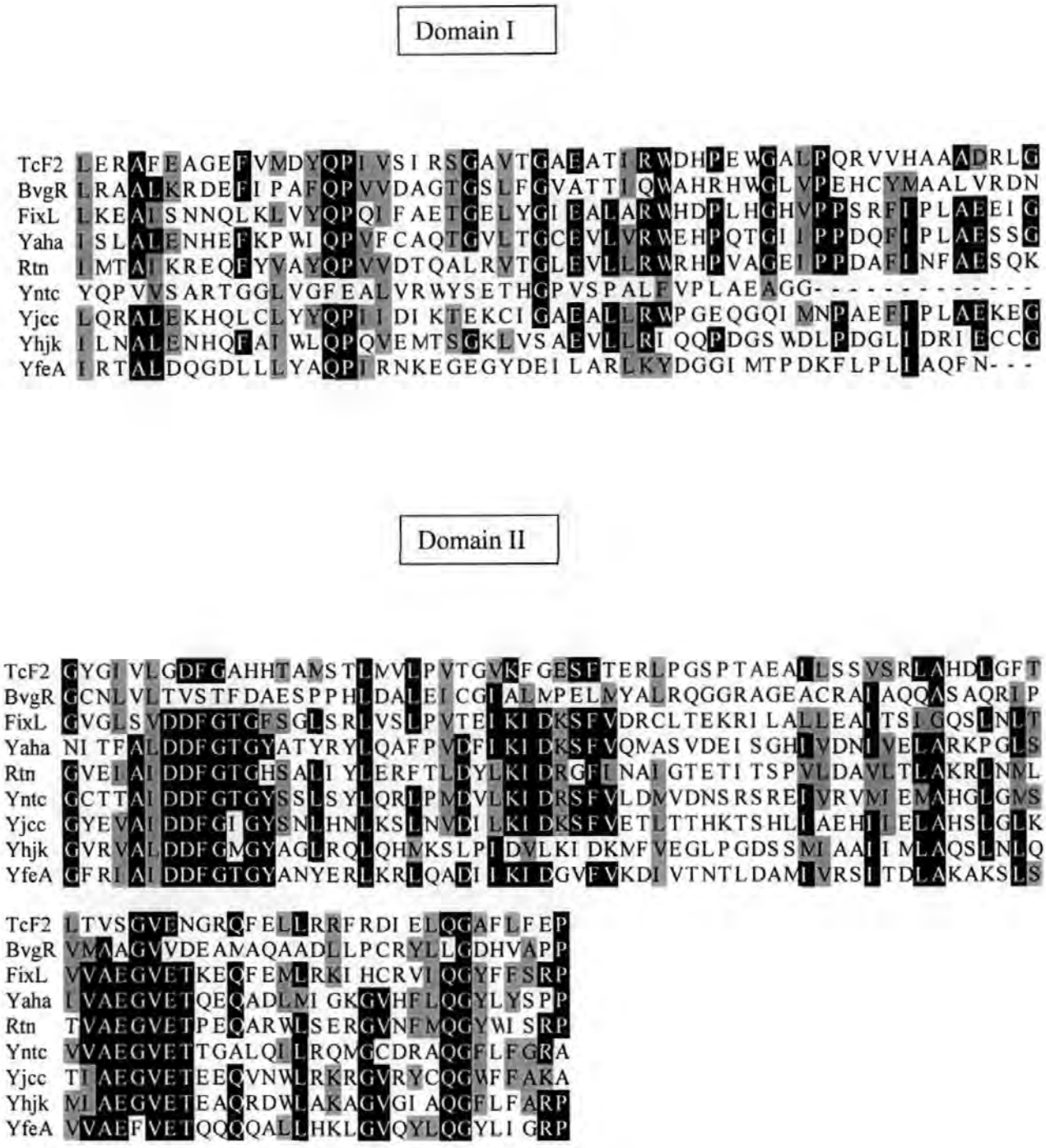


Fig 5.6: Sequence alignment using the Boxshade program for TcF2, encoded by ORFd, and homologous proteins. Black boxes indicate positions conserved in at least 65% of the aligned sequences.

or II but none appear to have all the conserved residues. Within this family, it was observed that in addition to the high degree of sequence conservation within the two domains, the spacing between the two domains, 90 amino acid residues, and the spacing between domain II and the carboxy-terminal ends of the proteins, 25 amino acid residues, are also conserved.

The second ORF downstream from *mrgRS* locus, ORFe, is predicted to start with a rare GTG translation initiation codon. The translation of ORFe reveals a protein of 160 amino acids with a molecular weight of 17,700 and pI 11.7. The protein was designated TcF3. Amino acid sequence analysis revealed that TcF3 has substantial similarity in some areas to FimZ protein, a transcriptional activator necessary for type 1 fimbrial expression in *S. typhi* and *E. coil* (**Table 3**, Appendix V). Multiple sequence alignment showed that the protein shares 43.3% similarity and 15.7% identity with FimZ from *S. typhi*. This increases to 73% similarity over a span of 60 amino acids towards the C-terminal end of the proteins (**Fig 5.7**). However, TcF3 lacks the helix-turn-helix DNA binding motif, which is characteristic of the FimZ regulatory protein. Moreover, scanning TcF3 using the Prosite program for conserved motifs revealed the presence of several possible conserved kinase phosphorylation sites SRR, SMR, SKR and TDR at positions 75-77, 82-84, 91-93 and 137-139, respectively.

There are two potential promoter sequences for ORFf with predicted probabilities of 0.89 and 0.83. Both of these promoters overlap in a region covering 31 nucleotides upstream from the rare translational start codon, GTG. ORFf codes for a basic protein, designated TcF4, which consists of 254 amino acids with a predicted molecular weight of 27,665 and pI 10.93. Sequence analysis of the deduced amino acid sequence revealed some homology to a range of hypothetical proteins, eukaryotic DNA binding proteins and to AlgP regulatory protein of *P. aeruginosa* (**Table 4**, Appendix V). AlgP is a histone like protein that is involved in regulating the synthesis of alginate, the viscous mucoid exopolysaccharide in

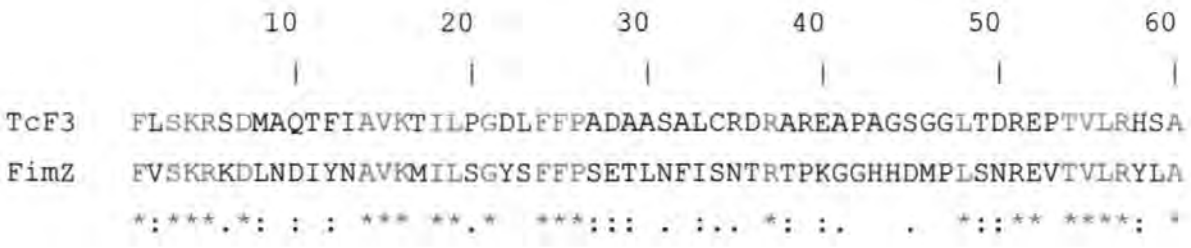


Fig 5.7: Clustal W multiple alignment of the TcF3 and *E. coli* FimZ over a span of 60 amino acids. Stars indicate identical residues.

Pseudomonads. AlgP has an unusual C-terminal region that resembles the tail of eukaryotic H1 histone and contains multiple repeats of Ala-Ala-Lys-Pro residues. TcF4 does not possess multiple repeats of this motif although it does have more alanine (11.81%), lysine (3.15 %) and proline (7.09 %) than are present in most of the other translated ORFs (**Table 5.3**), and over a span of 48 amino acid residues in the middle of the protein these 3 residues represent 40% of the amino acid composition.

On the reverse direction of the complementary strand three open reading frames have been identified, ORFg, ORFh, and ORFi. Unlike each of the other predicted open reading frames, ORFg has a relatively low mol% G+C (51%; **Table 5.1**). The protein encoded by ORFg, designated TcF5, consists of 356 amino acid residues and has a predicted molecular weight of 39,493 and pI 5.23. The deduced amino acid sequence of TcF5 contains the pentapeptide consensus sequence, G-H-S-L-G (position 187-191), with the active serine residue and invariant glycines in positions 3, 1 and 5, respectively. This G-X-S-X-G sequence (where X is any amino acid) is the consensus sequence for the active sites of esterases, lipases, and serine proteases (Brenner, 1988). Proteins that lack the active site serine are devoid of catalytic function. Many of these enzymes are preproenzymes that possess a signal peptide at the N-terminus; the mature enzyme is produced following cleavage of the N-terminal peptide. However, overall the TcF5 protein is most similar to the extracellular phospholipases A1 (PA1) from *Serratia liquefaciens* (50% similarity and 15% identity) and *Yersinia enterocolitica* (43.4% similarity and 12% identity) and also the ABC transporter ATP-binding proteins from a number of bacterial species (**Table 5**, Appendix V). Hydropathy predictions suggest that TcF5 does not contain a typical signal peptide that is often observed in exported prokaryotic proteins. Moreover, TcF5 was found to contain a number of other amino acid signatures for phosphorylation, which although shorter in size than the region of similarity described above, may provide some clues about the possible

function of the protein. These include protein kinase C phosphorylation sites, STK, TAK, TGK and SNK at positions 64, 151, 237, and 288, respectively, in addition to casein kinase phosphorylation sites, TELD, SLDE, TSND, SLTD and TADD at positions 3, 25, 99, 169 and 338, respectively.

ORFh commences one hundred and twenty three nucleotides downstream from ORFg. ORFh encodes a protein, TcF6, predicted to be 16,856 molecular weight and pI 8.6. Amino acid sequence analysis of TcF6 showed that it has some similarity to type III export proteins from a number of bacterial species, including PscD of *P. aeruginosa* (**Table 6**, Appendix V). Protein sequence alignment with PscD revealed 45.2% similarity with 19.2% identity over the 155 amino acid length of TcF6.

The predicted molecular weight of TcF7, encoded by ORFi, is 28,251. The first 22 amino acids of the N-terminus bear all the features of a prokaryotic signal peptide (**Fig 5.2**). The amino terminus contained two positively charged arginine residues following Met-1 and a hydrophobic core of 12 residues followed by a helix-breaking residue (proline) within five residues preceding the proposed cleavage site, which is located between Ala-22 and Arg-23. Following the signal sequence, another well-defined peptidase cleavage site was found between Ala-72 and Arg-73 (**Fig 5.2**). Analysis of the amino acid composition of TcF7 revealed a striking and unusually high content of the positively charged hydrophilic amino acid arginine (25%) as well as the hydrophobic nonpolar amino acids, alanine (18%) and proline (8%). This was reflected in the high pI value (12.28). Searches of the protein and nucleotide sequence databases under stringent parameters failed to reveal any similarity to other known proteins. Less stringent searches (expectation value, 1000) showed only that the protein has a limited homology to a range of hypothetical proteins. However, Using the Prosite program several candidates for protein kinase C phosphorylation sites were

identified, SRR, TRR, TRR, TRR, SAR, TDR, TTR and SGR at positions, 19, 30, 77, 97, 112, 126, 157, and 177.

5.3.3 SOUTHERN BLOT ANALYSIS

5.3.3.1 Development and preparation of oligonucleotide probes

Oligonucleotide probes covering the *mrgRS* locus and flanking regions were developed for PCR amplification by primers designed on the basis of the nucleotide sequence data. In addition to a probe specific for the 4.3 kb *Eco*RI fragment, RS7 (section 4.3.2.2), three pairs of PCR primers (**Table 5.4**), were designed (section 3.13.1) and used in PCR reactions to amplify and label oligonucleotide probes that span the 3' downstream sequence. The probes were labelled with DIG-dUTP (section 3.10.2.2) and purified (section 3.10.2.3). The probes were designed as follows: FR1 (flanking region), spanning 1276 bp over an *Eco*RI site and across the *mrgS* stop, FR2 spanning 1137 bp, and FR3 spanning 1781 bp of the 3' region of *mrgRS* (**Fig 3**, Appendix V). High stringency conditions, 68°C, were used for DNA hybridisation and washing steps.

Table 5.4: PCR primer sequences for DNA probes FR1, FR2 and FR3.

Forward primer (FR1F)	ATGATGGACGGTTTCCAGTTGCTC
Reverse primer (FR1R)	AACGTTAAATCAAGTCGCGGGTGG
Forward primer (FR2F)	AGAGCGCTGTCGCAACTGAATCTG
Reverse primer (FR2R)	TCGCTTCGCTTGCTGAGAAA
Forward primer (FR3F)	GGTCCGGGCCAAATATTACGATCC
Reverse primer (FR3R)	AGCGGAACCAATCCGAACTCACAG

5.3.3.2 Detection of the *mrgRS* locus in different *B. pseudomallei* strains

To confirm the presence of *mrgR* and *mrgS* in different *B. pseudomallei* strains, approximately 0.5 μg of genomic DNA isolated from a number of different *B. pseudomallei*

strains (section 3.4.1.2) from a wide geographical distribution, as well as other closely related *Burkholderia* species; *B. thailandensis*, *B. cepacia*, *B. vietnamiensis*, *B. plantarii*, *B. cocovenenans*, *B. vandii*, *Pseudomonas* species; *P. aeruginosa*, *P. fluorescens* and *E. coli* DH5α (**Table 5.5**) were completely digested with *Eco*RI restriction endonucleases (section 3.4.5.1). The resulting DNA fragments were separated by agarose gel electrophoresis (section 3.4.6.1) and transferred onto positively charged nylon membrane prior to DNA hybridisation as explained in section 3.10. *Eco*RI digested DNA from pMRG2 or pMRG4 (sections 4.3.7 and 4.3.15, respectively) was included in the membrane as a positive control. The blot was probed with RS7 (~ 10 ng ml^{-1}), the 4.3 kb *Eco*RI fragment labelled with DIG-11-dUTP, under high stringency conditions for hybridisation and washing (section 3.10.3). The DNA-DNA hybridisation was detected with anti-digoxigenin phosphatase conjugated antiserum and visualised by enzyme-substrate colour reaction (section 3.10.4). The molecular size of DNA bands were calculated with reference to bacteriophage λ DNA digested with *Hind*III and labelled with DIG-dUTP. Genomic DNA from all of the 19 environmental and clinical isolates of *B. pseudomallei*, originating from different areas of the world, possessed a single 4.3 kb *Eco*RI fragment. On the other hand, the probe did not hybridise with the DNA of any of the other closely related bacteria, including *B. thailandensis* (**Fig 5.8**).

5.3.3.3 Southern blot hybridisation analysis of the *mrgRS* flanking regions

Three oligonucleotide probes (FR1, FR2 and FR3) covering approximately 3 kb of the *B. pseudomallei* genome downstream from *mrgRS* were used to investigate *Eco*RI digests of genomic DNA from different *B. pseudomallei* strains as well as *B. thailandensis* (**Table 5.6**, **Fig 4**, Appendix V).

The FR1 probe spans an *Eco*RI site, and on the basis of the sequencing data obtained in section 4.3.15, was expected to hybridise two *Eco*RI fragments of genomic DNA of 684 and

Table 5.5: Isolates of *B. pseudomallei* and other bacterial species that were used in Southern hybridisation with probe RS7. NE, Northeast; C, Central.

Strain	Source	Sample	Geographical origin	Date
B. pseudomallei				
204	Human	Blood	Thailand	-
E8	Environment	Soil	NE Thailand	1990
25	Environment	Soil	Madagascar	1977
E25	Environment	Soil	Thailand	-
19	Environment	Soil	Singapore	1991
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	-
576	Human	Blood	Thailand	-
B. thailandensis				
E27	Environment	Soil	Thailand	-
E82	Environment	Soil	NE Thailand	1990
E255	Environment	Soil	NE Thailand	-
E256	Environment	Soil	Thailand	-
E260	Environment	Soil	C Thailand	1993
E38	Environment	Soil	Thailand	-
Closely related species				
B. cocovenenans LMG 11626				
B. plantarii LMG 10908				
B. vietnamiensis LMG 6998				
B. cepacia				
B. vandii				
P. aeruginosa				
P. fluorescens				
E. coli DH5				

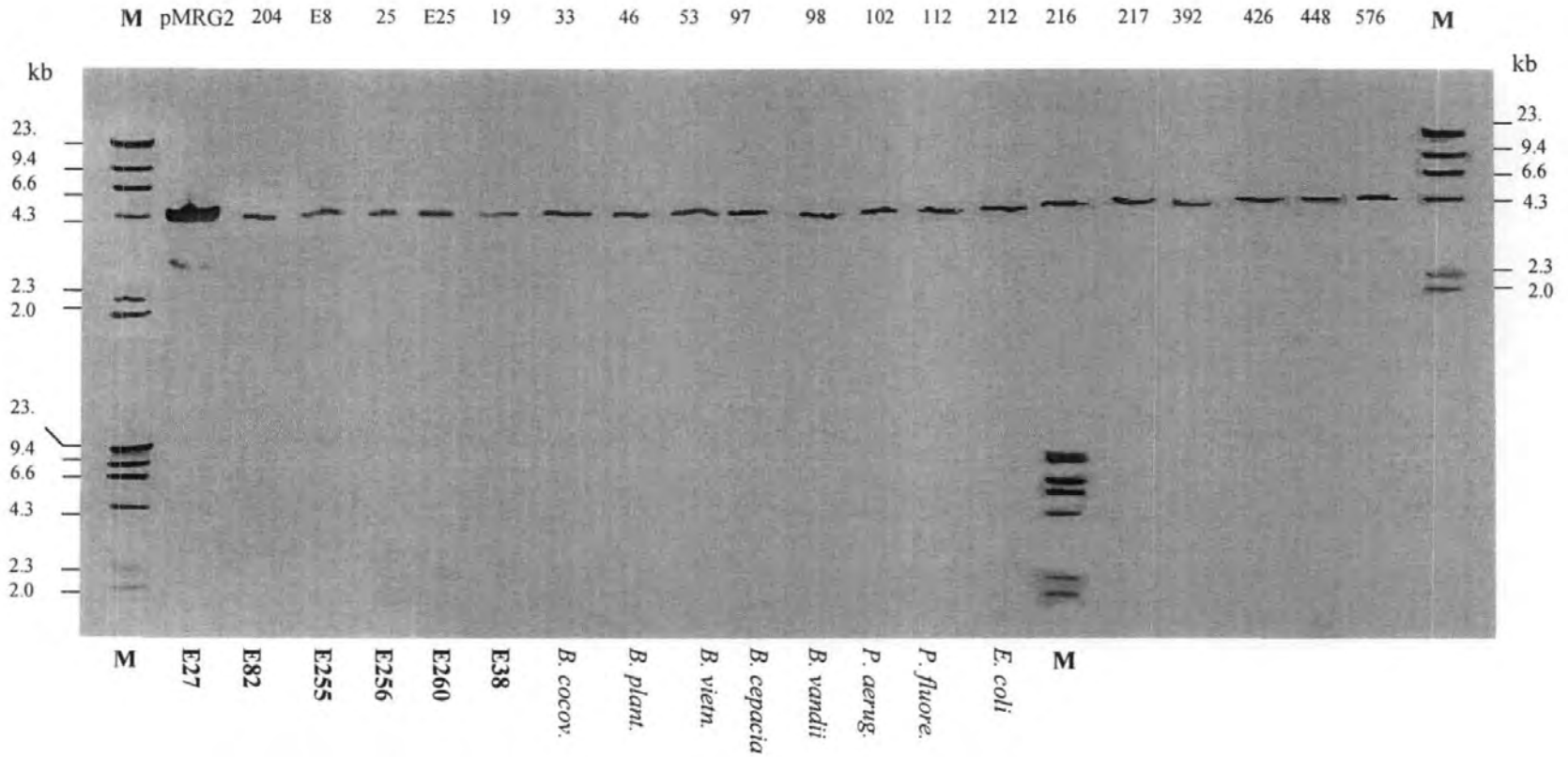


Fig 5.8: RS7 hybridisation analysis of the *Eco*RI-digested genomic DNA from different *B. pseudomallei*, *B. thailandensis* (in bold) and closely related bacteria. The RS7 DIG-labelled probe contains the 4.3 kb *Eco*RI fragment, i.e. the flanking region upstream from the *mrgRS* locus, *mrgR* and 2342 bp of the total length of *mrgS*. M: DNA molecular markers. The full description of each isolate can be found in Table 5.5.

Table 5.6: Isolates of *B. pseudomallei* and *B. thailandensis* that were used in Southern hybridisation analysis with probes FR1, FR2 and FR3. NE, Northeast.

Strain	Source	Sample	Geographical origin	Date
B. pseudomallei				
204	Human	Blood	Thailand	-
E8	Environment	Soil	NE Thailand	1990
19	Environment	Soil	Singapore	1991
22	Environment	Soil	Burkina Faso	1973
97	Environment	Soil	Australia	-
98	Environment	Soil	Australia	-
112	Human	Multiple	NE Thailand	1992
216	Environment	Soil	NE Thailand	1990
392	Human	Pus	NE Thailand	1989
426	Environment	Soil	Vietnam	-
448	Environment	Soil	Vietnam	-
576	Human	Blood	Thailand	-
H55	Human	Blood	China	1996
H706	Human	Parotitis abcesses	China	1996
H1	Human	Fossa abcesses	China	1996
H2D	Human	Lymph node abcesses	China	1996
B. thailandensis E82	Environment	Soil	NE Thailand	1990
B. thailandensis E255	Environment	Soil	NE Thailand	
B. thailandensis E256	Environment	Soil	Thailand	-

593 bases in length. The results demonstrated the presence of two bands of the expected sizes in most of the isolates of *B. pseudomallei*. However, two *B. pseudomallei* isolates, 392 and H55, lacked the 684 bp fragment and H55 possessed an additional unique band of approximately 4.5 kb. No signal was obtained from *B. pseudomallei* 576 or any of the tested *B. thailandensis* isolates (**Fig 5.9**).

The FR2 probe of 1137 bp without any internal *Eco*RI restriction sites demonstrated the presence of single bands in all of the *B. pseudomallei* isolates, except 576. The bands were about 1137 bp in size but in isolate H2D the band was smaller, 1000 bp. All *B. thailandensis* isolates were negative (**Fig 5.10**).

On the other hand, the FR3 probe of 1781 bp that lacked an internal *Eco*RI site revealed two banding patterns. One pattern indicated the presence of single bands of 1.7 kb in seven isolates, as expected from the available sequence information, while the other pattern demonstrated two bands of 1421 and 360 bp in nine isolates (**Fig 5.11**, **Table 5.7**). This suggests that some isolates of *B. pseudomallei* must possess an additional *Eco*RI site in the region of the genome spanned by FR3. To investigate this possibility two sets of PCR primers (**Table 5.8**) were used to amplify fragments of 611 bp and 1781 bp from one of the isolates thought to possess an additional *Eco*RI site in this region, *B. pseudomallei* 112, as well as isolate 204 from which the original sequence data had been obtained. The PCR products were purified (section 3.10.2.3) and the DNA was sequenced (section 3.8).

Table 5.8: PCR primers for the amplification of a polymorphic region within the *B. pseudomallei* genome.

Forward primer (FR3F)	GGTCCGGGCCAAATATTACGATCC
Reverse primer (FR3R)	AGCGGAACCAATCCGAACTCACAG
Forward primer (FR4F)	TGAATGCCAGGCGATAGTCCTTGT
Reverse primer (FR4R)	GCATCACTGCCATCATTCGAGCAA

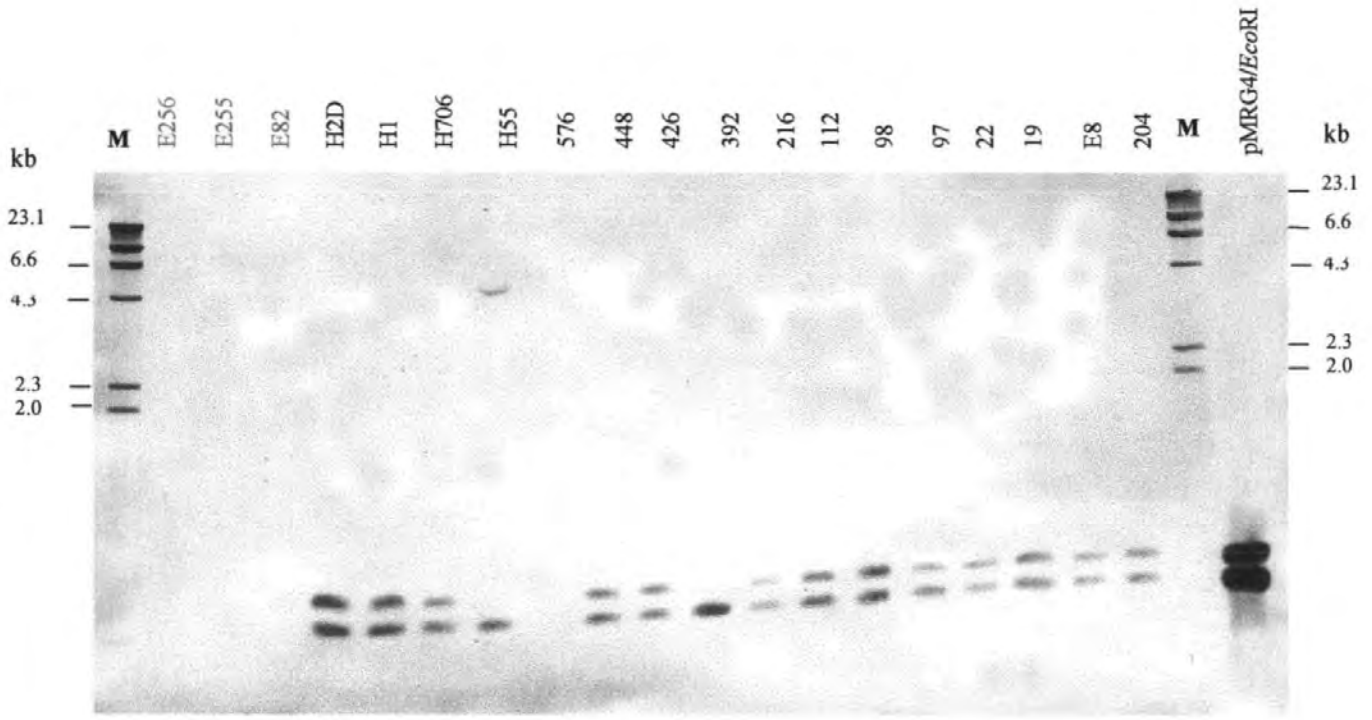


Fig 5.9: FR1 hybridisation analysis of the genomic DNA from different *B. pseudomallei* (in black) and *B. thailandensis* (in red) isolates. The FR1 DIG-labelled probe spans the distal end of *mrgS* and 285 bp further downstream from the *mrgS* stop codon. M: DNA molecular marker. The full description of each isolate can be found in Table 5.6.

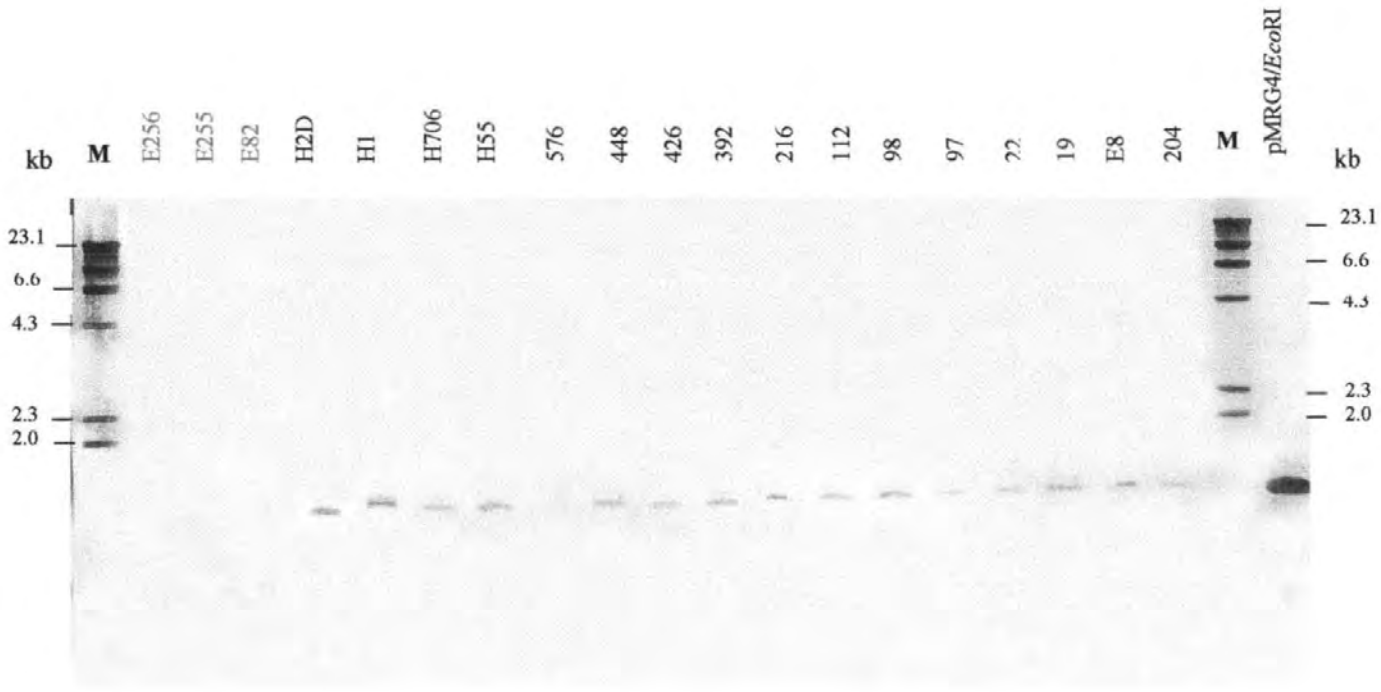


Fig 5.10: FR2 hybridisation analysis of *Eco*RI-digested genomic DNA from different *B. pseudomallei* (in black) and *B. thailandensis* (in red) isolates. The FR2 DIG-labelled probe spans 1137 bp downstream from the *mrgRS* locus with no internal *Eco*RI restriction sites. M: DNA molecular size markers. The full description of each isolate can be found in Table 5.6.

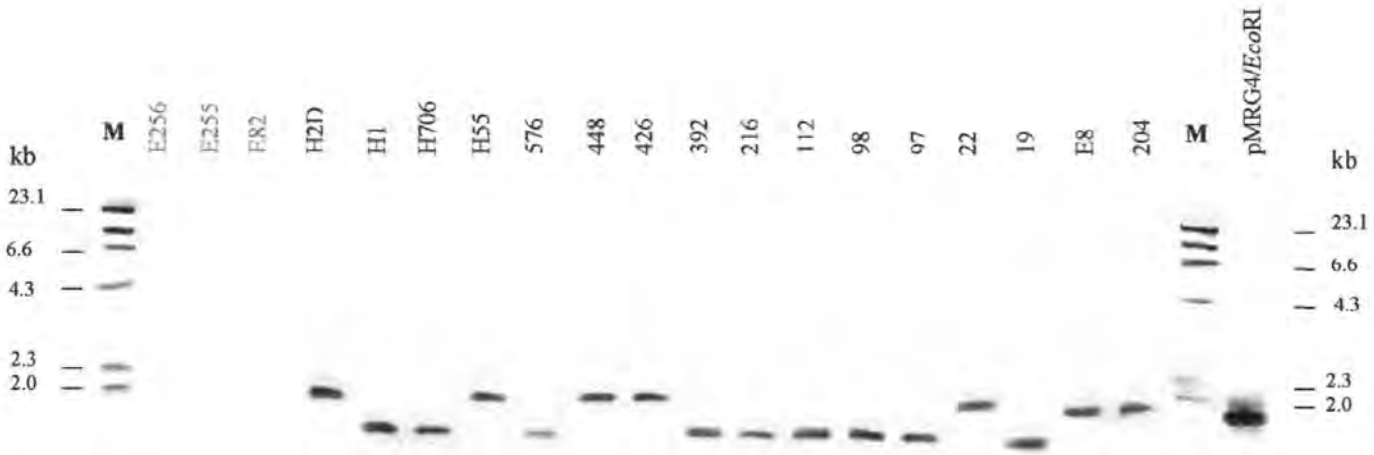


Fig 5.11: FR3 hybridisation analysis of the *Eco*RI-digested genomic DNA from different *B. pseudomallei* (in black) and *B. thailandensis* (in red) isolates. The FR3 DIG-labelled probe spans 1781 bp downstream from the *mrgRS* locus. Probe hybridisation divided the examined *B. pseudomallei* isolates into two groups. M: DNA molecular size markers. The full description of each isolate can be found in Table 5.6.

Table 5.7: FR3 hybridisation patterns of *B. pseudomallei* isolates with FR3

Pattern I	Source	Site	Geographical origin	Date
204	Human	Blood	Thailand	-
E8	Environment	Soil	Thailand	1990
22	Environment	Soil	Burkina Faso	1973
426	Environment	Soil	Vietnam	-
448	Environment	Soil	Vietnam	-
H55	Human	Abcesses	China	1996
H2D	Human	Abcesses	China	1996
Pattern II				
19	Environment	Soil	Singapore	1991
97	Environment	Soil	Australia	-
98	Environment	Soil	Australia	-
112	Human	Multiple	NE Thailand	1992
216	Environment	Soil	NE Thailand	1990
392	Human	Pus	NE Thailand	1989
576	Human	Blood	Thailand	-
H706	Human	Abcesses	China	1996
H1	Human	Abcesses	China	1996

Nucleotide sequence alignment of the two sequences from *B. pseudomallei* isolate 204 and isolate 112 revealed that the DNA was almost identical (98% identity). Base substitutions were detected in 19 positions over a span of 1472 nucleotides (**Fig 5**, Appendix V). One of these substitutions was found to cause a change in the nucleotide sequence from GAACTC to GAATTC, the recognition sequence for the restriction enzyme *Eco*RI. ORFg includes 16 substitutions within its' full length of 1071 bp. At the amino acid level, 10 substitutions were synonymous, indicating that the peptide sequences were invariable, while the other 6 substitutions resulted in changes to 6 amino acids in TcF5, which did not apparently alter the overall nature of the encoded protein (**Table 5.9**).

5.4 DISCUSSION

The sequence of the DNA region flanking the *B. pseudomallei mrgRS* locus was analysed and seven putative open reading frames were identified. The 1.319 kb DNA sequence upstream from the *mrgRS* locus contains a single open reading frame, while the 4.753 kb DNA sequence downstream from the *mrgRS* transcriptional termination harboured six putative genes, three in each orientation. Apart from ORFg, all the identified ORFs possessed mol% G+C (59-72%) that are within the overall G+C content of *B. pseudomallei* genome, which ranges between 55.6 and 74.0% with an average of 65.7% (Songsivilai & Dharakul, 2000). ORFg has a remarkably low G+C content of 51%. The putative proteins encoded by ORFc to ORFi were designated TcF1-7 for two-component flanking proteins, respectively.

The absence of a well-defined promoter region upstream from ORFc, suggests that it may be possible that ORFc may be transcribed as a polycistronic unit under the control of an upstream promoter. Consistent with this possibility, a strong promoter sequence with a predicted probability of 0.97 was identified in the 5' untranslated sequence of the upstream

Table 5.9: Nucleotide substitutions and amino acid changes in ORFg that encodes TcF5. Only the characters of amino acids that differ are shown.

Nucleotides		Amino acids		Mol wt		Hydropathy		Charge		Chemical	
From	To	From	To	From	To	From	To	From	To	From	To
G	A	E	E								
A	G	N	S	132	105	Phi	Phi	0	0	Amide	Hydrox
A	T	G	G								
C	T	Y	Y								
C	T	G	G								
C	T	N	N								
T	C	I	I								
G	A	V	I	117	131	Pho	Pho	0	0	Aliph	Aliph
A	G	E	G	143	75	Phi	Phi	-	0	Acidic	Aliph
C	T	F	F								
C	T	A	V	89	117	Pho	Pho	0	0	Aliph	Aliph
C	T	I	I								
C	T	A	A								
G	A	D	N	133	132	Phi	Phi	-	0	Acidic	Amide
C	G	Q	E	146	147	Phi	Phi	0	-	Amide	Acidic
A	G	Q	Q								

ORF (ORFx). The presence of a stem-loop transcriptional termination sequence downstream from ORFc strongly suggests that transcription does not extend further downstream, into the *mrgRS* locus. Sequence comparison revealed that TcF1 has some similarity to PssL, a polysaccharide export protein from *R. leguminosarum*. *R. leguminosarum* synthesises exopolysaccharide during the early stages of nodule invasion, to facilitate the process of root infection.

To further elucidate a possible function for TcF1, sequence information from the *B. pseudomallei* isolate K96243 genome sequencing project database at the Sanger Centre, was used. The ~ 1.3 kb region upstream of the *mrgRS* locus in isolate 204, containing ORFc, was found to be present in the genome sequence of isolate K96243 and to share 99% nucleotide sequence identity with that sequence. This region of the *B. pseudomallei* genome is therefore present and nearly identical in isolates K96243 and 204. Furthermore, the nucleotide sequence upstream of OFRc in K96243 contained an open reading frame, ORFx, the deduced translation of which was found to be substantially similar to outer membrane proteins from a number of bacterial species. This was supported by the presence of two hydrophobic stretches that are strongly reminiscent of transmembrane regions. In bacteria, genes with related functions are often located together in an operon transcribed from a single promoter in a single polycistronic mRNA molecule. Hence TcF1, a possible polysaccharide export protein, and TcFx, a potential outer membrane protein, may be cotranscribed under certain conditions and may perhaps share some functional relationship.

The strong termination sequence following the *mrgS* gene suggests that the downstream sequences are transcribed separately. Computer-generated sequence alignments placed TcF2 as a member of a family of proteins that are encoded by genes that either lie within or are very closely linked to functional operons. Most of the members of this family are characterised by two highly conserved domains that are located in the C-terminal portion

of the proteins and which are separated by approximately 130 residues. The predicted TcF2 protein contains two domains that are strongly conserved among a large number of predicted proteins from a variety of bacterial species. In addition, the spacing between the two domains and the distance between domain II and the carboxy-terminal ends of the predicted proteins are also conserved. This conservation may reflect the convergent evolution of successful structure-function motifs rather than the usual primary sequence similarity. The activity of one member of this family, BvgR, has been assigned recently. The *bvgR* gene, which encodes a 32-kDa repressor protein, resides immediately downstream from the well-characterised *bvg* locus. The *bvg* locus encodes a two-component regulatory system consisting of a sensor protein, BvgS, and a transcriptional activator, BvgA. This locus is known to regulate the *Bordetella* virulence regulon (Scarlato *et al.*, 1993). Inactivation of *bvgAS* results in the loss of virulence gene expression (Cotter & DiRita, 2000). It is known that BvgR binds to a conserved nucleotide sequence element found within the coding sequences of the *bvg*-repressed genes. The expression of *bvgR* is activated at the level of transcription by the products of *bvgAS* genes. This activation was proposed to result from the binding of phosphorylated BvgA to the *bvgR* promoter. Merkel *et al.* (1998) have shown that in *Bordetella pertussis*, the gene encoding BvgR resides on a separate regulatory unit from the unit containing the *bvgA* and *bvgS* two-component regulatory genes. The latter genes are transcribed as an operon that is under the control of a positively acting regulator. The close correlation between the genomic organisation in *bvgAS-bvgR* and *mrgRS*-ORFd together with the fact that both BvgAS and MrgRS are members of the two-component family allows the speculation that TcF2 may perform a similar function to that of BvgR, although absolute proof requires further experimentation.

The second ORF downstream of *mrgRS* locus, ORFd, codes for a protein designated TcF3. Amino acid sequence analysis of TcF3 suggests that this protein may be related to

prokaryotic transcriptional regulators. The most closely related proteins include the positive regulator, FimZ from *S. typhi* (Boyd & Hartl, 1999) and *E. coli* with which TcF3 shares considerable similarity (73%) over a region spanning 60-amino acids in the C-terminal portion of the protein. The type 1 pili encoded by *fim* genes are present in both *E. coli* and *Salmonella* species (Boyd & Hartl, 1999). The *fimZ* gene is an essential component of the regulatory cascade involved in fimbrial production, and its expression varies according to environmental conditions. The ability of FimZ to bind to the promoter region of the *S. typhimurium fimA* gene has been demonstrated and a *fimZ* mutant was found to be nonfimbriate (Yeh *et al.*, 1995). The C-terminal region of FimZ regulators contains sequence reminiscent of a helix-turn-helix DNA binding motif; in contrast, TcF3 exhibits only weak similarity to this structural motif. However, closer examination of the C-terminal region of TcF3 revealed the presence of several possible serine/threonine kinase phosphorylation sites (SRR, SMR, SKR, TDR), suggesting that the action of this protein could depend on phosphorylation. The apparent lack of a DNA binding motif that is a characteristic of transcriptional activators and the small size of the protein and lack of apparent transmembrane domains that characterise many phosphorylating sensors suggests a different role for TcF3. The TcF3 protein may interact with another protein in order to achieve DNA regulation or alternatively, it may act within a unique phosphorelay system.

ORFe codes for a putative protein, designated TcF4. Protein sequence analysis showed that it belongs to a long list of hypothetical proteins, eukaryotic DNA binding proteins and AlgP regulatory protein of *P. aeruginosa*. Eubacteria maintain their DNA in a negatively supercoiled form by DNA topoisomerase and by architectural elements called histone-like proteins, e.g. HU, HNS, and IHF (Pettijohn, 1988; Owen-Hughes *et al.*, 1992). These chromatin-associated proteins organise the bacterial chromosome and also exert regulatory influence on transcription, recombination, and DNA replication (Schmid, 1990). AlgP is a

histone-like protein that is involved in the regulation of alginate synthesis in Pseudomonads. Strains of *P. aeruginosa* isolated from the lung of cystic fibrosis patients frequently have a mucoid phenotype due to the synthesis of a slimy exopolysaccharide called alginate, the production of which is partially under the control of AlgP. Prokaryotic chromatin-associated proteins with homology to histone H1 have so far only been identified in pathogenic bacteria, including *P. aeruginosa* and *B. pertussis* (Goyard, 1996). The AlgP protein contains multiple repeats of Ala-Ala-Lys-Pro, as do several other proteins that resemble histones. Studies utilising synthetic oligopeptides have shown that these repeats bind to DNA *in vitro* (Medvedkin *et al.*, 1995). Although none of these repeats are present in TcF4, the relatively high content of these residues in the protein may be the reason for this homology.

ORFg codes for a member of the lipase family protein that possesses the lipase serine active site, GXSXG, which is conserved in all known lipases (Derewenda & Cambillau, 1991). Many different bacterial species produce and secrete lipases into the extracellular medium, which hydrolyse esters of glycerol with preferably long-chain fatty acids. They act at the interface generated by a hydrophobic lipid substrate in a hydrophilic aqueous medium. The secretion pathway is known for Pseudomonad lipases, *P. aeruginosa* lipase uses a two-step mechanism while *P. fluorescens* lipase uses a one-step mechanism. Additionally, some Pseudomonad lipases need specific chaperone-like proteins assisting their correct folding in the periplasm. The catalytic site of lipases is buried inside the protein and contains a serine-protease-like catalytic triad consisting of the amino acids serine, histidine, and aspartate (or glutamate). The serine residue is located in a strictly conserved beta-epsilon serine-alpha motif. The active site is covered by a lid-like alpha-helical structure, which moves away upon contact of the lipase with its substrate, thereby exposing hydrophobic residues at the protein's surface and mediating the contact between protein and substrate (Jaeger *et al.*,

1994). Lipases of pathogenic bacteria are known to contribute to their pathogenicity, for example, the lipases of *P. aeruginosa* (Wohlfarth *et al.*, 1992) and *S. aureus* are produced during infection and impair the function of different cell types involved in the human immune response, such as macrophages or platelets. The similarity of the TcF5 protein to phospholipase A1 in both *Serratia liquefaciens* and *Neisseria enterococci* may be attributed to the lipase serine active site. The phospholipase A1 of both of these species is known to be excreted to the outer environment, mainly by means of the N-terminal signal peptide, which upon membrane transfer, is cleaved by a specific signal peptidase, resulting in the liberation of the mature protein on the exterior of the membrane (Givskov *et al.*, 1988, Heijne & Abrahamsen, 1989). Inspection of the N-terminal sequence of TcF5 protein showed that it lacks the features usually associated with a signal peptide. However, the pathways for the secretion of extracellular enzymes in Gram-negative bacteria include at least two routes of transport for proteins. One depends on the induction of specific helper proteins to assist in the secretion of the enzyme, while the other is based on lysis through the enzymatic activity of the protein. For example, secretion of *S. liquefaciens* phospholipase A1 in *E. coli* was partially dependent on the induction of a helper protein, PhlB. Furthermore, the secretion of metalloprotease from *Serratia marcescens* and *Erwinia chrysanthemi* has been demonstrated to be totally dependent on the presence of three other genes encoding protease secretion functions (Letoffe *et al.*, 1991). Another example is the α-haemolysin found in haemolytic *E. coli* strains, for which secretion appears to be independent of the general protein secretion pathway. The secretion apparatus for this group of proteins consists of two inner-membrane proteins and one outer-membrane protein, which act in concert to direct the protein across both cell membranes, without accumulation in the periplasm (Lory 1992, Pugsley, 1993). Furthermore, exoenzyme S, an ADP-ribosyltransferase of *P. aeruginosa*, which also lacks

an N-terminal signal peptide is secreted without processing, by a type III secretion system (Yahr *et al.*, 1996).

The neighbouring open reading frame to ORFg, ORFh, lacks an obvious promoter and may be transcribed using the promoter located in the 5' untranslated region of ORFg. The ATG translational start is followed immediately by the base A (ATGA). It has been shown that this arrangement of bases causes the most efficient binding of tRNA, which would lead to enhanced initiation of translation (Taniguchi *et al.*, 1978). In addition, shifting of the ribosome by one base at this codon would lead to the reading of TGA stop codon. This may be a mechanism by which the organism shuts off translation at the earliest possible opportunity should such a shift occur.

ORFh codes for a protein, TcF6, with low homology to a type III export protein from *P. aeruginosa*, PscD. The low degree of homology between related proteins from different species is a common feature among components of type III secretion systems (Lee, 1997). Type III secretion systems consists of a secretion apparatus, made up of ~25 proteins, and an array of proteins may be released by this apparatus. Some of these released proteins are "effectors" which are delivered into the cytosol of the target cell, while others are "translocators" which help the effectors to cross the membrane of the eukaryotic cell (Bogdanove *et al.*, 1996; Cornelis & Gijsegem, 2000). All effectors that are delivered by type III secretion systems lack the classical N-terminal signal peptide sequence (Michiels *et al.*, 1990). In pathogenic bacteria, type III secretion systems allow bacteria adhering to the external surface of a host cell to inject specialised proteins across the plasma membrane. Many Gram-negative pathogens use a type III secretion system to translocate protein toxins across the bacterial cell envelope, including *P. aeruginosa*, *Bordetella bronchiseptica* and *B. pseudomallei* (Hueck, 1998; Winstanley *et al.*, 2000). The PscD protein is located in the inner membrane, where it mediates the transport of virulence factors through the bacterial

membrane of *P. aeruginosa*. The *pscD* gene is located in the *psc* locus of *P. aeruginosa*, which is homologous to the *psc* type III secretion system of Yersiniae that, in turn, is responsible for controlling the secretion of Yop virulence determinants *in vitro* (Michiels & Cornelis, 1991; Salmond & Reeves, 1993).

Taken together, the presence of a gene encoding a possible type III secretion protein, TcF6, downstream from a gene encoding a secretory phospholipase A1 homolog that lacks a classical N-terminal signal peptide, perhaps transcribed from the same promoter, may suggest a strategy for the secretion of TcF5.

The 23 amino acids constituting the N-terminal sequence of TcF7 strongly resemble a typical Gram-negative bacterial signal peptide sequence. The signal sequence, which distinguishes between exported and cytoplasmic proteins, is presumably required for secretion. The general feature of a signal peptide includes from one to three charged residues within the five amino acids and a core of at least nine hydrophobic residues that are believed to be sufficient to span a membrane, followed at the end by a helix-breaking residue (proline) within four to eight residues preceding the cleaving site, which very frequently is an alanine (Watson, 1984). The proposed signal sequence of TcF7 possesses all of these features. Interestingly, another peptidase cleavage site was identified 50 amino acids toward the C-terminus after the first signal peptide cleavage site. N-terminal signal sequences are usually found in precursors of proteins destined for extracellular localisation, as well as in membrane and periplasmic proteins. The release of such proteins from Gram-negative bacteria proceeds via a two-stage process. First, the N-terminal signal sequence directs the protein into the major secretion pathway before the signal peptide is cleaved by the leader peptidase. In the second stage of the export process, the protein is released from the cell (Lory, 1992). A wide range of bacterial enzymes can be exported with the help of a signal peptide including proteases, lipases and toxins. The presence of the second cleavage site

suggests that the 50 amino acid peptide following the first cleavage site may represent a propeptide. Propeptides are essential for proper folding of the enzymes to which they are attached and thus are defined as intramolecular chaperones (IMC) (Inouye, 1991; Shinde and Inouye, 1993). Propeptides are well represented in a growing family of proteases. Members of this family are generally synthesised as pre-proenzymes with an amino-terminal signal (pre) sequence, followed by a relatively long propeptide domain located between the signal and the mature enzyme sequence (Inouye, 1991). Proteases assigned to this family include proteinases that are mechanistically unrelated (e.g. serine and aspartate proteases), and show no striking homology which suggests a convergent evolution of this folding mechanism (Mclver *et al.*, 1995). A well-established member of this family is the elastase of *P. aeruginosa*, which synthesised as a pre-proenzyme containing a classical signal sequence and an amino-terminal propeptide. The signal sequence is removed upon passage through the inner membrane into the periplasm, where the propeptide is then rapidly cleaved off by autoproteolysis (Mclver *et al.*, 1993). The cleaved propeptide remains non-covalently associated with periplasmic elastase and appears to specifically inhibit its activity while cell bound, while the secretion of mature elastase requires a complex extracellular protein export apparatus (Bally *et al.*, 1992). Similarly, the pre-proenzyme structure that exists in TcF7 suggests that it may belongs to the IMC (protease) family. On the other hand, computerised searches of the GenBank and EMBL database using stringent matching criteria failed to identify other protein sequences with significant homology to the TcF7 protein. However, under less stringent conditions the protein has a limited similarity to a range of hypothetical proteins in both prokaryotes and eukaryotes.

One striking observation is that the amino acid composition of TcF7 contains an unusually high content of arginine (25%), alanine (18%), and proline (8%). Eukaryotic proteins that have a high arginine content, such as protamine, are known to be capable of

folding DNA. Protamine and DNA can form a compact structure, protecting DNA from digestion by intracellular enzymes (Yh *et al.*, 1999). Alternatively, the high proportion of arginine residues may be a result of the high frequency of the nucleotide G in the coding region of the gene, because all six arginine codons contain at least one G.

Nineteen *B. pseudomallei* isolates from various sources and recovered over 30 years, in addition to other closely related species, were analysed for the presence of the *mrgRS* locus by Southern blot analysis. In selecting the isolates, an attempt was made to represent the wide geographical distribution of *B. pseudomallei* and covered human and environmental sources from five different countries. Southern hybridisation using the RS7 DNA probe showed that a single copy of the 4.3 kb *Eco*RI fragment was present in the genome of all of these isolates. The 4.3 kb *Eco*RI fragment contains the flanking sequence upstream from the *mrgRS* locus (1.3 kb), *mrgR* and 2,342 bp of the total length of *mrgS* (3,237 bp). Moreover, the hybridization results demonstrated the high specificity of the cloned 4.3 kb insert in pMRG2 as a probe. It did not cross-hybridise with genomic DNA from closely related bacteria, such as *B. thailandensis* and *B. cepacia* nor did it hybridise with the DNA from the other Gram-negative bacteria that were tested, including *P. aeruginosa* and *E. coli*.

Hybridisation analysis of the region downstream from the *mrgRS* locus was performed using the FR1 DNA probe. This probe spans the distal end of *mrgS* and 285 nucleotides further downstream from the *mrgS* stop codon. The results of hybridisation analysis revealed that most of the *B. pseudomallei* isolates that were examined were the same in this region although two isolates, 392 and H55, lacked the 684 bp band (the fragment encoding 228 amino acids of the C-terminal end of *mrgS*), while one isolate, 576, lacks both bands. In addition, H55 has a 4.5 kb band. The result suggests that in these strains this region is subject to some variability and may be wholly or partly absent. Furthermore, hybridisation analysis using the FR3 DNA probe also indicated that the flanking region downstream from

the *mrgRS* locus is a region of genetic variation. Many studies have demonstrated considerable variation in the genetic structure of bacterial species although the reasons for this variation are sometimes unclear. This is also true for *B. pseudomallei* where many studies have demonstrated that substantial variation exists across the genome using a wide variety of methods (reviewed in section 2.2.14). However, few studies have been carried out to assess the extent of variation at specific loci in *B. pseudomallei*. Haase *et al.* (1995) speculated that the large number of organisms in tropical countries in which melioidosis is endemic and the frequent interactions that may be possible in that environment and within the human host creates opportunities for genetic change. However, a mutation that confers a selective advantage in one environment is unlikely to be a benefit in all environments (Cohan, 1994). Winstanley *et al.* (1998) assessed the extent of variation in the flagellin genes of *B. pseudomallei* by comparing the sequences that were obtained from environmental and clinical isolates of the pathogen. The sequences showed either 100% identity or differed by a single nucleotide demonstrating a remarkably conservative nature in these genes. In this study, a comparison of the sequences that were obtained from two isolates, one clinical and one environmental, in a small region of the *B. pseudomallei* genome downstream from the *mrgRS* locus revealed the presence of no less than 19 nucleotide substitutions. One of these substitutions resulted in a change in the sequence GAACTC to GAATTC, the recognition sequence for the restriction enzyme *Eco*RI. The consequent restriction fragment length polymorphism (RFLP) was shown to separate the *B. pseudomallei* isolates into two distinct groups of almost equal proportions (44% and 56%). Each group contains isolates from clinical and environmental sources and from different geographic origins. There was no apparent association between the source of the isolates, the date and place of isolation, and the RFLP pattern. Nevertheless, the substitutions represent a stable genetic modification since they have been retained over time. While it is possible that

these sequence differences were due to point mutations randomly arising in a bacterial population the almost equal distribution of the polymorphism among the isolates suggests that it is much more likely that some process of selection may have occurred. For example, differences in primary or secondary host factors may play a role in the selection for a mutation arising within a bacterial population.

Alignment of the sequences encoding TcF5, which were obtained from isolates 204 and 112 and contained 16 variable nucleotides, predicted the occurrence of 6 amino acid substitutions. These changes surround the lipase active site of TcF5. Interestingly, there was neither a change in the hydropathic character of the substituted residues nor a remarkable difference in the overall molecular weight of the protein. Substitutions occurring outside the active site of a protein have been suggested as a means by which adaptive changes arise in protein structure. Among five proteins studied in bacteria, polymorphic amino acids tended to occur on the exposed surfaces of a protein, their side chain accessible to solvents (Bustamante *et al.*, 2000). In *E. coli* IDH, switching cosubstrate from NAD^+ to $NADP^+$ involves changing five amino acids around the cosubstrate binding site. In mammalian examples, the major pancreatic enzymes; trypsin, chemotrypsin, and elastase, are kinetically very similar, catalysing the hydrolysis of peptides and synthetic ester substrates. The polypeptide backbones of all three are essentially the same. The difference in their specificities lies mainly in a pocket that binds the amino acid side chain (Perona & Craik, 1997). In trypsin, the residue at the bottom of the pocket is an aspartate instead of the serine of chymotrypsin, while in elastase, the two glycines at the mouth of the pocket in chemotrypsin are replaced by a bulky valine and threonine. Although the amino acid substitutions were not in the active site (GXSXG) of TcF5, it may nevertheless have some effect on its functional specificity. Studies utilising anti-catalytic antibodies have shown that amino acids replacements that are distant from the active sites of enzymes can dramatically

alter the activity of a molecule (Wedemayer *et al.*, 1997). Furthermore, substitutions outside the active site may alter the flexibility of the active site by increasing the conformational entropy of the enzyme. *B. pseudomallei* is able to adapt to and survive in two broadly distinct habitats, the internal environment of the host and the environment external to the host. In addition, it is possible that *B. pseudomallei* can survive in soil and water by parasitising other organisms including amoebae, fungi or plant (Inglis *et al.*, 2000; Pitt *et al.*, 2000). The observed genetic variation in ORFg among the *B. pseudomallei* isolates that were used in this study may represent an adaptation of the bacterium to two different survival strategies, such as secondary hosts.

CHAPTER 6

THE CONSTRUCTION AND EXPRESSION OF MrgR AND MrgS FUSION PROTEINS AND THEIR RECOGNITION BY POLYCLONAL ANTIBODIES

6.1 INTRODUCTION

Following the cloning and characterisation of the *mrgRS* locus and its flanking regions, it was desirable to verify gene expression, *in vitro*. Gene expression can be studied by a number of methods, either at the mRNA level or at the protein level. Investigation at the protein level can provide useful information about protein structure and function. Therefore, as a step toward understanding and verifying the expression of the MrgR and MrgS proteins in *B. pseudomallei*, polyclonal antibodies were raised against each of the two proteins. Specific antibodies are frequently the key to the characterisation of protein structure, function and distribution. They provide tools for the identification of the protein in ELISA, immunoblots and immunohistology, permit efficient purification by affinity chromatography and allow functional characterisation by blocking binding or active sites in the molecule. The production of specific antibodies usually requires the purification of the native protein from cells, tissues or fluids. This strategy is not possible for all proteins for reasons that include the scarcity of starting material, lack of a functional assay and the low yields of the pure protein obtained (Spiller et al., 1999). With the advent of molecular biology, it became possible to obtain a large amount of information about a protein without ever having purified it. Antibodies may be generated against peptides derived from the predicted sequence and although this approach may be problematic, short peptides are expected to yield linear epitopes, the antibodies can be useful for the detection of denatured protein by immunoblotting (Atassi, 1986; Getzoff et al., 1988). On the other hand, purification of a recombinant protein requires the efficient expression of the foreign gene in *E. coli* and is dependent on the presence of the appropriate signals that can be recognised by the bacterium, which provide instructions for the transcriptional and translational apparatus of the cell. These signals include a promoter, a terminator and a ribosome-binding site.

The objective of this study was to produce polyclonal antibodies that would be useful

for the further characterisation of the MrgR and MrgS two-component regulatory proteins, and to evaluate the use of these antibodies for the detection of the expressed proteins.

6.2 MATERIALS AND METHODS

6.2.1 Construction and expression of recombinant fusion proteins in *E. coli* K-12

6.2.1.1 pMAL-C2 expression vector

The pMAL-C2 plasmid vector (**Fig 6.1**) provides a means for expressing and purifying a protein produced from a cloned gene or open reading frame using *E. coli* as host. The cloned gene is inserted downstream from the *malE* gene of *E. coli* K-12, which encodes maltose-binding protein (MBP) resulting in the expression of an MBP fusion protein (Guan *et al.*, 1987). The method uses the strong "tac" promoter to sustain a high rate of transcription, and hence expression of the cloned sequence. The vector expresses the *malE* gene fused to the *lacZ*α gene. Insertion of a foreign gene into the multiple cloning site inactivates the β-galactosidase α-fragment activity and allows for blue/white screening (section 3.5.5). In addition, the vector carries the *lacI* gene, which codes for the Lac repressor. This keeps expression from P_{tac} low in the absence of IPTG induction. Furthermore, the multiple cloning site allows the insertion of a cloned gene in the correct orientation by using directional cloning thereby ensuring the expression of the gene.

6.2.1.2 Construction of *mrgR* and *mrgS* gene fusions

The procedure was performed according to the instructions provided by the manufacturer of the pMAL-C2 fusion vector (New England Biolabs), and to the general methods for restriction enzyme digestion (section 3.4.5), cloning into plasmid vectors (sections 3.5 & 3.6), restriction mapping (section 3.7), DNA purification (section 3.4.4) and DNA sequencing (section 3.8).

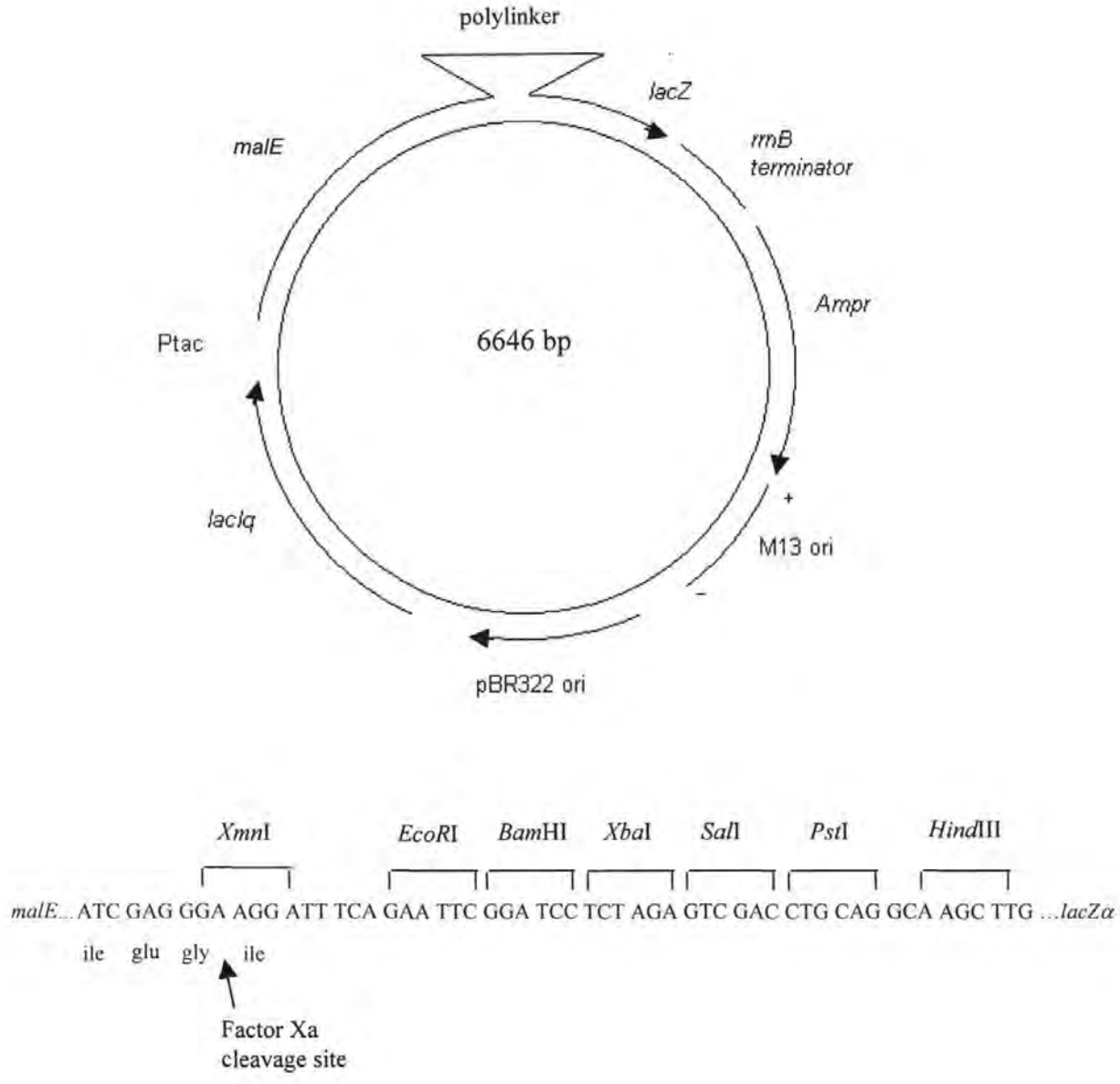


Fig 6.1: Schematic presentation of the pMAL-C2 cloning vector (New England Biolabs).

6.2.1.3 Expression and identification of recombinant fusion proteins

Following the isolation of *E. coli* DH5α clones in pure culture, 10 ml LB broths (Appendix III) supplemented with ampicillin (100 μg/ml) and 0.2% glucose were inoculated and incubated at 37°C with shaking (275 rpm) to A_{600} ~ 0.5, usually for 3-4 h. A 1 ml sample of the uninduced culture was removed, centrifuged (10,000 x g, 2 min, RT) and the cell pellet was resuspended in 50 μl of SDS-PAGE sample buffer (Appendix III). To induce the expression of fusion proteins produced by the clones, IPTG, 0.3 mM final concentration, was added to the remaining cultures and incubation was continued for a further 2 h. For each culture, a 0.5 ml volume was removed, centrifuged (10,000 x g, 2 min, RT) and the cell pellet was resuspended in 100 μl of SDS-PAGE sample buffer. The samples were then analysed by SDS-PAGE (sections 3.14.5 & 6.2.1.5).

To determine the behaviour of each MBP fusion protein, a small-scale pilot experiment was performed. Cultures were grown overnight at 37°C (200 strokes/min) in 10 ml of LB broth supplemented with ampicillin (100 μg/ml) and 0.2% glucose. To 80 ml of pre-warmed LB broth containing ampicillin (100 μg/ml) and 0.2% glucose, a 0.8 ml volume of the overnight culture was added, and incubation was continued at 37°C until an OD of A_{600} ~ 0.5 was reached. A 1 ml sample of the uninduced culture was removed, centrifuged (10,000 x g, 2 min, RT) and the cell pellet was resuspended in 50 μl of SDS-PAGE sample buffer, and placed on ice. To the remaining culture, IPTG was added to a final concentration of 0.3 mM and the incubation was continued for 2 h. A 0.5 ml sample was removed, centrifuged (10,000 x g, 2 min, RT) and the cell pellet was resuspended in 100 μl of SDS-PAGE sample buffer and placed on ice. Following the induction period, 40 ml of the culture was harvested by centrifugation (2,500 x g, 10 min, RT), and the cell pellet was resuspended in 5 ml of column buffer (Appendix III). Samples were frozen overnight at -20°C, thawed in an ice-water bath and sonicated on ice in short pulses of 3 s for 3 min (W-385 Ultrasonic

Processor, Heat Systems Ultrasonics). The samples were centrifuged (9,000 x g, 20 min, 4°C) and the supernatant (crude extract) was removed and retained on ice, while the pellet was resuspended in 5 ml of column buffer (insoluble matter). A 50 µl volume of amylose resin was washed twice with 1.5 ml volumes of column buffer and then mixed with 50 µl of the crude extract. The mixture was incubated on ice for 15 min, then centrifuged (9,000 x g, 1 min, RT) and the cell pellet was washed with 1 ml of column buffer. After centrifugation (9,000 x g, 1 min, RT), the pellet was resuspended in 50 µl of SDS-PAGE buffer. All of the samples were analysed by SDS-PAGE (sections 3.14.5 & 6.2.1.5).

6.2.1.4 SDS-PAGE analysis

SDS-PAGE gel electrophoresis was performed as stated in section 3.14.5. The resolving gel consisted of 10-12% T (Total acrylamide concentration) and a stacking gel of 5% T. The samples of uninduced cells, induced cells, crude extracts, insoluble material and resin purified protein, which had been mixed with SDS-PAGE sample buffer, were boiled for 5 min. Samples were separated on SDS-PAGE gels that included protein molecular weight standards and purified maltose-binding protein (New England Biolabs). Following electrophoresis, the gels were either stained for protein with Coomassie brilliant blue (section 3.14.4) or electro-blotted onto nitrocellulose and probed with the appropriate antiserum (section 3.14.5 & 6.3.4).

6.2.2 Development of polyclonal antibodies

6.2.2.1 Designing synthetic peptides

A number of factors should be considered when designing synthetic peptide antigens for the development of an antibody in order to improve the chances of a favourable outcome. Ideally, a synthetic peptide should be approximately 6 to 20 amino acid residues in length, with a high hydrophilicity value, and minimal homology with any other known protein sequences. Another useful consideration when selecting peptide sequences is the

presence of the candidate peptide in the C- or N-terminal end of the protein, to increase the probability of surface orientation (Hopp & Woods, 1986; Doolittle, 1986). To determine the degree of homology of the selected peptides to known protein sequences, the BLAST server at NCBI and Blitz was used to search the nr database (All non-redundant GenBank CDS translations + PDB + SwissProt + PIR + PRF). Several different methods for performing the search were used. Both the old (Ungapped) BLAST program and the BLAST2 (Gapped BLAST) program were used and the Expect parameter (E) was increased from the default of 10 to a value of at least 1000. This allows a wider search for less conserved sequence homology with low statistical significance. Following this selection process, the peptides were synthesised by Bethyl Laboratories Inc. (Montgomery, USA) and the purity of the peptides (>90%) was determined by HPLC and verified by mass spectroscopy.

6.2.2.2 Production of polyclonal antibodies

Following synthesis and purification of the peptides rabbit polyclonal antibodies were raised using the synthetic peptides as immunogens, at Bethyl Laboratories Inc. (Montgomery, USA). Typically, 2-3 mg of each peptide was conjugated to keyhole limpet hemacyanin (KLH), using maleimide chemistry, linking sulfhydryl-containing residues of the peptide to the carrier. The conjugated peptide was mixed with Freund's complete adjuvant (FCA) and injected subcutaneously at multiple sites into 2 New Zealand white rabbits. A "two-week" immunisation protocol was followed, with bleeds performed in alternate weeks (**Table 6.1**).

Following immunisation, the hyperimmune serum was harvested from rabbits immunised with peptide-KLH and the peptide-specific antibodies were purified by affinity-chromotography using a peptide-Sepharose coupled matrix. Finally the affinity-purified antibodies were filter sterilised and supplied at a concentration of about 1 mg/ml.

Table 6.1: Immunisation protocol for raising anti-peptide antibodies in rabbits

Day 0	Pre-bleed, Antigen injection
Day 14	Antigen injection
Day 28	Antigen injection
Day 35	Bleed
Day 42	Antigen injection
Day 49	Bleed
Day 56	Antigen injection
Day 63	Bleed

6.2.2.3 Western blot analysis

Samples separated by SDS-PAGE were transferred electrophoretically to nitrocellulose membrane by using the Bio-Rad Trans-Blot apparatus as described previously (section 3.14.5). Blots were washed (3 x 5 min in PBS), then placed in 1% casein blocking solution (Appendix III) at RT for 2 h. Primary and secondary antibodies were diluted in the blocking solutions to reduce non-specific binding. The blots were incubated overnight at 4°C in primary antibody, washed (3 x 5 min in 1% casein), and then incubated in secondary antibody conjugated to horseradish peroxidase (Dako A/S, Glostrup, Denmark) for 2 h at RT. A list of the primary and secondary antibodies and the concentrations that were used is illustrated in **Table 6.2**. Following antibody incubation, the blots were washed 3 x 10 min in PBS, and antibody binding was visualised (section 3.14.5).

Table 6.2: Different primary and secondary antibodies and the working dilutions that were used in this study.

Primary Antibody	Dilution	Secondary Antibody	Dilution
Anti-MrgR	1:500	Swine Anti-rabbit IgG peroxidase	1:2000
Anti-MrgS	1:300	Swine Anti-rabbit IgG peroxidase	1:1000
Anti-MBP	1:1000	Swine Anti-rabbit IgG peroxidase	1:2000
Convalescent melioidosis patient serum	1:300	Rabbit Anti-human IgG peroxidase	1:1000

6.2.3 Specificity of antibodies

For testing the specificity of the antibodies and to determine non-specific binding, rabbit polyclonal anti-peptide antisera were preabsorbed either with the specific peptide immunogen or with an irrelevant peptide. Peptides, supplied as lyophilised powder, were reconstituted with distilled water, in the case of MrgR, or 1% NH_3, for MrgS, to a stock concentration of 2 mg/ml. The preabsorption was performed by incubating 1 ml (1 mg/ml) rabbit serum with 50 µl of the peptide immunogen (2 mg/ml) for 60 min at 37°C, and then overnight at 4°C with agitation. The preabsorbed antibody was then centrifuged (10,000 x g, 10 min, RT) and stored at -20°C until use. Blots performed with preabsorbed antibody were developed as explained above (6.2.2.3).

6.2.4 Detection of MrgR and MrgS proteins in cell lysates and extracellular products of *B. pseudomallei*

Western blotting of cellular and extracellular components from *B. pseudomallei* cultures was used to identify MrgR and MrgS. The whole cell proteins from 24 isolates of *B. pseudomallei* were prepared from broth cultures and separated by SDS-PAGE (**Fig 1**, Appendix VI). Whole cell lysates from 6 isolates of *B. pseudomallei*, representing both clinical (112, 204, 576) and environmental sources (E8, 98, 216) (**Table 3.1**), were chosen

for further analysis, subjected to electrophoresis (section 3.14.3) and transferred to nitrocellulose membrane (section 3.14.5). In order to obtain cell lysates, cultures were grown statically within a Containment Level 3 Laboratory (section 3.2.2) in LB media (Appendix III) for 48-72 h, centrifuged (4,000 x g, 15 min, 4°C), and washed in PBS. The cell pellet was lysed in a solution containing 10% SDS, 0.5 M Tris-HCl (pH 6.8), mercaptoethanol and glycerol (Appendix III), which is typical of SDS sample buffer but lacking bromophenol blue. To confirm sterility, viability tests were performed before samples were released from the CL3 (section 3.2.3). Approximately 5 μg of crude lysate, measured by the microassay procedure (section 3.14.2), was boiled for 5 min, separated by SDS-PAGE on 10 or 12% gels (section 3.14.3) and processed for Western blot analysis as explained earlier (sections 3.14.5 & 6.2.2.3). For the recovery of extracellular products (ECPs) from culture supernatants, 10 ml volumes of *B. pseudomallei* broth cultures were grown as above and centrifuged (5,000 x g, 20 min, 4°C). The resulting supernatant was aliquoted into sterile 1.5 ml microfuge tubes and centrifuged again (15,000 x g, 10 min, 4°C). Following this step, the supernatant was carefully removed and concentrated by a factor of 250 using Amicon Minicon Miniplus units, 10,000 nominal mw cutoff (Millipore Ltd, Watford, UK). The protein concentration of the ECPs was determined by the microassay procedure (section 3.14.2). A 2 μg protein sample of the ECPs from 6 isolates of *B. pseudomallei* were boiled for 5 min, separated by SDS-PAGE and transferred to nitrocellulose membrane as described (sections 3.14.3, 3.14.5 & 6.2.2.3).

6.3 RESULTS

6.3.1 Construction of gene fusions

6.3.1.1 Construction of *malE-mrgR* fusion

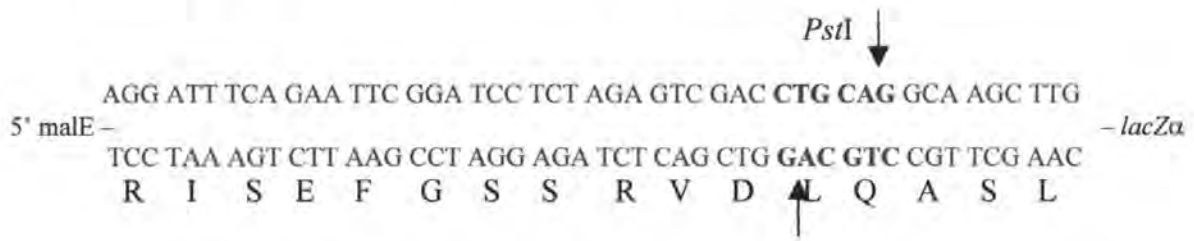
From the restriction map available for pMRG2 (section 4.3.10), suitable restriction

sites for the construction of an in-frame gene fusion were identified. A translational fusion of *malE* and *mrgR* was constructed in the pMAL-C2 expression vector (**Fig 6.2**). Plasmid pMRG2 was digested with *Pst*I and *Nsi*I (section 3.4.5.2), releasing a DNA fragment length of 701 bp, consisting of 472 bp of *mrgR* gene and 229 bp from the 3' downstream region. The fragment was separated and purified by agarose gel electrophoresis (sections 3.4.6.1 & 3.4.6.2), and ligated (section 3.5.3) to pMAL-C2, which had been previously digested with *Pst*I, dephosphorylated and purified. The stop codon for this construction is provided by the *mrgR* stop codon. *E. coli* XLI-Blue was used as the host for transformation with the recombinant plasmid DNA (section 3.5.4). Clones were selected on LB agar containing ampicillin (100 μg/ml) and small-scale plasmid DNA preparations were made (section 3.4.2.3), and cleaved with *Eco*RI and *Mlu*I, in single digests, and *Eco*RI-*Pst*I, in double digest, to confirm the correct orientation of the inserted DNA (**Fig 6.3**). The fidelity of each gene fusion was confirmed by DNA sequencing (section 3.8). One clone, designated pMFU1, was selected for further analysis. A complete restriction map of pMFU1 is shown in **Fig 6.4**.

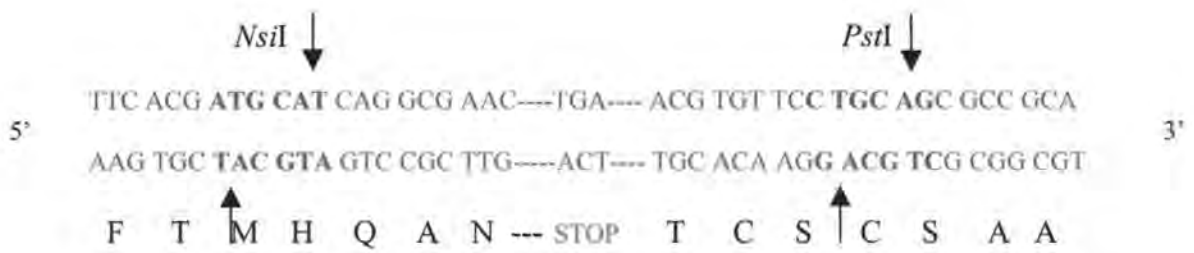
6.3.1.2 Construction of *malE-mrgS* fusion

For cloning purposes, the N-terminal region of *mrgS* was amplified by PCR, using pMRG2 as a template. A translational fusion of *malE* with this *mrgS* fragment was constructed in the pMAL-C2 expression vector (**Fig 6.5**). Two oligonucleotide primers (**Table 6.3**) were designed (section 3.13.1) to amplify a 749 bp fragment that consists of 48 nucleotides upstream from the *mrgS* start codon and a further 701 nucleotides downstream from the *mrgS* start codon. The 5' end of the forward primer included a 9 nucleotide addition containing an *Eco*RI endonuclease cleavage site plus 4 extra bases to allow efficient cleavage of the amplified product. This exogenous sequence was incorporated into the PCR product and facilitated subsequent cloning steps. The reverse primer contained a sequence

(a) The pMAL-C2 was digested with *Pst*I restriction enzyme.



(b) The *mrgR* gene was digested with *Nsi*I that cuts 192 bp downstream from the *mrgR* start codon and *Pst*I that cuts 228 bp downstream from the *mrgR* stop codon.



(c) On ligation, a translational fusion is produced

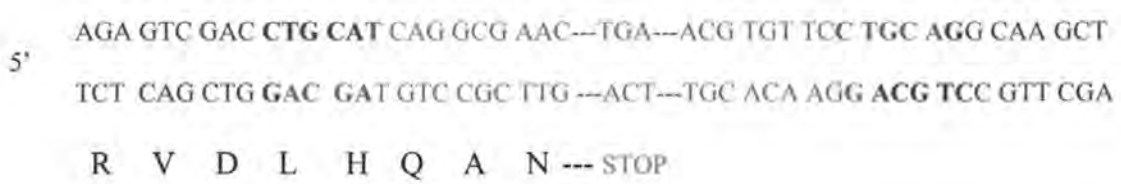


Fig 6.2: Cloning strategy for the translational fusion of *mrgR* and *malE* in pMAL-C2 expression vector.

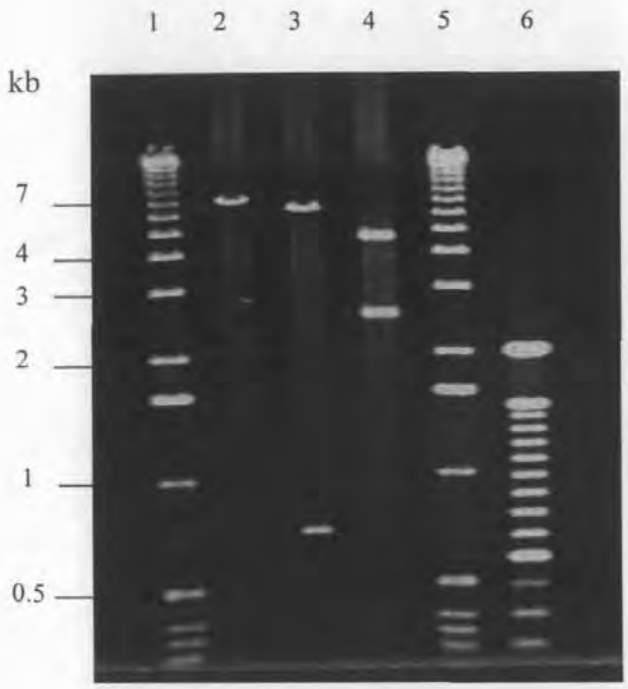


Fig 6.3: Agarose gel electrophoresis of restriction enzyme digests of pMFU1. The sizes of fragments were estimated by comparison with 1 kb (Lane 1 & 5) and 100 bp (Lane 6) DNA ladders.

Track	Enzyme (s)	No. of DNA fragments	No of cleavage sites on vector	insert	Approx. size of DNA fragments (kb)	Total size (kb)
2	*Eco*RI	1	1	0	7.347	7.347
3	*Eco*RI & *Pst*I	2	1	1	0.729, 6618	7.347
4	*Mlu*I	2	1	1	2.574, 4.773	7.347

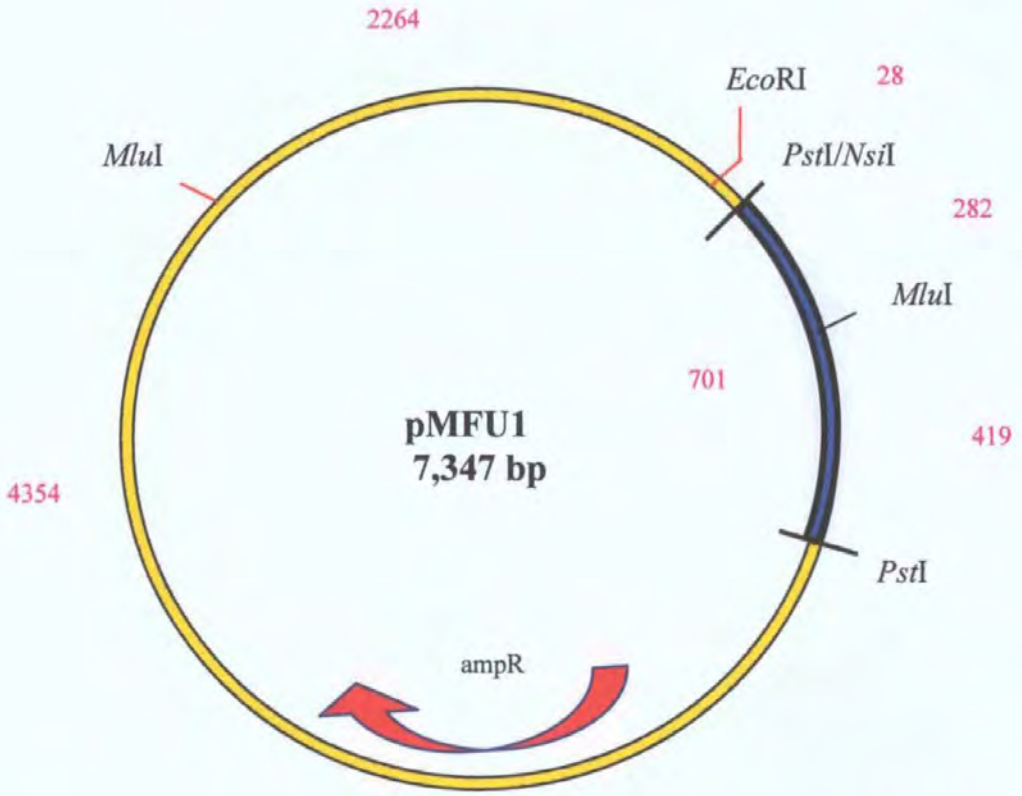
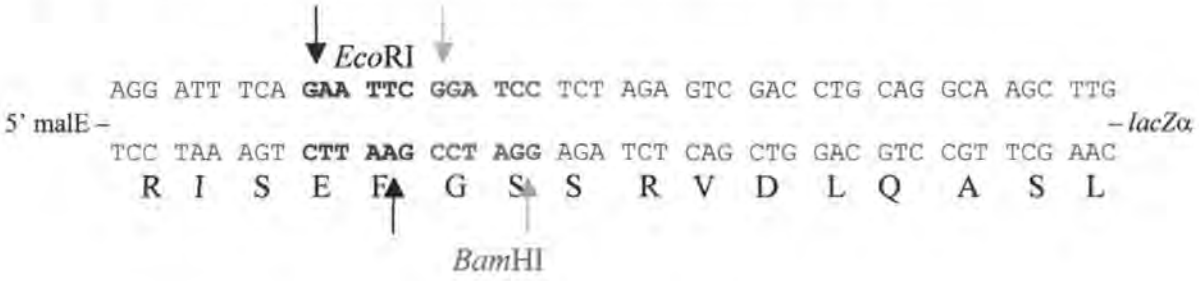
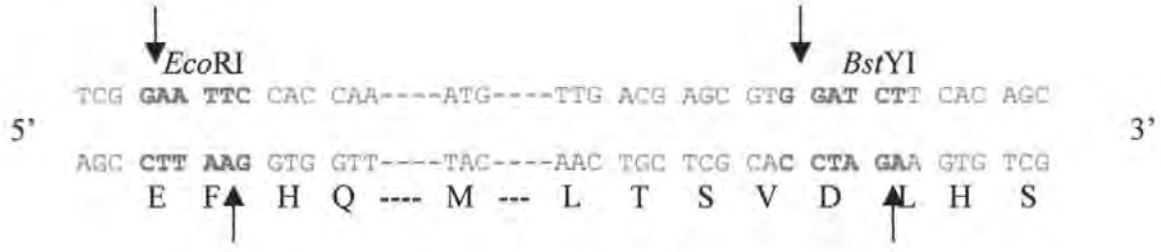


Fig 6.4: Physical map of pMFU1 deduced from restriction enzyme digests. The pMAL-C2 vector DNA is represented by the yellow area, while the *mrgR* insert is in blue. The *ampR* gene encodes ampicillin resistance, while numbers represent the number of base pairs between adjacent restriction sites.

(a) The pMAL-C2 expression vector was digested with *Eco*RI and *Bam*HI



(b) The proximal portion of the *mrgS* gene was amplified by PCR using a forward primer containing an *Eco*RI site and a reverse primer containing a *Bst*YI site. The amplicon was digested with both enzymes.



(c) On ligation, a translational fusion is produced

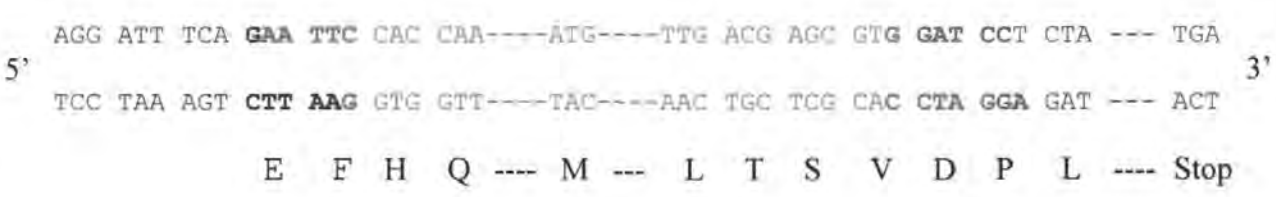


Fig 6.5: Cloning strategy for the translational fusion of the proximal part of *mrgS* and *malE* in pMAL-C2 expression vector.

that can be cleaved with the restriction enzyme *Bst*YI, thereby producing a cohesive end (R/GATCY) that is compatible with *Bam*HI cleaved DNA (G/GATCC). The amplified DNA fragment was purified (section 3.4.4.3), digested with *Eco*RI and *Bst*YI (section 3.4.5.2), re-purified (section 3.4.4.3) and then ligated (section 3.5.3) into *Eco*RI/*Bam*HI-digested and dephosphorylated pMAL-C2. The stop codon for this construct is provided by the vector, 30 nucleotides downstream from the *mrgS* insert. After transformation (section 3.5.4.1) into *E. coli* XLI-Blue, white colonies (recombinant clones) were recovered on LB agar plates (Appendix III) containing 100 μg/ml ampicillin. Plasmid DNA was extracted from selected colonies (section 3.4.2.3) and the presence of the correct insert was verified by restriction analysis with *Eco*RI, *Pst*I and *Eco*RV in single digests and *Eco*RI-*Hind*III in double digests (section 3.4.5.2) (**Fig 6.6**) and by DNA sequencing (section 3.8). One clone, containing plasmid DNA designated pMFU2, which contains the correctly oriented DNA insert, was selected for further analysis. A complete restriction map of pMFU2 is shown in **Fig 6.7**.

Table 6.3: PCR primers used for the amplification of a fragment encoding the N-terminus of *mrgS*.

Forward primer (FusF)	CTCGGAATTCCACCAACGTCGATCTCATCAACTG
Reverse primer (FusR)	GTGAAGATCCACGCTCGTCAAGACGAGCGCGTAG

6.3.2 Expression of fusion proteins

6.3.2.1 Expression of MBP-MrgR fusion protein

Following the construction of pMFU1, containing the translational fusion of *mrgR* to *malE*, the fusion protein was expressed in *E. coli* XL1-Blue. The *Pst*I-*Nsi*I truncated *mrgR* gene would encode a product consisting of 165 amino acids and 16,817 molecular weight. When fused to the C-terminal end of maltose-binding protein, which has a molecular mass of 42,700 daltons, a product of molecular weight 59,517 was expected. The expression of

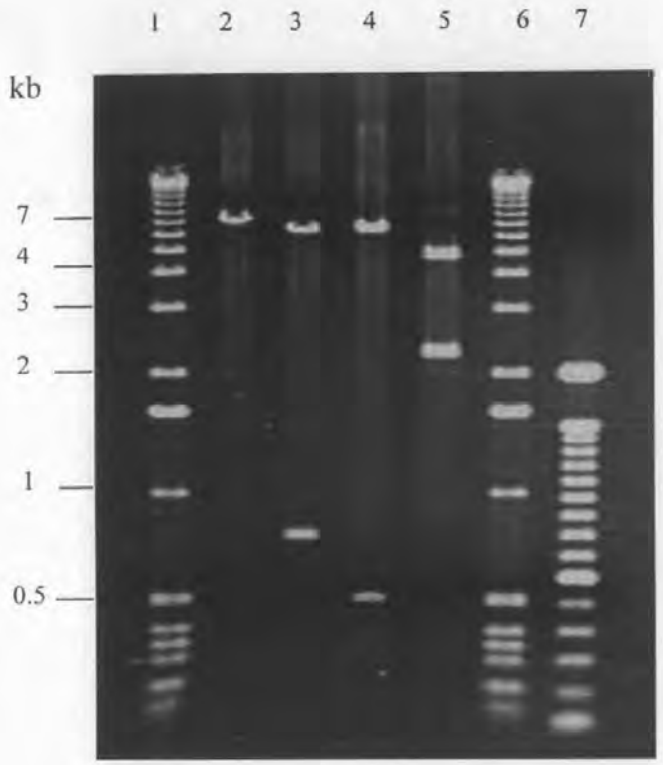


Fig 6.6: Agarose gel electrophoresis of restriction enzyme digests of pMFU2. The sizes of fragments were estimated by comparison with 1 kb (Lane 1 & 6) and 100 bp (Lane 7) DNA ladders.

Track	Enzyme (s)	No. of DNA fragments	No of cleavage Sites on vector insert	Approx. size of DNA fragments (kb)	Total size (kb)
2	*Eco*RI	1	1 0	7.395	7.395
3	*Eco*RI & *Hind*III	2	1 1	6620, 775	7.395
4	*Pst*I	2	1 1	6908, 487	7.395
5	*Eco*RV	2	1 1	5067, 2328	7.395

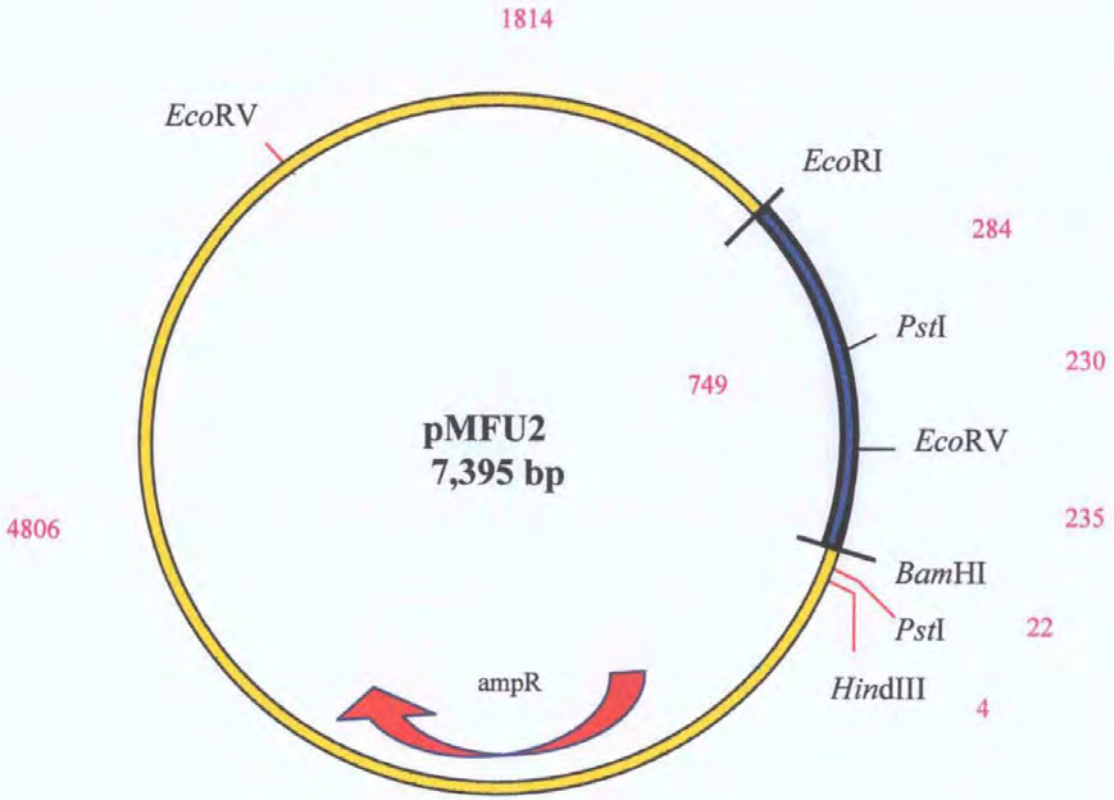


Fig 6.7: Physical map of pMFU2 deduced from restriction enzyme digests. The pMAL-C2 vector DNA is represented by the yellow area, while the *mrgS* insert is in blue. The *ampR* gene encodes ampicillin resistance, while numbers represent the number of base pairs between adjacent restriction sites.

MBP-MrgR was identified as a major protein band (molecular mass ~ 60 kDa) when cell extracts of bacteria containing pMFU1 that had been induced with IPTG, were separated on SDS-PAGE gels and stained with Coomassie brilliant blue (**Fig 6.8**). Hence, the size of the inducible band corresponds to the predicted size of the fusion. The small-scale pilot experiment revealed that the MBP-MrgR fusion protein was expressed at a high level, mainly in the soluble form.

6.3.2.2 Expression of MBP-MrgS fusion protein

The translational fusion of *mrgS* with *malE* was constructed in pMFU2 and expressed in *E. coli* XL1-Blue. The 749 bp *Eco*RI-*Bst*YI amplicon encoding the N-terminus of MrgS would yield a protein consisting of 249 amino acids of molecular weight 25,378. An additional 30 bp of vector DNA, encoding 10 amino acids of molecular weight 1,110, is included before the translational stop. When fused to the C-terminal end of maltose binding protein (molecular mass of 42,700 daltons) a product of molecular weight 69,188 was expected. The expression of MBP-MrgS was identified as a faint but discernible protein band (molecular mass ~ 70 kDa) when bacteria containing pMFU2, which had been induced with IPTG, were loaded onto SDS-PAGE gels and stained with Coomassie brilliant blue (**Fig 6.9**). Hence, the size of the inducible band corresponds to the predicted size of the fusion. The small-scale pilot experiment revealed that the MBP-MrgS fusion protein was expressed mainly in an insoluble form.

6.3.3 Development of polyclonal antibodies

6.3.3.1 Design and production of an anti-peptide antibody for the detection of MrgR

A synthetic peptide consisting of 9 amino acid residues, DTNVDLINC, corresponding to amino acids 206 to 214 located toward the C-terminus of MrgR was designed, synthesised and used as an immunogen for the production of polyclonal antibodies in rabbits. A glycine residue was incorporated into the carboxy terminus of the peptide to facilitate coupling to

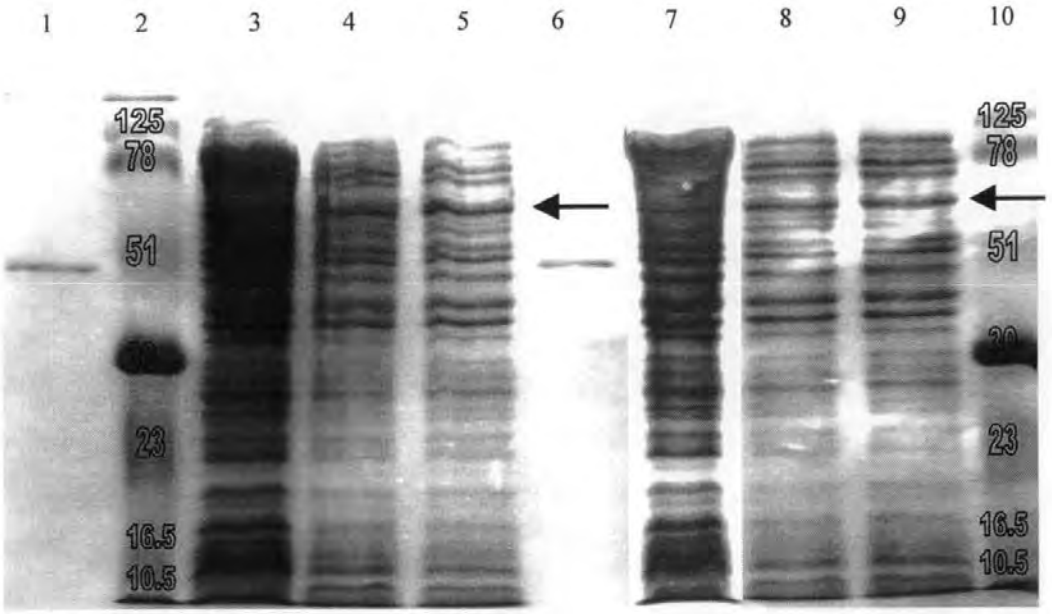


Fig 6.8: Expression of MBP-MrgR fusion protein in whole cell extracts of *E. coli* XL1-Blue shown on a 12% SDS-PAGE gel stained with Coomassie brilliant blue. Lanes 1 & 6: purified MBP; Lanes 2 & 10: protein molecular weight standards (kDa); Lanes 3 & 7: uninduced cells; Lanes 4 & 8: 1 h induction; Lanes 5 & 9: 2 h induction. The arrows indicate the location of the fusion bands.

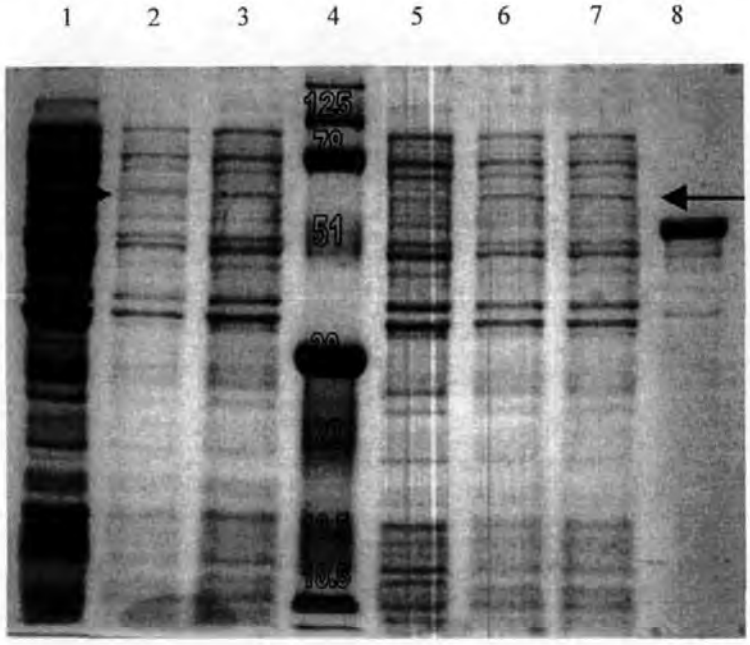


Fig 6.9: Expression of MBP-MrgS fusion protein in whole cell extracts of *E. coli* XL1-Blue shown on a 12% SDS-PAGE gel stained with Coomassie brilliant blue. Lanes 1 & 5: uninduced cells; Lanes 2 & 6: 1 h induction; Lanes 3 & 7: 3 h induction; Lane 4: protein molecular weight standards (kDa), and Lane 8: purified MBP. The arrows indicate the location of the fusion bands.

the KLH carrier protein. The peptide was designed to include a high percentage of hydrophilic residues, over 65%, to improve the chances of success. A predominantly hydrophilic peptide may be more likely to be soluble, easier to conjugate and exposed on the surface of the native protein. Rigorous computer-based sequence analysis and comparison with protein sequence databases demonstrated that the peptide sequence chosen for antibody productions was unique. It shared no exact identity when compared with the protein sequences contained in the GenBank, EMBL and Protein Identification Resources (PIR) databases. However, the peptide does possess up to 60% amino acid homology with a number of sequences in the databases (**Table 1**, Appendix VI). The peptide-KLH conjugate was purified prior to immunisation. Rabbit antisera were tested by enzyme-linked immunosorbent assay (ELISA) for their ability to specifically recognise the peptide and the titre of the antibody preparation was found to be 1/43,970. Approximately 20 mg of antibody was obtained from 20 ml of the antiserum.

6.3.3.2 Design and production of an anti-peptide antibody for the detection of MrgS

A synthetic peptide consisting of 9 amino acid residues, RKFYSLESN, corresponding to amino acids 14 to 22 located in the N-terminus of MrgS was designed and synthesised for immunisation. The peptide shared no exact sequence identity with the proteins contained in the GenBank, EMBL and PIR (**Table** 2, Appendix VI). A cysteine residue was attached to the N-terminus of the peptide to facilitate conjugation to the immunogenic carrier KLH through the sulfhydryl group of the cysteine. The peptide was designed so as to contain more than 75% of hydrophilic amino acids. The purified peptide-KLH conjugate was used to immunise rabbits and the resulting antisera possessed a specific titre of 1/203,240 when tested by ELISA. Approximately 20 mg of the antibody was obtained from 19 ml of the antiserum.

6.3.4 Immunological identification of MBP-MrgR and MBP-MrgS fusion proteins and antibody specificity

The identity of the 60-kDa MBP-MrgR fusion protein was confirmed by immunological assays. Polyclonal antibodies raised against either an MrgR peptide sequence (section 6.3.3.1) or MBP (New England Biolabs) were used separately to probe Western blots of cell lysates of recombinant *E. coli* XL1 Blue expressing MBP-MrgR. Both antibodies recognised bands of ~ 60 kDa, and although some other smaller bands were also evident, it is likely that these represent degradation products of the fusion protein (**Fig 6.10**). It can be seen that the expression of the 60 kDa protein band appears to be stronger in the IPTG-induced lanes, but there is also a 60 kDa band present in the uninduced cells. The latter was consistently observed in Western blots that were probed with the antibodies despite considerable care taken to ensure repression of the *tac* promoter prior to IPTG induction. This experiment was repeated a number of times with the same results. Similarly, and in contrast to the faintly discernible protein band which was present in Coomassie brilliant blue stained SDS-PAGE gels (section 6.3.2.2), both MBP antibody and the anti-peptide antibody produced for the detection of MrgS reacted specifically with MBP-MrgS fusion protein produced in XL1-Blue, and recognised a major immunoreactive band of ~70 kDa (**Fig 6.11** and **Fig 6.12**). However, the reaction with MBP antibody was much stronger than that observed with the MrgS antibody.

The specificity of antibody recognition was determined by blocking experiments in which each of the specific antibodies were preincubated with each of the peptide immunogens. Only antibodies that had been preincubated with the specific peptide that was used for immunisation were blocked from binding to the corresponding fusion protein (**Fig 2**, Appendix VI). Furthermore, no protein bands were detected when the primary antibody was omitted from the experiment.

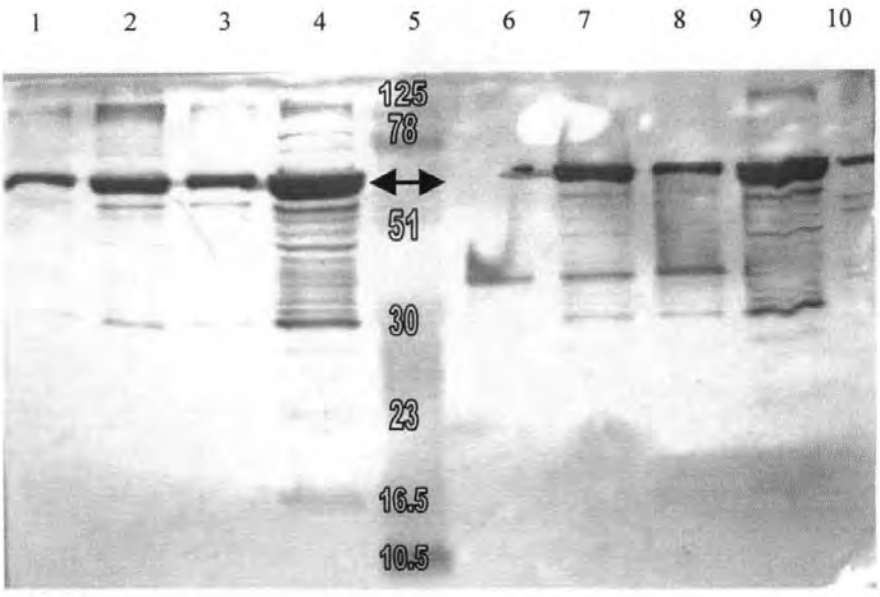


Fig 6.10: Western blot detection of MBP-MrgR fusion protein by anti-MrgR antibody. Lanes 1& 6: insoluble matter; 2&7: crude extract; 3&8 resin purified extract; 4&9: 3 h induced cells; Lane 10: uninduced cells and Lane 5: molecular weight markers (kDa). Arrows indicate the position of the fusion proteins.

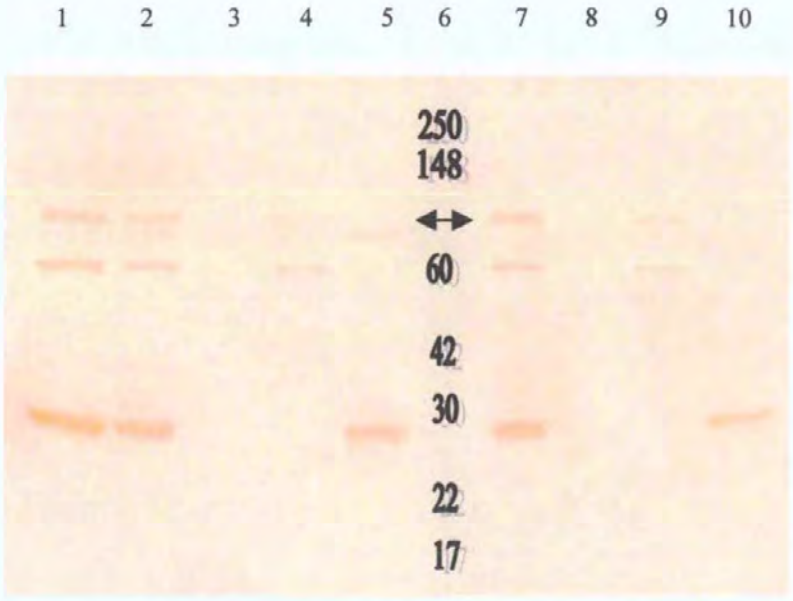


Fig 6.11: Western blot detection of MBP-MrgS fusion protein by anti-MrgS antibody. Lanes 1& 7: insoluble matter; 2 & 5: crude cell extract; 3 & 8: resin purified cell extract; 4 & 9: 3 h induced cells; and Lane 5: molecular weight markers (kDa). The arrow indicates the expected size of fusion protein.

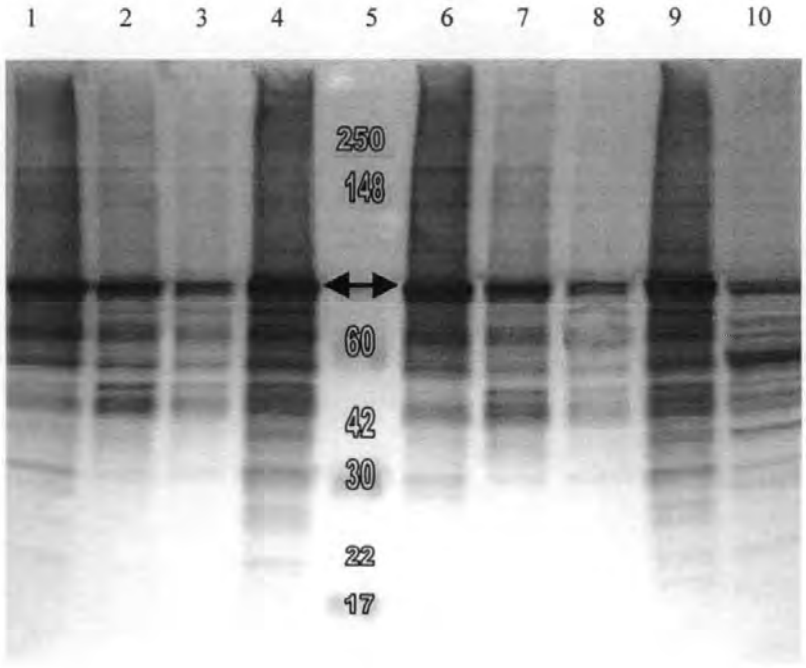


Fig 6.12: Immunoblotting of MBP-MrgS fusion protein with anti-MBP antibody. Lanes 1&6: insoluble matter; Lane 2 & 7: crude extract; Lanes 3 & 8: resin purified extract; Lanes 4 & 9: 3 h induced cells; Lane 10: uninduced cells; Lane 5: molecular weight markers (kDa). Arrows indicate the position of the fusion protein.

6.3.5 Recognition of MBP-MrgR fusion protein by convalescent patient serum

When cell lysates of *E. coli* harbouring the plasmid pMFU1 were electrophoresed (section 3.4.6.1), electroblotted (section 3.14.5) and incubated with serum from a patient with melioidosis it was found that a band of ~60-kDa that corresponds to the MBP-MrgR fusion protein was recognised in soluble, insoluble, and amylose resin purified extracts of bacteria that had been induced with IPTG. As with the blots probed with anti-MrgR antibody, a low level of recognition was evident in the uninduced cell extracts. The serum did not recognise the purified MBP, MBP-MrgS or any other *E. coli* components (**Fig 6.13**). A variety of components that were present on Western blots of *B. pseudomallei* whole cell extracts were very strongly stained when probed with the same patient serum (**Fig 6.14**).

6.3.6 Detection of MrgR and MrgS in *B. pseudomallei* whole cell lysates

The anti-MrgR peptide antibody detected a band of approximately 24 kDa on Western blots of the cell lysates of 6 isolates of *B. pseudomallei* cultured under standard growth conditions (**Fig 6.15**). This corresponds to the calculated molecular weight of MrgR, 23.860 kDa, which is based on the deduced translation of the nucleotide sequence of the *mrgR* gene (section 4.4.2). Unexpectedly, the MrgR antibody labels at least four other bands of varying intensities at ~150, ~90, ~80 and ~30 kDa. The most intensely stained band was located at ~80 kDa.

The anti-MrgS peptide antibody recognised a band of ~115 kDa in each of the isolates of *B. pseudomallei*, which corresponds with the calculated molecular weight, 118.227 kDa, of the deduced translation of the nucleotide sequence of the *mrgS* gene (section 4.4.2). However, in some samples this band was stained with greater intensity than in others. The antibody also recognised other bands that varied in both staining intensity and molecular mass (**Fig 6.16**). One band of ~30 kDa appears to be present in all of the isolates, while a band at ~90 kDa is present in all but isolate 576 and a band at 80 kDa is present only in

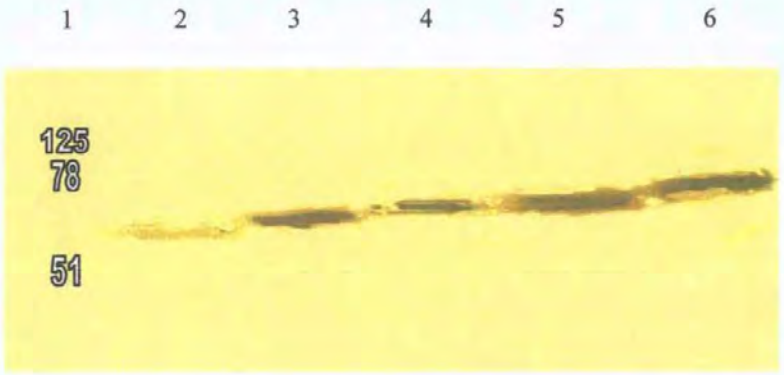


Fig 6.13: Western blot analysis showing the recognition of MBP-MrgR fusion protein by melioidosis patient serum. Lane 1: protein molecular weight markers (kDa); Lane 2: Uninduced cells; Lane 3: 3 h induced cells; Lane 4: resin purified cell extract; Lane 5: crude cell extract; and Lane 6: insoluble matter.

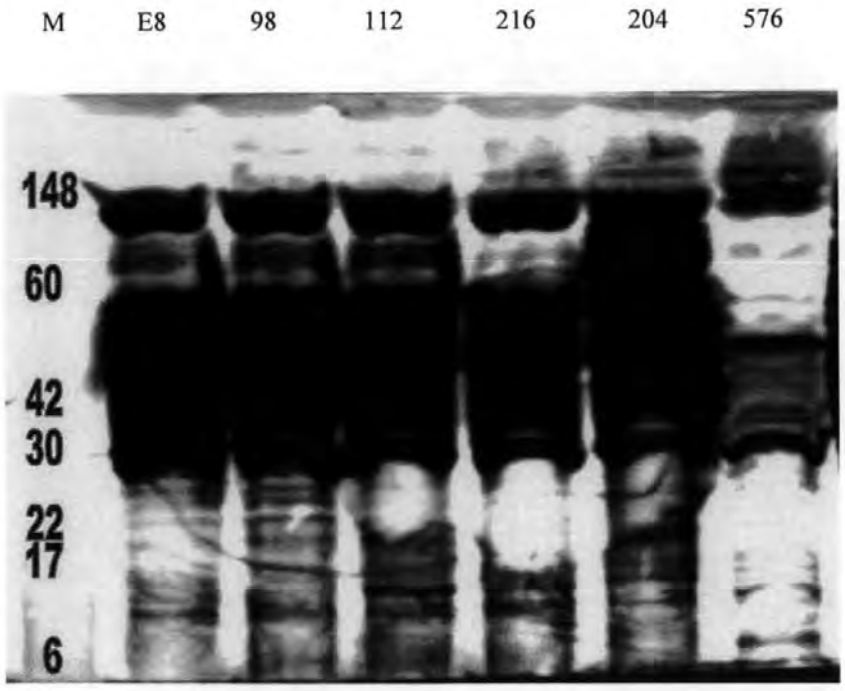


Fig 6.14: Western blot analysis of *B. pseudomallei* whole cell proteins with melioidosis patient serum. Isolate number is indicated above each lane and protein molecular weight markers are on the left (kDa).

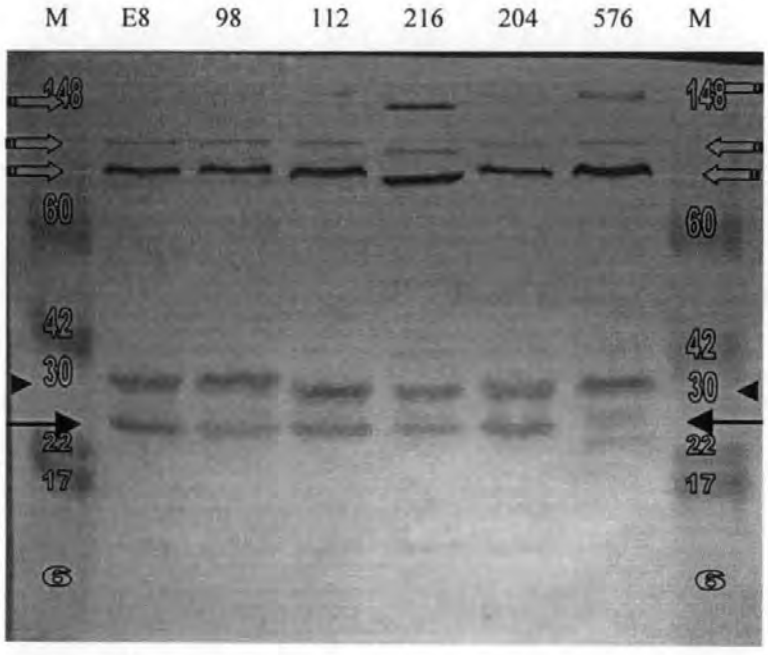


Fig 6.15: Western blot analysis of MrgR from *B. pseudomallei* whole cell lysates using anti-MrgR antibody. Protein molecular masses (M) are given in kilodaltons. The arrows (⟶) indicate the expected location of MrgR bands, ▶ indicates possible phosphorylated form of MrgR, while ⇨ denotes the putative isoforms of the protein, as explained in section 6.4.

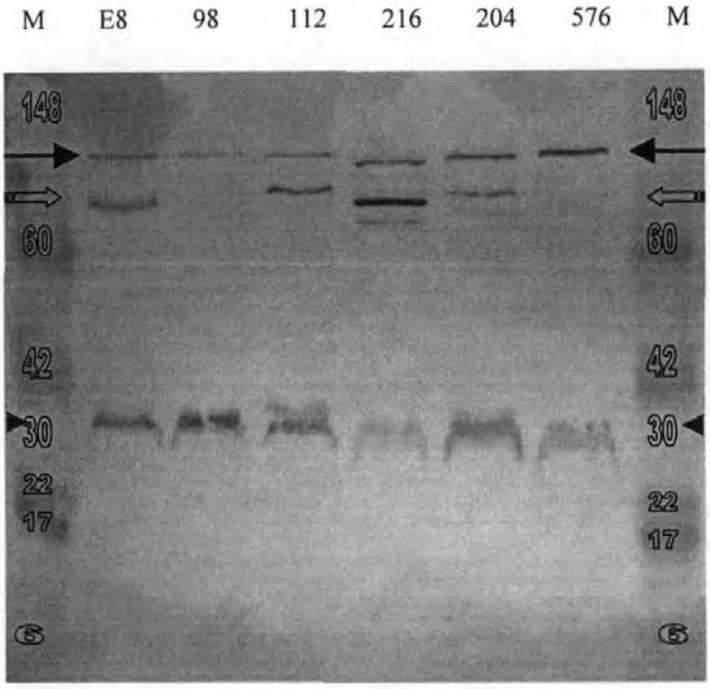


Fig 6.16: Western blot analysis of MrgS from *B. pseudomallei* whole cell lysates using anti-MrgS antibody. Protein molecular masses (M) are given in kilodaltons. The arrows (⟶)indicate the expected full size location of MrgS bands, while ► and ⇨ denote the putative processed forms of the protein.

isolate 216. It was noted that the intensity of the ~115 kDa band was stronger in those samples with weaker recognition of the ~30 kDa band and the inverse was also true.

The specific recognition shown by each of the anti-peptide antibodies was totally abolished by preincubation with specific peptide against which each antibody was raised. In contrast, preincubation with an irrelevant peptide had no effect on the staining intensity. To investigate the basis for this specific recognition of a number of proteins in *B. pseudomallei* whole cell lysates, the sequence of the synthetic peptides were compared with sequence information from the *B. pseudomallei* genome sequencing project database at the Sanger Centre. The search retrieved 6.2 Mb of overlapping contig sequences in the genome. Although no exact match with either peptide sequence was found a number of possible sequence matches were identified that could conceivably account for the recognition of more than a single band on Western blots (**Table 3**, Appendix VI).

In order to investigate the basis for the recognition of MBP-MrgR by serum from a convalescent melioidosis patient, Western blots of the ECPs derived from six isolates of *B. pseudomallei*, including the isolate from which the *mrgR* gene was cloned, were probed with anti-MrgR and anti-MrgS peptide antibodies. In this experiment, none of these ECPs produced an immunoreactive signal.

6.4 DISCUSSION

In order to gain a better understanding of the MrgR-MrgS two-component regulatory proteins and to further assess and visualise their expression in *B. pseudomallei* cells, specific antibodies were developed against synthetic peptide analogues to sequences derived from the C- and N-termini of MrgR and MrgS, respectively. The production of specific antibodies is a desirable prerequisite to the full characterisation of a newly cloned protein. However, this requires the expression and purification of the protein in sufficient quantity for the immunisation of animals. Even with the development of highly efficient protocols based

upon the incorporation of specific peptide tags, this can be a tedious and time-consuming process (Spiller *et al.*, 1999). Simple, efficient and rapid methods for the generation and characterisation of such antibodies would be of broad utility. A common approach is to use peptides derived from the predicted protein sequence (Geysen *et al.*, 1985). Synthetic peptides, when coupled to a carrier protein (e.g. KLH or BSA), normally elicit a strong humoral response. However, anti-peptide antibodies are often poorly reactive with the native protein, limiting their use to the detection of denatured protein by immunoblotting (Spiller et al., 1999). Nevertheless, the use of synthetic peptides as immunogens for the production of polyclonal antibodies has been widely accepted and relies on the careful selection of a peptide sequence and the use of a carrier protein in order to elicit a good immune response. This method has been successfully used to generate antibodies for the detection of many proteins including exotoxin A from *P. aeruginosa* (Elzaim *et al.*, 1998) and protease from *B. pseudomallei* (Lee & Liu, 2000).

To determine whether the anti-peptide antibodies could recognise the MrgR and MrgS proteins, chimaeric fusion proteins were constructed and expressed in *E. coli* XL1-Blue. The stability of the pMFU1 and pMFU2 recombinant plasmids in the *E. coli* host was demonstrated by the culture of a number of different clones and also by nucleotide sequencing. The MBP-MrgR fusion appeared to be expressed at high levels as a soluble molecule reflecting the predicted cytoplasmic nature of the protein. However, attempts to overexpress the MrgS fragment in pMAL-C2 demonstrated that MBP-MrgS was mainly in an insoluble form, an observation that was previously reported for other overexpressed two-component regulatory and sensory proteins (Schroder *et al.*, 1994). Furthermore, this provides some support for the computer-generated structural predictions for this region of MrgS, which suggest that it is a predominantly hydrophobic region, and may represent a transmembrane domain within the full length of the MrgS protein. The MrgR and MrgS

polyclonal antibodies bound to their cognate fusion proteins and the specificity of the antibodies for the peptide used for immunisation was demonstrated by the lack of recognition of MBP-MrgR and MBP-MrgS fusion proteins when the antibodies were blocked by preincubation with the immunising peptide. Furthermore, this specific recognition was not blocked by preincubation with an irrelevant peptide. The same result was recorded when *B. pseudomallei* cell lysates were probed with the antibodies.

An important consideration in the choice of the peptides that were used to raise antibodies to MrgR and MrgS was that exhaustive searches of protein databases revealed no exact match with either of the peptide sequences. Nevertheless, a number of bands were recognised by each of the polyclonal antibodies on Western blots of *B. pseudomallei* cellular extracts. The apparently unique nature of the primary amino acid sequences of the peptide immunogens together with the observation that immunoreactivity can be blocked by preincubation of the serum with the immunising peptide antigen further suggests that this recognition is specific. The lack of detectable quantities of either MrgR or MrgS in the ECPs of *B. pseudomallei* strongly suggests that under the conditions employed in this study neither of these proteins is secreted. This supports the structural predictions for these proteins that were made in Chapter 4 (section 4.4.2).

There are a number of possible explanations for the occurrence of multiple bands on Western blots of *B. pseudomallei* cells that have been probed with specific antibodies against MrgR and MrgS. These include the presence of multiple isoforms of proteins with identical epitopes, proteolytic degradation, dimerisation or complexing, and some level of antibody cross-reactivity with proteins with very similar epitopes. Some of these possibilities provide a more likely explanation for the observed banding patterns. It is important to appreciate that the inability to detect similar sequences in protein databases is no guarantee that antibodies raised against a peptide will not cross-react with other cellular

proteins (Reubi *et al.*, 1999). The lack of sequence identity between the peptides and *B. pseudomallei* sequence should not be taken as definitive evidence that the observed reactivity is not directed toward the product of another gene in the *B. pseudomallei* genome. The observed reactivity is presumably directed against a primary sequence feature, as it was detected using Western blotting, but the scoring matrixes used in the computer analyses have been optimised for detecting evolutionary relationships rather than features important for antibody recognition and thus could well have missed the relevant sequence. On the other hand, the nature and relationship between an antibody and its antigen is never exclusive; therefore, in addition to recognising the epitope against which it was elicited, an antibody has the potential to bind to a variety of related moieties, or antigenic sites, which share some structural features with the immunogen (Van Regenmortel, 1998). Antigenic sites of proteins were thought to consist of short stretches of about six amino acids that may be arranged in the primary structure either continuously or discontinuously (Atassi & Young, 1985). However, recent structural analyses suggest that antigenic epitopes in native proteins are mostly of the discontinuous type, that is they are composed of several short stretches, mainly loops exposed on the surface (Laver *et al.*, 1990). In accordance with that, the recognition of multiple bands by MrgR and MrgS antibodies may be due to this occurrence.

An alternative and perhaps stronger possibility is that inducing a specific signal(s) may promote autophosphorylation of MrgS. This would result in the phosphorylation of MrgR, which could then function in concert or sequentially with other components to activate or repress the transcription of other genes. Under some conditions, MrgS may not be activated, or phosphorylated MrgS may be unstable, and in these circumstances MrgR would not be phosphorylated. This proposal resembles what is known about the standard two-component paradigm in which phosphorylation-dephosphorylation activities follow stimulation (section

2.5). Each of the phosphorylated and non-phosphorylated proteins would migrate differentially resulting in two closely located bands, such as those present at 24 and 30 kDa on blots probed with anti-MrgR antibody. It has been established that although phosphorylation does not alter the overall folding of a protein or result in substantial changes in protein secondary structure, it can dramatically affect the molecular surface, hence altering both topological and electrostatic features of the protein (Stock *et al.*, 2000).

There is also a strong possibility that MrgR may dimerise or form a complex with another protein. Phosphorylation can promote dimerisation (McCleary, 1996), higher-order oligomerisation (Wyman *et al.*, 1997), or interactions with other proteins (Blat & Eisenbach, 1994) or DNA (Aiba *et al.*, 1989). Some proteins use a combination of these mechanisms (Anand *et al.*, 1998). The process of dimerisation is required for the control of transcription by all regulators whose three-dimensional structure is known, e.g. lambda repressor (Pabo & Lewis, 1982) and BvgA (Scarlato *et al.*, 1990). BvgA, a 23-kDa response regulatory protein that regulates virulence genes in *B. pertussis*, has been shown to dimerise and a cross-reacting band of 46 kDa was stained on Western blots probed with antibodies specific for BvgA (Scarlato et al., 1993). Furthermore, the formation of heterodimers of MrgR and another unidentified response regulatory protein is also possible. For example, in *E. coli,* under appropriate conditions RcsF and RcsC were hypothesised to facilitate the phosphorylation of the response regulator RcsB. Phosphorylated RcsB then interacts with RcsA to form a more stable complex or heterodimer in order to stimulate the expression of the capsule genes (Arricau *et al.*, 1998). In *P. syringae,* heterodimer formation has been suggested to occur between HrpR and HrpS, two response regulators that are required for transcriptional activation of *hrpL* (Xiao *et al.*, 1994) and also between CorP and CorR, two response regulators that are involved in the temperature-dependent biosynthesis of the *P. syringae* phytotoxin coronatine (Ullrich *et al.*, 1995). In addition, transcriptional regulators

that function as homodimers include NtrC (nitrogen regulatory protein) and DtxR (diphtheria toxin repressor) in *Corynebacterium diphtheriae* (Schmitt & Holmes, 1994) and LsR-like proteins CarR and AmpR in *Citrobacter freundii* (Bishop & Weiner, 1993). This may offer some explanation for the presence of high molecular weight bands such as those at ~80, ~90 and ~150 kDa on blots probed with anti-MrgR antibody. On the other hand, any of these higher molecular weight bands may represent either the failure to solubilize MrgR fully or aggregation of MrgR with other cellular proteins during electrophoresis at room temperature.

DNA sequencing of the *P. aeruginosa* genome has led to the identification of 64 and 63 genes encoding response regulators and histidine kinases, respectively. Given the highly conserved nature of two-component regulatory proteins and the adaptable nature of *B. pseudomallei* it may be that it is similar to *P. aeruginosa* in this respect. Hence it seems possible that *B. pseudomallei* possesses other regulatory proteins that may share substantial amino acid sequence identity with MrgR or MrgS and could cross-react with the peptide-specific antibodies.

The presence of bands of lower molecular weight than the predicted full size MrgS protein on Western blots of *B. pseudomallei* cell lysates, could arise as a consequence of proteolytic processing or degradation of the immunoreactive protein before or during sample preparations or during the bacterial growth cycle. This proposal is supported by the observation that the staining intensity of the ~30 kDa band seems to vary inversely with the staining intensity of the ~115 kDa band i.e. when more of one is present, less of the other is evident **(Fig 6.16)**.

Probing Western blots of bacterial preparations with convalescent patient serum is a widely used technique for identifying components that may be expressed by pathogens during the course of an infection (Watson *et al.*, 1998). Serum from a melioidosis patient

strongly stained Western blots of MBP-MrgR and *B. pseudomallei* whole cell lysates, but not purified MBP, MBP-MrgS or other *E. coli* components. Although this is not definitive proof that infected hosts specifically recognise the MrgR protein and that MrgR may be expressed by *B. pseudomallei* during melioidosis infection, it does provide some evidence for this possibility. Why MrgR should be recognised by patient antibodies remains unclear, particularly when an investigation of the ECPs produced by six isolates of *B. pseudomallei* revealed that MrgR is not present in detectable quantities in *B. pseudomallei* culture supernatant. This supports the suggestion that MrgR, as a member of the response regulator family, is probably not secreted and may well act as a cytoplasmically located protein. However, because two-component regulatory proteins are involved in mediating cellular responses to environmental changes, the possibility that MrgR might be more directly involved in the pathogenesis of *B. pseudomallei* in humans cannot be excluded.

In conclusion, the anti-peptide antibodies that have been generated in this project allowed the detection of the MrgR and MrgS proteins in *B. pseudomallei*. These antibodies will permit further studies of these proteins both *in vitro* and perhaps *in vivo*. Such work may help to resolve the reasons for the immune recognition of MrgR by an infected host. The antibodies may be useful in understanding how the interaction of the bacterium with the host modifies the expression of MrgR and MrgS. Understanding where and when during the course of infection these proteins are active may provide valuable clues about their function and how it may be disrupted.

CHAPTER 7

GROWTH OF *B. PSEUDOMALLEI* AND THE EXPRESSION OF MrgR AND MrgS PROTEINS UNDER DIFFERENT CONDITIONS OF TEMPERATURE, pH AND SALINITY

7.1 INTRODUCTION

The development of anti-peptide antibodies for detecting MrgR and MrgS proteins in *B. pseudomallei* permits further studies of the possible role of these proteins in the adaptive responses of the pathogen to environmental stimuli. In this chapter, the effects of variation in three environmental parameters on the expression of MrgR and MrgS in a number of different isolates of *B. pseudomallei* were investigated.

Most, if not all, bacteria are exposed to alterations in environmental conditions during their life cycle (Slauch *et al.*, 1996). Accordingly, bacteria regulate gene expression in order to ensure that only those genes whose protein products are required at a particular time are switched on, while those that are not needed are turned off. Understanding how environmental conditions regulate the expression of specific genes is one step towards understanding gene function. This is especially important for opportunist pathogens, such as *Burkholderia pseudomallei*, that are capable of survival in a wide range of environmental conditions. Environmental conditions can be established in a laboratory setting in order to mimic the microenvironment encountered outside or within a host organism (Heithoff *et al.*, 2000). Although mimicking these environments in the laboratory is problematic since the microenvironments within the host tissues are both complex and dynamic, any environment encountered by a bacterium during its' life cycle may be defined in terms of parameters such as temperature, osmotic pressure, oxygen status, pH and nutrient availability. Many *in vitro* experiments have been conducted for the purposes of detecting bacterial genes whose expression is modulated by changes in these parameters, and this work has yielded much information on environmentally controlled genes, some of which contribute to virulence (Dorman & Bhriain, 1992).

Temperature is one of a few environmental parameters that have been demonstrated, both *in vivo* and *in vitro*, as being an important regulator of bacterial genes (Smith, 1990). It

is one of the factors that are known to change as a consequence of an induced inflammatory response (Wang *et al.*, 1998). Moreover, the role of temperature in the expression of virulence genes in human and animal pathogens is well established (Mekalanos, 1992). Transcription of these genes is activated when microbes enter their warm-blooded hosts and elevated temperature might be an early signal heralding the entry of a pathogen into the host from an environmental source.

Regulation of gene expression by pH has been described in several pathogenic bacteria. If a bacterium follows the ingestion route of infection, it must survive the acidic conditions of the stomach and the alkaline environment of the small intestine. Likewise, when a microbe is engulfed by a macrophage, the phagocytic vesicle reaches pH 6 within minutes; after fusion to lysosomes, it reaches a pH of 4.5 (Alpuche-Aranda *et al.*, 1992). Hence, low pH has been shown to provide a signal for regulating the expression of bacterial genes that are presumed to function within the macrophage.

Osmolarity has been shown to provide an environmental signal controlling the expression of virulence genes in several bacterial species (Mekalanos, 1992). For example, in *V. cholerae* the expression of cholera toxin and other virulence determinants is optimal at osmolarities that fall within the physiological range of host tissues (Miller & Mekalanos, 1988). Furthermore, the expression of the virulence-associated alginate capsule of *P. aeruginosa* is affected by osmolarities that reflect the altered environment inside the 'salty' cystic fibrosis lung. Usually the influence of osmolarity on gene expression is studied *in vitro* by varying the concentration of either sucrose or NaCl in culture media.

The aim of this work was to examine the effects of variation in three environmental parameters, temperature, pH and salinity, on the expression of MrgR and MrgS proteins in different isolates of *B. pseudomallei* cultured *in vitro*, by using polyclonal antibodies directed against MrgR and MrgS that were reported in Chapter 6. In addition, the effects on

bacterial growth were monitored by measuring the total cellular protein concentration in the cultures. Temperatures of 25, 37 and 42°C were chosen to reflect the survival and growth of *B. pseudomallei* at ambient tropical temperatures as well as within mammalian and avian hosts. Similarly, a range of pH 5, 6.8 and 8, and salinities, 0.15, 0.5 and 2.2% NaCl, were chosen to reflect the possible range of conditions that exist within marine and freshwater environments and also within the cells, tissues, and organs of an infected mammalian host.

7.2 MATERIALS AND METHODS

7.2.1 Culture of *B. pseudomallei* under different conditions of temperature, pH and NaCl

Three clinical (112, 204, 576) and three environmental (E8, 98, 216) isolates of *B. pseudomallei* (**Table 5.5**) were inoculated into 10 ml volumes of LB broth (Appendix 1) and incubated for 24 h at 37°C within a Containment Level 3 Laboratory (section 3.2.2). Sets of 10 ml volumes of LB broth with different NaCl contents (0.15, 0.5, 2.2%) were adjusted to pH 5, 6.8 and 8 with HCl and then autoclaved. The broth was inoculated with 100 µl of the overnight culture then grown statically for 48 h at 25, 37 or 42°C, as detailed in **Table 7.1**.

Table 7.1: Different parameters that were used for the culture of *B. pseudomallei* growth

Temperature	25 or 37°C			42°C		
pH	5	6.8	8	5	6.8	8
Salinity %	0.15	0.15	0.15	0.15	0.15	0.15
pH	5	6.8	8	6.8	6.8	
Salinity %	0.5	0.5	0.5	0.5	2.2	

7.2.2 Protein estimation, SDS-PAGE and Western blotting analysis

A 1 ml volume of bacterial culture was removed, centrifuged (4,000 x g, 10 min, 4°C), and washed once in PBS. The pellet was resuspended and lysed in 50 µl of modified SDS-PAGE sample buffer (Appendix 1). Determination of total protein concentration in bacterial cell lysates was performed by a microassay procedure based on the method of Bradford using the Bio-Rad Protein Assay Kit (Bio-Rad, UK) (section 3.14.2). Approximately 5 µg of whole cell lysate was boiled for 5 min and the protein samples were separated on SDS-PAGE gels with protein molecular weight standards (section 3.14.3). Following electrophoresis, the proteins were electroblotted onto nitrocellulose and probed with either the MrgR or MrgS anti-peptide antiserum (sections 3.14.5 & 6.2.2.3).

7.3 RESULTS

7.3.1 Growth of *B. pseudomallei* under different conditions of temperature, pH and NaCl

Although all of the isolates of *B. pseudomallei* grew under each of the conditions of temperature, pH and salinity that were tested, some conditions were more favourable for growth (**Table 7.2 & Fig 7.1**). Optimal growth was assessed by the amount of protein produced and this was highest at 42°C, pH 6.8 and 0.5% NaCl. Interestingly, increasing the NaCl content to 2.2% under these conditions drastically reduced the amount of protein produced by 43-71% depending on the isolate. Generally, growth was better at pH 5 and lower at pH 8 and increasing the NaCl content to 0.5% led to reduced growth, although this is clearer at 37°C than at 25°C. Isolates 204 and 216, both from NE Thailand, show very similar patterns of protein production in contrast to the variation that is present in the other isolates (**Fig 7.1**).

Generally, isolates grew better at each pH under low NaCl conditions (0.15%, 20 mM).

Table 7.2: Growth of *B. pseudomallei* isolates under different conditions of temperature, pH and NaCl and the amount of total cellular protein produced (mg/ml)

Strain	Growth conditions																
	37°C & 0.15% NaCl			37°C & 0.5% NaCl			25°C & 0.15% NaCl			25°C & 0.5% NaCl			42°C & 0.15% NaCl			42°C & 0.5%	42°C & 2.2%
	5	6.8	8	5	6.8	8	5	6.8	8	5	6.8	8	5	6.8	8	6.8	6.8
E8	3.55	3.05	1.50	2.35	2.20	2.10	2.15	0.65	1.25	2.00	2.00	0.75	3.60	3.90	2.80	5.10	2.90
98	5.20	4.10	2.35	3.75	3.65	2.65	1.90	0.75	1.20	4.10	1.60	1.05	4.20	6.10	1.90	6.60	3.00
112	3.50	2.65	1.30	3.15	2.00	1.30	1.05	2.05	1.05	2.20	2.05	1.05	3.70	2.10	3.00	3.90	1.80
216	2.85	2.40	0.85	2.50	1.20	0.55	1.00	2.10	1.05	1.05	1.10	0.50	2.80	3.00	1.60	5.60	1.90
204	7.65	6.15	5.65	5.20	3.30	2.60	1.45	1.30	0.85	1.90	1.70	1.00	8.80	5.00	5.70	6.90	2.90
576	4.15	3.60	3.45	3.05	2.95	2.50	3.60	2.00	1.90	3.70	1.65	0.85	6.10	4.40	6.00	8.90	2.60

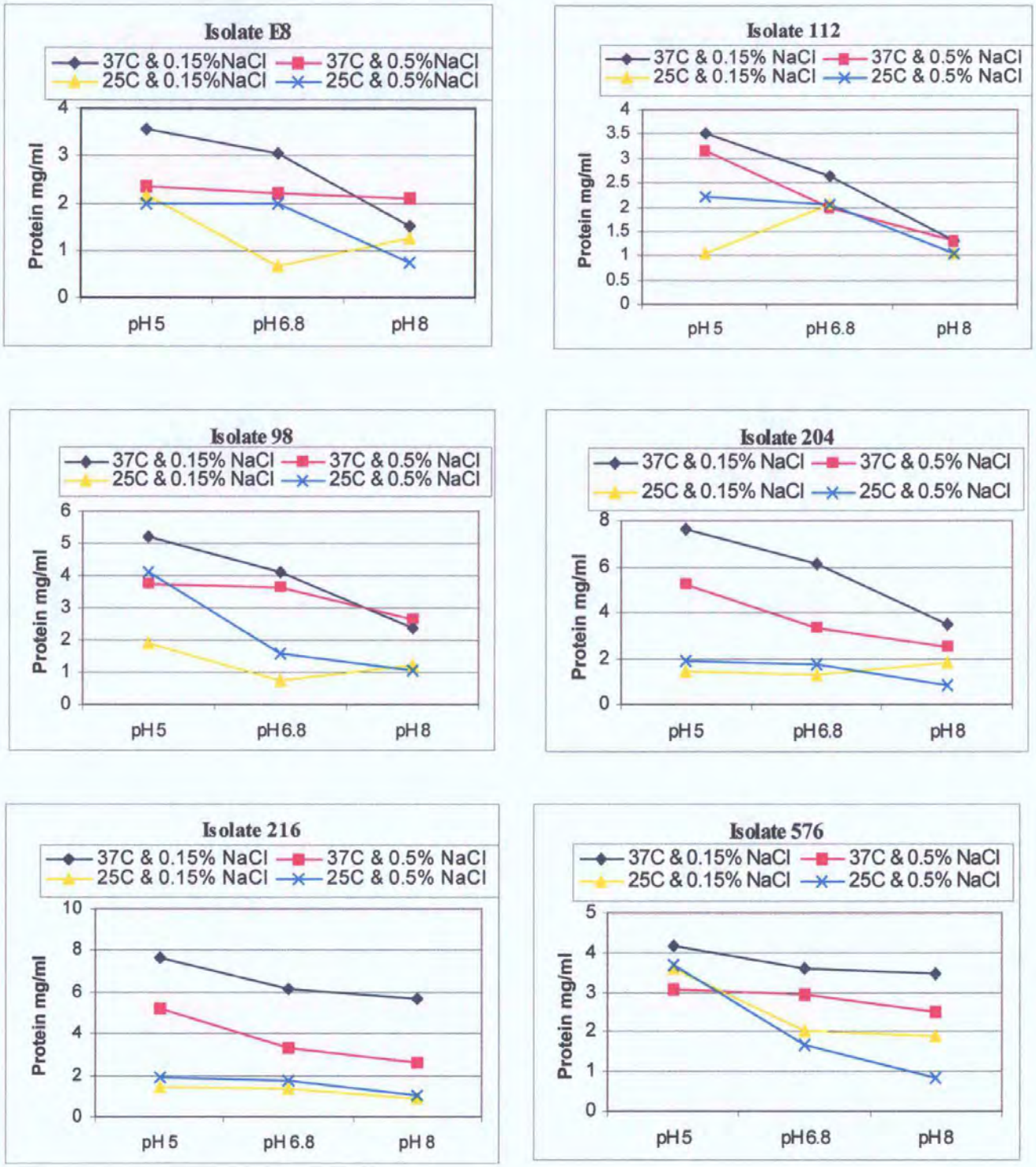


Fig 7.1: Growth, as indicated by the production of total cellular proteins (mg/ml), for 6 isolates of *B. pseudomallei* grown under different combinations of temperature and NaCl, including 37°C & 0.15% NaCl, 37°C & 0.5% NaCl, 25°C & 0.15% NaCl, and 25°C & 0.5% NaCl. All were grown under three different pH values 5, 6.8 and 8.

When the isolates were cultured in the presence of 0.5% NaCl (68 mM) the maximum protein production occurred at pH 5 and the lowest at pH 8, regardless of whether cultures were grown at 25 or 37°C. When the bacteria were cultured using 0.15% NaCl a similar pattern was observed but only at 37°C, while growth at 42°C or 25°C resulted in different patterns. At 25°C, maximum protein production occurred at pH 5 for four strains; E8, 98, 204 and 576 and at pH 6.8 for strains 112 and 216. On the other hand, at 42°C, the 3 clinical isolates (112, 204, 576) all showed a similar pattern of protein production with a peak at pH 5 and a trough at pH 6.8. The 3 environmental isolates (E8, 98, 216) were typified by another pattern of protein production that peaked at pH 6.8 (**Fig 7.2**).

7.3.2 Western blot analysis and the effect of different culture conditions on the expression of MrgR and MrgS proteins

MrgR and MrgS antibodies each recognised more than one band on Western blots of whole cell lysates of *B. pseudomallei* grown under each of the culture conditions, which is consistent with previous observations (section 6.3.6). In addition to the 24 kDa band, the MrgR antibody highlighted 4 other components, two major bands at 80 and 30 kDa that were present in all samples, and two faint bands in some samples at 150 and 90 kDa. The MrgS antibody mainly recognised the ~115 kDa band in addition to two other major bands, one at 30 kDa and the other at 90 kDa. The latter band is present in all but isolate 576.

Figures 7.3-7.8 and **7.9-7.14** show the expression of MrgR and MrgS, respectively, under each condition of pH, salt and temperature. Generally, the growth of *B. pseudomallei* at higher temperature (42°C) resulted in a progressive increase in immunoreactive components, suggesting that increasing temperature may coordinately up-regulate these antigenically-related molecules. This was strikingly obvious on blots that were probed with MrgS antibodies where the 90 kDa band is intensely stained in isolates 98, 112, 204 and 216 grown at 42°C in the presence of 2.2% NaCl. This band is also very obvious in isolate 216

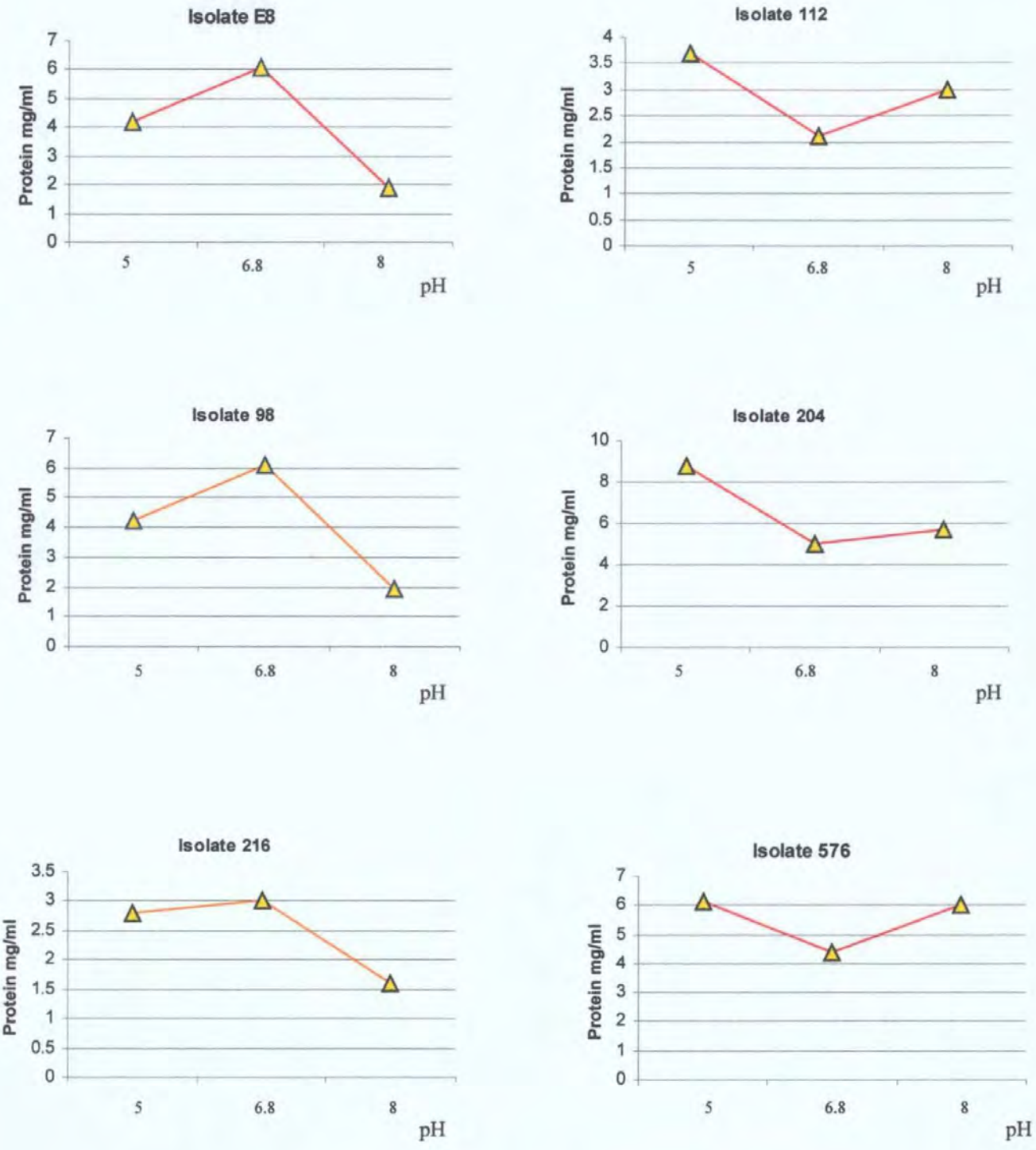


Fig 7.2: Growth, as indicated by the production of total cellular proteins (mg/ml), for six *B. pseudomallei* isolates grown at 42°C and under different pH and NaCl contents reveals two distinct patterns, 'V' for clinical isolates and 'Λ' for environmental isolates.

Fig 7.3: Expression of MrgR for 6 isolates of *B. pseudomallei* grown under different combinations of temperature, pH and NaCl. Western blots of whole cell lysates were probed with anti-MrgR antibody. The identities of the symbols (⟶, ▶, ⇨) are as described in Fig 6.15 and as discussed in section 6.4, while the identity of each lane is indicated above each figure. M: indicates protein molecular weight markers (kDa).

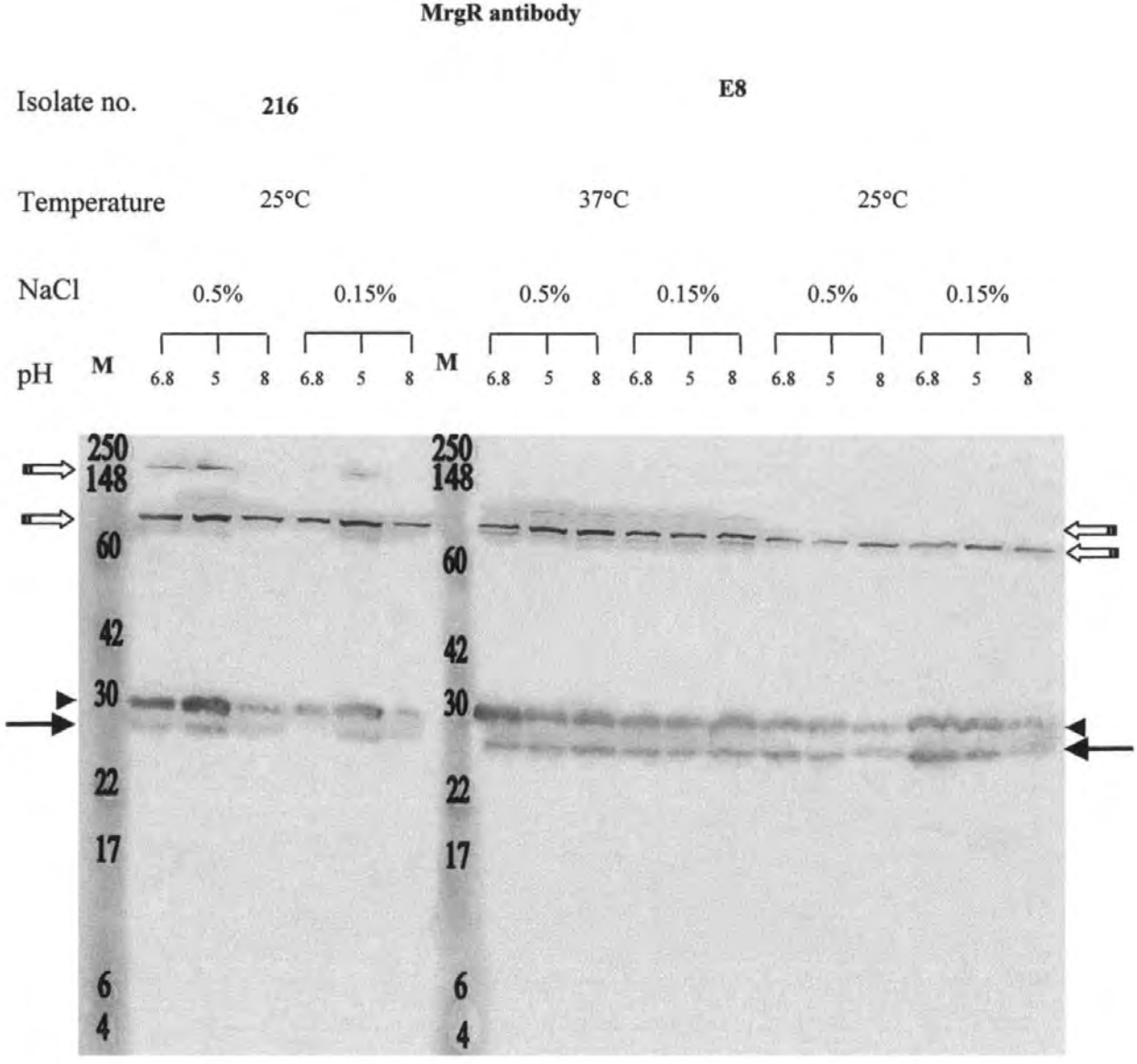


Fig 7.3 continued

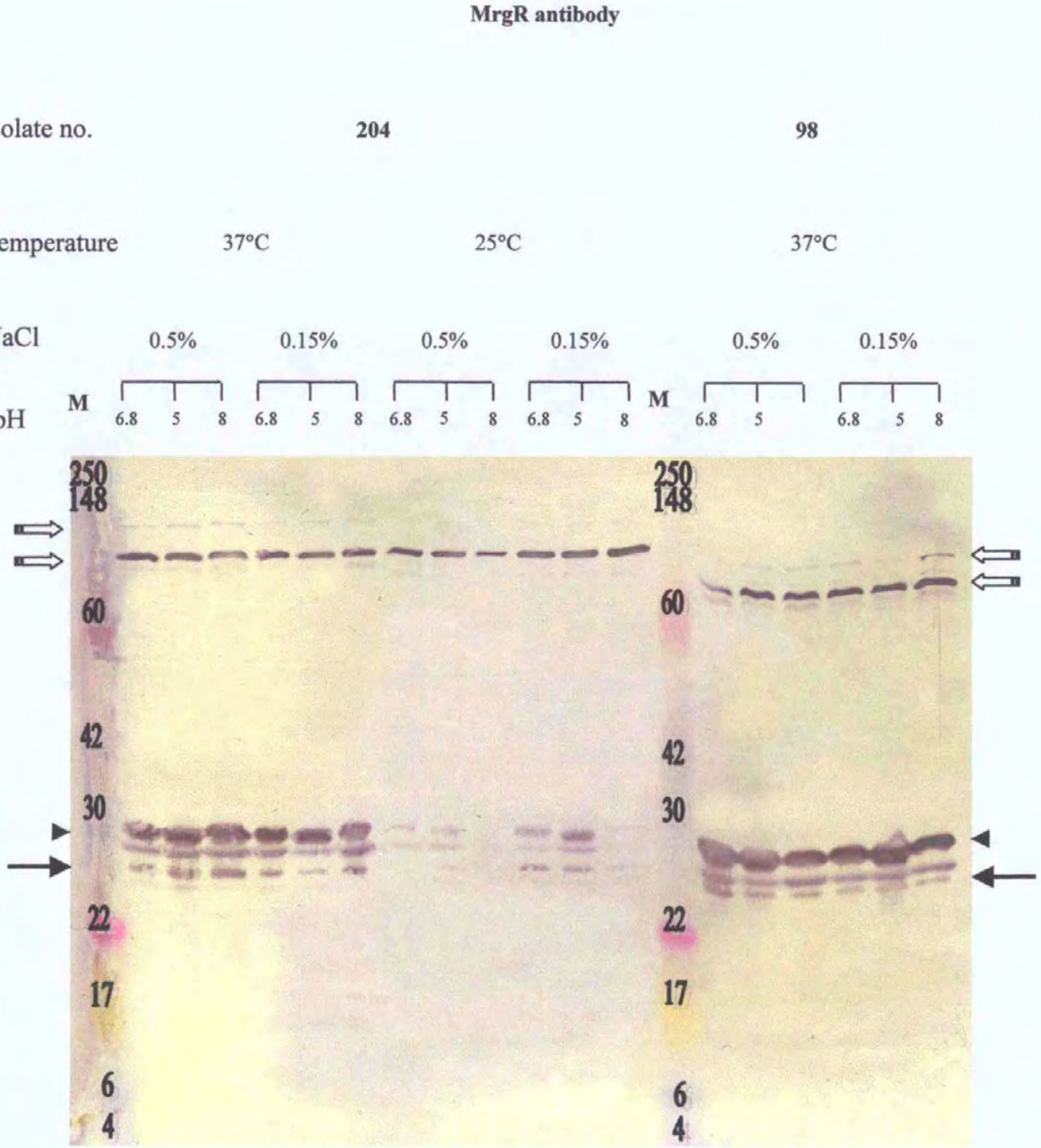


Fig 7.3 continued

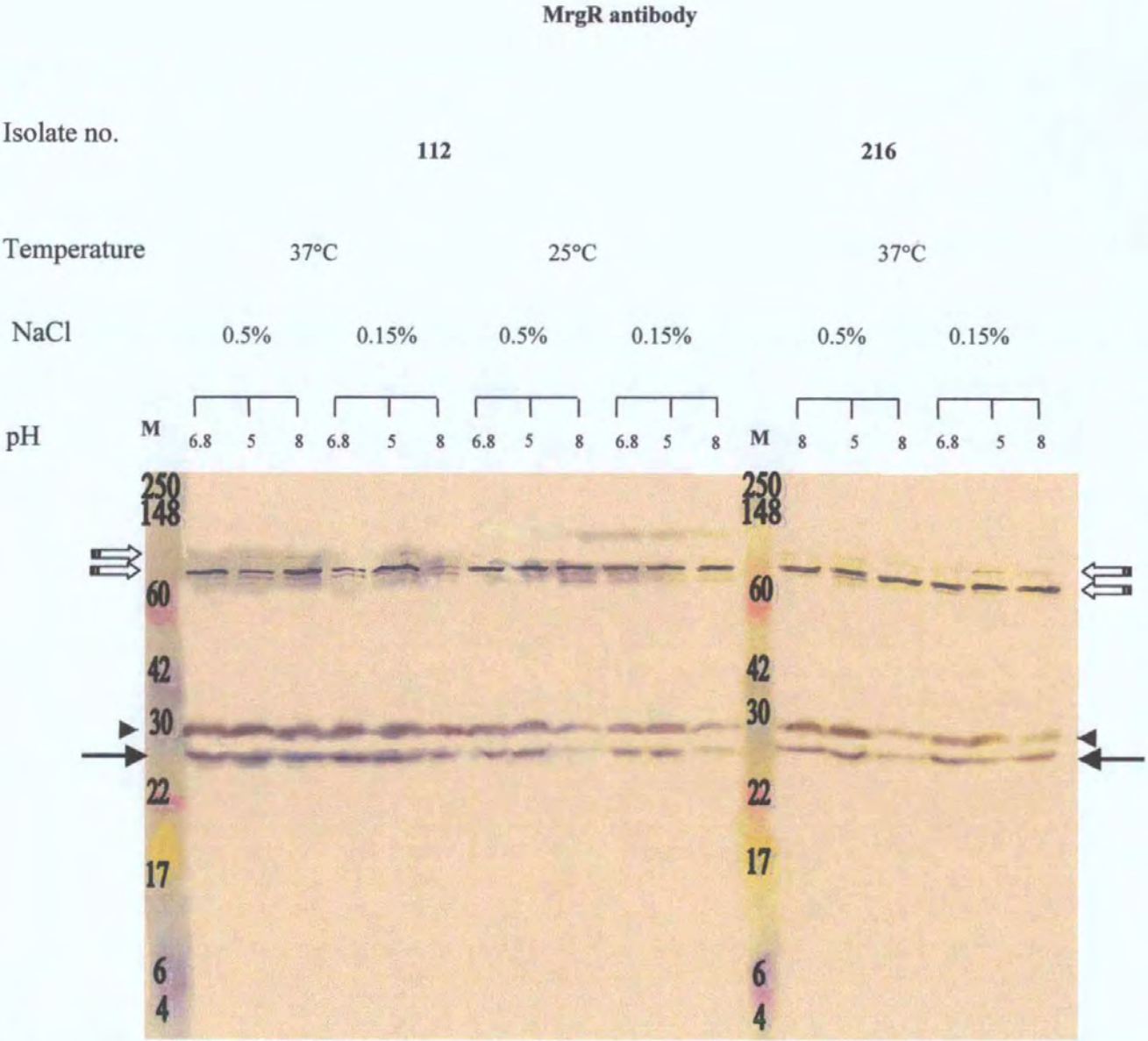


Fig 7.3 continued

MrgR antibody

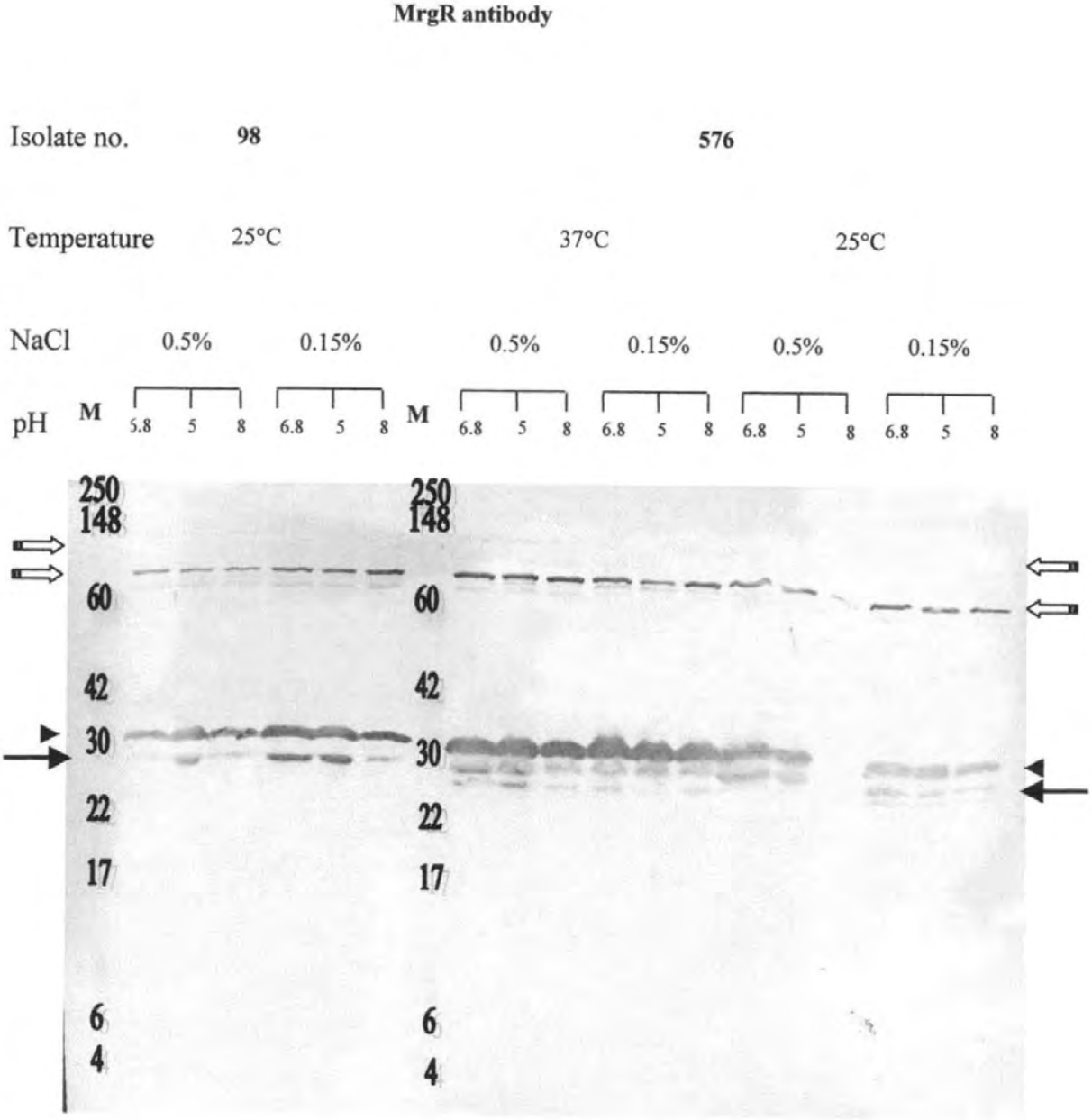


Fig 7.3 continued

MrgR antibody

0.15% NaCl & 42°C

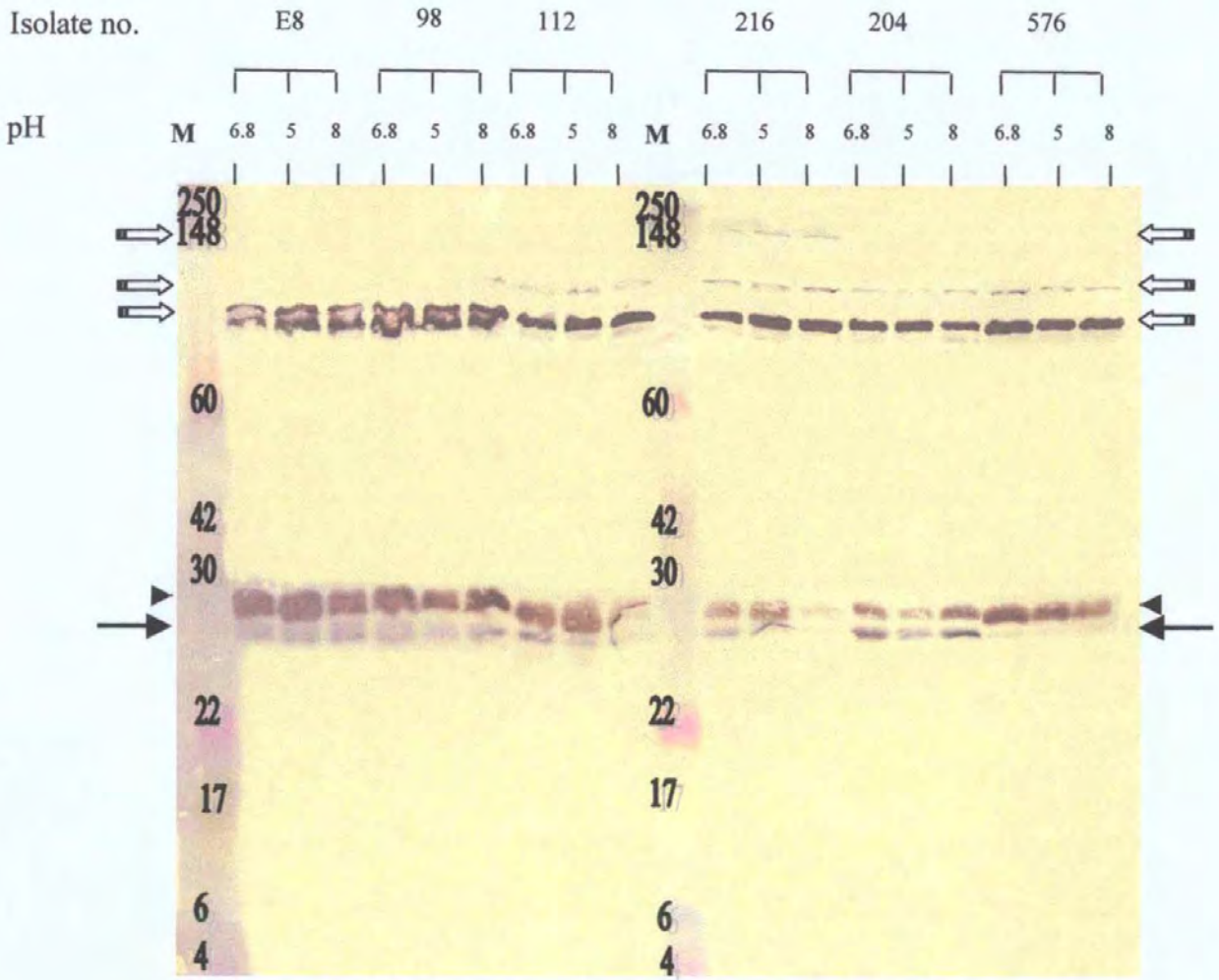


Fig 7.3 continued

MrgR antibody

Temperature 42°C

pH 6.8

NaCl 2.2% 0.5%

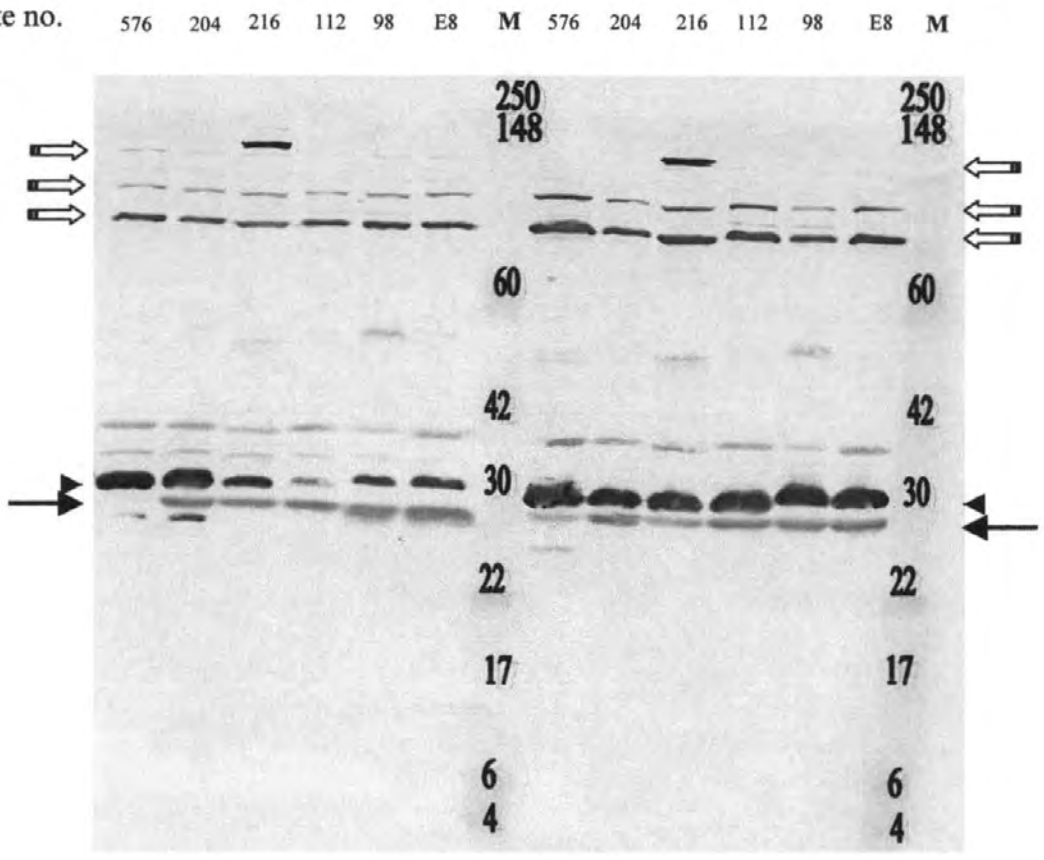


Fig 7.4: Expression of MrgS for 6 isolates of *B. pseudomallei* grown under different combinations of temperature, pH and NaCl. Western blots of whole cell lysates were probed with anti-MrgS antibody. The identities of the symbols (→, ⇨, ▶) are as described in Fig 6.16 and as discussed in section 6.4, while the identity of each lane is indicated above each figure. M indicates protein molecular weight markers (kDa)

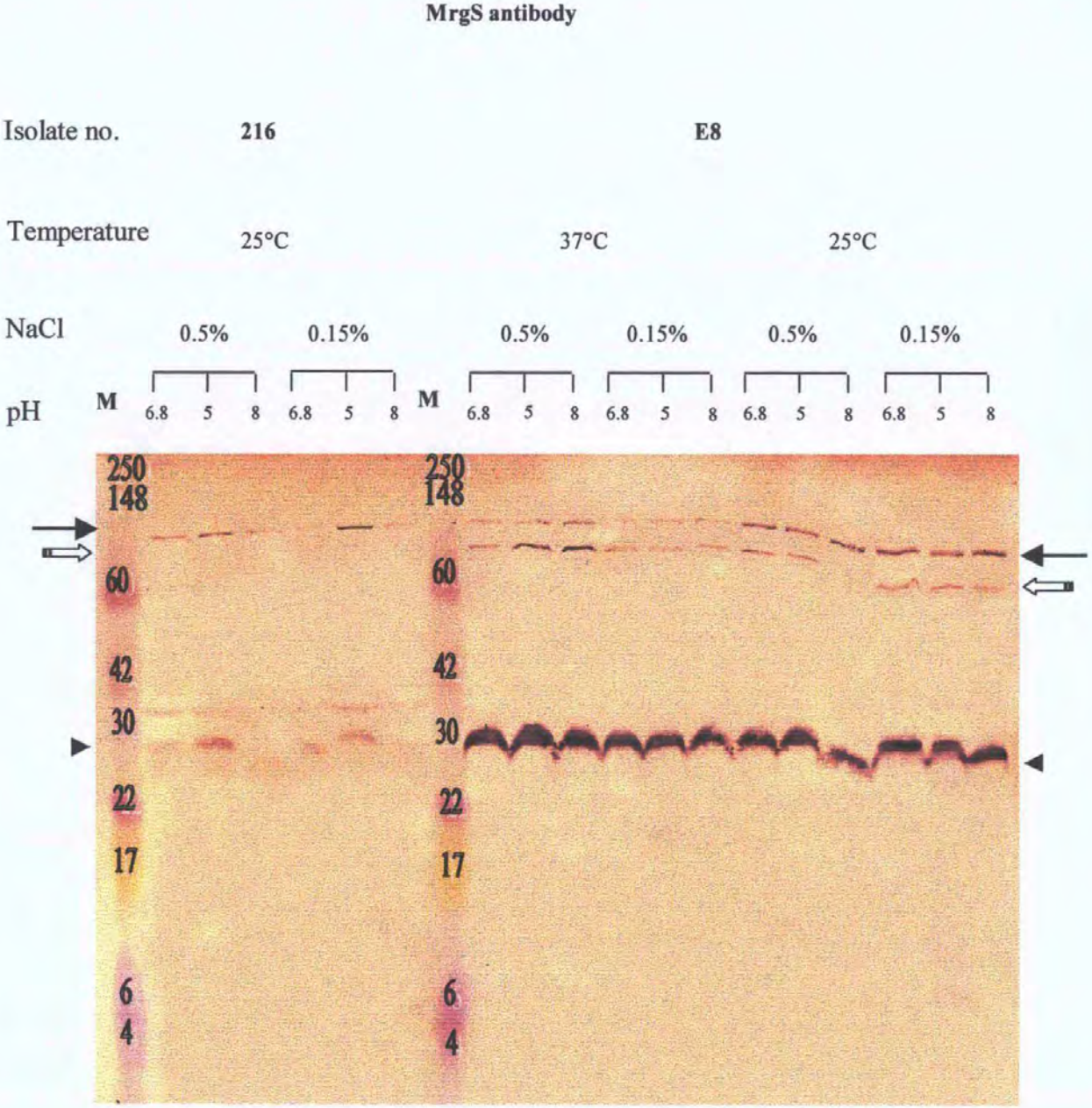


Fig 7.4 continued

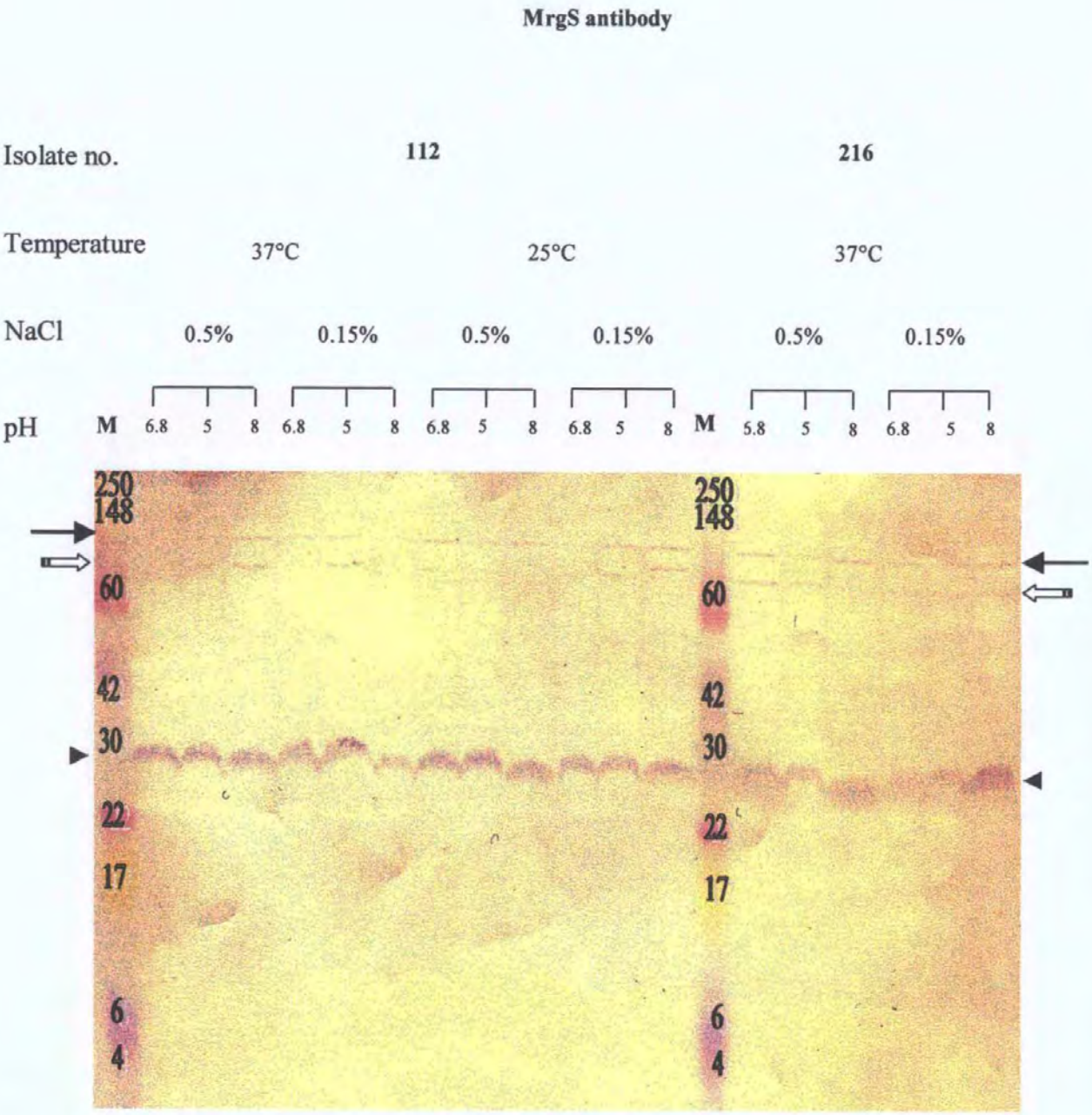


Fig 7.4 continued

MrgS antibody

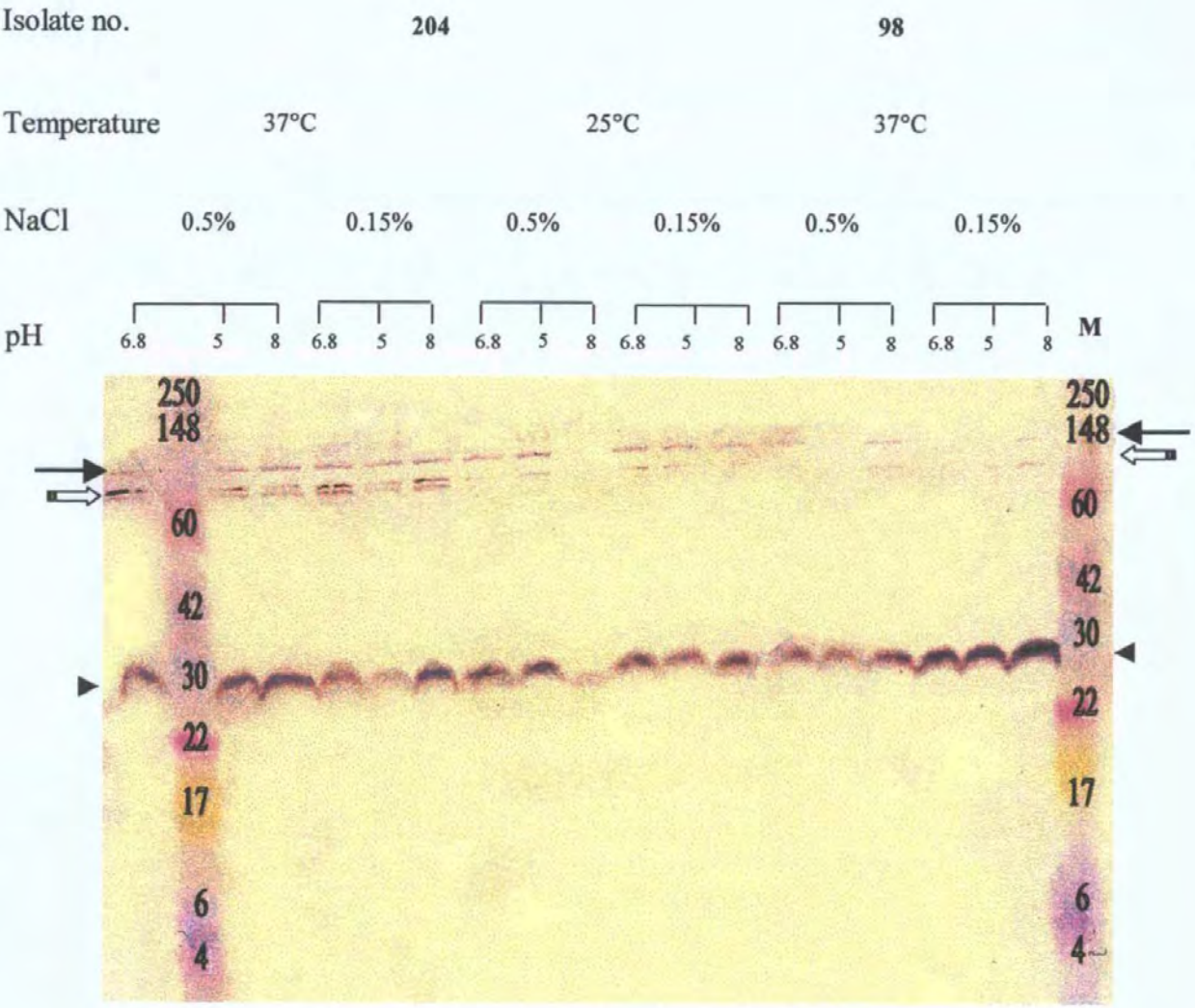


Fig 7.4 continued

MrgS antibody

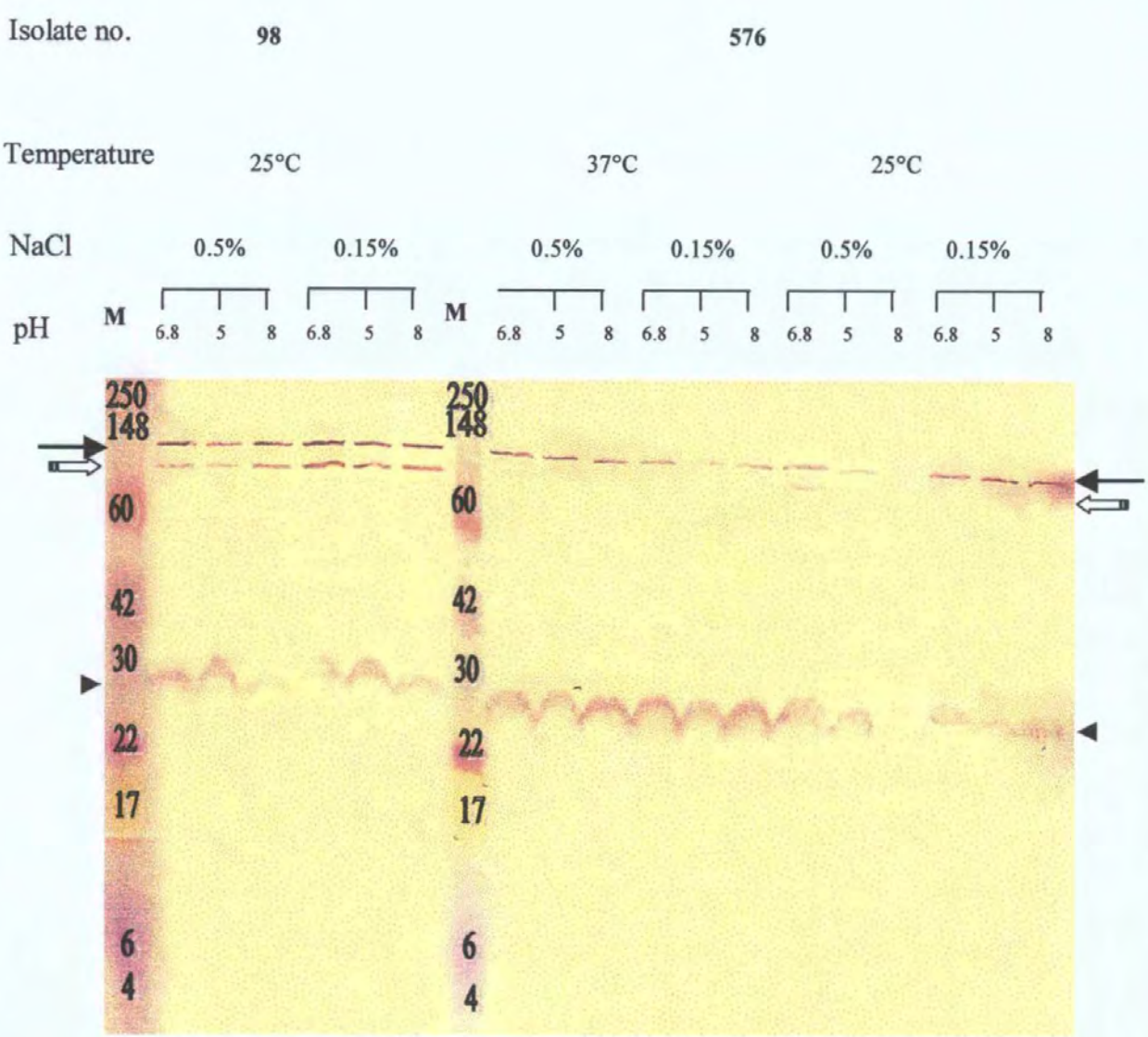


Fig 7.4 continued

MrgS antibody

0.15% NaCl, 42°C

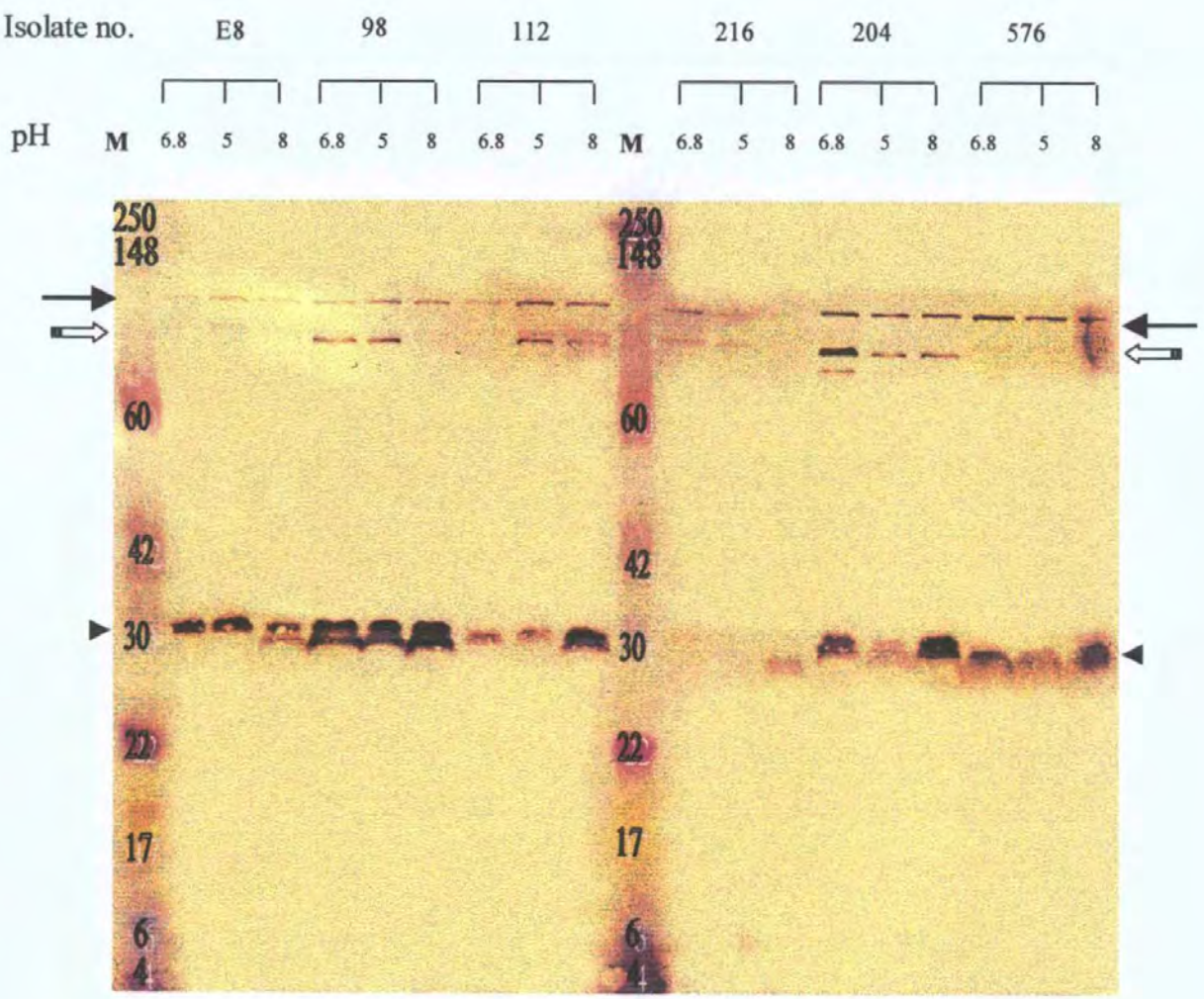


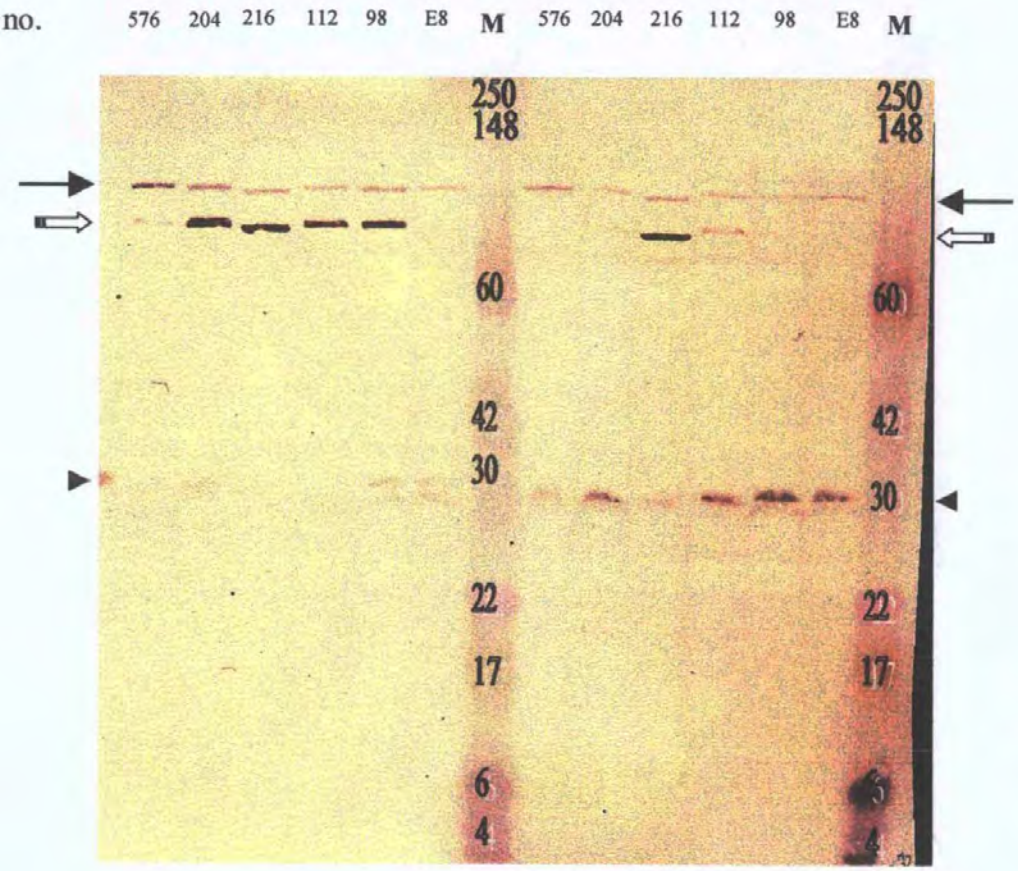
Fig 7.4 continued

MrgS antibody

Temperature 42°C

pH 6.8

NaCl 2.2% 0.5%



grown in medium containing 0.5% NaCl and in 204 at pH 6.8 and 0.15% NaCl. In contrast, at the lower growth temperatures the 90 kDa band is either absent or extremely weak. In addition, the 90 kDa band is absent in isolates E8 and 576 under any of the growth conditions. On the other hand, some bands that were recognised by anti-MrgS were less intensely stained following growth at 42°C. Compared with growth at 25 or 37°C, expression of the 30 kDa band was markedly reduced in cultures grown at 42°C, and absent when using 2.2% NaCl. Although the 115 kDa band stained with varying intensity in all isolates under different growth conditions, it was slightly stronger when the environmental isolates (E8, 98, 216) were cultured at 25°C in comparison with culture at 37°C, regardless of pH or salt content.

For blots probed with anti-MrgR antibodies the 150 kDa band was very distinct in isolate 216 grown at 42°C but was weak or absent at 25 and 37°C. This band was not obviously present in any of the other isolates. In all of the isolates, the anti-MrgR antibody showed a decrease in the staining intensity of the 24 kDa band and also the 30 kDa band when cultures were grown at 25°C and pH 8. While in comparison, the quantity of the 80 kDa band appeared to be unaffected by these conditions. However, at 42°C, the 80 kDa band is much more abundant, particularly at 0.15% NaCl, while the 30 kDa band is also more abundant but this is more noticeable at 0.5% NaCl. Similarly, the 90 kDa band was somewhat more distinct at 42°C, especially at 0.5% NaCl and to lesser extent at 2.2% NaCl while the same band was either absent or extremely weak in each of the other conditions that were tested.

7.4 DISCUSSION

Many bacteria are commonly exposed to dramatic changes in their environment and must respond appropriately in order to survive. This is also true for *B. pseudomallei*, an

opportunist that can survive intracellularly and adapts well to a variety of growth conditions (Dejsirilert *et al.*, 1991). In an attempt to improve our understanding of this adaptation, *B. pseudomallei* was grown under varying conditions of pH, salinity and temperature that reflected the different environments to which the bacterium may be exposed. Both the growth of the bacteria and the expression of *mrgR* and *mrgS* were studied using *B. pseudomallei* that had been isolated from either clinical or environmental samples as a means of determining whether isolates from these sources differ in their response to the variables studied. This study has confirmed earlier work showing that *B. pseudomallei* can grow over a wide range of temperatures (Yabuuchi *et al.*, 1993), from 25-42°C, and will tolerate extremes of pH and wide variation in NaCl content.

The interaction between *B. pseudomallei* and its' environment have been studied by a number of authors. In the work presented here, it has been demonstrated that *B. pseudomallei* grows well at pH 5 and under some conditions this growth is at an optimum e.g. 37°C and 0.15% NaCl, conditions that reflect those encountered within the mammalian host. It is known that *B. pseudomallei* can survive and grow under low pH conditions both *in vitro* and *in vivo* (Mitsuchi *et al.*, 1986; Kanai & Kondo, 1994). To some extent this reflects the acid soil conditions in endemic areas such as north-east Thailand ranging from pH 4.38-7.7. However, acid conditions (pH 4.7-5.5; Sprich, 1956) also exist inside the phagolysosomes of mononuclear cells, the effects of which *B. pseudomallei* can withstand (Wongwanich *et al.*, 1996).

According to the results of the protein determination assay, *B. pseudomallei* can grow under conditions covering a wide range of pH, NaCl concentration and temperature. However, maximal protein synthesis occurred at 42°C, pH 6.8 and 0.5% NaCl. A dramatic decrease in cellular protein under these conditions, up to 71%, occurred in response to an elevation in NaCl content to 2.2%. The increased synthesis of protein at 42°C may reflect a

survival response by the bacterium to heat-shock or stressful conditions, and a consequent reduction or loss of control of its' ability to osmoregulate. On the other hand, lowering the temperature to 25°C was most effective in restricting the growth of *B. pseudomallei*, particularly at pH 8. There is some evidence that alkaline pH may present substantial difficulties for some isolates of *B. pseudomallei* and that this may also be related to the control of osmoregulation. The combination of growth at 25°C, pH 8 and 0.5% NaCl resulted in the lowest levels of protein synthesis in all of the isolates of *B. pseudomallei* that were examined. Once again, this is consistent with the known distribution of the organism and its' persistence in the environment and may represent an intolerance to less optimal environmental conditions at low temperatures.

A striking observation was that at 42°C, the 6 isolates of *B. pseudomallei* that were examined showed two consistent patterns of response that apparently reflected the source of isolation. The first pattern, which was displayed by the 3 clinical isolates, is 'V' shaped with peaks of protein production at pH 5 and pH 8 and a trough at pH 6.8. The other pattern, which was shared by the 3 environmental isolates, is 'Λ' shaped with a single peak of protein produced at pH 6.8. It is tempting to speculate that this finding may reflect the response of clinical isolates to stressful conditions similar to those that might occur during inflammation and the acute phases of infection i.e. higher than normal mammalian body temperatures and active phagocytic mechanisms. It may be that the sequential passage of clinical isolates through the host has acted to positively select for an increase capacity to grow under these conditions.

This study has demonstrated that MrgR and MrgS are expressed by different isolates of *B. pseudomallei* under different growth conditions. The proteins are therefore required to meet the needs of cellular metabolism under most circumstances. Environmental conditions are known to modulate the expression of two-component regulatory proteins, although the

response may differ from one species to another. For example, the expression of OmpR, a response regulatory protein in *E. coli*, is induced in media of high osmolarity and repressed under low osmolarity, as part of the adaptive response to conditions within the mammalian gut. However, osmolarity changes do not affect the expression of OmpR in *S. typhi* (Puente *et al.*, 1991). On the other hand, osmotic upshift is one of the environmental signals shown to induce the transcription of capsule genes in *E. coli* through the RcsB-RcsC pathway (Sledjeski & Gottesman, 1996).

Temperature controls virulence gene expression in *Bordetella* species via the BvgA/BvgS two-component signal transduction system and in *V. cholerae* by another two-component system, the ToxR system (Miller *et al.*, 1989). In *Listeria monocytogenes*, the expression of many of the genes that are crucial to the infection process, including the genes encoding invasin and listeriolysin as well as many other heat-shock or stress proteins, are controlled by temperature via the transcriptional regulator *prfA* (Leimeisterwachter *et al.*, 1992; Lingnau *et al.*, 1995).

Many bacteria are known to respond to pH changes. Most investigators have focused on adaptive responses to a decrease in pH. Acidification induces expression of specific genes in several bacterial pathogens, such as *S. typhimurium* (Miller *et al.*, 1991) and *V. cholera* (DiRita & Mekalanos, 1989). At least 18 proteins are altered in expression as a result of acid shift in *S. typhimurium*. The bacteria possess an acid tolerance response, which is triggered at pH 5.5-6, which protects the cell against even lower pH values. The two-component regulatory system PhoP-PhoQ negatively regulates the expression of *S. typhimurium* invasion genes and positively regulates the expression of genes required for the intracellular survival of the bacteria within acidified macrophages (Miller *et al.*, 1989; Pegues *et al.*, 1995). In *E. coli*, CadC, the membrane-attached DNA-binding protein, undergoes a conformational change in response to external pH changes that results in

modulation of its transcriptional activation capacity (Watson *et al.*, 1992).

In many cases, the combined effects of more than one environmental signal may activate two-component regulatory systems. For example, BvgAS, which controls the expression of virulence factors in *Bordetella* species, is modulated by the presence of $MgSO_4$ or nicotinic acid and growth at temperatures below 26°C (Coote, 1991). When *Bordetella* is grown at 37°C in the relative absence of $MgSO_4$ and nicotinic acid, BvgAS is active (Bvg^+ phase), hence a large set of genes encoding protein virulence factors are expressed. The Bvg^- phase is expressed when *Bordetella* is grown at room temperature or in the presence of >10 mM nicotinic acid or >40 $MgSO_4$. Similarly, *Yersinia* has a collection of genes that are temperature regulated but require another signal to initiate their full range of regulatory activities (Straley & Perry, 1995). Many of the genes that are induced by high temperature also require the presence of low calcium levels for high level induction in culture medium, with cell contact probably providing the equivalent cue *in vivo*.

The differences seen in the expression of MrgR and MrgS might reflect differences in the growth phase of the *B. pseudomallei* cultures, since the bacteria were grown statically under various combinations of temperature, pH and NaCl concentration. However, in the discussion to Chapter 6 it was proposed that the observed pattern of bands on Western blots of *B. pseudomallei* that had been probed with anti-MrgR antibody, may represent the formation of complexes between the 24 kDa MrgR and other components. This complex pattern of expression of the protein under the variable conditions of pH, salt and temperature necessitates the prediction that MrgR may be integrated into more than one regulatory signal transduction pathway and a simple two-state model does not adequately explain the observed patterns of expression. Bacterial signalling systems are rarely being as simple as a system composed of two proteins (Corbell & Loper, 1995). Many regulators are classed as 'global' because they have the potential to affect the expression of large subsets of genes in

the cell (Dorman, 1994). The RcsB response regulator of *E. coli* and other species, which shares substantial sequence homology with MrgR, is reported to be a global regulon involved in a variety of functions including extracellular polysaccharide synthesis and the growth of the bacterial cell under certain conditions (Gervais *et al.*, 1992). It is likely that the activation of MrgR, by phosphorylation, and the interactions with other cellular components, would enhance its binding to other proteins with varying degrees according to the stimulus or combined stimuli encountered. Such a process could provide a means to optimise the activity of MrgR at critical points of the life cycle. When *B. pseudomallei* are grown under various conditions of osmolarity, pH and temperature, a basal level of expression of the 24 kDa MrgR band, and its assumed phosphorylated form the 30 kDa band, was observed, although there is evidence for a lower level of expression at pH 8. Similarly, the 80 kDa band that was observed to be expressed under different conditions was highly expressed at 42°C. In contrast, the 90 kDa band seems to be weakly expressed under all growth conditions, apart from growth at 42°C, particularly at higher NaCl content (0.5 and 2.2%). These results suggest that *B. pseudomallei* has at least two kinds of temperature-related mechanisms involved in MrgR expression, one working at 25 or 37°C and the other at 42°C, which may represent a stress response. On the other hand, the 150 kDa band was only expressed in isolate 216 grown at 42°C, especially at 0.5 and 2.2% NaCl. In *E. coli*, although NaCl was shown to enhance the ability of RcsB to activate its own synthesis, high NaCl concentration was found to affect the expression of *rcsB*. Very little or no activation of *rcsB* expression was observed when NaCl concentration increased to 0.3 M (Gervais *et al.*, 1992).

The expression of the 115 kDa band, which closely approximates the deduced size of the full-length of MrgS, was variable between isolates and growth conditions. Nevertheless, the expression of this molecule was more apparent at 25°C in the environmental isolates of

B. pseudomallei regardless of pH and salt conditions and this may represent a lower level of proteolytic processing or turnover of cell wall components. Similarly, the apparent lack of the 90 kDa band in isolates E8 and 576 may arise as a consequence of the proteolysis of the full length 115 kDa band. The possibility exists that recycling or degradation of the MrgS molecule occurs by a specific proteolytic mechanism and in these isolates that mechanism may be absent or the cleavage site(s) within MrgS may be altered. A more striking change was observed with the increased expression of the 90 kDa band and the reduction in expression of the 30 kDa band at 42°C and its' absence in cultures grown with 2.2% NaCl. This change accompanies the increased level of cellular protein synthesis at 42°C and also the dramatic decrease in cellular protein production that occurs at this temperature when NaCl content is increased to 2.2%. This reduced expression is not apparent at other growth temperatures. The conclusion is, therefore, that expression of the 30 kDa band is sensitive to higher temperatures but is largely independent of NaCl content. It is possible to speculate that the 30 kDa band may arise as a specifically cleaved fragment of the 90 kDa band and that this process may be much less efficient at 42°C.

CHAPTER 8

GENERAL DISCUSSION

The subject of the research presented in this thesis, the saprophyte *B. pseudomallei*, is nutritionally diverse and is capable of resisting a variety of environmental extremes. In addition, the bacterium causes melioidosis, an emerging infection that affects the lives of many millions of people, particularly in the tropical regions of south-east Asia. This pathogen has the potential to adapt to and colonise other areas of the world. A thorough review of the scientific literature (Chapter 2) indicates that although there have been extensive studies of the biology of *B. pseudomallei*, very few of these have investigated the influence of environmental factors on gene expression and the adaptive responses of the bacterium. Recent studies have emphasised the importance of two-component regulatory systems in controlling the adaptive responses of many other bacterial species to environmental stimuli. This study set out to identify and characterise the genes encoding a two-component regulatory system in *B. pseudomallei* and to examine some of the possible environmental conditions that may modify the expression of these genes and control the adaptive responses of this pathogen.

The first step in fulfilling these objectives was the development of an appropriate cloning strategy. Genomic libraries were constructed in λGEM-11 and λZAP phage vectors. Both libraries, as well as a previously constructed λGT-11 library, were screened with separate DNA probes that were complementary to known two-component gene sequences. A number of recombinants were identified, and the fragments of cloned *B. pseudomallei* genomic DNA were sequenced. The initial analysis of these clones was described in Chapter 4. Nucleotide sequencing of about 4 kb of cloned genomic DNA revealed two open reading frames that were designated *mrgR* and *mrgS*. Because the deduced amino acid sequences of *mrgR* and *mrgS* are most similar to the sequences of genes encoding known response regulators and sensors, it has been inferred that these two loci constitute a two-component signal transduction system. All of the invariant amino acids that are found in RR and HPK

proteins are also found in the predicted protein sequences of MrgR and MrgS. The molecules are most similar to the RcsB and RcsC proteins that are involved in the regulation of capsular polysaccharide synthesis in *E. coli* and other bacterial species. In accordance with the gene organisation of the majority of the known bacterial two-component systems, *mrgR* and *mrgS* are physically linked on the chromosome. Furthermore, they are likely to be transcribed from the same promoter. MrgS, like BvgS of *Bordetella pertussis*, is a member of a family of complex sensor proteins, hybrid kinases, which contain multiple cytoplasmic domains that are believed to participate in a phosphorylation cascade in signal transduction. Sequence analysis, hydrophilicity profiles and protein expression studies highlighted the transmembrane association of MrgS, with a 232 amino acid domain protruding into the periplasm, possibly as a means of monitoring environmental parameter(s). The probable membrane location of MrgS is expected for a sensor regulatory protein that lends itself to signal transduction. On the other hand, the location of MrgR was predicted to be cytoplasmic which is consistent with the nature of a DNA-binding protein.

The copy number for *mrgRS* was determined by Southern blot analysis. The hybridisation pattern indicated that *mrgRS* is a single copy locus in *B. pseudomallei*, and it is widely distributed among different isolates that have a broad geographical derivation. The 4 kb *mrgRS* gene locus has the potential for usage as a *B. pseudomallei*-specific DNA probe, and hence as a tool for diagnosis. The probe was highly specific for *B. pseudomallei* DNA although the sensitivity of the probe has yet to be determined, particularly when extracting DNA from a wide variety of clinical and environmental samples.

The regions of the genome flanking the *mrgRS* locus of *B. pseudomallei* were characterised by DNA sequencing and Southern hybridisation analysis. The results of DNA sequencing revealed the presence of a single ORF upstream and six ORFs downstream of *mrgRS*. The first gene downstream of the 3' end of *mrgS*, ORFd, was of special interest

because of its' predicted regulatory function and similarity to BvgR which is similarly located downstream of the regulatory *bvgAS* locus of *B. pertussis*. The deduced amino acid sequences of MrgR and MrgS place them in the families of FixJ response regulators and hybrid kinases, respectively, and this is also true of BvgA and BvgS of *B. pertussis*. In this way, the *mrgRS* locus seems to mirror the *bvgAS* locus. It seems possible that a strategy similar to that used by the *bvg* signal transduction system of *B. pertussis* could be used by the *mrgRS* locus of *B. pseudomallei*. If this were the case, then *mrgRS* and ORFd would perform antagonistic regulatory functions.

Also downstream of *mrgRS* but reading in the opposite direction lies a sequence containing three potential genes, two of which are possibly transcribed from a single promoter (Chapter 5). One of these, ORFg, may encode a member of the lipase family, possibly phospholipase A1. Although the deduced protein apparently lacks a signal peptide for secretion, the downstream ORF, ORFh, is similar to a type III export protein and this may provide a possible strategy for the secretion of the putative phospholipase. Intriguingly, the third ORF in the series, ORFi, codes for a novel protein that possesses all of the general features of a signal peptide but lacks any significant homology to protein sequences in any of the public databases. Southern blot analysis revealed restriction fragment polymorphisms in the *mrgRS* downstream flanking region, one of these was located within ORFg and separated isolates of *B. pseudomallei* into two almost equally sized groups. The biological significance of this genetic variation remains to be established experimentally.

In an attempt to elucidate the effects of environmental variation on the expression of *mrgR* and *mrgS*, a study was made of the growth of *B. pseudomallei* under differing conditions of temperature, initial pH and NaCl concentration. Although the bacterium grows over a wide range of variation in these parameters, the highest quantities of protein were produced, as an indication of optimal growth, at 42°C, pH 6.8 and 0.5% NaCl. At either 37

or 42°C, and pH 5 and 0.15% NaCl, substantial growth was recorded, whilst at pH 8 all the cultures grew poorly. The present finding agrees with earlier studies showing that *B. pseudomallei* is highly adaptable to acidic environments (Yabuuchi *et al.*, 1970; Dejsirilert *et al.*, 1991) and its optimum growth temperature *in vitro* is between 37-42°C (Dance, 2000abc). While the *in vitro* situation can only attempt to mimic that of the natural environment, these results draw attention to the finding that *B. pseudomallei* grows exceptionally well at 42°C under some conditions. The possible association of *B. pseudomallei* with an intermediate host has been previously suggested. Such associations can provide many benefits for the bacterium, such as nutrient supply and freedom from competition, especially if the bacterium can enter a privileged niche like the internal compartments of host cells. Associations with protozoa (Inglis *et al.*, 2000) or plants (Pitt *et al.*, 2000) have been postulated to explain the prevalence and the uneven occurrence of bacterial infections especially in endemic areas (Dance, 2000). Birds are considered to be relatively resistant to *B. pseudomallei* infection (Lajudie & Brygoo, 1953) and the preferable growth at 42°C of all of the isolates of *B. pseudomallei* raises the possibility that birds may also be candidates for consideration as intermediate hosts. This is consistent with the studies of Vesselinova *et al.* (1996) who postulated that hens might act as a reservoir of *B. pseudomallei* and pose a risk for the transmission of active melioidosis to other species.

Although many bacterial two-component regulatory systems have been identified the sensory stimuli to which the systems respond are largely unknown. In an attempt to identify the nature of the environmental signal that is transduced by the MrgR/MrgS system, the expression of MrgR and MrgS was tested under a range of environmental conditions that were considered to mimic those that may be encountered in the host. Western blot analysis using anti-MrgR and anti-MrgS antibodies was carried out to examine the expression of MrgR and MrgS proteins. The results suggest that this system is under multifactorial control,

being expressed under different combinations of temperature, pH and salinity but especially at low pH. One possibility is that the MrgR/MrgS system transmits information about low pH, perhaps reflecting phagolysosomal pH.

In this study, the growth of *B. pseudomallei* under different conditions of pH, salt concentration and temperature was not apparently correlated with their ability to express MrgR-MrgS two-component regulatory proteins. Although *B. pseudomallei* responds to certain combinations of the tested parameters by reducing protein synthesis, the expression of MrgR and MrgS is not greatly affected suggesting that these genes may indeed have multiple biological functions. There are a number of possible explanations for the requirement for constitutive expression of MrgR and MrgS proteins but the most likely seems to be related to the ability of *B. pseudomallei* to survive in widely varied circumstances, from rice paddy fields to mammalian host tissues. Under such rapidly changing conditions the bacterium needs to respond quickly in a coordinate manner. MrgR and MrgS constitute a two-component regulatory system, where each of the two proteins may be coupled to more than one regulatory circuit. Cross-talk, which refers to the non-specific interactions between non-partner sensor and regulator proteins that collectively influence several regulatory responses (Gervais & Drapeau, 1992) (section 2.5.3.6), is common between members of the two-component family. This would provide a useful way to optimise expression of different sets of genes at any particular time, for example those that are needed to survive in the intracellular rather than extracellular situation or vice versa. Furthermore, it provides opportunities for signal integration and/or amplification, thereby fine-tuning adaptive responses. The EnvZ/OmpR system in *E. coli* and *S. typhimurium* is an example of a regulon that is networked to several other regulons and plays a role in both commensal and pathogenic processes. In *S. typhimurium* different sets of genes are regulated by the PhoPQ two-component regulatory system, in order to effect differential expression

during infection (Finlay & Falkow, 1997). Alternatively, such apparently constitutive expression may point to the existence of other regulatory factors, whether host-derived or endogenous, that come into play *in vivo* but which are not observed *in vitro*. On the other hand, it may be that other as yet untested parameters will effect substantial changes in the expression of MrgR and MrgS *in vitro*. e.g. iron, nutrients, oxygen, metabolites.

Fusion proteins, MBP-MrgR and MBP-MrgS, were constructed and expressed in *E. coli* and serum from a convalescent melioidosis patient reacted with MBP-MrgR by immunoblotting (section 6.3.5). Although this does not provide definitive proof for the specific recognition of MrgR by infected hosts it does raise the possibility that MrgR may be expressed by *B. pseudomallei* during melioidosis infection. It is not yet obvious why a cytoplasmic regulatory protein should be exposed to the host immune system and highlights the need for further investigation of the *mrgRS* regulatory locus.

There are a variety of studies that should be undertaken in order to fully investigate the ways in which MrgR and MrgS interact with each other, with other regulatory systems and with the transcriptional elements that regulate the expression of other genes. In particular, the antibodies that were developed in this study could be used to examine the *in vivo* expression of MrgR and MrgS. Evidence for the direct interaction between MrgR and MrgS has yet to be established. The isolation and purification of the MrgR and MrgS proteins will allow studies of the predicted biochemical properties of the proteins and their interactions. Studies of the phosphorylation of MrgS and the transfer of the phosphate moiety to MrgR would support the classification of these proteins as partner molecules of a two-component regulatory system.

In order to facilitate studies into the possible regulatory functions of the *mrgRS* locus, an attempt was made to construct a *mrgR* null mutant by gene disruption. The loss of function as a result of specific mutagenesis, and the restoration of the function by gene

complementation, has become the classic way of demonstrating the role of a cloned gene. The strategy required the insertion of an antibiotic cassette into the reading frame of *mrgR*, and the cloning of this construct into a non-replicative delivery system or suicide vector (Appendix VII). The mutant is generated by the replacement of the wild type allele with the disrupted *mrgR* suicide vector construct. Collaborators are still attempting to achieve this goal. The availability of such a mutant will allow further characterisation of the gene locus. For instance, differences in the biochemical and protein profiles of *B. pseudomallei* wild type and *mrgR* mutant strains could then be examined, e.g. by two-dimensional gel electrophoresis patterns, to evaluate the genes that are affected by *mrgR*, particularly under different growth conditions. Additionally, the effect on virulence could be assessed in experimental models of infection. Bacterial pathogenicity is multifactorial and pathogens require a number of mechanisms that may act individually or in concert to produce infection and disease. Nevertheless the elimination of a single component may yield valuable data about the contribution of the gene to these processes.

A total of four potential two-component regulatory systems have to date been identified in *B. pseudomallei*. i.e. *irlR/irlS* (Jones *et al.*, 1997), *pho* (Sangiambut *et al.*, 2000), *mrgR/mrgS* and *ntrB* (this study, Appendix 4). The nature of *B. pseudomallei* suggests that many more of these genes will be identified in this species particularly because the analysis of the complete bacterial genome sequences that are currently available indicates that the number of two-component systems is related to genome complexity (Msadek, 1999). The annotation of the recently completed *B. pseudomallei* genome sequence will facilitate the identification of the genes encoding two-component regulatory systems. Given the important regulatory functions of two-component signalling systems, an improved understanding of their function in *B. pseudomallei* could lead to the design of new classes of antibiotics, vaccines, or broad-spectrum antimicrobials that can block these signalling pathways.

APPENDICES

APPENDIX I

ABBREVIATIONS, MOLECULAR WEIGHTS AND CLASSIFICATIONS OF AMINO ACIDS

Amino acid	Abbreviation	Mol Wt	Codon	Functional	Hydrophobicity	Charge	Chemical
Alanine	Ala (A)	89	GCA GCC GCG GCU	HNP	Pho	0	Aliphatic
Arginine	Arg (R)	174	AGA AGG CGA CGC CGG CGU	+	Phi	+	Basic
Asparagine	Asn (N)	132	AAC AAU	POU	Phi	0	Amide
Aspartic acid	Asp (D)	133	GAC GAU	-	Phi	-	Acidic
Cysteine	Cys (C)	121	UGC UGU	POU	Phi	0	Sulfur
Glutamic acid	Glu (E)	147	GAA GAG	POU	Phi	-	Acidic
Glutamine	Gln (Q)	146	CAA CAG	POU	Phi	0	Amide
Glycine	Gly (G)	75	GGA GGC GGG GGU	POU	Phi	0	Aliphatic
Histidine	His (H)	155	CAC CAU	+	Phi	+	Basic
Isoleucine	Ile (I)	131	AUA AUC AUU	HNP	Pho	0	Aliphatic
Leucine	Leu (L)	131	CUA CUC CUG CUU UUA UUG	HNP	Pho	0	Aliphatic
Lysine	Lys (K)	146	AAA AAG	+	Phi	+	Basic
Methionine	Met (M)	149	AUG	HNP	Pho	0	Sulfur
Phenylalanine	Phe (F)	165	UUC UUU	HNP	Pho	0	Aromatic
Proline	Pro (P)	115	CCU CCC CCA CCG	HNP	Pho	0	Imino
Serine	Ser (S)	105	AGC AGU UCA UCC UCG UCU	POU	Phi	0	Hydroxyl
Threonine	Thr (T)	119	ACA ACC ACG ACU	POU	Phi	0	Hydroxyl
Tryptophan	Trp (W)	204	UGG	HPH	Pho	0	Aromatic
Tyrosine	Tyr (Y)	181	UAC UAU	POU	Pho	0	Aromatic
Valine	Val (V)	117	GUA GUC GUG GUU	HPN	Pho	0	Aliphatic
Termination	End		UAA UAG UGA				

HNP = Hydrophobic nonpolar, POU = Polar uncharged, Pho = Hydrophobic, Phi = Hydrophilic

APPENDIX II

DATA RELEVANT TO CHAPTER 2

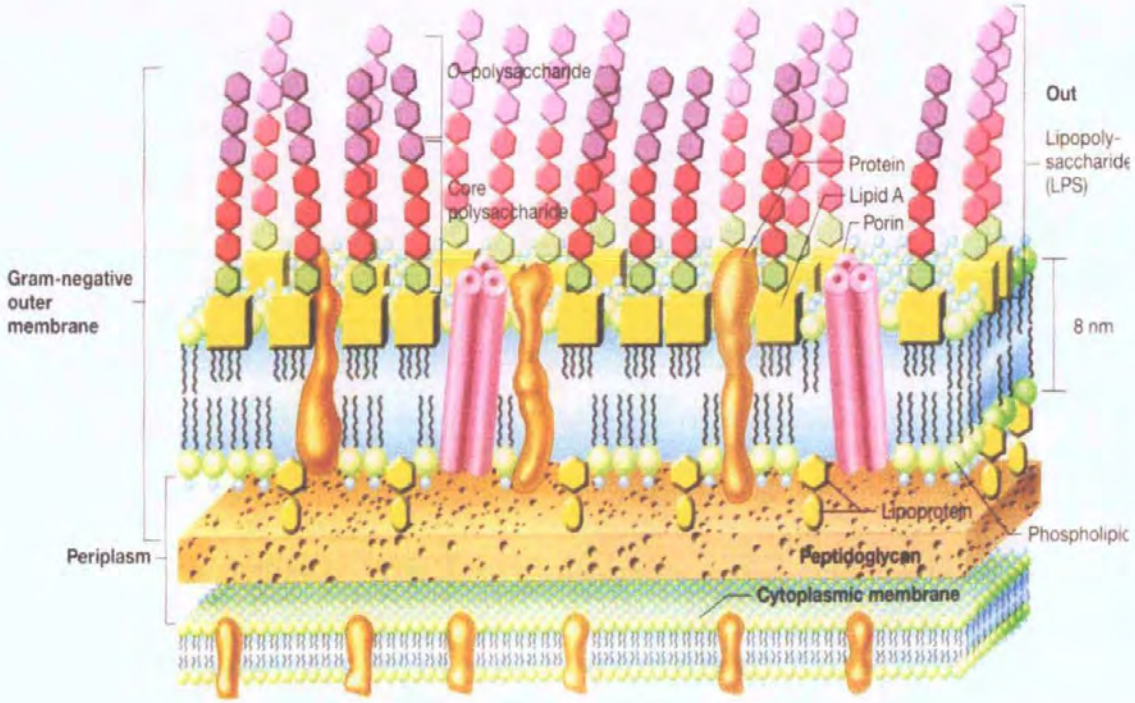


Fig 1: The structure of the cell wall of a typical Gram-negative bacterium. Adapted from Madigan *et al.* (1999).

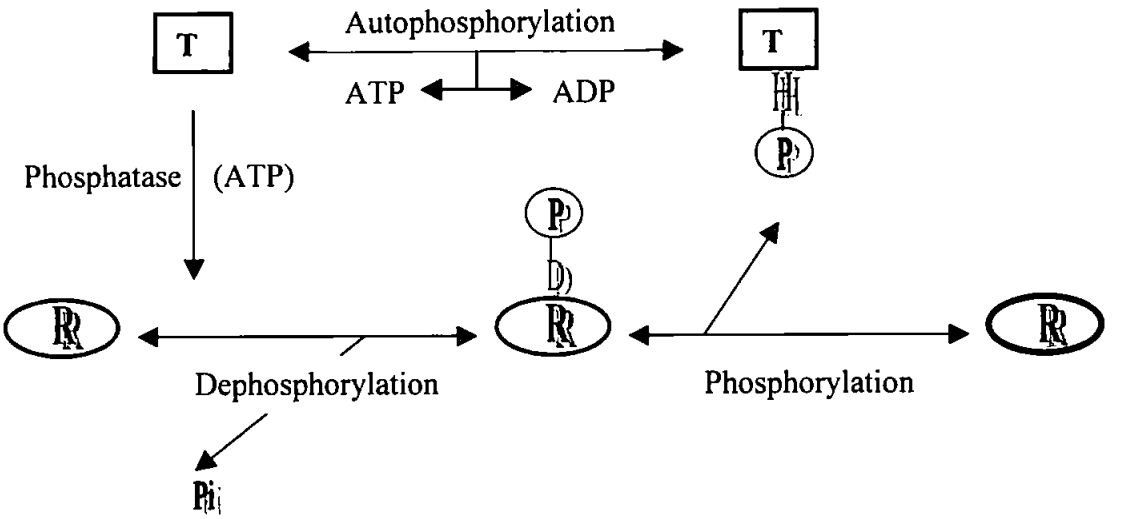


Fig 2: Sequence features and phosphorylation activities of the communication modules of two component regulatory proteins. Abbreviations: T, transmitter; R, receiver; H, histidine; D, aspartate; Pi, inorganic phosphate. (Adapted from Parkinson, 1993)

Table 1: Examples of two-component regulatory systems from various bacterial species and the adaptive responses they regulate.

System	Histidine protein kinase	Response regulator(s)	Adaptive response	Organism
Che	CheA	CheB/CheY	Chemotaxis	Ec/St
NR = Ntr	NRII=NtrB	NRI=NtrC	Nitrogen regulation	Ec/Kp/St
Nar	NarX/NarQ	NarL	Nitrate regulation	Ec
Pho	PhoR	PhoB	Phosphate regulation	Ec/Bs
Omp	EnvZ	OmpR	Porin gene regulation	Ec/St
Spo	KinA/KinB	SpoOA/SpoOF	Sporulation	Bs
Dct	DctB	DctD	Dicarboxylate transport	Rl
Hyd	HydH	HydG	Hydrogenase	Ec
Pgt	PgtB	PgtA	Phosphoglycerate transport	St
Arc	CpxA	ArcA	Respiratory control	Ec
Cre	CreC	CreB	Catabolite repression	Ec
Vir	VirA	VirA/VirG	*Agrobacterium* virulence	At
Pho	PhoQ	PhoP	*Salmonella* virulence	St
Uhp	UhpB	UhpA	Hexose phosphate uptake	Ec
Fix	FixL	FixJ	Nitrogen fixation	Rm
Nar	NarX	NarL	Nitrate reductase	Ec
Deg	DegS	DegU	Secretion of enzymes	Bs
Frz	FrzE	FrzE	Motility/development	Mx
Vir = Bvg	VirS = BvgS	VirR = BvgR	*Bordetella* virulence	Bp
Rcs	RcsC	RcsB	Capsule synthesis	Ec/Kp/Vc
Alg	AlgR2	AlgR1	Capsule synthesis	Pae
Kdp	KdpD	KdpD	Potassium transport	Ec
Irl	IrlS	IrlR	Heavy metal resistance	Bpm

Abbreviations: Ec, *Escherichia coli*; St, *Salmonella typhi*; Kp, *Klebsiella pneumoniae;* Bs, *Bacillus subtilis*; Rl, *Rhizobium leguminosarum*; At, *Agrobacterium tumefaciens*; Rm, *Rhizobium meliloti*; Mx, *Myxococcus xanthus*; Bp, *Bordetella pertussis*; Vc, *Vibrio cholerae*; Pae, *Pseudomonas aeruginosa*; Bpm, *B. pseudomallei.*

Table 2: A list of the ORFs that are considered to be involved in phosphotransfer signalling in *E. coli*. Adapted from Mizuno (1998).

Sensor/regulator	**Relative adaptive systems,** [putative]
OmpR-family	
PhoR/PhoB	Phosphate regulation
f480/f227	[Cops/R, Pseudomonas, copper response]
KdpD/KdpE	Potassium transport
TorS/TorR	Trimethylamine metabolism
PhoQ/PhoP	Stress situations
RstB/RstA	?
f452/f239	[CzcS/C, Alcaligenes, heavy metal homeostasis]
BaeS/BaeR	[AFQ2/1, Streptomyces, secondary metabolism]
o449/o219	[YgiH/X, Haemophilus, unknown]
ArcB/ArcA	Respiratory control
EnvZ/OmpR	Osmotic regulation
CpxA/CpxR	Multiple systems
BasS/BasR	Virulence
CreC/CreB	Catabolite repression
NarL-family	
NarX/NarL	Nitrate regulation
RcsB/RcsC	Capsule synthesis
EvgS/EvgA	[Bvgs/A, Bordetella, virulence]
NarQ/NarP	Nitrate regulation
UhpB/UhpA	Hexose phosphate uptake
-/FimZ	[-/FimZ, Salmonella, fimbrial expression]
-/UvrY	?
NtrC-family	
NtrB/NtrC	Nitrogen regulation
AtoS/AtoC	Acetoacetate metabolism
HydH/HydG	Labile hydrogenase activity
f496/f444	?
CheY-family	
CheA/CheB/CheY	Chemotaxis
Others	
o552/o226	[-/CriR, Shigella, ipa genes expression]
YjdH/YjdG	[CitA/B, Klebsiella, citrate metabolism]
YehU/YehT	[Lyts/T, Bacillus, autolysin response]
o565/o244	[-/MrkE, Klebsiella, fimbrial expression]
-/RssB	Sigma-S degradation
YojN	?
BarA/-	Pilus adherence

The annotations in parentheses are putative, based on similarities to the members of other species, as indicated.

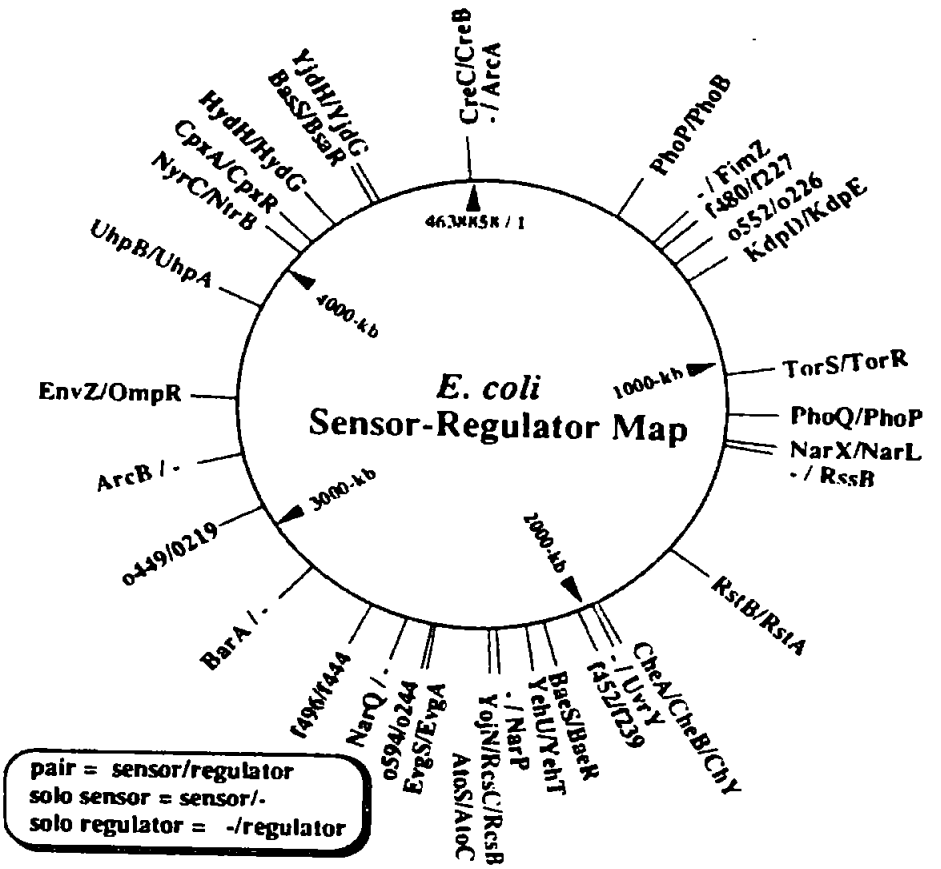


Fig 3: Chromosomal map of *E. coli* showing the positions of ORFs that are predicted to encode one or both of the signal transducer proteins. Adapted from Mizuno (1997).

APPENDIX III

FORMULATION AND PREPARATION OF MEDIA, BUFFERS AND SOLUTIONS

Unless otherwise stated all buffers and media were autoclaved at 121°C (15-pounds/square inch) for 15 min. The recipes contained the following per litre of distilled and deionised water.

MEDIA FOR BACTERIAL CULTURING

Luria-Bertani (LB) broth (pH 7.2)

Bacto-tryptone	10 g
Bacto-yeast extract	5 g
NaCl	5 g

LB agar

The same as LB broth plus 1.5% (w/v) agar.

LB soft agar overlay

The same as LB broth plus 0.75% (w/v) agar.

Table 1: ANTIBIOTIC STOCK SOLUTIONS (Filter sterilised, 4°C)

Antibiotic	Stock concentration	Solvent	Working conc.
Ampicillin	50 mg/ml	ddH_2O	50-100 µg/ml
Chloramphenicol	25 mg/ml	100% ethanol	25 µg/ml
Kanamycin	25 mg/ml	ddH_2O	25 µg/ml
Tetracycline	15 mg/ml	50% v/v ddH_2O/ethanol	15 µg/ml

PHAGE MEDIA AND BUFFERS

NZCYM broth (pH 7.0)

NZ amine (casein hydrolysate)	10 g
Bacto-yeast extract	5 g
NaCl	5 g
Casamino acids	1 g
$MgSO_4.7H_2O$	2 g

NZCYM agar

The same as NZCYM broth plus 1.5% (w/v) agar

NZCYM agarose overlay (soft)

The same as NZCYM broth plus 0.75% (w/v) agarose

Bacteriophage propagation medium

Consists of the following separately autoclaved solutions:

Casamino acid	10 g
M9 salts (10X)	100 ml
20% Maltose	20 ml
5 M $MgCl_2$	5 ml
1 M $CaCl_2$	0.1 ml

M9 salts (10X)

Na_2HPO_4	60 g
KH_2PO_4	30 g
NaCl	5 g
NH_4Cl	10 g

SM buffer

NaCl	5.8 g
$MgSO_4.7H_2O$	2 g
1 M Tris (pH 7.5)	50 ml
2% Gelatine	5 ml

EXTRACTION AND PURIFICATION OF DNA

B. PSEUDOMALLEI GENOMIC EXTRACTION SOLUTIONS

STE solution (pH 8.0)

NaCl	0.1 M
Tris-HCl	10 mM
EDTA	1 mM

Solution A

Sucrose	25% (w/v)
Tris-HCl (pH 8.0)	0.15 M

Solution B

Lysozyme	15 mg/ml
Tris-HCl (pH 8.0)	0.05 M

Solution C

Na_2EDTA (pH 8.0)	0.5 M

Solution D

Triton X100	8% (w/v)
Na_2EDTA	0.25 M

PLASMID DNA EXTRACTION SOLUTIONS

Birnboim and Doly (B & D) (1979) plasmid extraction

B & D Solution I (Resuspension solution) (pH 8.0)

Tris-HCl	25 mM
EDTA	10 mM
Glucose	50 mM

B & D Solution II (Lysis solution)

NaOH	0.2 M
SDS	1%

B & D Solution III (Neutralisation solution)

Na acetate	5 M

The pH was adjusted to 4.8 with glacial acetic acid

B & D Solution IV

Tris-HCl	50 mM (pH 8.0)
Na acetate	100 mM (pH 4.6)

RNase solution

Pancreatic RNase 1 mg/ml in 50 mM Tris-HCl (pH 8.0), heated at 95°C for 10 min, then stored at 20°C.

Lysozyme solution

10 mg/ml in 10 mM Tris-HCl (pH 8.0)

Tris-EDTA (TE) buffer (pH 8.0)

Tris-HCl	10 mM (pH 8.0)
EDTA	1 mM

Phenol: chloroform: IAA mixture (25:24:1)

Chloroform: iso amyl alcohol (IAA) 24:1, water saturated

Liquefied phenol equilibrated with Tris buffer (pH 7.4): chloroform IAA 1:1

AGAROSE GEL ELECTROPHORESIS OF DNA

Tris-borate-EDTA (TBE) buffer (10 X)

Tris base	108 g
Boric acid	55 g
EDTA 0.5 M (pH 8.0)	40 ml

This stock solution was diluted to prepare a 1 X TBE solution.

Gel loading buffer (6X)

Per 10 ml ddH_2O

Bromophenol blue	0.25 g
Xylene cyanol FF	0.025 g
Ficoll 400	1.5 g

Ethidium bromide solution

Ethidium bromide 10 mg/ml in ddH_2O, stored in dark bottles.

Table 2: Range of DNA separation using different agarose concentrations

Amount of agarose in gel (%w/v)	Efficient range of separation of linear DNA molecules (kb)
0.3	5 - 60
0.6	1 - 20
0.7	0.8 – 10
0.9	0.5 – 7
1.2	0.4 – 6
1.5	0.2 – 3
2.0	0.1 - 2

LIGATION CONDENSING AGENT

Hexaminecobalt chloride (HEXCO) made up as 10 mM stock solution in water and stored at -20°C.

TRANSFORMATION OF RECOMBINANT DNA INTO *E. coli* K-12

RF1 medium

RbCl	12 g
$MnCl_2.4H_2O$	9.9 g

Potassium acetate (1M, pH 7.5)	30 ml
$CaCl_2.2\ H_2O$	1.5 g
Glycerol	15 ml

The pH was adjusted to 5.8 with 0.2 M acetic acid and filter sterilised.

RF2 medium

RbCl	1.2 g
MOPS (0.5 M, pH 6.8)	20 ml
$CaCl_2.2\ H_2O$	11 g
Glycerol	150 ml

The pH was adjusted to 6.8 with NaOH and filter sterilised.

BLUE/WHITE SCREENING

IPTG solution

Isopropyl-β, D-thiogalactopyranoside (IPTG) made up as 24 mg/ml stock solution in dH_2O. Filter sterilised and stored at 4°C.

X-gal solution

5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) made up as 20 mg/ml stock solution in dimethylformamide (DMF). Filter sterilised and stored at 4°C.

HYBRIDISATION SOLUTIONS

SSC (20 X) (pH 7.0)

$Na_3C_6H_5O_7$	88.24 g
NaCl	175.32 g

Prehybridisation buffer

Per litre of SSC (5 X) solution

Blocking reagent (salmon sperm DNA)	0.5% (w/v),
SDS	0.02% (w/v)
N-lauroylsarcosine, Na-salt	0.1% (w/v)

The buffer was made fresh, dissolved and kept at 65°C.

Hybridisation buffer

Prehybridisation solution

Digoxygenin-11-dUTP labelled DNA probe

Wash buffer 1

SSC (20X)	100 ml
SDS (10%)	10 ml
ddH_2O	890 ml

Wash buffer 2

SSC (20X)	5 ml
SDS (10%)	10 ml
ddH_2O	985 ml

Detection buffer 1 (pH 7.5)

Tris-HCl	100 mM
NaCl	150 mM

Detection buffer 2

Buffer 1

Blocking reagent	0.5%

The buffer was made fresh, dissolved and kept at 65°C.

Detection buffer 3 (pH 9.5)

Tris-HCl	100 mM
NaCl	100 mM
$MgCl_2$	50 mM

Detection buffer 4 (pH 8.0)

Tris-HCl	10 mM
EDTA	1 mM

Depurination solution

HCl	0.2 M

Denaturation solutions

NaCl	1.5 M
NaOH	0.5 M

Neutralisation solution

NaCl	1.5 M
Tris-HCl (pH 8.0)	1 M

Colour development solution

Per 100 ml of Detection buffer 3

NBT	0.033 g

X-phosphate	0.0165 g

Transfer buffer

SSC (20X)	500 ml
ddH_2O	500 ml

PROTEIN ANALYSIS SOLUTIONS

SDS-PAGE preparations

Buffer A

Tris-HCl (pH 8.8)	0.75 M
SDS	0.2% (w/v)

Buffer B

Tris-HCl (pH 6.8)	0.25 M
SDS	0.2% (w/v)

Ammonium persulfate (APS)

Made up as 10 mg/ml stock solution in dH_2O

Resolving gel (12.5%)

Buffer A	5.2 ml
dH_2O	1.4 ml
Acrylamide (30% w/v)	4.4 ml
APS (10 mg/ml)	0.380 ml
TEMED	30 µl

Stacking gel (5%)

Buffer B	2 ml
dH_2O	1.3 ml
Acrylamide (30% w/v)	0.7 ml
APS (10 mg/ml)	0.1 ml
TEMED	15 µl

SDS-PAGE loading buffer (2X)

Per 10 ml

0.5 M Tris-HCl (pH 6.8)	2.5 ml
SDS 10% (w/v)	0.2 ml
β-mercaptoethanol	0.5 ml
bromophenol blue (0.1%)	0.5 ml

Glycerol	2.0 ml
ddH_2O	4.3 ml

SDS-PAGE modified loading buffer

The same as SDS-PAGE loading buffer minus bromophenol blue

Electrophoresis buffer (pH 8.3)

Tris-HCl	3 g
Glycine	14.4 g
SDS	1 g

STAINING OF POLYACRYLAMIDE GELS

Coomassie brilliant blue stain

Coomassie blue R250	0.25%
Methanol	30%
Acetic acid	10%

Destaining solution

Methanol	30%
Acetic acid	10%

WESTERN BLOTTING

Transfer buffer (pH 8.3)

Tris-HCl	3.03 g
Glycine	14.4 g
Methanol	100 ml

Phosphate-buffered saline (PBS)

One PBS tablet (Oxoid, Basingstoke) dissolved in 100 ml of ddH_2O.

Blocking solution

Casein 1% (w/v) in PBS buffer.

DAB Development solution

Per 100 ml PBS buffer

NaCl	1.17 g
Tris-Hcl	0.6 g
DAB	0.06 g
$NiCl_2$	0.03 g
H_2O_2	100 µl

Fusion protein column buffer (pH 7.4)

Tris-HCl	20 mM
NaCl	200 mM
EDTA	1 mM

PCR

Table 3: PCR reaction mixture (50 µl)

Components	Amount
Buffer	5 µl
dNTP mix (200 µM)	0.4 µl
Sense primer 20 pmol/µl	3 µl
Anti-sense primer 20 pmol/µl	3 µl
Taq polymerase 5U/µl	0.2 µl
Template DNA 20 ng/l	1-3 µl
HPLC H_2O	up to 50 µl

PCR buffer (1X) for *Taq* polymerase (Promega)

Trsi-HCl pH 9 at 20°C	11.2 mM
KCl	56 mM
Triton-X-100	0.112%

APPENDIX IV

DATA RELEVANT TO CHAPTER 4

Table 1: Complete restriction map of the 4.3 kb *Eco*RI fragment

Enzyme	Cuts No.	Positions of sites	Recognition sequence
AatII	4	137 505 561 3596	gacgt/c
Acc16I	2	1920 4290	tgc/gca
AccB1I	5	139 1173 2293 3198 4013	g/gyrcc
AccB7I	1	3088	ccannnn/ntgg
AccBSI	9	448 462 2065 2653 2764 2920 3556 3733 4126	gagcgg
AccI	4	154 1896 2199 2799	gt/mkac
AccIII	2	123 2790	t/ccgga
AcsI	3	1 421 4322	r/aatty
AcyI	13	134 140 502 558 1174 1431 1714 1803 2294 2396 3101 3593 4242	gr/cgyc
AfeI	2	1908 2519	agc/gct
AflIII	4	601 1795 2202 2301	a/crygt
AgeI	1	3021	a/ccggt
AhdI	1	3256	gacnnn/nngtc
Alw21I	7	601 1684 2007 2309 2514 3633 4191	gwgcw/c
Alw44I	1	2305	g/tgcac
AlwNI	1	2377	cagnnn/ctg
Ama87I	10	999 1683 1884 2039 2386 2513 3062 3560 4211 4237	c/ycgrg
Aor51HI	2	1908 2519	agc/gct
ApaI	1	2863	gggcc/c
ApaLI	1	2305	g/tgcac
ApoI	3	1 421 4322	r/aatty
AscI	2	114 2268	gg/cgcgcc
AspEI	1	3256	gacnnn/nngtc
AspHI	7	601 1684 2007 2309 2514 3633 4191	gwgcw/c
AspI	1	1892	gacn/nngtc
AtsI	1	1892	gacn/nngtc
AvaI	10	999 1683 1884 2039 2386 2513 3062 3560 4211 4237	c/ycgrg
AviII	2	1920 4290	tgc/gca
BalI	2	2381 4223	tgg/cca
BanI	5	139 1173 2293 3198 4013	g/gyrcc
BanII	4	1654 1684 2007 2863	grgcy/c
BanIII	3	3482 4071 4284	at/cgat
BbeI	3	143 1177 2297	ggcgc/c
BbiII	13	134 140 502 558 1174 1431 1714 1803 2294 2396 3101 3593 4242	gr/cgyc
BbrPI	1	1931	cac/gtg
BbsI	5	653 773 823 895 2074	gaagac

Table 1 continued

BbuI	4	316 588 1027 3314	gcatg/c
Bbv12I	7	601 1684 2007 2309 2514 3633 4191	gwgcw/c
Bbv16II	5	653 773 823 895 2074	gaagac
BcgI	7	327 507 1173 1383 2981 3179 4025	cgannnnnntgc
BclI	1	1618	t/gatca
BcoI	10	999 1683 1884 2039 2386 2513 3062 3560 4211 4237	c/ycgrg
BglI	4	1639 2961 3787 4256	
gccnnnn/nggc			
BlpI	3	3218 3662 3836	gc/tnagc
BpiI	5	653 773 823 895 2074	gaagac
BpmI	3	1820 2484 3142	ctggag
Bpu1102I	3	3218 3662 3836	gc/tnagc
Bpu14I	3	239 1096 3228	tt/cgaa
BpuAI	5	653 773 823 895 2074	gaagac
Bsa29I	3	3482 4071 4284	at/cgat
BsaAI	2	1931 3336	yac/gtr
BsaBI	4	642 753 2081 3614	gatnn/nnatc
BsaHI	13	134 140 502 558 1174 1431 1714 1803 2294 2396 3101 3593 4242	gr/cgyc
BsaI	1	3816	ggtctc
BsaMI	4	97 311 1187 1793	gaatgc
BsaOI	26	58 169 218 254 296 379 611 646 1113 1135 1375 2079 2421 2433 2508 2706 2719 2736 3108 3162 3706 3772 4195 4251 4271 4284	cgry/cg
BsaWI	7	123 702 869 2790 2885 3006 3021	w/ccggw
BscI	3	3482 4071 4284	at/cgat
Bse118I	18	298 793 1101 1112 1747 2132 2290 2714 2731 3021 3103 3195 3705 3884 3896 3993 4016 4270	r/ccggy
Bse8I	4	642 753 2081 3614	gatnn/nnatc
BseAI	2	123 2790	t/ccgga
BseCI	3	3482 4071 4284	at/cgat
BsePI	19	114 302 1707 1863 2174 2268 2469 2555 2710 3055 3164 3269 3301 3542 3572 3599 3880 3900 4038	g/cgcgc
BseRI	4	1670 2167 2611 3679	gaggag
BsgI	3	1931 2374 2728	gtgcag
Bsh1285I	26	58 169 218 254 296 379 611 646 1113 1135 1375 2079 2421 2433 2508 2706 2719 2736 3108 3162 3706 3772 4195 4251 4271 4284	cgry/cg
Bsh1365I	4	642 753 2081 3614	gatnn/nnatc
BshNI	5	139 1173 2293 3198 4013	g/gyrcc
BsiEI	26	58 169 218 254 296 379 611 646 1113 1135 1375 2079 2421 2433	cgry/cg

Table 1 continued

		2508 2706 2719 2736 3108 3162 3706 3772 4195 4251 4271 4284	
BsiHKAI	7	601 1684 2007 2309 2514 3633 4191	gwgcw/c
BsiI	9	289 343 1093 1168 2248 2503 2767 2980 3523	ctcgtg
BsiMI	2	123 2790	t/ccgga
BsiWI	2	1567 3336	c/gtacg
BsmBI	1	3595	cgtctc
BsmI	4	97 311 1187 1793	gaatgc
BsoBI	10	999 1683 1884 2039 2386 2513 3062 3560 4211 4237	c/ycgrg
Bsp106I	3	3482 4071 4284	at/cgat
Bsp119I	3	239 1096 3228	tt/cgaa
Bsp120I	1	2859	g/ggccc
Bsp13I	2	123 2790	t/ccgga
Bsp143II	9	143 347 1177 1296 1910 2218 2282 2297 2521	rgcgc/y
Bsp1720I	3	3218 3662 3836	gc/tnagc
Bsp68I	1	3734	tcg/cga
BspCI	10	169 218 611 646 1375 2079 2706 3162 3772 4284	cgat/cg
BspDI	3	3482 4071 4284	at/cgat
BspEI	2	123 2790	t/ccgga
BspMI	2	925 3206	acctgc
BspXI	3	3482 4071 4284	at/cgat
BsrBI	9	448 462 2065 2653 2764 2920 3556 3733 4126	gagcgg
BsrBRI	4	642 753 2081 3614	gatnn/nnatc
BsrDI	3	501 905 1081	gcaatg
BsrFI	18	298 793 1101 1112 1747 2132 2290 2714 2731 3021 3103 3195 3705 3884 3896 3993 4016 4270	r/ccggy
BssAI	18	298 793 1101 1112 1747 2132 2290 2714 2731 3021 3103 3195 3705 3884 3896 3993 4016 4270	r/ccggy
BssHII	19	114 302 1707 1863 2174 2268 2469 2555 2710 3055 3164 3269 3301 3542 3572 3599 3880 3900 4038	g/cgcgc
BssSI	9	289 343 1093 1168 2248 2503 2767 2980 3523	ctcgtg
BssT1I	1	3456	c/cwwgg
BstBI	3	239 1096 3228	tt/cgaa
BstD102I	9	448 462 2065 2653 2764 2920 3556 3733 4126	gagcgg
BstDSI	2	230 2590	c/crygg
BstEII	1	4312	g/gtnacc
BstH2I	9	143 347 1177 1296 1910 2218 2282 2297 2521	rgcgc/y

Table 1 continued

BstMCI	26	58 169 218 254 296 379 611 646 1113 1135 1375 2079 2421 2433 2508 2706 2719 2736 3108 3162 3706 3772 4195 4251 4271 4284	cgry/cg
BstPI	1	4312	g/gtnacc
BstSFI	2	2210 2924	c/tryag
BstX2I	2	2680 3414	r/gatcy
BstYI	2	2680 3414	r/gatcy
BstZI	11	55 251 293 376 1110 1132 2418 2505 2716 4248 4268	c/ggccg
Bsu15I	3	3482 4071 4284	at/cgat
CciNI	1	293	gc/ggccgc
CelII	3	3218 3662 3836	gc/tnagc
Cfr10I	18	298 793 1101 1112 1747 2132 2290 2714 2731 3021 3103 3195 3705 3884 3896 3993 4016 4270	r/ccggy
Cfr42I	2	233 2593	ccgc/gg
Cfr9I	2	999 4237	c/ccggg
CfrI	19	55 251 293 376 1110 1132 1357 1912 2118 2379 2418 2505 2716 3240 3453 4182 4221 4248 4268	y/ggccr
ClaI	3	3482 4071 4284	at/cgat
CpoI	2	2788 3193	cg/gwccg
Csp45I	3	239 1096 3228	tt/cgaa
CspI	2	2788 3193	cg/gwccg
DraII	1	2860	rg/gnccy
DraIII	2	1427 2326	cacnnn/gtg
DrdI	3	1941 2429 4241	gacnnnnn/nngtc
DsaI	2	230 2590	c/crygg
EaeI	19	55 251 293 376 1110 1132 1357 1912 2118 2379 2418 2505 2716 3240 3453 4182 4221 4248 4268	y/ggccr
EagI	11	55 251 293 376 1110 1132 2418 2505 2716 4248 4268	c/ggccg
Eam1104I	1	3967	ctcttc
Eam1105I	1	3256	gacnnn/nngtc
EarI	1	3967	ctcttc
Ecl136II	2	1682 2005	gag/ctc
EclHKI	1	3256	gacnnn/nngtc
EclXI	11	55 251 293 376 1110 1132 2418 2505 2716 4248 4268	c/ggccg
Eco130I	1	3456	c/cwwgg
Eco24I	4	1654 1684 2007 2863	grgcy/c
Eco31I	1	3816	ggtctc
Eco32I	3	52 787 2444	gat/atc
Eco47III	2	1908 2519	agc/gct

Table 1 continued

Eco52I	11	55 251 293 376 1110 1132 2418 2505 2716 4248 4268	c/ggccg
Eco57I	3	2056 3853 3955	ctgaag
Eco64I	5	139 1173 2293 3198 4013	g/gyrcc
Eco72I	1	1931	cac/gtg
Eco88I	10	999 1683 1884 2039 2386 2513 3062 3560 4211 4237	c/ycgrg
Eco91I	1	4312	g/gtnacc
EcoICRI	2	1682 2005	gag/ctc
EcoNI	2	3206 3671	cctnn/nnnagg
EcoO109I	1	2860	rg/gnccy
EcoO65I	1	4312	g/gtnacc
EcoRI	2	1 4322	g/aattc
EcoRV	3	52 787 2444	gat/atc
EcoT14I	1	3456	c/cwwgg
EcoT22I	1	1513	atgca/t
EheI	3	141 1175 2295	ggc/gcc
ErhI	1	3456	c/cwwgg
Esp1396I	1	3088	ccannnn/ntgg
Esp3I	1	3595	cgtctc
FauNDI	2	1210 2616	ca/tatg
FbaI	1	1618	t/gatca
FriOI	4	1654 1684 2007 2863	grgcy/c
FseI	1	4274	ggccgg/cc
FspI	2	1920 4290	tgc/gca
GsuI	3	1820 2484 3142	ctggag
HaeII	9	143 347 1177 1296 1910 2218 2282 2297 2521	rgcgc/y
HinlI	13	134 140 502 558 1174 1431 1714 1803 2294 2396 3101 3593 4242	gr/cgyc
HincII	7	155 1801 1897 2200 2229 2800 3622	gty/rac
HindII	7	155 1801 1897 2200 2229 2800 3622	gty/rac
Hsp92I	13	134 140 502 558 1174 1431 1714 1803 2294 2396 3101 3593 4242	gr/cgyc
KasI	3	139 1173 2293	g/gcgcc
Kpn2I	2	123 2790	t/ccgga
Ksp22I	1	1618	t/gatca
Ksp632I	1	3967	ctcttc
KspI	2	233 2593	ccgc/gg
LspI	3	239 1096 3228	tt/cgaa
MamI	4	642 753 2081 3614	gatnn/nnatc
MflI	2	2680 3414	r/gatcy
MluI	3	601 1795 2301	a/cgcgt
MluNI	2	2381 4223	tgg/cca
Mph1103I	1	1513	atgca/t

Table 1 continued

MroI	2	123 2790	t/ccgga
MroNI	6	298 2290 2714 3884 3896 4270	g/ccggc
MscI	2	2381 4223	tgg/cca
MslI	7	44 311 1259 1641 2312 3389 4250	caynn/nnrtg
Msp17I	13	134 140 502 558 1174 1431 1714 1803 2294 2396 3101 3593 4242	gr/cgyc
MspA1I	6	232 1628 2020 2545 2592 3508	cmg/ckg
Mva1269I	4	97 311 1187 1793	gaatgc
NaeI	6	300 2292 2716 3886 3898 4272	gcc/ggc
NarI	3	140 1174 2294	gg/cgcc
NdeI	2	1210 2616	ca/tatg
NgoAIV	6	298 2290 2714 3884 3896 4270	g/ccggc
NgoMI	6	298 2290 2714 3884 3896 4270	g/ccggc
NotI	1	293	gc/ggccgc
NruI	1	3734	tcg/cga
NsiI	1	1513	atgca/t
NspBII	6	232 1628 2020 2545 2592 3508	cmg/ckg
NspI	7	316 416 588 908 1027 1847 3314	rcatg/y
NspV	3	239 1096 3228	tt/cgaa
PaeI	4	316 588 1027 3314	gcatg/c
PaeR7I	4	2039 2513 3062 4211	c/tcgag
Pfl23II	2	1567 3336	c/gtacg
PflMI	1	3088	ccannnn/ntgg
PinAI	1	3021	a/ccggt
Ple19I	10	169 218 611 646 1375 2079 2706 3162 3772 4284	cgat/cg
PmaCI	1	1931	cac/gtg
PmlI	1	1931	cac/gtg
Ppu10I	1	1509	a/tgcat
PshAI	1	3104	gacnn/nngtc
Psp124BI	2	1684 2007	gagct/c
Psp1406I	1	1327	aa/cgtt
PspAI	2	999 4237	c/ccggg
PspALI	2	1001 4239	ccc/ggg
PspEI	1	4312	g/gtnacc
PspLI	2	1567 3336	c/gtacg
PspOMI	1	2859	g/ggccc
PstI	2	2214 2928	ctgca/g
PvuI	10	169 218 611 646 1375 2079 2706 3162 3772 4284	cgat/cg
PvuII	1	1628	cag/ctg
RsrII	2	2788 3193	cg/gwccg
SacI	2	1684 2007	gagct/c
SacII	2	233 2593	ccgc/gg
SalI	4	153 1895 2198 2798	g/tcgac
SfcI	2	2210 2924	c/tryag
Sfr274I	4	2039 2513 3062 4211	c/tcgag
Sfr303I	2	233 2593	ccgc/gg

Table 1 continued

SfuI	3	239 1096 3228	tt/cgaa
SgrAI	2	298 793	cr/ccggyg
SmaI	2	1001 4239	ccc/ggg
SphI	4	316 588 1027 3314	gcatg/c
SplI	2	1567 3336	c/gtacg
SrfI	1	1001	gccc/gggc
SstI	2	1684 2007	gagct/c
SstII	2	233 2593	ccgc/gg
StyI	1	3456	c/cwwgg
SunI	2	1567 3336	c/gtacg
Tth111I	1	1892	gacn/nngtc
Van91I	1	3088	ccannnn/ntgg
VneI	1	2305	g/tgcac
XcmI	1	3463	ccannnnn/nnnntgg
XhoI	4	2039 2513 3062 4211	c/tcgag
XhoII	2	2680 3414	r/gatcy
XmaI	2	999 4237	c/ccggg
XmaIII	11	55 251 293 376 1110 1132 2418 2505 2716 4248 4268	c/ggccg
Zsp2I	1	1513	atgca/t

The following endonucleases don't cut the sequence:

AatI, Acc113I, Acc65I, AclNI, AflII, AocI, AseI, AsnI, Asp700I, Asp718I, AvrII, BamHI, BfrI, BglII, BlnI, Bse21I, Bsp1407I, Bsp19I, BspHI, BspLU11I, BspTI, BsrGI, Bst1107I, Bst98I, BstI, BstSNI, BstXI, Bsu36I, CvnI, DraI, Eco105I, Eco147I, Eco255I, Eco81I, HindIII, HpaI, KpnI, MfeI, MspCI, MunI, NcoI, NheI, PacI, Pme55I, PmeI, PpuMI, PshBI, Psp5II, PstNHI, RcaI, SapI, SbfI, ScaI, SexAI, SfiI, SgfI, SmiI, SnaBI, SpeI, Sse8387I, SseBI, SspBI, SspI, StuI, SwaI, Vha464I, VspI, XbaI, XmnI.

Table 2: Codon usage deduced from ORFa (*mrgR*)

Codon	Count	%	Codon	Count	%	Codon	Count	%	Codon	Count	%
TTT-Phe	2	0.09	TCT-Ser	0	0.00	TAT-Tyr	0	0.00	TGT-Cys	0	0.00
TTC-Phe	3	1.35	TCC-Ser	2	0.90	TAC-Tyr	2	0.90	TGC-Cys	3	1.35
TTA-Leu	0	0.00	TCA-Ser	0	0.00	TAA-***	0	0.00	TGA-***	1	0.45
TTG-Leu	5	2.25	TCG-Ser	11	4.95	TAG-***	0	0.00	TGG-Trp	0	0.00
CTT-Leu	1	0.45	CCT-Pro	0	0.00	CAT-His	3	1.33	CGT-Arg	1	0.45
CTC-Leu	7	3.15	CCC-Pro	1	0.45	CAC-His	4	1.80	CGC-Arg	10	4.50
CTA-Leu	0	0.00	CCA-Pro	0	0.00	CAA-Gln	2	0.09	CGA-Arg	0	0.00
CTG-Leu	12	5.40	CCG-Pro	5	2.25	CAG-Gln	4	1.80	CGG-Arg	3	1.35
ATT-Ile	0	0.00	ACT-Thr	0	0.00	AAT-Asn	4	1.80	AGT-Ser	0	0.00
ATC-Ile	9	4.05	ACC-Thr	4	1.80	AAC-Asn	5	2.25	AGC-Ser	5	2.25
ATA-Ile	0	0.00	ACA-Thr	1	0.45	AAA-Lys	1	0.45	AGA-Arg	1	0.45
ATG-Met	9	4.05	ACG-Thr	9	4.05	AAG-Lys	5	2.25	AGG-Arg	2	0.90
GTT-Val	2	0.90	GCT-Ala	1	0.45	GAT-Asp	4	1.80	GGT-Gly	1	0.45
GTC-Val	12	5.40	GCC-Ala	7	3.15	GAC-Asp	7	3.15	GGC-Gly	10	4.50
GTA-Val	1	0.45	GCA-Ala	1	0.45	GAA-Glu	1	0.45	GGA-Gly	1	0.45
GTG-Val	12	5.40	GCG-Ala	11	4.95	GAG-Glu	10	4.50	GGG-Gly	4	1.80

Table 3: Codon usage deduced from ORFb (*mrgS*)

Codon	Count	%	Codon	Count	%	Codon	Count	%	Codon	Count	%
TTT-Phe	4	0.37	TCT-Ser	0	0.00	TAT-Tyr	10	0.92	TGT-Cys	1	0.09
TTC-Phe	35	3.24	TCC-Ser	10	0.92	TAC-Tyr	12	1.11	TGC-Cys	13	1.20
TTA- Leu	0	0.00	TCA-Ser	1	0.09	TAA-***	0	0.00	TGA-***	1	0.09
TTG- Leu	17	1.57	TCG-Ser	33	3.05	TAG-***	0	0.00	TGG-Trp	14	1.29
CTT-Leu	5	0.46	CCT-Pro	2	0.18	CAT-His	6	0.55	CGT-Arg	3	0.27
CTC-Leu	46	4.26	CCC-Pro	9	0.83	CAC-His	15	1.39	CGC-Arg	46	4.26
CTA-Leu	3	0.27	CCA-Pro	1	0.93	CAA-Gln	5	0.09	CGA-Arg	1	0.09
CTG-Leu	62	5.74	CCG-Pro	44	4.07	CAG-Gln	28	2.95	CGG-Arg	33	3.05
ATT-Ile	3	0.27	ACT-Thr	1	0.09	AAT-Asn	9	0.83	AGT-Ser	1	0.09
ATC-Ile	40	3.70	ACC-Thr	10	0.92	AAC-Asn	19	1.76	AGC-Ser	30	2.78
ATA-Ile	0	0.00	ACA-Thr	2	0.18	AAA-Lys	3	0.27	AGA-Arg	1	0.09
ATG-Met	32	2..96	ACG-Thr	39	3.61	AAG-Lys	20	1.85	AGG-Arg	5	0.46
GTT-Val	3	0.27	GCT-Ala	2	0.18	GAT-Asp	19	1.76	GGT-Gly	6	0.55
GTC-Val	28	2.59	GCC-Ala	34	3.15	GAC-Asp	36	3.33	GGC-Gly	45	4.17
GTA-Val	1	0.09	GCA-Ala	6	0.55	GAA-Glu	26	2.40	GGA-Gly	11	1.01
GTG-Val	49	4.54	GCG-Ala	83	7.69	GAG-Glu	43	3.98	GGG-Gly	12	1.11

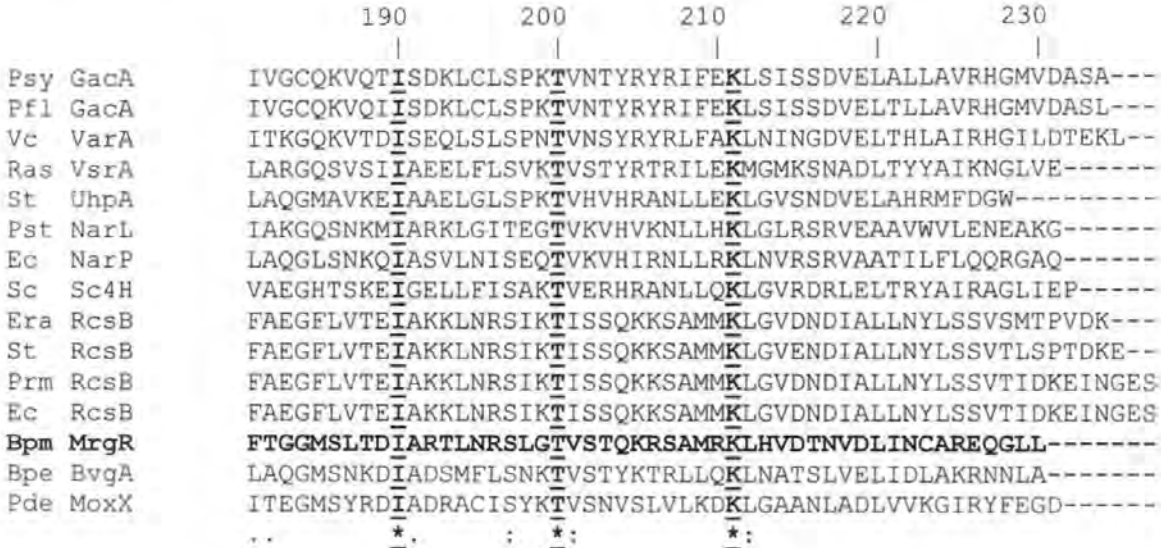
Fig 1: Multiple sequence alignment of proteins with >59% homology with the response regulator MrgR. Highly conserved amino acids are in bold and underlined.

```
                        10        20        30        40        50        60
                        |         |         |         |         |         |
Psy GacA    ----------MIKVLVVDDHDLVRTGITRMLADIDGLQVVGQADSGEESLKKARELK
Pfl GscA    ----------MIRVLVVDDHDLVRTGITRMLADIDGLQVVGQAESGEESLLKARELK
Vc  VarA    ----------MISVFLVDDHELVRTGIRRIIEDVRGMKVAGEADSGEEAVKWCRTNH
Ras VsrA    ----------MIRVLIADDHEIVRAGLRQFLSEERDIEVAGEAASGEEVMEQLRTEA
St  UhpA    ----------MITVALIDDHLIVRSGFAQLLGLEPDLQVVAEFGSGREALAGLPGRG
Pst NarL    --------MNEGTTTRILLVDDHPMMRRGLRDLLDLEDDLEMVGEAGNGEDAIRLALQIE
Ec  NarP    --------MPEATPFQVMIVDDHPLMRRGVRQLLELDPGSEVVAEAGDGASAIDLANRLD
Sc  Sc4H    --------MSARPPIRVLLADDHTLVRRGVRLILDGEPDLTVVAEAGDGAEAVAAARATE
Era RcsB    ----------MNNLNVIIADDHPIVLFGIRKSLEQIEWVNVVGEFEDSTALINSLSKLD
St  RcsB    ----------MNNMNIIIADDHPIVLFGIRKSLEQIEWVNVVGEFEDSTALINNLPKLD
Prm RcsB    ----------MNNLNIIIADDHPIVLFGIRKSLEQIEWVNVVGEFEDSTSLINNLSKLN
Ec  RcsB    ----------MNNLNIIIADDHPIVLFGIRKSLEQIEWVNVVGEFEDSTSLINNLSKLN
Bpm MrgR    ------MSNVALHTRKVVVADDHPIVLRAVTDYVNSLPGFHVVASVSSGDALLSAMREQE
Bpe BvgA    ----------MYNKVLIIDDHPVLRFAVRVLMEKE-GFEVIGETDNGIDGLKIAREKI
Pde MoxX    MQAALKLQSDDARPLQILIVDDHPVVAEGWGRIIRTKTACEIASAP-SASEGWRAWRQAR
                          : : *** ::  .   :       : .   ..

                        70        80        90        100       110       120
                        |         |         |         |         |         |
Psy GacA    PDVVLMDVKMP----GIGGLEATRKMLRSHPDIKVVAVTVCEEDPFPTRLLQAGAAGYMT
Pfl GscA    PYVVLMDVKMP----GIGGLEATRKLLRSHPDIKVVAVTVCEEDPFPTRLLQAGAAGYLT
Vc  VarA    ADVILMDMNMP----GIGGLEATKKLLRVNPDIKIIVLTVHTENPFPTKVMQAGAAGYLT
Ras VsrA    FDVVVLDISMP----DRNGIDTLKLVRQRHPDLPVLILSTFPEDQYAINLIRAGASGYLT
St  UhpA    VQVCICDISMP----DISGLELLSQLPKG---MATIVLSVHDSPALVEQALNAGARGFLS
Pst NarL    PDLILMDLNMP----GIDGLETTRRMRDADIDARIIMFTVSDEQSHVLEALRNGADGYLL
Ec  NarP    IDVILLDLNMK----GMSGLDTLNALRRDGVTAQIIILTVSDASSDVFALIDAGADGYLL
Sc  Sc4H    IDLAVLDVAMP----RMTGLQAARELSRRLPELRILILTMYDNEQYFFEALKAGACGYVL
Era RcsB    ANVLVTDLSMPGEKYG-DGITLIKYIKRHYPDLSIIVLTMNNNPAILSAVLDLDIEGIVL
St  RcsB    AHVLITDLSMPGDKYG-DGITLIKYIKRHFPSLSIIVLTMNNNPAILSAVLDLDIEGIVL
Prm RcsB    ADALITDLSMPGDKYG-DGITLIKYIKRHYPNLSIIVLTMNNNPAILSAVLDLEIGGIVL
Ec  RcsB    ADALITDLSMPGDKYG-DGITLIKYIKRHYPNLSIIVLTMNNNPAILSAVLDLEIGGIVL
Bpm MrgR    VNLVVTDFTMHQANDDKDGLRLISHLMRAYERTPIIVFTMLTNSGVISQLCRMGVAGLVG
Bpe BvgA    PNLVVLDIGIP----KLDGLEVIARLQSLGLPLRVLVLTGQPPSLFARRCLNSGAAGFVC
Pde MoxX    PDLMVVDLSIG--RNKIAGIRLIERLRRVDPDLPILVFTMHRSPVLARRALMFGANGIIN
                : *. :         *:      :          : .:                * :

                        130       140       150       160       170       180
                        |         |         |         |         |         |
Psy GacA    KGAGLAEMVQAIRLVFAGQ-RYISPQIAQQLALKSFQPQVNN--SPFDLLSEREIQIALM
Pfl GacA    KGAGLNEMVQAIRLVFAGQ-RYISPQIAQQLVFKSFQPSSD---SPFDALSEREIQIALM
Vc  VarA    KGAAPDEMVNAIRIVHSGQ-RYISPEIAQQMALSQFSPASE---NPFADLSERELQIMLM
Ras VsrA    KESAPDELVKAIRTVSQGR-RYVSPTVAELLIGGLEKPTDQ---PLHQTLSKREFQIFCK
St  UhpA    KRCSPDELIAAVHTVATGG-CYLTPDIAVKLAAGRQDPLTK-----------RERQVAEK
Pst NarL    KDMDAEQLIEQIRIAATGR-MALSPELTQVLAEAIRVRPKPSGQVQFSHLTKREKEVLRL
Ec  NarP    KDSDPEVLLEAIRAGAKGS-KVFSERVNQYLREREMFGAEED---PFSVLTERELDVLHE
Sc  Sc4H    KSVADRDLVEACRAAVRDE-PFIYPGAERALVRSYLDRVQGGGGLPERPITEREEEILKL
Era RcsB    KQGAPTDLPKALAALQKGK--KYTPESVAKLLEK--ISAGG---YGDKRLSPKESEVLRL
St  RcsB    KQGAPTDLPKALAALQKGK--KFTPESVSRLLEK--ISAGG---YGDKRLSPKESEVLRL
Prm RcsB    KQGAPADLPKALAALQKGK--KFMPESVAKLLEK--VNANG---YGDKRLSPKESEVLRL
Ec  RcsB    KQGAPADLPKALAALQKGK--KFMPESVAKLLEK--VNANG---YGDKRLSPKESEVLRL
Bpm MrgR    KEEEIAELGRVCVSVARGVSQSLSPGMAHRLAAVGSIRPGE---AAFNALTPKELEVVRL
Bpe BvgA    KHENLHEVINAAKAVMAGY-TYFPSTTLSEMRMGDNAKSDS---TLISVLSNRELTVLQL
Pde MoxX    KDSPPAEICAAFTEVARGG-NYVDSRLAMQIALLNVSRPGT----SAPRLTGREEEILGM
                *       :         .           :                     :*  :
```

Fig 1 continued



Abbreviations: Psy, *Pseudomonas syringae*; Pfl, *Pseudomonas fluorescens*; Vc, *Vibrio cholerae*; Ras, *Ralstonia solanacearum*, St, *Salmonella typhi*; Ec, *Escherichia coli*; Sc, *Streptococcus coelicolor*; Era, *Erwinia amylovora*; Prm, *Proteus mirabilis*; Bpm, *Burkholderia pseudomallei*; Bpe, *Bordetella pertussis*; Pde, *Paracoccus denitrificans*.

Fig 2: Multiple sequence alignment of the C-terminal of proteins with >34% homology with the MrgS sensor regulator. Highly conserved amino acids are in bold and underlined.

```
                        10        20        30        40        50        60
                        |         |         |         |         |         |
Ec  ArcB      --TFISTISHELRTPLNGIVGLSRILLDTELTAEQEKY-LKTIHVSAVTLGNIFNDIIDM
Vha LuxQ      --DFLAKMSHEIRTPINGILGVAQLLKDSVDTQEQKNQ-IDVLCHSGEHLLAVLNDILDF
Pfl ApdA      KSEFLANMSHEIRTPLNGILGFTHLLQKSELTPRQLDY-LGTIEKSADSLLGIINEILDF
St  SpiR      KSEFLANMSHEIRTPLNGILGFTHLLQKSELTPRQLDY-LGTIEKSADSLLGIINEILDF
Pto RtpA      KSEFLANMSHEIRTPLNGILGFTHLLQKSELSPRQLDY-LGTIEKSADNLLGIINEILDF
Psy GacS      KSEFLANMSHEIRTPLNGILGFTHLLQKSELTPRQFDY-LGTIEKSADNLLSIINEILDF
Psy rx10      ----LANMSHEIRTPLNGILGFTHLLQKSELTPRQFDY-LGTIEKSADNLLSIINEILDF
Pae lemA      KSEFLANMSHEIRTPLNGILGFTNLLQKSELSPRQQDY-LTTIQKSAESLLGIINEILDF
Vc  VieS      KSEFLANMSHEIRTPLNGIIGFTNLLQKCELTPRQHDY-LNTIQKSADSLLDIINEILDF
Erc RpfA      KSEFLANMSHELRTPLNGLIGFTRQTLKTPLNTTQTDY-LLTIERSANNLLNIINDVLDF
Ec  BarA      KSEFLANMSHELRTPLNGVIGFTRLTLKTELTPTQRDH-LNTIERSANNLLAIINDVLDF
Ec  RcsC      --MFLATVSHELRTPLYGIIGNLDLLQTKELPKGVDRL-VTAMNNSSSLLLKIISDILDF
Prm RcsC      ---FLATVSHELRTPLYGIIGNLELLQSLTQNEETARL-LKTMDNSSSLLLKIISDILDF
Bpm MrgS      -VAFFASMSHEIRTPLSSLVGNIELIARGPLAPEQQAR-VKAMETSARGLMQIVNDVLDF
Bbr BvgS      -TTFLATMSHEIRTPMNAIIGMLELALLRPADQEPDRQSIQVAYDSARSLLELIGDILDI
                  ::.:***:***: .::*                 : .  *.  *  :..:::*:

                        70        80        90        100       110       120
                        |         |         |         |         |         |
Ec  ArcB      DKMERRKVQLDNQPVDFTSFLADLENLSALQAQQKGLRFNLEPTLPLPHQVITDGTRLRQ
Vha LuxQ      SKIEQGKFNIQKHPFSFTDTMRTLENIYRPICTNKGVELVIENELDPNVEIFTDQVRLNQ
Pfl ApdA      SKIEAGKLVLDSIPFNLRDLLQDTLTILAPAAHAKQLELVSLVYRDTPLSLVGDPLRLKQ
St  SpiR      SKIEAGKLVLDSIPFNLRDLLQDTLTILAPAAHAKQLELVSLVYRDTPLSLVGDPLRLKQ
Pto RtpA      SKIEAGKLVLDSVPFNLRDLLQDTLTILAPAAHAKQLELVSLVYRDTPLALVGDPLRLKQ
Psy GacS      SKIEAGKLVLDNIPFNLRDLLQDTLTILAPAAHAKQLELVSLVYRDTPLALSGDPLRLRQ
Psy rx10      SKIEAGKLVLDNIPFNLRDLLQDTLTILAPAAHAKQLELVSLVYRDTPLALSGDPLRLRQ
Pae lemA      SKIEAGKLVLENLPFNLRDLIQDALTMLAPAAHEKQLELVSLVYRDTPIQLQGDPQRLKQ
Vc  VieS      SKIEAGKLVLEHTPFNLRELIQDTLTILAPAAHAKHLELVSLIYRDTPLALNGDPQRLRQ
Erc RpfA      SKLEAGKLVLEDIPFSLHNTLDDVVMLLAHTAHEKGLELTLSIQNDVPEQFVGDPLRIQQ
Ec  BarA      SKLEAGKLILESIPFPLRSTLDEVVTLLAHSSHDKGLELTLNIKSDVPDNVIGDPLRLQQ
Ec  RcsC      SKIESEQLKIEPREFSPREVMNHITANYLPLVVRKQLGLYCFIEPDVPVALNGDPMRLQQ
Prm RcsC      SKIESKQLKIDVKPFNCRQVVSFVISNYLPLIAKKDLAIYCFIEPDVPKIVNNDPVRVQQ
Bpm MrgR      SKIDVGELSLMEEWSNIAELLDRLALSHAPLATQQGLKFYMVFDRSLPARLYFDPIRVSQ
Bbr  BvgS     AKIEAGKFDLAPVRTALRALPEGAIRVFDGLARQKGIELVLKTDIVGVDDVLIDPLRMKQ
               *::  :. :                          : : :              .  *  *: *

                       130       140       150       160       170       180
                        |         |         |         |         |         |
Ec  ArcB      ILWNLISNAVKFTQQGQVTVRVRYD-----EGDMLHFEVEDSGIGIPQDELDKIFAMYYQ
Vha LuxQ      ILFNLVSNAVKFTPIGSIRLHAELEQFYGAENSVLVVELTDTGIGIESDKLDQMFEPFVQ
Pfl ApdA      ILTNLVSNAIKFTREGTIVARAMLEE-EHEDSVQLRISIQDTGIGLSNQDVRALFQAFSQ
St  SpiR      ILTNLVSNAIKFTREGTIVARAMLEE-EHEDSVQLRISIQDTGIGLSNQDVRALFQAFSQ
Pto RtpA      ILTNLISNAIKFTREGTIVARAMVED-EEEDSVQLRISVQDTGIGLSNQDVRALFQAFSQ
Psy GacS      ILTNLVSNAIKFTREGTIVARAMLED-ETEEHAQLRISVQDTGIGLSSQDVRALFQAFSQ
Psy rx10      ILTNLVSNAIKFTREGTIVARAMLED-ETEEHAQLRISVQDTGIGLSSQDVRALFQAFSQ
Pae lemA      ILTNLVGNAIKFTQGGTVAVRAMLED-ESDDRAQLRISVQDTGIGLSEEDQQALFKAFSQ
Vc  VieS      VLTNLISNAIKFTNEGTIAVRAMVED-IGEEFAQLRISVHDTGIGLSQEDLNALFQAFSQ
Erc RpfA      IITNLLGNAIKFTEQGNIDIRVEKRR-QELHQVLLEVQIRDTGIGIAELQQSQLFQAFRQ
Ec  BarA      IITNLVGNAIKFTENGNIDILVEKRA-LSNTKVQIEVQIRDTGIGIPERDQSRLFQAFRQ
Ec  RcsC      VISNLLSNAIKFTDTGCIVLHVRADG------DYLSIRVRDTGVGIPAKEVVRLFDPFFQ
Prm RcsC      VISNLLNNSIKFTTTGCITLHLYRDE------YYLYFDIHDTGRGINEKMLHQLFEPFFQ
Bpm MrgS      IVNNLLSNALKFTPSGKIVLRAGWR------AGALEISVTDSGIGIPDDLKHRLFLPFTQ
Bbr  BvgS     VLSNLVGNAIKFTTEGQVVLTVTARP--DGEAAHVQFSVSDTGCGISEADQRQLFKPFSQ
              :: **:.*::***  * :                 : . : *:* *:       :*  : *
```

Fig 2 continued

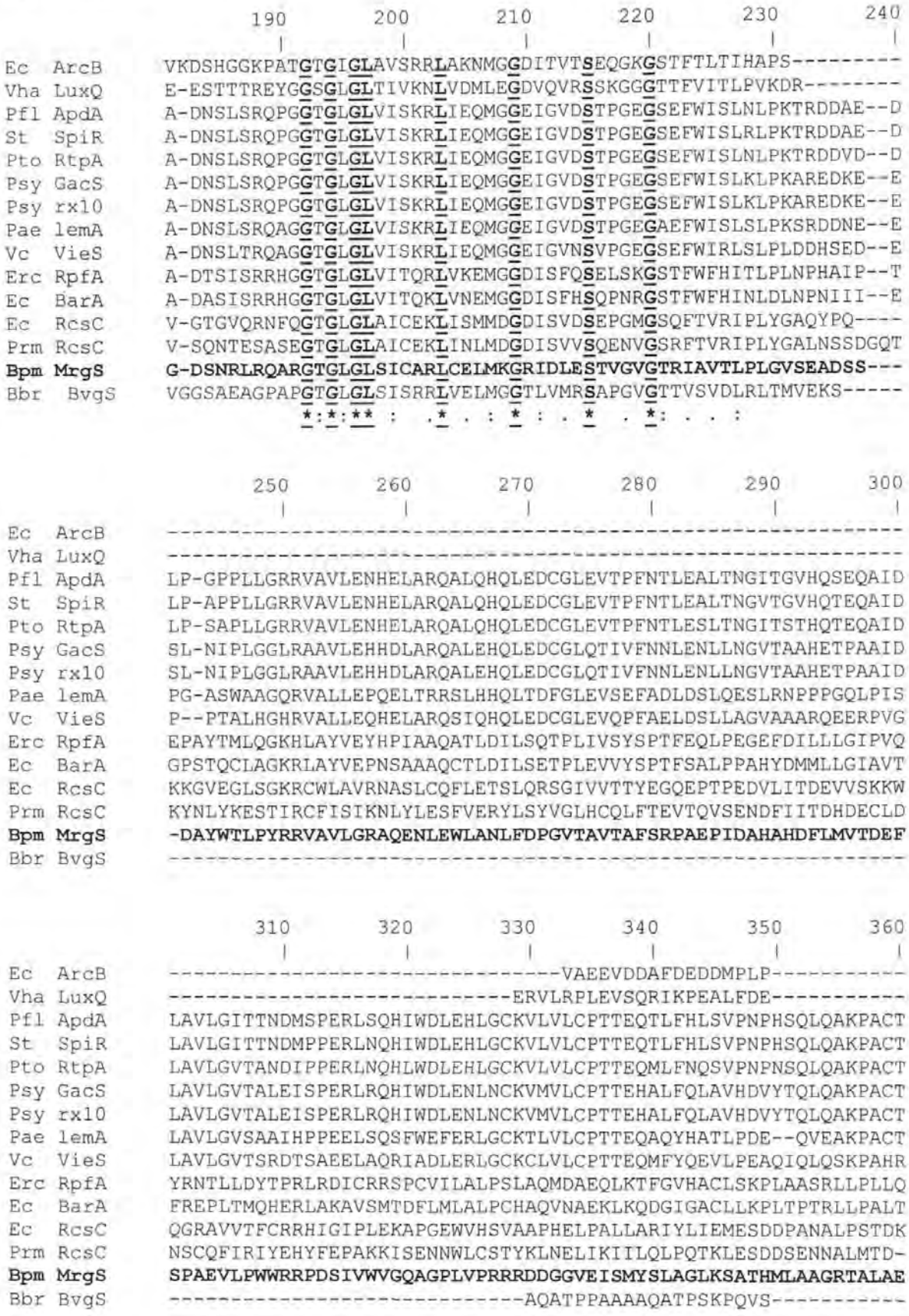


Fig 2 continued

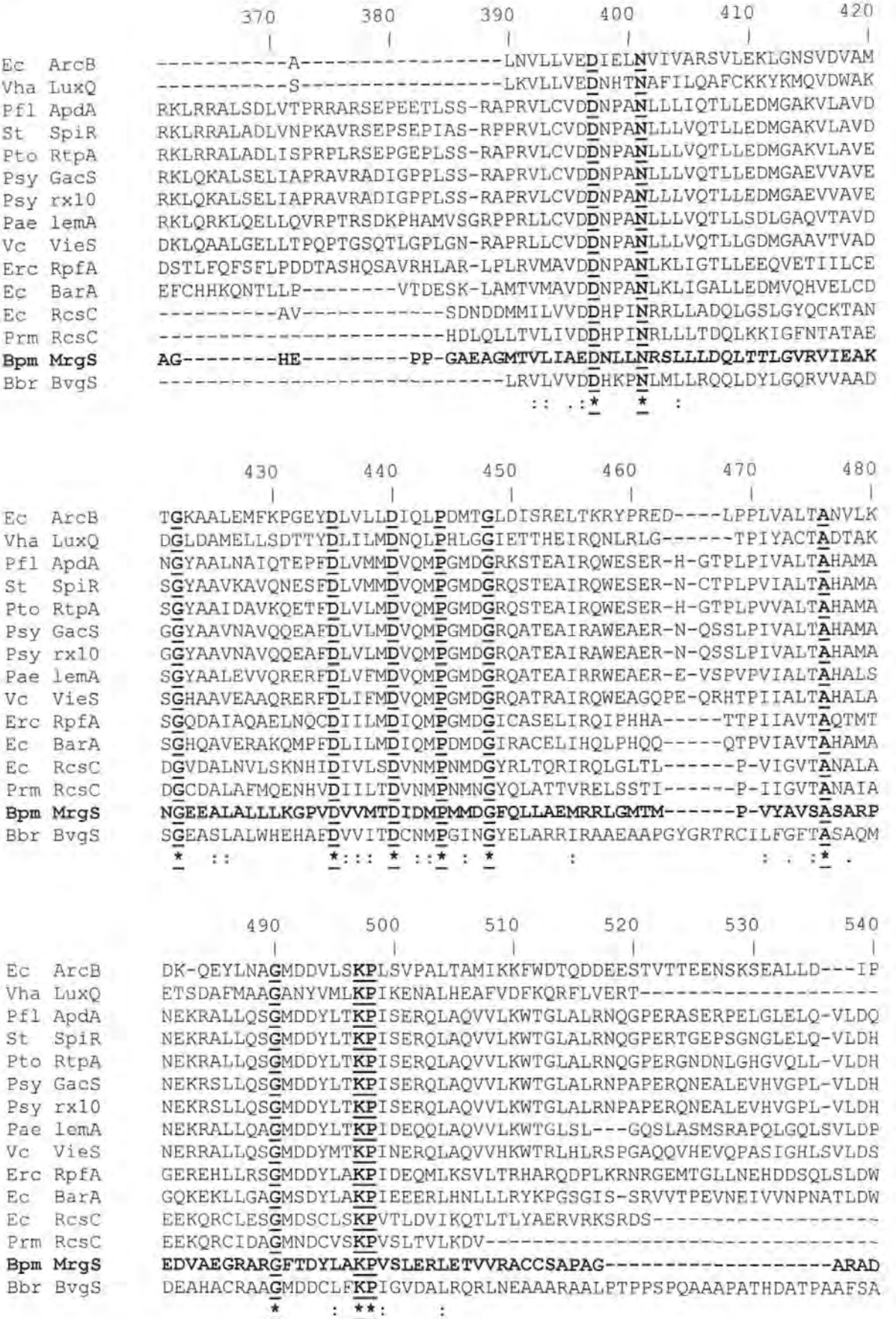


Fig 2 continued

```
                          550        560        570        580        590        600
                           |          |          |          |          |          |
Ec  ArcB     MLEQYLELVGPKLITDGLAVFEKMMPGYVSVLESNLTAQDKKGIVEEGHKIKGAAGSVGL
Vha LuxQ     ------------------------------------------------------------
Pfl ApdA     DEGLRLAAGKADLAADMLAMLLASLDADREAIKAARAANDQNALIERVHRLHGATRYCGV
St  SpiR     DEGLRLAAGKADLAADMLAMLLASLEADREAIRAARAANDHNALIERVHRLHGATRYCGV
Pto RtpA     EEGLRLAANKADLAADMLAMLLASLEADRLAITVAREAKDNNALIERVHRLHGATRYCGV
Psy GacS     EEGLRLAAGKADLAADMLAMLLASLDADREAIRVARANQDVHALIERIHRLHGATRYCGV
Psy rx10     EEGLRLAAGKADLAADMLAMLLASLDADREAIRVARANQDVHALIERIHRLHGATRYCGV
Pae lemA     EEGLRLAAGKADLAADMLAMLLASLAADRQAIRQARDNDDRTALLERVHRLHGATRYCGV
Vc  VieS     EEGLRLAAGKPDLAADMLSMLLASLDADRQAIHQARRDGDRAALIERVHRLHGATRYCGV
Erc RpfA     ALAQQQAANKPELGRDLLQMLLDFLPEVQRKIENVLKGQTDDDIIELVHKLHGSCSYSGV
Ec  BarA     QLALRQAAGKTDLARDMLQMLLDFLPEVRNKVEEQLVGENPEGLVDLIHKLHGSCGYSGV
Ec  RcsC     ------------------------------------------------------------
Prm RcsC     ------------------------------------------------------------
Bpm MrgS     EDAQDELPGLPDVPPAYASAFVAQAGSEIAEFDAILRERACRNCGGGCTAYRAASRSSGL
Bbr BvgS     ESILALTQNDEALIRQLLEEVIRTNRADVDQLQKLHQQADWPKVSDMAHRLAGGARVVDA

                          610        620        630        640        650
                           |          |          |          |          |
Ec  ArcB     RHLQQLGQQIQSPDLP-A-WEDNVGEWIEEMKEEWRHDVEVLKAWVAKATKK------
Vha LuxQ     ------------------------------------------------------------
Pfl ApdA     PQLRAACQRSETLLKQ---EDAKAFAALDELDHAIGRLAAEARTNA------------
St  SpiR     PQLRAACQRSETLLKQ---EDIKAFAALDELERAINRLATEARINA------------
Pto RtpA     PQLRAACQRAETLLKQ---DDAKAMAALDELDMAIARLASEARVNA------------
Psy GacS     PQLRSACQRAETLLKQ---NAPHTEEALNDLDKAIIRLEAEARVMA------------
Psy rx10     PQLRSACQRAETLLKQ---NAPHTEEALNDLDKAIIRLEAEARVMA------------
Pae lemA     PQLRAACQTSETLLKQ---NDPAAAAALDELDKAIEALADTASATTHLSSTSLDSSEL
Vc  VieS     PQLRAACQHSETLLKQ---HAPAADTALDELDAAILRLAGEAKFGA------------
Erc RpfA     P-LQPICRYLEQQLRKGV-HASDLEPEWLELLDEIDNVNRAAQPHINPRLP-------
Ec  BarA     PRMKNLCQLIEQQLRSGT-KEEDLEPELLELLDEMDNVAREASKILG------------
Ec  RcsC     ------------------------------------------------------------
Prm RcsC     ------------------------------------------------------------
Bpm MrgS     PHCMSNARSFEPTRAN---PANGIAKSNCRRWPFGTRWSEWSRR--------------
Bbr BvgS     KAMIDTALALEKKAQGQAGPSPEIDGMVRTLAAQSAALETQLRAWLEQRPHQGQP---
```

Abbreviations: Ec, *Escherichia coli*; Vha, *Vibrio harveyi*, Pfl, *Pseudomonas fluorescens*; St, *Salmonella typhi*; Pto, *Pseudomonas tolassi*, Psy, *Pseudomonas syringae*; Pae, *Pseudomonas aeruginosa*, Vc, *Vibrio cholerae*; Erc, *Erwinia carotovora*; Prm, *Proteus mirabilis*; Bpm, *Burkholderia pseudomallei*; Bbr, *Bordetella bronchiseptica.*

Fig 3: The complete nucleotide sequence of pMRG1 containing the partial sequences of two ORFs, ORFI (in red) and ORFII (in blue). The deduced translation of ORFI encodes a protein with high homology (up to 92%) with glutamine synthetase, GlnA, while ORFII encodes a protein with high homology (up to 63%) to sensor histidine kinase, NtrB. The putative stop and start codons are underlined.

Table 4: Examples of the NCBI BLAST matches for ORFI identified on a 1.8 kb *B. pseudomallei* DNA fragment (pMRG1)

Accession number	Match sequence title	% homology
AF082873	Histidine kinase- *Herbaspirillum seropedicae*	63
D83006	Two-component sensor NtrB- *Pseudomonas aeruginosa*	52
AF072440	Two-component sensor NtrB- *Enterobacter gergoriae*	53
L08499	Two-component sensor NtrB- *Vibrio alginolyticus*	50
A24114	Two-component sensor NtrB- *Klebsiella pneumoniae*	53
P41788	Two-component sensor NtrB- *Salmonella typhimurium*	51
S23900	Two-component sensor NtrB- *Proteus vulgaris*	51
C82038	Two-component sensor NtrB- *Vibrio cholerae*	49
P06712	Two-component sensor NtrB- *Escherichia coli*	50
AB038389	Two-component sensor NtrB- *Shewanella violacea*	50
AAC41398.1	Two-component sensor NtrB- *Acidithiobacillus ferrooxidans*	47
P45670	Two-component sensor NtrB- *Azospirillum brasilense*	44
Q52977	Two-component sensor NtrB- *Rhizobium meliloti*	42
P09431	Two-component sensor NtrB- *Rhodobacter capsulatus*	43
P14377	Sensor protein HydH- *Escherichia coli*	40
P08401	Sensor protein CreC- *Escherichia coli*	40
AF3059141	Sensor protein HydH- *Klebsiella pneumoniae*	39
AAB07807.1	Probable sensor protein MtrB- *Mycobacterium tuberculosis*	38

Table 5: Examples of the NCBI BLAST matches for ORFII identified on a 1.8 kb *B. pseudomallei* DNA fragment (pMRG1)

Accession number	Match sequence title	% homology
AF082873	Glutamine synthetase GlnA- *Herbaspirillum seropedica*	92
AAF40802.1	Glutamine synthetase- *Neisseria meningitidis*	85
AAB61772.1	Glutamine synthetase - *Neisseria gonorrhoeae*	85
AAF95885.1	Glutamine synthetase - *Vibrio cholerae*	83
BAB11930.1	Glutamine synthetase- *Shewanella benthica*	82
BAB11926.1	Glutamine synthetase- *Shewanella violacea*	83
AB03004.1	Glutamine synthetase- *Escherichia coli*	82
P06201	Glutamine synthetase- *Salmonella typhi*	82
P28786	Glutamine synthetase- *Proteus vulgaris*	80
Q59747	Glutamine synthetase- *Rhizobium meliloti*	79
P10583	Glutamine synthetase- *Azospirillum brasilense*	79
P43794	Glutamine synthetase- *Haemophilus influenzae*	77
Q10377	Glutamine synthetase- *Mycobacterium tuberculosis*	72

APPENDIX V

DATA RELEVANT TO CHAPTER 5

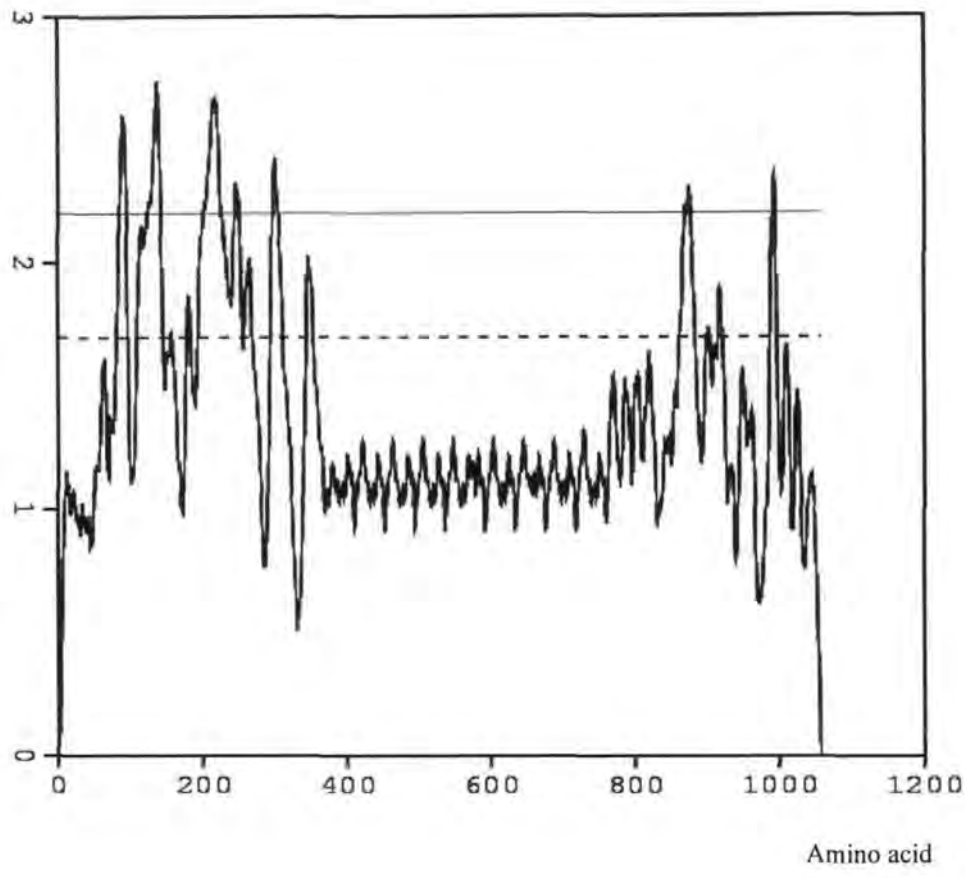


Fig 1: A computer hydropathy prediction showing possible transmembrane regions in the TcFx protein. Values above the horizontal dashed axis indicate regions of hydrophobic character and those below the axis indicate hydrophilic character.

Table 1: Examples of the NCBI BLAST matches for TcFx encoded by ORFx

Accession number	Match sequence title
A82615	Surface protein XF1981- *Xylella fastidiosa*
D82671	Surface protein XF1529- *Xylella fastidiosa*
AE005587	Putative adhesin- *Escherichia coli*
CAC14217	Autotransporter protein SapB- *Salmonella typhi*
AAC43721	Adhesin Hia- *Haemophilus influenzae*
C82672	Surface-exposed outer membrane protein- *Xylella fastidiosa*
NP052433	Adhesin YadA- *Yersinia enterocolitica*
AF181072	Outer membrane protein- *Moraxella catarrhalis*
AAC44560	Hsf protein- *Haemophilus influenzae*
P31489	Invasin precursor YadA- *Yersinia enterocolitica*
P10858	Invasin precursor YadA- *Yersinia pseudotuberculosis*
G81133	Adhesin protein- *Neisseria meningitidis*
AF226366	Outer membrane protein GNA992- *Neisseria meningitidis*
A81888	Surface fibril protein NMA1200- *Neisseria meningitidis*
S04911	Outer membrane protein YopA- *Yersinia pseudotuberculosis*
AJ277636	Putative autotransporter protein MapB- *Pasteurella multocida*
U86135	Outer membrane protein- *Moraxella catarrhalis*
AF187008	Serum resistance membrane protein- *Haemophilus ducreyi*

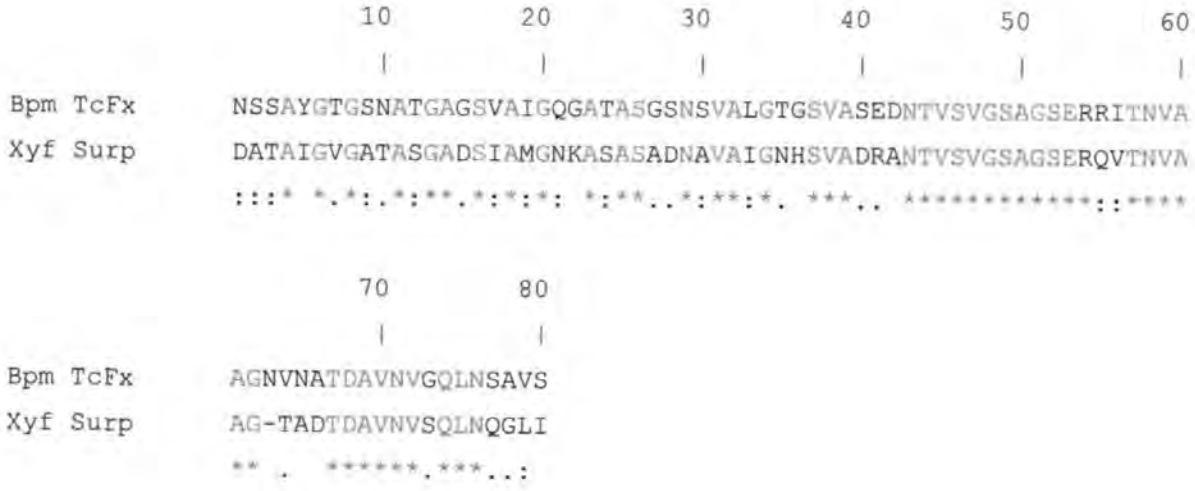


Fig 2: Clustal W multiple alignment of Tcfx with *Xylella fastidiosa* surface protein over a span of 80 amino acids located toward the C-terminal end of the proteins

Table 2: Examples of the NCBI BLAST matches for TcF2 encoded by ORFd

Accession Number	Match sequence title
AAC23902	BvgR protein- *Bordetella pertussis*
D64902	Hypothetical protein- *Escherichia coli*
BAA15144	Sensor protein FixL- *Escherichia coli*
P76129	Hypothetical protein YddU- *Escherichia coli*
P76129	Hypothetical protein *Pseudomonas tolassi*
S74609	Hypothetical protein Slr1102- *Synechocystis sp.*
CAB69812	Hypothetical protein YkoW protein- *Bacillus cereus*
CAA77326	Hypothetical protein Orf-2- *Pseudomonas aeruginosa*
AAC38223	Hypothetical protein Urf2- *Pseudomonas stutzeri*
P55552	Hypothetical 91.8 kDa protein Y4ll- *Rhizobium sp*
B70302	Sensory box/ GGDFE family protein- *Deinococcus radiodurans*
Q11024	Hypothetical 67.7 kDa protein Yd54- *Mycobacterium tuberculosis*
P77334	Hypothetical 74.7 kDa protein YciR- *Escherichia coli*
T09113	Probable response regulator VieA- *Vibrio cholerae*
P32701	Hypothetical 60.8 kDa protein Yjcc- *Escherichia coli*
Q11027	Hypothetical 33.9 kDa protein Yd57- *Mycobacterium tuberculosis*
P74101	Hypothetical 79.7 kDa protein Yi95- *Synechocystis sp*
E64862	Probable membrane protein B1168- *Escherichia coli*
P37649	Hypothetical 73.1 kDa protein YhjK- *Escherichia coli*
Q04855	Hypothetical 80.5 kDa protein YntC- *Azorhizobium caulinodans*
P77473	Hypothetical 58.9 kDa protein YlaB- *Escherichia coli*
P21514	Hypothetical 40.7 kDa protein YahA- *Escherichia coli*
AF071413	Hypothetical protein YfeA- *Escherichia coli*
P76446	Rtn protein- *Escherichia coli*
AAA50754	Rtn protein- *Proteus vulgaris*

Table 3: Examples of the NCBI BLAST matches for TcF3 encoded by ORFe

Accession number	Match sequence title
P26319	Expression regulatory protein FimZ- *Salmonella typhimurium*
P21502	Expression regulatory protein FimZ- *Escherichia coli*
AE0044432	Probable response regulator- *Pseudomonas aeruginosa*
AAF82336.1	Putative response regulator YsrR- *Yersinia enterolitica*
CAA35973	Putative two-component protein- *Streptomyces coelicolor*
B82043	Transcription regulator LuxR family- *Vibrio cholerae*
NB106300	Response regulator- *Mesorhizobium loti*
BAB01271.1	DNA binding protein- *Arabidopsis thaliana*

Table 4: Examples of the NCBI BLAST matches for TcF4 encoded by ORFf

Accession number	Match sequence title
AAF12376.1	Adenine deaminase-related protein- *Deinococcus radiodurans*
Q9SMV7	DNA repair- DNA binding protein- *Arabidopsis thaliana*
CAB69856.1	Hypothetical protein- *Arabidopsis thaliana*
P24563	Hypothetical 57.4 kDa protein- *Pseudomonas aeruginosa*
AAF35965.1	Hypothetical protein- *Caenorhabditis elegans*
B75276	DNA binding response regulator- *Deinococcus radiodurans*
Q62755	Transcriptional regulator-DNA binding protein- *Rattus norvegicus*
BAA83713.1	RNA binding protein- *Homo sapiens*
NP057417.1	RNA binding protein- *Homo sapiens*
A36128	Regulatory protein AlgP- *Pseudomonas aeruginosa*
BAB01996.1	DNA repair protein MutS- *Arabidopsis thaliana*

Table 5: Examples of the NCBI BLAST matches for TcF5 encoded by ORFg

Accession number	Match sequence title
P18952	Extracellular phospholipase A1- *Serratia liquefaciens*
AAC31479.1	Extracellular phospholipase A1- *Yersinia enterocolitica*
AAB09557.1	Putative ABC transporter ATPase BldkD- *Streptomyces coelicolor*
P18765	Oligopeptide transport ATP-binding protein- *Streptococcus pneumoniae*
CAB38590.1	Putative lipase- *Streptomyces coelicolor*
AAC26305.1	Hypothetical lipase- *Caenorhabditis elegans*
CAA35216.1	AimE transport protein- *Streptococcus pneumoniae*
Q07954	Low density lipoprotein- *Homo sapiens*
AAG51101.1	Putative lipase- *Arabidopsis thaliana*
AAC26305.1	Putative lipase- *Caenorhabditis elegans*

Table 6: Examples of the NCBI BLAST matches for TcF6 encoded by ORFh

Accession number	Match sequence title
AAC44775.1	Type III export protein PscD- *Pseudomonas aeruginosa*
AAC12650.1	Glycosidase OleR- *Streptomyces antibioticus*
F70111	V-type ATPase, subunit A homologous- *Borrelia burgdorferi*
AAG03817.1	Probable ATP-dependent RNA helicase- *Pseudomonas aeruginosa*
CAB84378.1	Putative isocitrate dehydrogenase- *Neisseria meningitidis*
E70839	Probable PE protein- *Mycobacterium tuberculosis*
G75477	Probable general secretion pathway protein- *Deinococcus radiodurans*

Fig 3: The complete nucleotide sequence of the cloned 10 kb fragment showing the location of DNA probes, FR1, FR2 and FR3 used in Southern hybridisation analysis. Sequence in blue indicates *mrgRS* flanking regions, while the PCR primer sequences that were used to generate each probe are in red bold.

GAATTCTAACTTCTTTGATGTTCGTCCGATGCGGGTTCGCATATAAATGGATATCGGCCGCAGTCAGG
CCGCCGCGCGAACGCCGGGCGACGCATTCGAGCAAGAAAATAGCGGCGCGCCCGTCCGGATGCGGACG
TCGGCGCCAACAGGATGTCGACGGTGGCGCGATCGTTCGGCGGCGGATGCGCTTTCGCCGCACGCGCG
TTCGGACAGTCGATCGGGAATCGTCCCGCGGCTTTCGAACGAGAGGCGGCCGTCGAGCGCGATCCCGC
AGGCGCGTTTTCTCGTGGGGCGGCCGCCGGCGCGCATTCGCATGCAAGCGTTTCGTCGATTTTCGCTC
GTGGCGCTTTTCGCAACGCTATGATTTTCGCCGCCCGGCCGCGCATCTATTTTCGTGAAAGGCCGATG
TGGACATGCGTGAAATTTCGAGTCACGCCGATTCCCGCTCGATTCGGCGAGCGGACACTTTTTCTCTC
TGGATCGGGAGCGTATTGGGCAATGACGTCGCGCGACTTTCGCGGCGGAGCGATGGGGGATTGATTTT
GATGGAACACACGACGTCGATTGGATGGGCGGCTCGCTCGCATGCGCGATCCGAGCACGCGTGCGATC
GGATTTCTCAATGGATTCTTGTTCTGATGCCGATCGTCTTCGTCGCGTTCTGCGTACTCCTGATTTCG
CTCTGGTACATCGACATCTTGACCGGATGGCTGATCGCGGCTTTTGCGCTGCTCATTGAAGCCATGCT
GATCGGCATCGGGTTTTTTGTCTTCGTGATGGTGCGGATATCGCACCGGCGGCGGCGCGAGCGCGAGG
CGAAGACGGCGTTGTGCCGCGCGAATCAGCCCGACGATGCCGATGTGCCGTTTCCGGTGGGATTCGAC
GGCGTGAAGACGGTTCATTGCATGTTTGGATCGGGGCAGGTCGGTGGGTTTCGGGTAAGCTGTTACAA
CACGAATTGCCCGTTGATGCGTGGAGCAACGGAAACCGAGCGATTGCCCGGGCAGGCGGAGCGAGCGT
AGGCATGCGACGCGCCATGACATGAAAAGCGAAAGCATATTGAAACCTTGAATATGCAATGGATGCGC
TCGTGATTCGAAGCCGGTTCACGGCCGGTTTACAGCCGATTCACGGCCGATTCGCGGCATGGCGAGGC
GGGGCGCACGAGCGTCGGCGCCGTTGCATTCGAGCGAATCGCGCTCGCAAGGCATATGCAGGAACGCA
GGCAATCGTTATCGTTTTGCATGGAACGCGCATGAGCGTGCGCCGGACCTCGTTCGTGCGGCACGGAG
GCGCTTTCCGCTTTTTTCGGAGCCGCCATGTCCAACGTTGCCCTGCATACAAGAAAAGTCGTCGTGGC
CGACGATCATCCGATCGTCCTGCGTGCGGTGACGGATTACGTCAATTCGCTGCCGGGTTTTCACGTCG
TGGCGTCGGTTTCGTCGGGGGACGCGCTGTTGTCCGCGATGCGGGAGCAGGAGGTCAATCTCGTCGTG
ACCGATTTCACGATGCATCAGGCGAACGACGACAAGGACGGCTTGCGCTTGATCTCGCACCTGATGAG
GGCGTACGAGCGCACGCCGATCATCGTGTTCACGATGCTGACCAATAGCGGCGTGATCAGCCAGCTGT
GCCGCATGGGCGTGGCGGGGCTCGTCGGCAAGGAGGAGGAGATCGCCGAGCTCGGGCGCGTATGCGTG
AGCGTGGCGCGCGGCGTCAGCCAATCGTTGTCGCCCGGCATGGCTCACCGGCTTGCCGCCGTCGGCAG
CATCAGGCCGGGCGAGGCGGCATTCAACGCGTTGACGCCGAAGGAACTGGAGGTCGTGCGGCTGTTTA
CGGGCGGCATGTCGCTCACGGACATCGCGCGCACGCTGAATCGCTCGCTCGGGACCGTGTCGACGCAG
AAGCGCTCGGCCATGCGCAAGCTGCACGTGGACACCAACGTCGATCTCATCAACTGCGCCCGCGAGCA
AGGACTGCTCTGATGCAAGGACTTCTGCAAGAGCTCGACGGATCGCCGCTGAGGAAGTTCTATTCGCT
CGAGTCGAATCTGAAGCGCGAGCGGCGGGTCTTCACGATCGTCATCGTGCTGCTCGTCTGCGCGGCCC
TCAGCATCGCGGCCATGACCGTCACCGGCTTGTTCCAGACCGCTTTCCGGCAGGAGGAGCAATCCGCG
CGCATCCACGAAAAGGAAGTGGTCGACGTGTTCCTGCAGCGCCGCATGATGTTGACGACGGCAAGCCT
CGTGCTGCAACTGCGGATGAACGGCGCGCCTTCGGCGCTGAACGTGCCGGCGCCGAACGCGTGCACGC
CGATGGCCCACAATGTGCGCGACGATGCGATCCTGCGCGAGAGCTGCGATTACACGGTGCAGTTGCTG
GCCAACTCGGGGCAGACGCCGAGCGTCGAAATGGTGACGGCCGACGGTTCGGTCGGCTATGGATATCT
GTTTCCGACGGGCGACCTGAGCGCGCTACGCTCCAGCACGCCGTCCGAACTCGTGTCGGCCGTGCTCG
AGCGCTACGGCAAGCGCGGCCTGGACCCGCTGGAAGCCGCGCGCAAGAAGCGGATTCTCTGGTTCGCG
GTGGGCCGCGGCGGGCGCGGCGAGGAGCTGCATATGATCGGCGCGTCAGTCGTGTTCAAGGACGAGCG
GCTCTACGCGCTCGTCTTGACGAGCGTGGATCTTCACAGCCTCGTTTCGCCGATCGAGCGCGCCGGCC
GCGTGCAGCAGCCGGTCGTCGTGGATTCCGAGGGCGTGCCGCTCGTGAACGCGGACGACGCGGAAACG
GTCCGGAAGGTCGACGGGCGGCTCGCCGGACAACAGGATGGCCTGTATCACTGGATTCCCGGCTTCGG
GTGGGCCCTACGCCGTCCCGCGCCGTTTTCCGGTTTCGGGCACATGACGTATCTGCTTCCGCTCGATC
TGCAGTTGCGCTCGATGCGCTACGAGTTGAGCCTCGTCGGCGGCGCGACGCTCGTGCTGATCGTGTTG
CTGTTCGTCGCGTTCCGGTACTGGAATTACCGGTTCTTGACGCGCATCTACGAGGAAGCGTCGCGCGC
GCTCGAGAGCGAAATGCTCAACCATCTGCTGGTTCATGCGACGCCGGTCGGGTTGTGCATCGTGCGGC
GGGCGACACTGGAGATCGTCGTCGCCAACCCGATCGCGCGCACGATGCTCGGCTTGCGGCTGTCGGAC
CGGCACCTGCCGCAGGAATTGCTGAGCGCGTTCGAATCGTCGCTGGCCGAGCAGGACACCCAGTCCGA
CGACGCGCGCATTTTCCAGTTCCCGTTCACGCTGTCGCGCGCCGGGCATGCGGCGGTCCATATCGAAA
TCACGTACGCGCCCGCGATGCTGAACGCGCGGGAGGTGTTCTTTTGCGCGATCACGGACATGACGGCG
CACCACCAGGCGGAGATCCTGCTGCGCGAGGCGAAGCTGACGAGCGACGCGGCGGCCAAGGCGAAGGT
GGCGTTCTTCGCATCGATGAGCCATGAAATCCGCACGCCGCTGTCGTCGCTCGTGGGCAACATCGAGC
TGATCGCGCGCGGGCCGCTCGCGCCCGAGCAGCAGGCGCGCGTGAAGGCGATGGAGACGTCGGCGCGC

Fig 3 continued

GGCTTGATGCAGATCGTCAACGATGTGCTCGATTTCTCGAAGATCGACGTGGGCGAGCTGAGCCTCAT
GGAGGAGTGGTCGAACATCGCCGAGCTGCTCGACCGGCTCGCGCTCTCGCACGCGCCGCTCGCGACGC
AGCAGGGTTTGAAGTTCTACATGGTGTTCGATCGCAGCCTGCCCGCGCGGCTCTACTTCGATCCGATC
CGGGTCTCGCAGATCGTGAACAATCTGCTGAGCAACGCGCTGAAGTTCACGCCGTCCGGCAAGATCGT
GCTGCGCGCCGGCTGGCGTGCCGGCGCGCTCGAAATCAGCGTGACGGATTCCGGCATCGGCATCCCCG
ATGACCTGAAGCACCGCCTCTTCCTGCCCTTCACGCAGGGCGACAGCAACCGGCTGCGGCAGGCACGC
GGCACCGGCCTCGGATTGTCGATCTGCGCGCGTCTGTGCGAGCTGATGAAAGGGCGCATCGATCTGGA
AAGCACCGTCGGCGTGGGAACCCGGATCGCGGTGACACTGCCGCTCGGCGTGTCGGAGGCCGATTCGA
GCGATGCGTACTGGACGCTTCCGTATCGGCGCGTGGCCGTGCTCGGTCGCGCACAGGAAAATCTCGAG
TGGCTGGCCAACCTGTTCGACCCGGGCGTCACGGCCGTGACGGCTTTCTCGCGGCCGGCCGAGCCGAT
CGATGCGCACGCGCACGATTTCCTGATGGTCACCGAC**GAATTC**TCGCCGGCCGAGGTGCTGCCGTGGT
GGAGGCGGCCGGACTCGATCGTGTGGGTCGGGCAGGCCGGCCCGCTCGTGCCGAGACGGCGCGACGAC
GGCGGAGTGGAAATCAGCATGTATAGCCTCGCGGGGCTGAAATCCGCGACTCACATGCTCGCGGCCGG
CCGCACGGCGCTCGCCGAAGCGGGGCACGAGCCGCCGGGAGCCGAGGCGGGAATGACGGTGCTGATCG
CCGAGGACAATCTGCTCAACCGCAGCCTGCTGCTCGATCAGCTGACGACGCTGGGCGTGCGGGTCATC
GAGGCGAAGAACGGCGAGGAGGCGCTCGCGTTGCTGTTGAAGGAGCCGGTGGACGTCGTGATGACCGA
CATCGACATGCCG**ATGATGGACGGTTTCCAGTTGCTC**GCCGAGATGAGGCGGCTCGGCATGACGATGC
CGGTGTACGCGGTGAGTGCGAGCGCGCGGCCGGAAGATGTGGCGGAAGGCCGGGCGCGCGGCTTTACC
GACTATCTCGCGAAGCCGGTTTCGCTCGAGCGGCTCGAGACGGTGGTACGCGCATGTTGCAGCGCGCC
GGCGGGCGCGCGCGCCGACGAAGACGCGCAGGACGAACTGCCGGGCCTACCCGACGTGCCGCCCGCCT
ATGCGAGCGCGTTCGTCGCGCAGGCCGGCAGCGAAATCGCG**GAATTC**GACGCGATCCTGCGCGAACGC
GCCTGCCGAAACTGCGGCGGTGGCTGCACGGCGTATCGGGCGGCATCGCGGTCCTCGGGCCTTCCGCA
CTGCATGAGCAATGCCAGGAGCTTCGAGCCTACGCGCGCGAATCCGGCGAATGGAATCGCGAAATCGA
ACTGCAGGCGCTGGCCATTCGGAACGCGCTGGAGCGAATGGTCGCGGCGC**TGA**CGAGCGCGTGATCGG
CGCGGGCGCGGCCCTGCCGGAGCGGGCCCGCGCCCGCCATGGACGGCGCGGCGCGGGGGTGTGCGCGG
CGTCGCACGTTCCGGCGCGGATTCGAAGATTTGTGATTCTCC**CCACCCGCGACTTGATTTAACGTT**TC
CTGACGATTCGCGGGGCGATCGGGCCGGCGGGCGTCGGCGTCGTGTGACGAGGCCGTCCTGCGGCGAT
CGCCAAGCCGCGACGAGCGATGCTGCGACGGCGGCGTCTTTGGGAGAGCGATATGGAAGCGGGCATCA
CGCGTCGTAGGGGTGCATTCTTTGCCCGGGAACACGAACACGGACTGTCGTTGGACGAGGCCGATCTC
GAGCGGGCATTCGAGGCCGGA**GAATTC**GTCATGGATTATCAGCCGATCGTCTCGATCCGCTCCGGGGC
CGTGACGGGCGCCGAAGCGACGATACGATGGGATCATCCGGAATGGGGCGCGCTGCCTCAGCGTGTGG
TCCACGCCGCCGCGGACCGATTGGGCGCGGCGTCCCGGATCGCCGGCCACGCGTTCGGCGAGGCGTGC
AGGCAGCTCGCGCGCTGGAAGCAGCAGGGCGCCGGCGTGTCGGCGCTGTCGATGCGGCTTTCGGGCGC
CCAGTTGGGCGCCGAAAGCGTATTCGAGCGACTCGCGCTGTCGATCGAGGCGTTCGATATCTCGAGCA
GCCAGTTCACGCTCGAAGTGCCGGAAATGTCGGCGCCGGAGGAATCGCTGAGTCTTCTCGATCGCCTG
AAGCGGCTGAGGCAGAAGGGTTACGGCATCGTGCTCGGCGATTTCGGAGCGCATCACACGGCGATGTC
GACGTTGATGGTGCTGCCCGTCACGGGCGTCAAATTCGGCGAATCGTTCACGGAGCGATTGCCCGGTT
CGCCGACTGCGGAGGCGATCCTGTCGAGCGTCTCGCGTCTCGCGCACGATCTGGGTTTCACGTTGACG
GTGTCCGGCGTCGAAAACGGCCGGCAGTTCGAATTGCTGCGTCGCTTTCGCGATATCGAGCTGCAGGG
CGCCTTTCTGTTCGAGCCGATGCGCGCCGAGGTGTGGCACGAGCGCGTGAATCCG**AGAGCGCTGTCGC**
AACTGAATCTGCGTGAGCCGCCGTTCTGACGCGGCGAGCCTGGCCGTTCGCGTTTCATCCCGAGGTGC
GCGGTCGGCCCGCCCGCATGGCGAGCCGGGTTGCGCAGCCGCCGCGCGTGCGCCGACGCGAGGACGGC
CGCGAGTCGCGGCCGGTGAGATCCGCGCCGCGCTCGCGCGCGCCGATGTCTGGCGCGGGCGTCGCGGA
TGGCGTTTATGCCGCTCATGCTTTCGCTCGCCAGTCTTGGGCGATCCGGTGGCCCGCCGAGCGCGCAT
CCGATGCGCTCGCGTGTCCGTGCGCGCATGCGCGGCGGCACGGCGCGTTCGACGTGGCCCGAGCGGCG
CGATGCGGCCGACGACGGCATACGCGATCGCGCCCGGCGCGGGGCCGTTCGCCTGCCGATATCGCGCG
CCGCGGCCCGACGGGCCGGCTCATTGCGAACCCACGTCCAGACATTGCCTCGCTTGTCTTCGCCCGGA
AATCACAATTGTTAGAGACGCGGCCTCGTGTCGTTCTTGCACGCGCTTGCATCGTCGAGCGCGCGCGT
CGCGCCGGCTGCGCGGACATTTCTAAGAACTTTGATTCAGTCGCGCACCCGCGATCTGTACAGTCGAA
TCCCAACGAGCGAGGCAGGCCGCGCCGCGCGCTTTCACGATTCGTCGCGATGGCGCGCCGGGCGGGGC
GCCTCCGCGTCGCGCATGCCCGGCCGTGCGCATTCGAGCGCATTCGAAGGCGCATGCGCGGACGCCGC
GTGAAAGGTCGGCGAGCGTGTCCCGGGTGCGGGCGGCGGAGTGGGCGCGGTGGCGCGACAATCGGCGC
GCCGTGGATCGCGGAGCGCGGCATGTGTCGCTTCTGCATGCGTCGCGCGGGCGGCGGTTTCAAGCCCC
GCCGTGTGCCGCCATGTGCCGCCCGCCGTGCGCGTCGGCTCGGCGCGATGCGCGCGGCATGGGCGCTC
GCACGCGCTATCTCGCGTTCCCGCATGAAACGCTTTCGCGCCGGTGCGGCGCTTCGAGTATGCGGCGC
CGGCTCGAT**TTTCTCAGCAAGCGAAGCGA**CATGGCACAGACCTTCATCGCGGTGAAGACCATCCTGCC
CGGCGATCTGTTTTTTCCGGCCGACGCCGCATCGGCGCTTTGCCGCGATCGCGCGCGGGAAGCCCCTG
CCGGAAGCGGCGGGCTGACGGACCGCGAGCCGACGGTGCTGCGCCACAGCGCCG**GAATTC**GAGCGGGG
CGTTTTTTTATCGGGTCCGCGTGAGCAAGCTTTTTGACGGCGATGCAAGACTGCGCGAGCGACGAATT
TCAAGCGAATGGCAAAAAATGTGAGATCAGAAATTTGAGAGTGTGACGAAATGATTGCATATCAGCGA
TTCAAGGAACTCAAAATCGAACTCGAGGCCCGTTTGGCCGACGAGCGCAGGGCGGCCAAGGAAGCCGT

FR1

FR2

Fig 3 continued

CCTCACCGAGATACGCGCGTGTGTCGACGAGTTCGGCTTCTCCCTGGACGACGTATTCCCTCCCGACG
GCGGGCACGGCAGACGAAA**GGTCCGGGCCAAATATTACGATCC**CGCCAGTGGCGCGACCTGGTCGGGC
GTCGGGCGCGAGCCGGTGTGGATTCGCGGCAAAGATCGCGAGCAATTCGAATTGCGGTGCAGCGGCGG
CGTACAGGAGTCGTCCTGAATCGATGCGCCCGCGCGTGAGGTTGAAGAAGCCGGCGATCATCCGCAGC
GTCTGCCGTCTTGCCGCAGCCCGACGGGCCGAGCAAGCGACACGATACGGCGAAATGAAAAGGCCCGC
GCGAAGAAAACACGTCATGCCGGGAGGTTGGCGATATTTTCCGCAACCGCCCGCATGGCCCGCGCGTG
GCCTCGTTTCGGGGCGCGCAACTTCGCAAGGCTCGTACGGAACCTTGCGTGCCAACGCCCGCTGTCGT
GGGTCAAGCGAGGGTGCCAAGGCTCGGTTGCCACGGCTGGCCCTCGTTCAGATCATCGGCTGTCAAGG
AAAGCGGCGCCAGATATCCGAACGTGTGTTGCCAGCCGGCTTCATCCGTGAACACCTTGAATCCGGAA
TGATATGCCCCTTCCAATTCGAATTGCTGGGAGTGATCCAGCGATCCGGGCTGGGCGTAATAGATATG
CTTGTTGGAGAAGTAATAGACGFATTTTTCTTATCTTGCGGATGAAGTCCTCCGTTGGGAAAAGATTG
GGTTCGTAAATATACTTGATTCTGTTCATCGTGGATAAATTCCACGAGGATGGAATGACGTCATACCG
ATGCCAAAGTCGCTTCATTTTTCCTGTCAACGAGCCGTCTGAATATGCGGCGAATGCGACGTTTCCGG
CGGTGGCGGCTGCAAAAGAGTGGCAGAAAACGTTTTTGAATGGGTAACCCGAGGATTGCGAGAAATAT
AGCGCGAAGGTGGATGCCAGAACGCCGCCGAGGCTATGCCCCGTCACGGTGAGGGTGGATTCTTTGGG
TAAATTTTTCAGGAAATCGGTGAGCGATGCCGAGCTATTTGAGGGGAGGGTTTGAAGGGTGACGTCCA
ATGCCTTGGCTGTCCCGTTTGCTATGAGCGCATCCGGCGGGCAATCGCTGACATATTTCGAGAACGGG
GTCAGATTTTTGACGTCATTGTCTTCAATGACTATATCGGCAAA**CTGTGAGTTCGGATTGGTTCCGCT**
GAATGCCAGGCGATAGTCCTTGTCGTTGCTGGTGTTTTTGAAGATCGCGGTTACGTTGGCGGCATATC
CGGAAAGAGGGTCGACATAGACCGCGGGACCCCAGACAATGGAAAATCTATTGTTTTTCGTCTTCGTT
GATTTGCCGAGGTAGCCGATAATTAATTTTCCCAGGATGTCCGTGGTGGGATTCTTGATATTCTGGCC
CCAGTCCATCGCGCACGATAACGTGAAATTCGTTTGTCGCTCGTCGAGAGAGTCGAATAAATCGTATG
AGTCGATCGAATCGAATGGATCGAATGAATCGGGAGGATCGAGTTCGGTGGCCATGATGGTGTTCCTT
GAATTCGGCGGGGTGGGGGCATGCAATTCAAGATATTCGCAAGCTCTTCGTGAGTGTGCTTGCCCGAA
TAAAATATATGGGTATCGCCGCGCGATCAATTACAGGGGATTACGCGCAATATAAATTCCGGATCGCC
GTCTCCCCGCTGGACGAGCCGTGGCAAAGCGGTTCCGGGAGGGATGCCGACTCGACCGTATCACGATG
TGCAGTCCGCTCGAGGGGATGACAGTCGCGCAAAAAGTATTCTTGCTCGAATGATGGCAGTGATGCAG
CTGAACGGGCCGCATCGGCGAATGAGCCGGGCTGGATCACTCGTTTTTCTCGCTCGCGAAGAACATTC
CGTCTGATGATCTATTCGTCGTCGGATTGCCGGATGGGGAAGAGAAAGAAGGCGCGGGGTCGATGGAG
TTTGTGACGCTCTGTCTCGAAGACTGGCAAGCAGCGCTGAATCGCGGGCTCGATGTGCTGTCGTTCTC
GGCCCCGCCGGCGTTCCTGACGGTGTGCCCGAATGTTTTTCAGCCACTTTTCGCACAAGCGGCCGACG
ATCTGGACGCCGTACACGGGCACGCCGGGTATGCGGTGAATCTGTCACTGCTGCGGCGCGATCCGAAC
GCGGCGTCGGAGTATTTTCTGGCTCGTCGCTATGGTCCGGGGCTCGATGTAGGTGACCCAGTTCTCGT
GGTCACTTCCCGAATGACGAGGCCGTGATCA

FR3

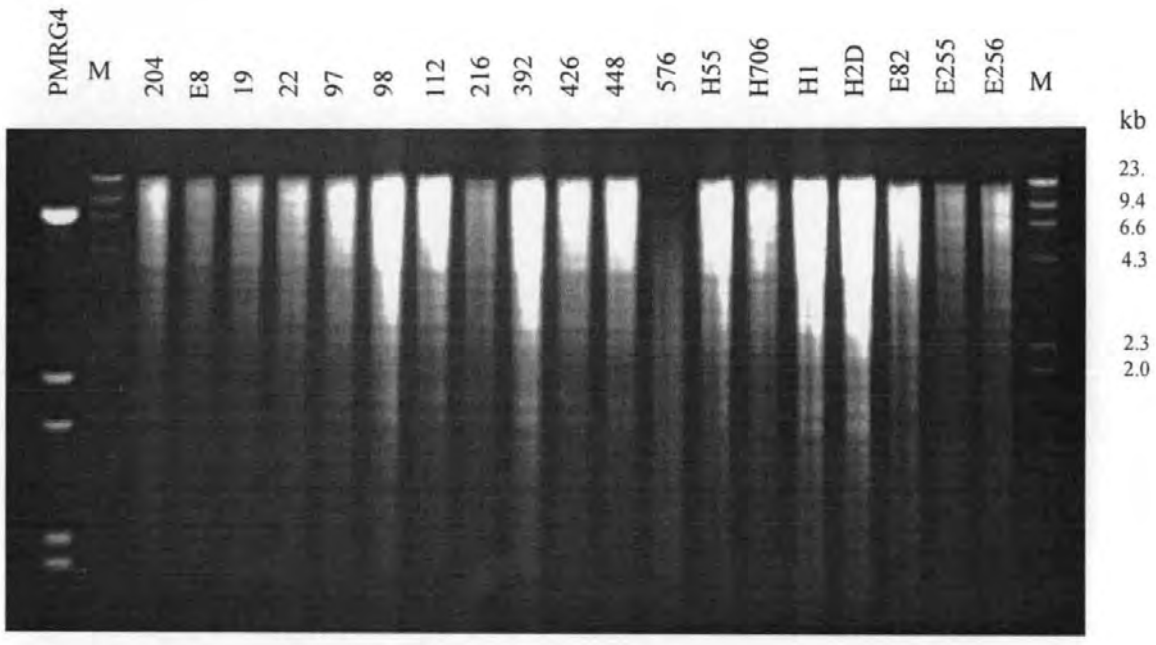


Fig 4: Agarose gel electrophoresis of *Eco*RI digested genomic DNA from a number of isolates of *B. pseudomallei* and *B. thailandensis*, as well as pMRG4.

Fig 5: Polymorphic sites within *mrgRS* downstream region of two clinical isolates of *B. pseudomallei*, 204 and 112. Atypical nucleotides are in bold and underlined.

```
204: 1    ggtccgggccaaatattacgatcccgccagtggcgcgacctggtcgggcgtcgggcgcga 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 1    ggtccgggccaaatattacgatcccgccagtggcgcgacctggtcgggcgtcgggcgcga 60

204: 61   gccggtgtggattcgcggcaaagatcgcgagcaattcgaattgcggtgcagcggcggcgt 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 61   gccggtgtggattcgcggcaaagatcgcgagcaattcgaattgcggtgcagcggcggcgt 120

204: 121  acaggagtcgtcctgaatcgatgcgcccgcgcgtgaggttgaagaagccggcgatcatcc 180
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
112: 121  acaggagtcgtcctgaatcgatgcgcccgcgcgtgaggttgacgaagccggcgatcatcc 180

204: 181  gcagcgtctgccgtcttgccgcagcccgacgggccgagcaagcgacacgatacggcgaaa 240
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
112: 181  gcagcgtctgccgtcttgccgcagcccgacgggccgagcaagcgacacgctacggcgaaa 240

204: 241  tgaaaaggcccgcgcgaagaaaacacgtcatgccgggaggttggcgatattttccgcaac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 241  tgaaaaggcccgcgcgaagaaaacacgtcatgccgggaggttggcgatattttccgcaac 300

204: 301  cgcccgcatggcccgcgcgtggcctcgtttcggggcgcgcaacttcgcaaggctcgtacg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 301  cgcccgcatggcccgcgcgtggcctcgtttcggggcgcgcaacttcgcaaggctcgtacg 360

204: 361  gaaccttgcgtgccaacgcccgctgtcgtgggtcaagcgagggtgccaaggctcggttgc 420
          || |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
112: 361  gagccttgcgtgccaacgcccgctgtcgtgggtcaagcgagggtgccaaggctcggctgc 420

204: 421  cacggctggccctcgttcagatcatcggctgtcaaggaaagcggcgccagatatccgaac 480
          ||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
112: 421  cacggctcgccctcgttcagatcattggctgtcaaggaaagcggcgccagatatccgaac 480

204: 481  gtgtgttgccagccggcttcatccgtgaacaccttgaatccggaatgatatgccccttcc 540
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
112: 481  gtgtgttgccagccagcttcatccgtgaacaccttgaatccggaatgatatgccccttcc 540

204: 541  aattcgaattgctgggagtgatccagcgatccgggctgggcgtaatagatatgcttgttg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 541  aattcgaattgctgggagtgatccagcgatccgggctgggcgtaatagatatgcttgttg 600

204: 601  gagaagtaatagacgattttcttatcttgcggatgaagtcctccgttgggaaaagattg 660
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
112: 601  gagaagtaatagacgattttcttatcttgcgaatgaagtcctccgttgggaaaagattg 660
```

Fig 5 continued

```
204: 661  ggttcgtaaatatacttgattctgttcatcgtggataaattccacgaggatggaatgacg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 661  ggttcgtaaatatacttgattctgttcatcgtggataaattccacgaggatggaatgacg 720

204: 720  tcataccgatgccaaagtcgcttcatttttcctgtcaacgagccgtctgaatatgcggcg 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
112: 721  tcataccgatgccaaagtcgcttcatttttcctgtcaacgagccgtctgaatatgcgaca 780

204: 780  aatgcgacgtttccggcggtggcggctgcaaaagagtggcagaaaacgtttttgaatggg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 781  aatgcgacgtttccggcggtggcggctgcaaaagagtggcagaaaacgtttttgaatggg 840

204: 840  taacccgaggattgcgagaaatatagcgcgaaggtggatgccagaacgccgccgaggcta 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 841  taacccgaggattgcgagaaatatagcgcgaaggtggatgccagaacgccgccgaggcta 900

204: 900  tgccccgtcacggtgagggtggattctttgggtaaatttttcaggaaatcggtgagcgat 960
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
112: 901  tgccccgtcacggtgagggtggatcctttgggtaaatttttcaggaaatcggtgagcgat 960

204: 960  gccgagctatttgaggggagggtttgaagggtgacgtccaatgccttggctgtcccgttt 1020
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
112: 961  gccgagctatttgaggggagggtttgaagggtgatgtccaatgccttggctgtcccgttt 1020

204: 1020 gctatgagcgcatccggcgggcaatcgctgacatatttcgagaacggggtcagatttttg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 1021 gctatgagcgcatccggcgggcaatcgctgacatatttcgagaacggggtcagatttttg 1080

204: 1080 acgtcattgtcttcaatgactatatcggcaaactgtgagttcggattggttccgctgaat 1140
          |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||
112: 1081 acgtcattgtcttcgatgactatatcggcaaactgtgaattcggattggttccgctgaat 1140

204: 1140 gccaggcgatagtccttgtcgttgctggtgttttgaagatcgcggttacgttggcggca 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 1141 gccaggcgatagtccttgtcgttgctggtgttttgaagatcgcggttacgttggcggca 1200

204: 1200 tatccggaaagagggtcgacatagaccgcgggaccccagacaatggaaaatctattgttt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 1201 tatccggaaagagggtcgacatagaccgcgggaccccagacaatggaaaatctattgttt 1260

204: 1260 ttcgtcttcgttgatttgccgaggtagccgataattaatttcccaggatgtccgtggtg 1320
          |||||||||||||||| ||||| |||||||||||||||| ||||||||||||||||||
112: 1261 ttcgtcttcgttgatttaccgagatagccgataattaatttacccaggatgtccgtggtg 1320

204: 1320 ggattcttgatattctggccccagtccatcgcgcacgataacgtgaaattcgtttgtcgc 1380
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 1321 ggactcttgatattctggccccagtccatcgcgcacgataacgtgaaattcgtttgtcgt 1380
```

Fig 5 continued

```
204: 1380 tcgtcgagagagtcgaataaatcgtatgagtcgatcgaatcgaatggatcgaatgaatcg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
112: 1381 tcgtcgagagagtcgaataaatcgtatgagtcgatcgaatcgaatggatcgaatgaatcg 1440

204: 1440 ggaggatcgagttcggtggccatgatggtgtt 1472
          ||||||||||||||||||||||||||||||||
112: 1441 ggaggatcgagttcggtggccatgatggtgtt 1472
```

APPENDIX VI

DATA RELEVANT TO CHAPTER 6

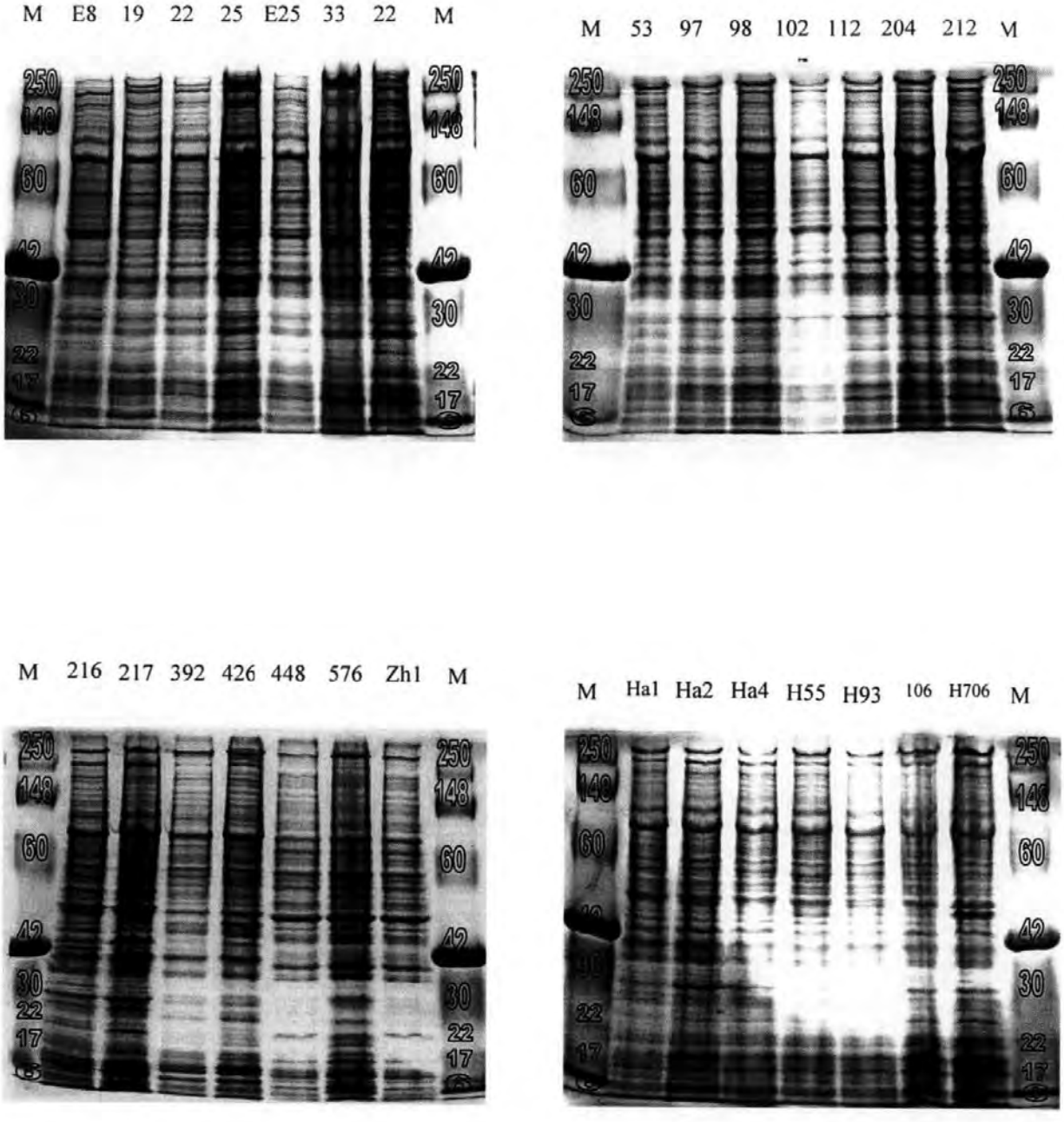


Fig 1: A comparison of the whole cell proteins from different *B. pseudomallei* isolates. M indicates molecular weight markers (kDa).

Table 1: Homology of DTNVDLINC peptide used for the production of anti-MrgR antibody to sequences in database. Identical amino acid residues are in red bold.

Sequence	Description	Homologous sites
SW: FMK1- *E. coli*	P04740 Fimbrillin precursor	**DTNVDLINC**
SW: MT21- *S. pneumoniae*	P04043 Methylase	**DTNVDLINC**
SW: YHEV- *E. coli*	P56622 Hypothetical protein	**DTNVDLINC**
SW: PAPA- *E. coli*	Pap fimbrial protein	**DTNVDLINC**
SW: YRBE- *E. coli*	Hypothetical protein	**DTNVDLINC**
SW: PERA- *E. coli*	Transcriptional activator	**DTNVDLINC**
SW: FER- *C. butyricum*	Ferrodoxin	**DTNVDLINC**
SW: YJEB- *V. parahaemolyticus*	Hypothetical protein	**DTNVDLINC**
SW: MALD- *S. pneumoniae*	Maltodextrin transport	**DTNVDLINC**
SW: HSTO- *V. cholerae*	Enterotoxin	**DTNVDLINC**
SW: HST- *V. cholerae*	Heat shock protein	**DTNVDLINC**
SW: DMA7- *E. coli*	DNA adenine methylase	**DTNVDLINC**
SW: DMA- *H. influenzae*	DNA adenine methylase	**DTNVDLINC**
SW: SP4A- *B. subtilis*	sporulation protein	**DTNVDLINC**
SW: YQJM- *B. subtilis*	NADH dependent flavin	**DTNVDLINC**
SW: YTFM- *H. influenzae*	Hypothetical protein	**DTNVDLINC**
SW: ALGE- *P. aeruginosa*	Alignate production protein	**DTNVDLINC**
SW: G3P- *P. aeruginosa*	Glyceraldhyde 3-phosphate	**DTNVDLINC**
SW: YDGJ- *E. coli*	Hypothetical oxidoreductase	**DTNVDLINC**
SW: DNIV- *E. coli*	DNA integrase	**DTNVDLINC**
SW: EUTH- *E. coli*	Ethanolamine utilizing protein	**DTNVDLINC**
SW: MALD- *S. pneumoniae*	Maltodextrin transport	**DTNVDLINC**
SW:Y390- *M. genitalis*	Hypothetical ATP-binding	**DTNVDLINC**
SW: YJDB- *S. typhimurium*	Hypothetical protein	**DTNVDLINC**
SW:MT21- *S. pneumoniae*	Methylase	**DTNVDLINC**

Table 2: Homology of RKFYSLESN peptide used for the production of anti-MrgS antibody to sequences in database. Identical amino acid residues are in red bold.

Sequence	Description	Homologous sites
SW: YCBK- *H. influenzae*	Hypothetical protein	RKFYSLESN
SW: RL27- *E. coli*	50S Ribosomal protein	RKFYSLESN
SW: YWFA- *B. subtilis*	Hypothetical protein	RKFYSLESN
SW: CLPR- *H. influenzae*	Glycerol 3-phosphate regulator	RKFYSLESN
SW: YDGT- *E. coli*	Hypothetical protein	RKFYSLESN
SW: RFAP- *E. coli*	Lipopolysaccharide core	RKFYSLESN
SW: BINR- *S. aureus*	DNA invertase	RKFYSLESN
SW:DDLB- *E. coli*	D-alanine ligase	RKFYSLESN
SW: YPB3- *E. coli*	Hypothetical protein	RKFYSLESN
SW: Y074- *H. influenzae*	Hypothetical protein	RKFYSLESN
SW: RIBR- *B. subtilis*	RibR regulatory protein	RKFYSLESN
SW: SP5M- *B. subtilis*	Sporulatoin protein	RKFYSLESN
SW: FMF1- *E. coli*	Fimbrial protein	RKFYSLESN
SW: LEP3- *V. cholerae*	Prepilin-like protein	RKFYSLESN
SW:YQEL- *B. subtilis*	Hypothetical protein	RKFYSLESN
SW: PARP- *E. coli*	DNA replication protein	RKFYSLESN
SW:TEHB- *H. influenzae*	Resistance protein	RKFYSLESN
SW: SAPC- *E. coli*	Peptide permease	RKFYSLESN
SW:SAPC- *E. coli*	Peptide transport protein	RKFYSLESN
SW: YBDJ- *K. pneumoniae*	Hypothetical protein	RKFYSLESN
SW:YAAC- *B. subtilis*	Hypothetical protein	RKFYSLESN

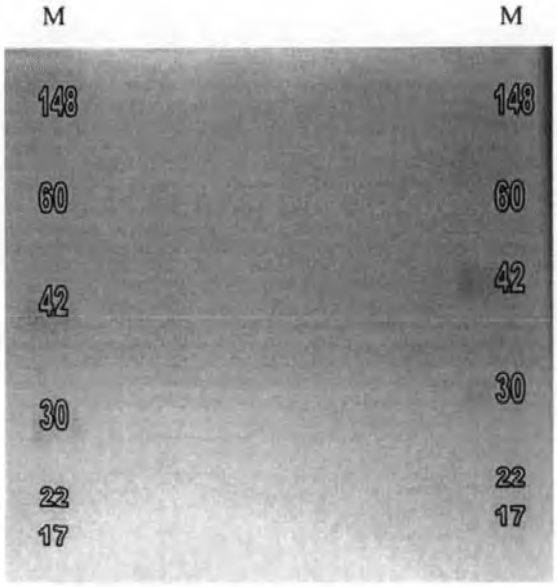


Fig 2: Specificity of anti-MrgR antibody. Failure of recognition of MBP-MrgR fusion after blocking with specific peptide. M indicates protein molecular weight markers (kDa).

Table 3: Examples of the homology of DTNVDLINC and RKFYSLESN peptides to potential proteins deduced from the genome sequence of *B. pseudomallei* isolate K96243. Identical amino acid residues are in red bold.

Homologous sites	
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN
DTNVDLINC	RKFYSLESN

APPENDIX VII

CONSTRUCTION OF A GENE MUTANT

Creating a *mrgR* deficient mutant of *B. pseudomallei*

Further investigation of the role of MrgR in the adaptive responses of *B. pseudomallei* would be greatly facilitated by the construction of a mutant deficient in the expression of the *mrgR* gene. The strategy used to create such a mutant is illustrated in **Fig 1**.

pCVD442 suicide vector

The pCVD442 plasmid vector (Donnenberg and Kaper, 1991) (**Fig 1**) is based on R6K, which means it will only replicate in permissive *E. coli* strains in which the π protein encoded by the *pir* gene is provided in trans (Kolter *et al.*, 1978). In most cases this is achieved by using *E. coli* strains that are lysogenic for a λ *pir* transducing phage (Simon *et al.*, 1983), e.g. *E. coli* SY327. In addition, pCVD442 contains the *mob* region of RP4 such that strains with the RP4 *tra* functions inserted on the chromosome are able to mobilise the plasmids into the target bacteria (Simon *et al.*, 1983). The suicide vector is used as the basis to develop positive selection that permits marker exchange mutagenesis and introduction of the reporter gene fusion onto the chromosome. The basis of the positive selection is the presence of a cloned *Bacillus subtilis sacB* gene; this encodes levensucrase, which catalyses the hydrolysis of sucrose as well as the synthesis of levans (Gay *et al.*, 1983). Expression of the *sacB* gene is toxic for Gram-negative bacteria when they are growing in the presence of 5% sucrose. This greatly facilitates the introduction of defined deletion mutations back into the chromosome of the bacteria (Simon *et al.*, 1991). Once a single homologous recombination event has been achieved, it is possible to directly select for the second recombination event by growing the strains in the presence of sucrose. Only those colonies that have lost the vector sequences will be able to grow on this medium. The pCVD442 suicide vector was kindly provided by Dr G. Dougan, Imperial college, London, UK.

Disruption of *mrgR* by the insertion of a gene encoding kanamycin resistance

It was decided that the most suitable antibiotic to use for the mutagenesis of *B. pseudomallei* would be kanamycin. The kanamycin resistance gene (*kanR*) was cleaved from pGD103 (**Fig 1**) using *Pvu*II (section 3.4.5.1), which cleaves sites located 5' and 3' to the gene. The gene fragment was checked on 0.8% agarose gel (section 3.4.6) and purified with the Prep-A-Gene kit (section 3.4.4.3). The resultant blunt-ended DNA fragment was ligated (section 3.5.2) into the *Pml*I restriction site in pMRG2 (section 4.3.7) located at nucleotide 612 in the *mrgR* coding sequence. Recombinants were selected on LB agar containing kanamycin (25 μg/ml) and ampicillin (100 μg/ml). One clone, pMMUT1, was selected for further analysis.

Insertion of the disrupted *mrgR* gene into pCVD442

The *mrgR*::*kan*::*mrgR* fragment was obtained by *Eco*RV digestion of the pMMUT1, which produced a fragment carrying the disrupted *mrgR* gene plus a flanking region of 532 bp upstream and 459 bp downstream sequences. This fragment was purified then ligated into the unique *Sma*I site of pCVD442 before transforming *E. coli* SY327 (section 3.5.4.1). Recombinants were selected on LB agar containing both ampicillin (100 μg/ml) and kanamycin (25 μg/ml). The construction, in the correct orientation, was confirmed by restriction site mapping of the recombinant plasmid DNA, and a positive clone, referred to as pMMUT2, was chosen for further studies.

Fig 1: Strategy for the construction of *mrgR* mutant in pCVD442 suicide vector.

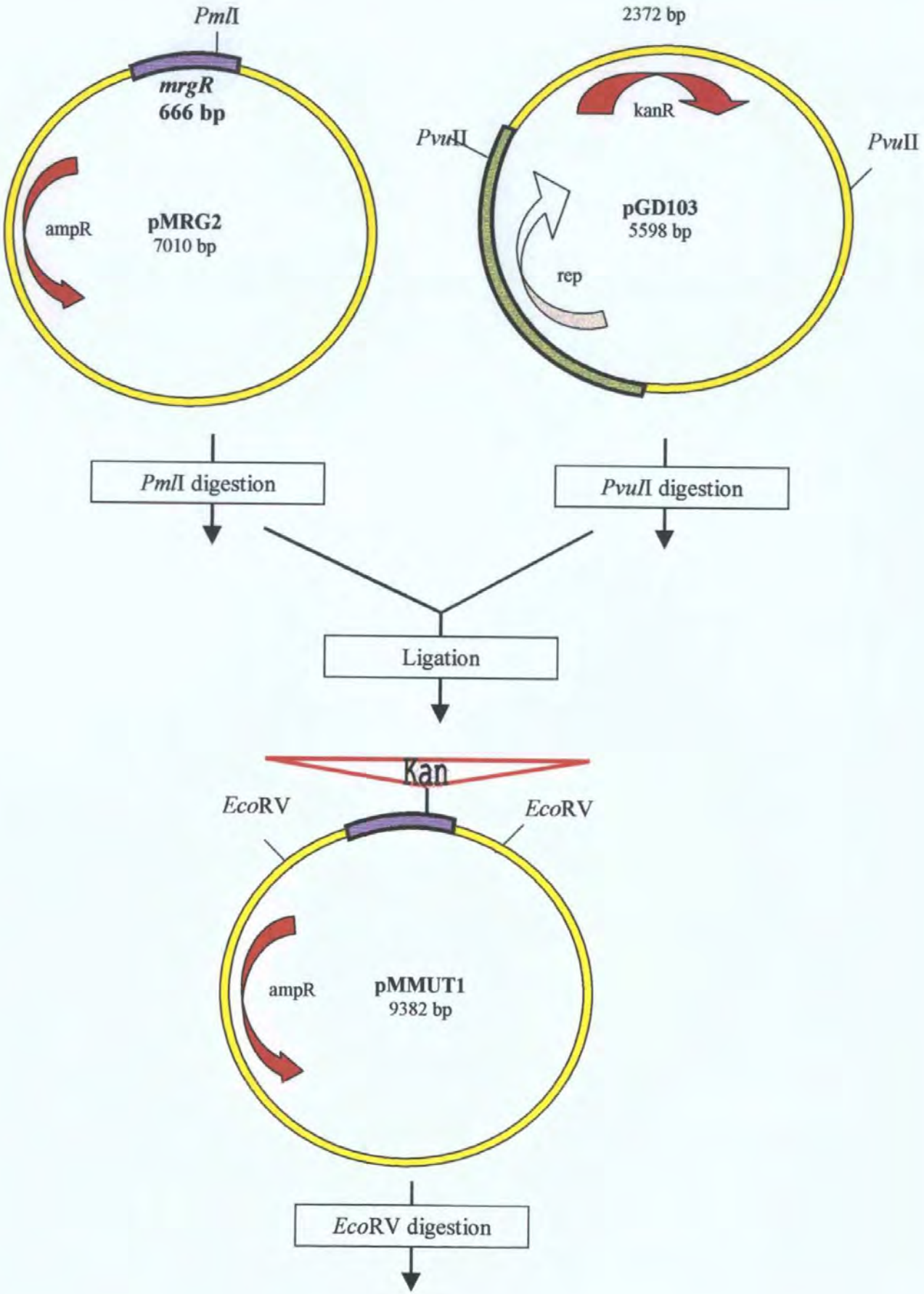
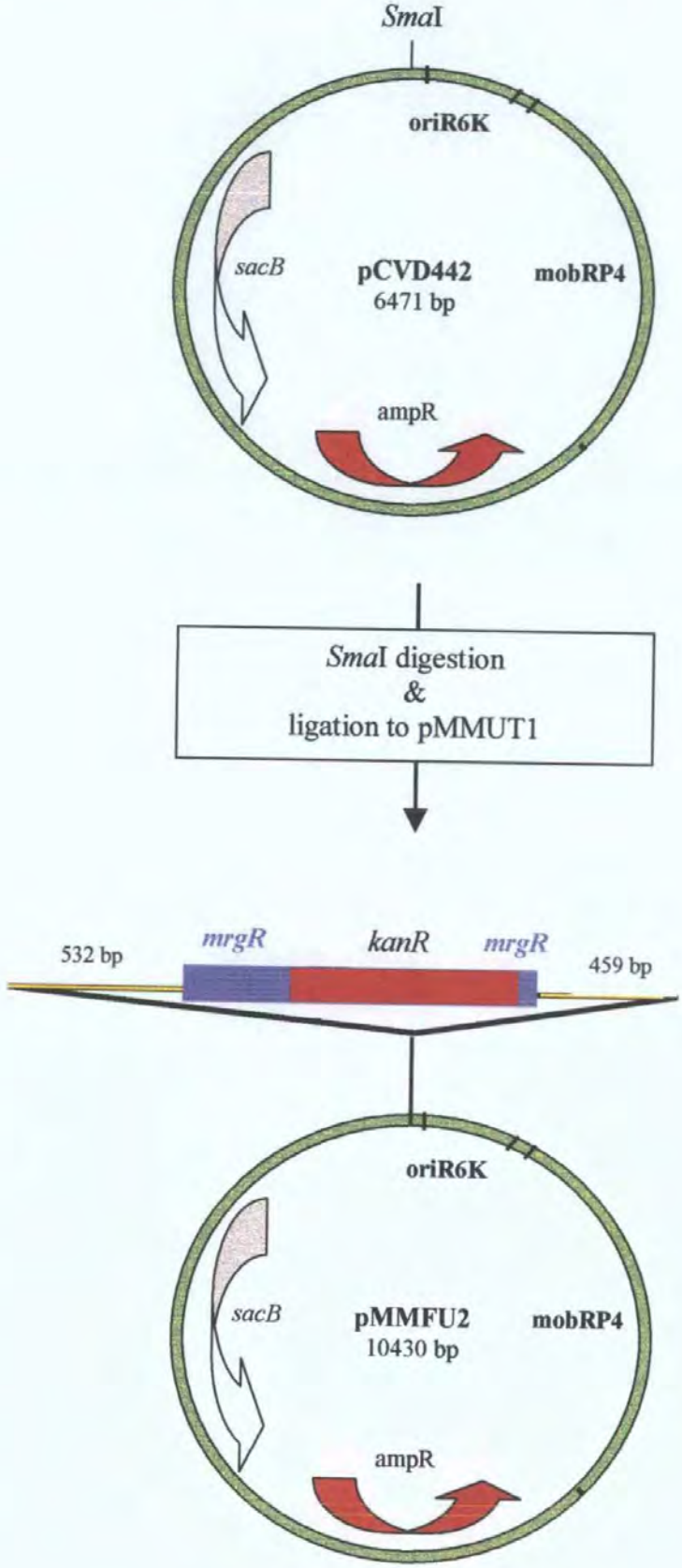


Fig 1 continued



REFERENCES

Achana, V., K. Silpapojakul, W. Thinita and S. Kalnaowakul. 1985. Acute *Pseudomonas pseudomallei* and septicaemia following aspiration of contaminated water: a case report. *Southeast Asian J. Trop. Med. Public Health* **16** 500-504.

Ahmed, K. H., D. R Enciso, H. Masaki, M. Tao, A. Omori, P. Tharavichikul and T. Nagatake. 1999. Attachment of *Burkholderia pseudomallei* to pharyngeal epithelial cells: a highly pathogenic bacteria with low attachment ability. *Am. J. Trop. Med. Hyg.* **60** 90-93.

Aiba, H., F. Nakasai, S. Mizushima and T. Mizuno. 1989. Phosphorylation of a bacterial activator protein, OmpR, by a protein-kinase, EnvZ, results in stimulation of its DNA-binding ability. *J. Biochem.* **106** 5-7.

Aiba, H., M. Nagaya and T. Mizuno. 1993. Sensor and regulator proteins from the cyanbacterium *Synechococcus* species PCC7942 that belong to the bacterial signal-transduction protein families: implication in the adaptive response to phosphate limitation. *Mol. Microbiol.* **8** 81-91.

Alex, L. A. and M. I. Simon. 1994. Protein histidine kinase and signal transduction in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* **10** 133-138.

Alex, L. A., K. A. Borkovich and M. I. Simon. 1996. Hyphal development in *Neurospora crassa*: involvement of a two-component histidine kinase. *Proc. Natl. Acad. Sci. USA* **93** 3416-3421.

Allen, P., C. A. Hart and J. R. Saunders. 1987. Isolation from *Klebsiella* and characterization of two *rcs* genes that activate colanic acid capsular biosynthesis in *Escherichia coli*. *J. Gen. Microbiol.* **133** 331-340.

Alpuche-Aranda, C. M., J. A. Swanson, W. P. Loomis and S. I. Miller. 1992. *Salmonella typhimurium* activates virulence gene transcription within acidified macrophage phagosomes. *Proc. Natl. Acad. Sci. USA* **89** 10079-10089.

Altschul, S., F. W. Gish, E. F. Myers and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215** 403-410.

Anand, G. S., P. N. Goudreau and A. M. Stock. 1998. Activation of methylesterase CheB: evidence of a dual role for the regulatory domain. *Biochemistry* **37** 14038-14047.

Anonymous. 1995. Melioidosis in Singapore. *Epidemiol. News Bull. (Singapore)* **21** 69-72.

Anuntagool, N., P. Rugdech and S. Sirisinha. 1993. Identification of specific antigens of *Pseudomonas pseudomallei* and evaluation of their efficacious for diagnosis of melioidosis. *J. Clin. Microbiol.* **31** 1232-1236.

Anuntagool, N., T. Panichakul, P. Aramsri and S. Sirisinha. 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.

Appleby, J. L., J. S. Parkinson and R. B. Bourret. 1996. Signal transduction via the multi-step phosphorelay: not necessarily a road less travelled. *Cell* **86** 845-848.

Arricau, N., D. Hermant, H. Waxin, C. Ecobichon, P. S. Duffey and M. Y. Popoff. 1998. The RcsB-RcsC regulatory system of *Salmonella typhi* differentially modulates the expression of invasion proteins, flagellin and Vi antigen in response to osmolarity. *Mol. Microbiol.* **29** 835-850.

Asche, V. 1991. Melioidosis: a disease of all organs. *Today's Life Science.* **June** 34-40.

Ashdown, L. R. 1979. An improved screening technique for isolation of *Pseudomonas pseudomallei* from clinical specimens. *Pathology* **11** 293-297.

Ashdown, L. R. 1987. Indirect haemagglutination test for melioidosis. *Med. J. Aust.* **147** 314-316.

Ashdown, L. R. and J. M. Koehler. 1990. Production of haemolysin and other extracellular enzymes by clinical isolates of *Pseudomonas pseudomallei. J. Clin. Microbiol.* **28** 2331-2334.

Ashdown, L. R. and R. W. Guard. 1984. The prevalence of human melioidosis in north Queensland. *Am. J. Trop. Med. Hyg.* **33** 474-478.

Atassi, M. Z. 1986. Preparation of monoclonal antibodies to preselected protein regions. *Methods Enzymol.* **121** 69-95.

Atassi, M. Z. and C. R. Young. 1985. Discovery and implications of the immunogenicity of free small synthetic peptides- powerful tools for manipulating the immune system and for production of antibodies and T-cells of preselected submolecular specificities. *Crit. Rev. Immunol.* **5** 387-409.

Atthasampunna, P., H. E. Noyes, P. E. Winter, D. L. Brendlinger, P. Busapathamrong, P. Pongpradit, P. Sukomol, P. Tanticharoenyos, S. Taranatti, S. Wongsathuaythong, W. L. Wooding, T. L. Hensarling, Y. Kasemsanta, W. Panlom, Y. Raengpradub and T. Tamaarr. 1969. Annual progress report of SEATO medical research study on melioidosis. Bangkok.

Baikalov, I., I. Schroder, M. Kaczor-Grzeskowiak, K. Grzeskowiak, R. P. Gunsalus and R. E. Dickerson. 1996. Structure of the *Escherichia coli* response regulator NarL. *Biochemistry* **35** 11053-11060.

Bally, M., A. Filloux, M. Akrim, G. Ball, A. Lazdunski and J. Tommassen. 1992. Protein secretion in *Pseudomonas aeruginosa*: characterisation of seven *xcp* genes and processing of secretory apparatus components by prepilin peptidase. *Mol. Microbiol.* **6** 1121-1131.

Barer, M., L. T. Gribbon, C. R. Harwood and C. E. Nwoguh. 1993. The viable but non-culturable hypothesis and medical bacteriology. *Rev. Med. Microbiol* **4** 183-191.

Barnes, W. M. and E. Tuley. 1983. DNA sequence changes of mutations in the histidine operon control region that decrease attenuation. *J. Mol. Biol.* **165** 443-459.

Barrett, J. F. and J. A. Hoch. 1998. Two-component signal transduction as a target for microbial anti-infective therapy. *Antimicrob. Agents Chemother.* **42** 1529-1536.

Barrett, J. F., R. M. Goldschmidt, L. E. Lawrence, B. Foleno and R. Chan. 1998. *Proc. Natl. Acad. Sci. USA* **95** 5317-5322.

Bayles, K. W. 1993. The use of degenerate, sensor gene-specific, oligodeoxyribonucleotide primers to amplify DNA fragments from *Staphylococcus aureus*. *Gene* **123** 99-103.

Behr, M. A., M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolink, S. Rane and P. M. Small. 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284** 1520-1523.

Bernhard, F., K. K. Potter and D. L. Coplin. 1990. The *rcsA* gene from *Erwinia amylovora*: identification, nucleotide sequence, and regulation of exopolysaccharide biosynthesis. *Mol. Plant-Microbe Interact.* **3** 429-437.

Berswill, S. and K. Geider. 1997. Characterization of the *rcsB* gene from *Erwinia amylovora* and its influence on exopolysaccharide synthesis and virulence of the fire blight pathogen. *J. Bacteriol.* **179** 1354-1361.

Bessen, D. E. and S. K. Hollingshead. 1994. Allelic polymorphism of *emm* loci provides evidence for horizontal gene spread in group A streptococci. *Proc. Natl. Acad. Sci. USA* **91** 3280-3284.

Bianciotto, V., C. Bandi, D. Minerdi, M. Sironi, H. V. Tichy and P. Bonfante. 1996. An obligately endosymbiotic mycorrhizal fungus itself harbors obligately intracellular bacteria. *Appl. Environ. Microbiol.* **62** 3005-3010.

Birnboim, H. C and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7** 1513-1523.

Bishop, R. E. and J. H. Weiner. 1993. Overproduction, solubilisation, purification and DNA-binding properties of AmpR from *Citrobacter freundii*. *Eur. J. Biochem.* **213** 405-412.

Blat, Y. and M. Eisenbach. 1994. Phosphorylation-dependent binding of the chemotaxis signal molecule CheY to its phosphatase, CheZ. *Biochemistry* **33** 902-906.

Bliska, J. B., J. E. Galan and S. Falkow. 1993. Signal transduction in the mammalian cell during bacterial attachment and entry. *Cell* **73** 903-920.

Bogdanove, A. J., Z. M. Wei, L. Zhao and S. V. Beer. 1996. *Erwinia amylovora* secretes hairpin via a type III pathway and contains a homolog of *yop* of *Yersinia* spp. *J. Bacteriol.* **178** 1720-1730.

Bokman, A. H., H. B. Levine and M. Lusby. 1957. Glucose catabolism in *Malleomyces pseudomallei*. *J. Bacteriol.* **73** 649-654.

Bonfante, P., R. Balestrini and K. Mendgen. 1994. Storage and secretion processes in the spore of *Gigaspora margarita* Becker & Hall as revealed by high-pressure freezing and freeze substitution. *New Phytol.* **128** 93-101.

Boonsawat, W., P. Boonma and T. Tangdajahiran. 1990. Community acquired pneumonia in adults at Srinagarind Hospital. *J. Med. Assoc. Thailand.* **73** 345-352.

Bourret, R. B., J. F. Hess and M. I. Simon. 1990. Conserved aspartate residues and phosphorylation in signal transduction by the chemotaxis protein CheY. *Genetics* **87** 41-45.

Bourret, R. B., K. A. Borkovich and M. I. Simon. 1991. Signal transduction pathways involving protein phosphorylation in prokaryotes. *Annu. Rev. Biochem.* **60** 401-441.

Boyd, E. F. and D. L. Hartl. 1999. Analysis of the type 1 pilin gene cluster *fim* in *Salmonella*: its distinct evolutionary histories in the 5' and 3' regions. *J. Bacteriol.* **181** 1301-1308.

Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Anal. Biochem.* **72** 248-252.

Brenner, S. 1988. The molecular evolution of genes and proteins: a tale of serines. *Nature* **334** 528-530.

Brett, P. J. and D. E. Woods. 1996. Structural and immunological characterization of *Burkholderia pseudomallei* O-polysaccharide-flagellin protein conjugates. *Infect. Immun.* **64** 2824-2828.

Brett, P. J. and D. E. Woods. 2000. Pathogenesis of and immunity to melioidosis. *Acta Tropica* **74** 201-210.

Brett, P. J., D. C. Mah and D. E. Woods. 1994. Isolation and characterization of *Pseudomonas pseudomallei* flagellin proteins. *Infect. Immun.* **62** 1924-1919.

Brett, P. J., D. Deshazer and D. E. Woods. 1997. Characterization of *Burkholderia pseudomallei* and *Burkholderia pseudomallei*-like strains. *Epidemiol. Infect.* **118** 137-148.

Brett, P. J., D. Deshazer and D. E. Woods. 1998. *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. *Int. J. Syst. Bacteriol.* **48** 317-320.

Brill, J. A., C. Q. Walshe and S. Gottesman. 1988. Fine - structure mapping and identification of two regulators of capsule synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **170** 2599-2611.

Brindle, C. S. and S. T. Cowan. 1951. Flagellation and taxonomy of Whitmore's bacillus. *J. Pathol. Bact.* **63** 571-575.

Brissette, R. E., K. L. Tsung M. Inouye. 1991. Intramolecular second-site revertants to the phosphorylation site mutation in OmpR, a kinase-dependent transcriptional activator in *Escherichia coli. J. Bacteriol.* **173** 3749-3755.

Brook, M. D., B. Currie and P. M. Desmarchelier. 1997. Isolation and identification of *Burkholderia pseudomallei* from soil using selective culture techniques and the polymerase chain reaction. *J. Appl. Microbiol.* **82** 589-596.

Brown, A. E., D. A. B. Dance, Y. Suputtamonkol, W. Chaowagul, S. Kongcharoen, H. K. Webster and N. J. White. 1991. Immune cell activation in melioidosis: Induced serum level of interferon-γ and soluble interleukin-2 receptors without change in soluble CD8 protein. *J. Infect. Dis.* **163** 1145-1148.

Brown, M. and R. N. Thin. 1986. Melioidosis. In: A.I.Braude, C. E. Davis and J. Fierer (Eds.) *Infectious Diseases and Medical Microbiology*. 2nd Edn, pp 802-804. 1986. Philadelphia, W. B. Saunders.

Brown, N. F. and I. R. Beacham. 2000. Cloning and analysis of genomic differences unique to *Burkholderia pseudomallei* by comparison with *B. thailandensis. J. Med. Microbiol.* **49** 993-1001.

Brown, T. A. 1999. Genomes. In: F.Kingston (Ed). 1st Edn. Oxford, BIOS Scientific Publishers Ltd.

Bryan, L. E., S. Wong, D. E. Woods, D. A. B. Dance and W. Chaowagul. 1994. Passive protection of diabetic rats with antisera specific for the polysaccharide portion of the lipopolysaccharide isolated from *Pseudomonas pseudomallei. Can. J. Infect. Dis.* **5** 170-178.

Bullen, J. J. and E. Griffiths. 1987. Iron and Infection: Molecular, Physiological and Clinical Aspects. J.Wiley, and Sons. New York.

Bustamante, C. D., J. P. Townsend and D. L. Hartl. 2000. The nature of purifying selection: structural constraints on amino acid variation within *Escherichia coli* and *Salmonella enterica. Mol. Biol. Evol.* **17** 301-308.

Carsiotis, M. D., L. Weinstein, H. Karch, I. A. Holder and A. D. Q'Brien. 1984. The flagella of *Salmonella typhimurium* are a virulence factor in infected C57BL/6J mice. *Infect. Immun.* **46** 814-818.

Casadevall, A. and L. A. Pirofski. 1999. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. *Infect. Immun.* **67** 3703-3713.

Chambon, L. 1955. Isolement du bacilli de Whitmore a partir du milieu exterieur. *Ann. Inst. Pasteur* **89** 229-235.

Chambon, L. and J. Fournier. 1956. Constitution antigenique de *Malleomyces pseudo-mallei.* II. Demonstration de l'existence des antigenes M, K, et O et etudes de leurs proprietes immunologiques. *Ann. Inst. Pasteur* **91** 472-485.

Chambon, L. and J. Fournier. 1956. Constitution antigenique de *Malleomyces pseudo-mallei.* 1. Caracteres morphologiques, culturaux, biochimiques et variations de type immunologique. *Ann. Inst. Pasteur* **91** 335-362.

Chaowagul, W. N., J. White, D. A. B. Dance, Y. Wattanagoon, P. Naigowit, T. M. E. Davis, S. Looaresuwan and N. Pitakwatchara. 1989. Melioidosis: a major cause of community-acquired septicaemia in northeastern Thailand. *J. Infec. Dis.* **159** 890-899.

Chaowagul, W., Y. Supputtamongkol, D. A. B. Dance, A. Rajchanuvong, J. Pattara-arechai and N. J. White. 1993. Relapse in melioidosis: incidence and risk factors. *J. Infect. Dis.* **168** 1181-1185.

Chodimella, U. W., L. Hoppes, S. Whalen, A. J. Ognibene and G. W. Rutecki. 1997. Septicaemia and suppuration in a Vietnam veteran. *Hospital Practice* **15** 219-221.

Choy, J. L., M. Mayo, A. Janmaat and B. J. Currie. 2000. Animal melioidosis in Australia. *Acta Tropica* **74** 153-158.

Clarke, L. and J. Carbon. 1976. A colony bank containing synthetic ColE1 hybrid plasmids representative of the entire *Escherichia coli* genome. *Cell* **9** 91-99.

Clavel, T. J., C. Lazzaroni, A. Vianney and R. Portalier. 1996. Expression of the *tolQRA* genes of *Escherichia coli* K-12 is controlled by the RcsC sensor protein involved in capsule synthesis. *Mol. Microbiol.* **19** 19-25.

Cohan, F. M. 1994. Genetic exchange and evolutionary divergence in prokaryotes. *Trends Ecol. Evol.* **9** 175-180.

Coleman, M., R. Pearce, E. Hitchin, F. Busfield, J. W. Mansfield and I. S. Roberts. 1990. Molecular cloning, expression and nucleotide sequence of the rcsA gene of *Erwinia amylovora*, encoding a positive regulator of capsule expression: evidence for a family of related capsule activator proteins. *J. Gen. Microbiol.* **136** 1799-1806.

Congakpoon, C., W. Susaengrat, V. Yaisoongren and P. Mongkoltananant. 1989. Melioidosis. A laboratory study. In: S.Punyagupta, T. Sirisanthana and B. Stapatayavong (Eds). Melioidosis. pp 158-169. Bangkok, Thailand, Bangkok Med. Pub.

Cook, D., E. Barlow and L. Sequeira. 1989. Genetic diversity of *Pseudomonas solanacearum*: detection of restriction fragment length polymorphisms with DNA probes that specify virulence and the hypersensitive response. *Mol. Plant-Microbe Interact.* **2** 113-121.

Coote, J. G. 1991. Antigenic switching and pathogenicity: environmental effects on virulence gene expression in *Bordetella pertussis*. *J. Gen. Microbiol.* **137** 2493-2503.

Coplin, D. L., K. Poetter and D. R. Majerczak. 1990. Regulation of exopolysaccharide synthesis in *Erwinia stewartii* by *rcsB* and *rcsC*. *Phytophathology* **81** 1220.

Corbell, N. and J. E. Loper. 1995. A global regulator of secondary metabolite production in *Pseudomonas fluorescens*. *J. Bacteriol* **177** 6230-6236.

Cornelis, G. R. and F. V. Gijsegem. 2000. Assembly and function of type III secretory systems. *Annu. Rev. Microbiol.* **54** 735-774.

Costerton, J. W., K. J. Cheng, G. G. Gessey, T. I. Ladd, J. C. Nickel and M. Dasgupta. 1987. Bacterial biofilms in nature and disease. *Annu. Rev. Microbiol.* **41** 435-464.

Cotter, P. A. and V. J. DiRita. 2000. Bacterial virulence gene regulation: an evolutionary perspective. *Annu. Rev. Microbiol.* **54** 519-565.

Cottew, G. S. 1950. Melioidosis in sheep in Queensland. A description of the casual organism. *Austral. J. Exp. Biol. Med. Sci.* **28** 677-683.

Cottew, G. S. 1952. Melioidosis. *Aust. Vet. J.* **31** 155-159.

Crichton, R. R. and M. Charloteaux-Wauters. 1987. Iron transport and storage. *Eur. J. Biochem.* **164** 485-506.

Crotty, J. M., A. F. Bromwitch and J. V. Quinn. 1963. Melioidosis in the Northern Territory: a report of two cases. *Med. J. Aust.* **1** 274-275.

Cserzo, M., E. Wallin, I. Simon, G. von Heijne and A. Elofsson. 1997. Prediction of transmembrane α-helices in prokaryotic membrane proteins: the Dense Alignment Surface method. *Protein Eng.* **10** 673-676.

Currie, B., D. Howard, V. T. Nguyen, K. Withnall and A. Merianos. 1993. The 1990-1991 outbreak of melioidosis in the Northern Territory of Australia: clinical aspects. *Southeast Asian J. Trop. Med. Public Health* **24** 436-443.

Currie, B., H. Smithvaughan and C. Golledge. 1994. *Pseudomonas pseudomallei* isolates collected over 25 years from a nontropical endemic focus show clonality of the basis of ribotyping. *Epidemiol. Infect.* **113** 307-312.

Currie, B., J. D. A. Fisher, D. M. Howard, J. N. C. Burrow, S. Selvanayagam, P. L. Snelling, N. M. Anstey and M. J. Mayo. 2000a. The epidemology of melioidosis in Australia and Papua New Guinea *Acta Tropica* **74** 121-127.

Currie, B., J. D. A. Fisher, D. M. Howard and J. N. C. Burrow. 2000b. Neurological melioidosis. *Acta Tropica* **74** 145-151.

Currie, B., J. D. A. Fisher, N. M. Anstey and S. P. Jacups. 2000c. Melioidosis: acute and chronic disease, relapse and re-activation. *Trans. R. Soc. Trop. Med. Hyg.* **94** 301-304.

Dance, D. A. B. 1989. Identification of *Pseudomonas pseudomallei* in clinical practice: use of simple screening tests and API 20NE. *J. Clin. Pathol.* **42** 645-648.

Dance, D. A. B. 1990. Melioidosis. *Rev. Med. Microbiol.* **1** 143-150.

Dance, D. A. B. 1991. Melioidosis: the tip of the Iceberg ? *Clin. Microbiol. Rev.* **4** 52-60.

Dance, D. A. B. 2000a. Melioidosis as an emerging global problem. *Acta Tropica* **74** 115-119.

Dance, D. A. B. 2000b. Ecology of *Burkholderia pseudomallei* and the interactions between environmental *Burkholderia* spp. and human-animal hosts. *Acta Tropica* **74** 159-168.

Dance, D. A. B. 2000c. Melioidosis. In: N.J.Beeching, C. A. Hart and B. I. Duerden (Eds). Tropical and exotic infections. *J. Med. Microbiol.* 49, 5-27.

Dance, D. A. B., T. M. E. Davis and Y. Wattanagoon. 1989. Acute suppurative parotitis caused by *Pseudomonas pseudomallei. J. Infect. Dis.* 159, 654-660.

Dance, D. A. B., V. Wuthiekanun, P. Naigowit and N. J. White. 1989. Identification of *Pseudomonas pseudomallei* in clinical practice: use of simple screening test and API 20NE. *J. Clin. Pathol.* **42** 645-648.

Dannenberg, A. M. and E. M. Scott. 1958. Melioidosis: Pathogenesis and immunity in mice and hamsters. I. Studies with virulent strains of *Malleomyces pseudomallei. J. Exp. Med.* **107** 153.

Das, A. C. Yanofsky. 1984. A ribosome binding site is necessary for efficient expression of the distal gene of a translationally-coupled gene pair. *Nucleic Acids Res.* **12** 4757-4768.

Dejsirilert, S., E. Kondo, D. Chiewslip and K. Kanai. 1991. Growth and survival of *Pseudomonas pseudomallei* in acidic environments. *Jpn. J. Med. Sci. Biol.* **44** 63-74.

Dejsirilert, S., R. Butraporn, D. Chiewsilp, E. Kondo and K. Kanai. 1989. High activity of acid phosphatase of *Pseudomonas pseudomallei* as a possible attribute relating to its pathogenicity. *Jpn. J. Med. Sci. Biol.* **42** 39-49.

Deretic, V., C. D. Mohr and D. W. Martin. 1991. Mucoid *Pseudomonas aeruginosa* in cyctic fibrosis: signal transduction and histone-like elements in the regulation of bacterial virulence. *Mol. Microbiol.* **5** 1577-1583.

Derewenda, Z. S. and C. Cambillau. 1991. Effects of gene mutations in lipoprotein and hepatic lipases as interpreted by a molecular model of the pancreatic lipase. *J. Biol. Chem.* **34** 23112-23119.

Deshazer, D. and D. E. Woods. 1999. Pathogenesis of Melioidosis: use of Tn5-OT182 to study the molecular basis of *Burkholderia pseudomallei* virulence. *J. Infect. Dis. Antimicrob. Agents* 91-96.

Deshazer, D., P. J. Brett, R. Carylon and D. E. Woods. 1997. Mutagenesis of *Burkholderia pseudomallei* with Tn5-OT182: isolation of motility mutants and molecular characterization of the flagellin structural gene. *J. Bacteriol.* **179** 2116-2125.

Deshazer, D., P. J. Brett and D. E. Woods. 1997. The type II O-antigenic polysaccharide moiety of *Burkholderia pseudomallei* lipopolysaccharide is required for serum resistance and virulence. *Mol. Microbiol.* **30** 1081-1100.

Deshazer, D., P. J. Brett, M. N. Burtnick and D. E. Woods. 1999. Molecular characterization of genetic loci required for secretion of exoproducts in *Burkholderia pseudomallei. J. Bacteriol.* **181** 4661-4664.

Desmarchelier, P. M., D. A. B. Dance, W. Chaowagul, Y. Suputtamongkol, N. J. White and T. L. Pitt. 1993. Relationships among *Pseudomonas pseudomallei* isolates from with recurrent melioidosis. *J. Clin. Microbiol.* **31** 1592-1596.

Dharakul, T. and S. Songsivilai. 1996. Recent developments in the laboratory diagnosis of melioidosis. *J. Infect. Dis. Antimicrob. Agents* **13** 77-80.

Dharakul, T. and S. Songsivilai. 1999. The many facets of melioidosis. *Trends Microbiol.* 7 138-139.

Dharakul, T., S. Songsivilai, S. Viriyachitra, V. Luangwedchakarn, B. Tassaneetritap and W. Chaowagul. 1996. Detection of *Burkholderia pseudomallei* DNA in patients with septicemic melioidosis. *J. Clin. Microbiol.* **34** 609-614.

Dharakul, T., S. Vejbaesya, W. Chawagul, P. Luangtrakool, H. A. F. Stephens and S. Songsivilai. 1998. HLA -DR and -DQ associations with melioidosis. *Human Immunol.* **59** 580-586.

DiRita, V. J. 1998. Regulation of virulence properties. In: M.Schaechter, N. C. Engleberg B. I. Eisenstein and G. Medoff (Eds). Mechanism of Microbial Disease. 3rd Edn., pp 205-216. Maryland, Williams & Wilkins.

DiRita, V. J. and J. Mekalanos. 1989. Genetic regulation of bacterial virulence. *Annu. Rev. Genet.* **23** 455-482.

Dodd, I. B. and J. B. Egan. 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *Nucleic Acids Res.* **18** 5019-5026.

Dodin, A. and J. Fournier. 1970. Precipitating and agglutinating antigens of *Pseudomonas pseudomallei* (Whitmore's bacillus). 1. Thermostable and thermolabile complexes-serological typing. *Ann. Inst. Pasteur Microbiol.* **119** 211-221.

Domenighini, M., M. Pizza and R. Rappuoli. 1995. Bacterial ADP-ribosyltransferases. In: J.Moss, B. Iglewski M. Vaughn and A. Tu (Eds). Bacterial Toxins and Virulence Factors in Disease. pp 59-80. New York, Marcel Dekker.

Donnenberg, M. S. and J. B. Kaper. 1991. Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect. Immun.* **59** 4310-4317.

Donohue-Rolfe, A., M. Jacewicz and G. T. Keusch. 1988. Shiga toxins as inhibitors of protein synthesis. *Methods Enzymol.* **165** 231-235.

Dorman, C. J. 1994. Genetics of Bacterial Virulence. 1st Edn. London, Blackwell Scientific Publications.

Dorman, C. J. and N. Ni Bhriain. Global regulation of gene expression during environmental adaptation: implications for bacterial pathogens. In: C.E.Hormaeche, C. W. Penn and C. J. Smyth (Eds). Molecular biology of bacterial infection. pp 193-230. 1992. Cambridge, CambridgeUniversityPress.

Dougan, G. 1989. Molecular characterization of bacterial virulence factors and the consequences for vaccine design. *J. Gen. Microbiol.* **135** 1397-1406.

Eaton, K. A., D. R. Morgan and S. Krakowa. 1992. Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori*. *J. Med. Microbiol* **37** 123-127.

Ebel, W. and J. E. Trempy. 1999. *Escherichia coli* RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions to activate its own expression. *J. Bacteriol.* **181** 577-584.

Ebel, W., G. J. Vaughn, H. K. Peters and J. E. Trempy. 1997. Inactivation of *mdoH* leads to increased expression of colanic acid capsular polysaccharide in *Escherichia coli*. *J. Bacteriol.* **179** 6858-6861.

Edelstein, P. H. 1986. Evaluation of the Gen-Probe DNA probe for the detection of *Legionellae* in culture. *J. Clin. Microbiol.* **23** 481-484.

Edmond, K., B. Currie, D. Brewster and C. Kilburn. 1998. Pediatric melioidosis in tropical Australia. *Pediatric infect. dis. J.* **17** 77-80.

Egan, A. M. and D. L. Gordon. 1996. *Burkholderia pseudomallei* activates complement and is ingested but not killed by polymorphonuclear leukocytes. *Infect. Immun.* **64** 4952-4959.

Egger, L. A., H. Park and M. Inouye. 1997. Signal transduction via the histidyl-aspartyl phosphorelay. *Genes to Cells* **2** 167-184.

Eggert, H., K. Bergemann and H. Saumweber. 1998. Molecular screening for *p*-element insertions in a large genomic region of *Drosophila melanogaster* using polymerase chain reaction mediated by the vectorette. *Genetics* **149** 1427-1434.

Elzaim, H. S., A. K. Chopra, J. W. Peterson, R. Goodheart and J. P. Heggers. 1998. Generation of neutralizing antipeptide antibodies to enzymatic domain of *Pseudomonas aeruginosa* exotoxin A. *Infect. Immun.* **66** 2170-2179.

Engels, W. R. 1993. Contributing software to the Internet: the Amplify program. *TIBS* **18** 448-450.

Everett, E. D. and R. A. Nelson. 1975. Pulmonary melioidosis. Observations in thirty-nine cases. *Am. Rev. Resp. Dis.* **112** 331-340.

Falkow, S. 1997. What is a pathogen? *ASM News* **63** 359-365.

Feng, J., M. R. Atkinson, W. McCleary, J. B. Stock, B. L. Wanner and A. J. Ninfa. 1992. Role of phosphorylated metabolic intermediates in the regulation of glutamine synthetase synthesis in *Escherichia coli. J. Bacteriol.* **174** 6061-6070.

Finlay, B. B. and S. Falkow. 1989. Common themes in microbial pathogenicity. *Microbiol.Rev.* **53** 210-230.

Finlay, B. B. and S. Falkow. 1997. Common themes in microbial pathogenicity revisited. *Microbiol. Mol. Biol. Rev.* **61** 136-169.

Fischer, E. H. and E. G. Krebs. 1989. Commentary on 'the phosphorylase *b* and *a* converting enzyme of rabbit skeletal muscle'. *Biochim. Biophys. Acta* **1000** 297-301.

Flamma, R. J., F. G. DiVincenti, L. N. Dotin and B. A. Pruitt. 1969. Pulmonary melioidosis. A diagnostic dilemma and increasing threat. *Ann. Thorac. Surg.* **7** 491-499.

Forst, S. and M. Inouye. 1988. Environmentally regulated gene expression for membrane proteins in *Escherichia coli. Annu. Rev. Cell Biol.* **4** 21-42.

Forst, S. A. and D. L. Roberts. 1994. Signal transduction by the EnvZ-OmpR phosphotransfer system in bacteria. *12th forum in Microbiology* 363-373.

Fournier, J. 1965. La melioidose et le b. de Whitmore. Controverses epidemiologiques et taxonomiques. *Bull. Soc. Pathol. Exot.* **58** 753-765.

Friedland, J. S., Y. Supputtamongkol, D. G. Remich, W. Chaowagul, R. M. Strieter, S. L. Kunkel and N. J. White. 1992. Prolonged elevation of interleukin-8 and interleukin-6 concentrations in plsma and leukocyte interleukin-8 mRNA levels during septicemic and localized *Pseudomonas pseudomallei* infection. *Infect. Immun.* **60** 2402-2408.

Galimand, M. and A. Dodin. 1982. Le point sur la melioidose dans le monde. *Bull. Soc. Pathol. Exot.* **75** 375-383.

Gauthier, Y. P., F. M. Thibault, J. C. Paucod and D. R. Vidal. 2000. Protease production by *Burkholderia pseudomallei* and virulence in mice. *Acta Tropica* **74** 215-220.

Gay, P., D. Le Coq, M. Steinmetz, E. Ferrari and J. A. Hoch. 1983. Cloning the structural gene *sacB*, which encodes for exoenzyme levansucrase of *Bacillus subtilis* : expression of the gene in *E. coli*. *J. Bacteriol* **153** 11424-11431.

Gervais, F. G. and G. R. Drapeau. 1992. Identification, cloning, and characterization of *rcsF*, a new regulator gene for exopolysaccharide synthesis that suppresses the division mutation *fts Z84* in *Escherichia coli* K-12. *J. Bacteriol.* **174** 8016-8022.

Gervais, F. G., P. Phoenix and G. R. Drapeau. 1992. The *rcsB* gene, a positive regulator of colanic acid biosynthesis in *Escherichia coli,* is also an activator of *ftsZ* expression *J. Bacteriol.* **174** 3964-3971.

Getzoff, E. D., J. A. Tainer, R. A. Lerner and H. M. Geysen. 1988. The chemistry and mechanism of antibody binding to protein antigens. *Adv. Immunol.* **43** 1-98.

Geysen, H. M., S. J. Barteling and R. H. Meloen. 1985. Small peptides induce antibodies with a sequence and structural requirement for binding antigen comparable to antibodies raised against the native protein. *Proc. Natl. Acad. Sci. USA* **82** 178-182.

Givskov, M., L. Olsen and S. Molin. 1988. Cloning and expression in *Escherichia coli* of the gene for extracellular phospholipase-A1 from *Serratia liquefaciens*. *J. Bacteriol* **170** 5855-5862.

Godfrey, A. J., S. Wong, D. A. B. Dance, W. Chaowagul and L. E. Bryan. 1991. *Pseudomonas pseudomallei* resistance to beta-lactam antibiotics due to alterations in the chromosomally encoded beta-lactamase. *Antimicrob. Agents Chemother.* **35** 1635-1640.

Golledge, C. L., W. S. Chin and A. E. Tribe. 1992. A case of human melioidosis originating in south-west Western Australia. *Med. J. Aust.* **157** 332-334.

Gottesman, S. and V. Stout. 1991. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. *Mol. Microbiol.* **5** 1599-1606.

Gottesman, S., P. Trisler and A. Torres-Cabassa. 1985. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K-12: characterization of three regulatory genes. *J. Bacteriol.* **162** 1111-1119.

Govan, J. R. W. and V. Deretic. 1996. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiol. Rev.* **60** 539-574.

Goyard, S. 1996. Identification and characterisation of BpH2, a novel histone H1 homolog in *Bordetella pertussis*. *J. Bacteriol* **178** 3066-3071.

Grant, C. C. R., M. E. Konel, W. Cieplak and L. S. Tompkins. 1993. Roles of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in non-polarized and polarized epithelial-cell cultures. *Infect. Immun.* **61** 1764-1771.

Gray, L. and A. Kreger. 1979. Microscopic examination of rabbit lung damage produced by *Pseudomonas aeruginosa* proteases. *Infect. Immun.* **23** 150-159.

Green, R. N. and P. G. Tuffnell. 1968. Laboratory acquired melioidosis. *Am. J. Med.* **44** 599-605.

Guan, C., P. Li, P. D. Riggs and H. Inouye. 1987. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* **67** 21-30.

Guard, R. W., F. A. Khafagi, M. C. Brigden and L. R. Ashdown. 1984. Melioidosis in far north Queensland: a clinical and epidemiological review of twenty cases. *Am. J. Trop. Med. Hyg.* **33** 467-473.

Gupte, G., C. Woodward and V. Stout. 1997. Isolation and characterization of *rcsB* mutations that affect colanic acid capsule synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **179** 4328-4335.

Haase, A., H. S. Vaughan, A. Melder, Y. Wood, A. Janmaat, J. Gilfedder, D. Kemp and B. Currie. 1995. Subdivision of *Burkholderia pseudomallei* ribotypes into multiple types by random amplified polymorphic DNA analysis provides new insights into epidemiology. *J. Clin. Microbiol.* **33** 1687-1690.

Haase, A., J. Janzen, S. Barrett and B. Currie. 1997. Toxin production by *Burkholderia pseudomallei* strains and correlation with severity of Melioidosis. *J. Med. Microbiol* **46** 557-563.

Halder, D., N. Zainal, C. M. Wah and J. A. Haq. 1998. Neonatal meningitis and septicaemia caused by *Burkholderia pseudomallei*. *Ann. Trop. Paediatr.* **18** 161-164.

Harley, C. B. and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.* **15** 2343-2361.

Hassett, D. J., P. A. Sokol, M. L. Howell, J. F. Ma, H. T. Schweizer, U. Ochsner and M. L. Vasil. 1996. Ferric uptake regulator (Fur) mutants of *Pseudomonas aeruginosa* demonstrate defective siderophore-mediated iron uptake, altered aerobic growth and decreased superoxide dismutase and catalase activities. *J. Bacteriol.* **178** 3996-40003.

Haubler, S., M. Nimtz, T. Domke, V. Wray and I. Steinmetz. 1998. Purification and characterization of a cytotoxic exolipid of *Burkholderia pseudomallei. Infect. Immun.* **66** 1588-1593.

Havell, E. A. and P. B. Sehgal. 1991. Tumor necrosis factor independent IL-6 production during murine listeriosis. *J. Immunol.* **146** 756-761.

Hawley, D. and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11** 2237-2255.

Haynes, W. C. 1957. Bergey's Manual of Determinative Bacteriology. Baltimore: Williams and Wilkins. 7th, p 100.

Heckly, R. J. and C. Nigg. 1958. Toxins of *Pseudomonas pseudomallei.* II characterization. *J. Bacteriol.* **76** 427-436.

Heijne, V. G. and L. Abrahamsen. 1989. Species-specific variation in signal peptide design. *FEBS Lett.* **244** 439-446.

Heinzel, F. P., D. S. Schoenhaut, R. M. Rerko, L. E. Rosser and M. K. Gately. 1993. Recombinant interleukin I2 cures mice infected with *Leishmania major. J. Exp. Med.* **177** 1505-1509.

Heithoff, D. M., R. L. Sinsheimer, D. A. Low and M. J. Mahan. 2000. *In vivo* gene expression and the adaptive response: from pathogenesis to vaccines and antimicrobials. *Phil. Trans. R. Soc. Lon.* **335** 633-642.

Hess, J. F., R. B. Bourret and M. I. Simon. 1988. Histidine phosphorylation and phosphoryl group transfer in bacterial chemotaxis. *Nature* **336** 139-143.

Hibbs, J. B., R. R. Taintor and Z. Vavrin. 1987. Macrophage cytotoxicity: role for L-arginine deaminase and amino nitrogen oxidation to nitrite. *Science* **197** 473-476.

Higgins, D. G. and P. M. Sharp. 1988. CLUSTAL: a package for performing multiple sequences alignment on a microcomputer. *Gene* **73** 237-244.

Ho, M., T. Schollaardt, M. D. Smith, M. B. Perry, P. J. Brett, W. Chaowagul and L. E. Bryan. 1997. Specificity and functional activity of anti-*Burkholderia pseudomallei* polysaccharide antibodies. *Infect. Immun.* **65** 3648-3653.

Hobbs, M. and J. S. Mattick. 1993. Common components in the assembly of type 4 fimbriae, DNA transfer system, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol. Microbiol.* **10** 233-143.

Hoch, J. A. and T. J. Silhavy. 1995. Two-component signal transduction. Washington, DC, American Society for Microbiology.

Holder, I. A. and A. N. Neely. 1989. Pseudomonas elastase acts as a virulence factor in burned hosts by hageman factor-dependent activation of the host kinin cascade. *Infect. Immun.* **57** 3345-3348

Hopp, T. P. and K. R. Woods. 1986. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA* **78** 3824-3828.

Houng, H. H., K. F. Noon, J. T. Ou and L. S. Baron. 1992. Expression of Vi antigen in *Escherichia coli* K-12: characterization of ViaB from *Citrobacter freundii* and identity of ViaA with RcsB. *J. Bacteriol.* **174** 5910-5915.

Howe, C., A. Sampath and M. Spotniz. 1971. The pseudomallei group: a review. *J. Infect. Dis.* **124** 598-606.

Hueck, C. J. 1998. Type III protein secretion in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62** 379-433.

Hunter, T 1995. Protein kinase and phosphatase: the yin and yang of protein phosphorylation and signalling. *Cell* **80** 225-236.

Ieven, M. 1998. Detection, specification and identification: 2. Detection. In: P.Williams, J. Ketley and G. Salmond (Eds). Methods in Microbiology: Bacterial Pathogenesis. pp 41-50. London,AcademicPress.

Inglis, T. J., J. D. Chiang, G. S. H. Lee and L. C. Kiang. 1998. Potential misidentification of *Burkholderia pseudomallei* by API 20NE *Pathology* **30** 62-64.

Inglis, T. J., J. P. Rigby, T. A. Robertson, N. S. Dutton, M. Henderson and B. J. Chang. 2000. Interaction between *Burkholderia pseudomallei* and *Acanthamoeba* species results in cloning phagocytosis, endamebic bacterial survival, and escape. *Infect. Immun.* **68** 1681-1686.

Inglis, T. J., J. S. C. Garrow, C. Adams, M. Henderson, M. Mayo and B. J. Currie. 1999. Acute melioidosis outbreak in Western Australia. *Epidemiol. Infect.* **123** 437-443.

Inouye, M. 1991. Intermolecular chaperone: the role of the pro-peptide in protein folding. *Enzyme* **45** 314-321.

Ip, M., G. Osterberg, P. Y. Chau and T. A. Raffin. 1995. Pulmonary melioidosis. *Chest* **108** 1420-1424.

Ishige, K., S. Nagasawa, S. Tokishita and T. Mizuno. 1994. A novel device of bacterial signal transducers. *EMBO J.* **13** 5195-5202.

Ishihama, A. 1993. Protein-protein communication within the transcription apparatus. *J. Bacteriol.* **175** 2483-2489.

Ismail, G., M. N. Embi, O. Omar, N. Razak, J. C. Allen and C. J. Smith. 1987. Enzyme immunoassay for the detection of antibody to *Pseudomonas pseudomallei* exotoxin in mice. *FEMS Microbiol. Lett.* **40** 27-31.

Ismail, G., M. N. Embi, O. Omar and N. Razak. 1987. Toxigenic properties of *Pseudomonas pseudomallei* extracellular products. *Tropical Biomedicine* **4** 101-109.

Ismail, G., N. Razak, R. Mohamed, N. Embi and O. Omar. 1988. Resistance of *Pseudomonas pseudomallei* to normal human serum bactericidal action. *Microbiol. Immunol.* **32** 645-652.

Iwasa, S., W. Petkanchanopong, P. Naigowit, E. Kondo and K. Kanai. 1992. Endotoxin of *Pseudomonas pseudomallei* detected by the body weight-decreasing reaction in mice and comparison of it with those of *P. Cepacia* and *P. aeruginosa. Jpn. J. Med. Sci. Biol.* **45** 35-47.

Jacobson, S. H., K. Tullus and B. Wretlind. 1988. Aerobactin-mediated uptake of iron by strains of *Escherichia coli* causing acute pyelonephritis and bacteremia. *J. Infection* **16** 147-152.

Jacobson, S. H., M. Hammarlind, K. J. Lidefeldt, E. Osterberg, K. Tullus and A. Brauner. 1988. Incidence of aerobactin-positive *Escherichia coli* strains in patients with symptomatic urinary tract infection. *Eur. J. Clin. Microbiol.* 7 630-634.

Jaeger, K. E., S. Ransac, B. W. Dijkstra, C. Colson, M. Vanheuvel and O. Misset. 1994. Bacterial lipases. *FEMS Microbiol. Rev.* **15** 29-63.

Jann, B. and K. Jann. 1990. Structure and biosynthesis of the capsular antigens of *Escherichia coli. Curr. Top. Micobiol.* **150** 19-42.

Jann, K. and B. Jann. 1987. Polysaccharide antigens of *Escherichia coli. Rev. Infect. Dis.* **9** S517-S525.

Jayanetra, P. 1989. Melioidosis in Thailand-update, microbiological aspect of melioidosis. *Proc. 1st Nat. Cong. Pathol. Chieng Mai Univ.* 71-72.

Jayanetra, P., M. Vorachit and S. Bhatarakamol. 1975. *Pseudomonas pseudomallei*: II. Laboratory and experimental studies in animals. *Southeast Asian J. Trop. Med. Public Health* **6** 10-17.

Jayaratne, P., W. J. Keenleyside, P. M. MacLachlan, C. Dodgson and C. Whitfield. 1993. Characterization of *rcsB* and *rcsC* from *Escherichia coli* O9:K30:H12 and examination of the role of the *rcs* regulatory system in expression of group I capsular polysaccharides. *J. Bacteriol.* **175** 5384-5394.

Jepson, M. and R. Titball. 2000. Structure and function of clostridial phospholipase C. *Microbes Infect.* **2** 1277-1284.

Jin, S. G., R. K. Prusti, T. Roitsch, R. G. Ankenbauer and E. W. Nester. 1990. Phosphorylation of the VirG protein of *Agrobacterium tumefaciens* by the autophosphorylated VirA protein: essential role in biological activity of VirG. *J. Bacteriol.* **172** 4945-4950.

Jin, T. and M. Inoue. 1993. Ligand binding to the receptor domain regulates the ratio of kinase to phosphatase activities of the signalling domain of the hybrid *Escherichia coli* transmembrane receptor Taz 1. *J. Mol. Biol.* **232** 484-492.

Jonas, D., B. Schultheis, C. Klas, P. H. Krammer and S. Bhakdi. 1993. Cytocidal effects of *Escherichia coli* hemolysin on human T-lymphocytes. *Infect. Immun.* **61** 1715-1721.

Johnson, A. P. and T. J. Inzana. 1986. Loss of ciliary activity in organ cultures of rat trachea treated with lipo-oligosaccharide isolated from *Haemophilus influenzae*. *J. Med. Microbiol* **22** 265-268.

Johnston, R. B., Jr. H. P. Misra, J. E. Lehmeyer, L. S. Webb, R. L. Baehner and K. V. Rajagopalan. 1975. The role of superoxide anion generation in phagocytic bactericidal activity. Studies with normal and chronic granulomatous disease leukocytes. *J. Clin. Invest.* **55** 1357-1372.

Jones, A. L., D. Deshazer and D. E. Woods. 1997. Identification and characterization of a two-component regulatory system involved in invasion of Eukaryotic cells and heavy-metal resistance in *Burkholderia pseudomallei*. *Infect. Immun.* **65** 4972-4977.

Jones, A. L., T. J. Beveridge, and D. E. Woods. 1996. Intracellular survival of *Burkholderia pseudomallei. Infect. Immun.* **64** 782-790.

Joys, T. M. 1988. The flagellar filament protein. *Can. J. Microbiol.* **34** 452-458.

Kahn, C. R. 1997. Insulin receptors and insulin signalling in normal and disease states. In: K.G.M.M.Alberti, P. Zimmet R. A. DeFronzo and H. Keen (Eds). International Textbook of Diabetes Mellitus, 437-458. John Wiley & Sons Ltd.

Kan, S. K. P. and R. W. W. Kay. 1978. Melioidosis presenting as prostatitis. A case report from Sabah. *Trans. Roy. Soc. Trop. Med. Hyg.* **72** 522-524.

Kanai, K., Y. Suzuki, E. Kondo, Y. Maejima, D. Miyamoto, T. Suzuki and T. Kurata. 1997. Specific binding of *Burkholderia pseudomallei* cells and their cell-surface acid phosphatase to ganglio-tetraosylceramide (Asialo gm1) and ganglotriaosylceramide (Asialo gm2). *Southeast Asian J. Trop. Med. Public Health.* **28** 781-790.

Kanai, K. and E. Kondo. 1991. Substrate response in acid phosphatase avtivity of *Pseudomonas pseudomallei* and *Pseudomonas cepacia*, with special reference to tyrosine phosphatase. *Jpn. J. Med. Sci. Biol.* **44** 225-237.

Kanai, K. and E. Kondo. 1994. Recent advances in biomedical sciences of *Burkholderia pseudomallei* (Basonym : *Pseudomonas pseudomallei*). *Jpn. J. Med. Sci. Biol.* **47** 1-45.

Kanai, K., E. Kondo and T. Kurata. 1996. Affinity and response of *Burkholderia pseudomallei* and *Burkholderia cepacia* to insulin. *Asian J. Trop. Med. Public Health.* **27** 584-591.

Kanaphun, P., N. Thirawattanasuk and Y. Suputtamongkol. 1993. Serology and carriage of *Pseudomonas pseudomallei* - a prospective study in 1000 hospitalized children in north-east Thailand. *J. Infec. Dis.* **167** 230-233.

Karkkainen, U., R. Ikaheimo, M. L. Katila and P. Mantijarvi. 1991. P-fimbriation of *Escherichia coli* strains from patients with urosepsis demonstrated by a commercial agglutination-test (PE-Test). *J. Clin. Microbiol.* **29** 221-224.

Kaufmann, S. H. E. 1993. Immunity to intracellular bacteria. *Annu. Rev. Immunol.* **11** 129-163.

Kaufmann, S. H. E. 2000. Is the development of a new tuberculosis vaccine possible? *Nature Med.* **9** 955-960.

Kaufmann, S. H. E. and M. J. Reddehase. 1989. Infection of phagocytic cells. *Curr. Opin. Immunol.* **2** 43-49.

Kaufmann, S. H. E. and I. E. A. Flesch. 1992. Life within phagocytic cells. In: C.E.Hormaeche, C. W. Penn and C. J. Smyth (Eds). Molecular Biology of Bacterial Infections: Current Status and Future Prespective. New York, Cambridge University Press.

Kawahara, K., S. Dejsirilert and T. Ezaki. 1998. Characterization of three capsular polysaccharides produced by *Burkholderia pseudomallei. FEMS Microbiol. Lett.* **169** 283-287.

Keener, J. and S. Kustu. 1988. Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. *Proc. Natl. Acad. Sci. USA* **85** 4976-4980.

Kelley, W. L. and C. Georgopoulos. 1997. Positive control of the two-component RcsC/B signal transduction network by DjlA: a member of the DnaJ family of molecular chaperones in *Escherichia coli. Mol. Microbiol.* **25** 913-931.

Kennelly, P. J. and M. Potts. 1996. Fancy meeting you here! a fresh look at prokaryotic protein phosphorylation. *J. Bacteriol.* **178** 4759-4764.

Ketterer, P. J., B. Donald and R. J. Rogers. 1975. Bovine melioidosis in South Eastern Queensland. *Aust. Vet. J.* **51** 395-398.

Ketterer, P. J., W. R. Webster, J. Shield, R. J. Arthur and P. J. Blackall. 1986. Melioidosis in intensive piggeries in South-Eastern Queenslands. *Aust. Vet. J.* **63** 146-149.

Khaira, B. S., W. B. Young and V. P. DE Hart. 1959. Melioidosis. *British Medical J.* 949-956.

Khomenko, A. G., V. I. Litvinov, V. P. Chukanova and L. E. Pospelov. 1990. Tuberculosis in patients with various HLA phenotypes. *Tubercle* **71** 187-192.

Knirel, Y. A., N. A. Paramonov and A. S. Shashkov. 1992. Structure of the polysaccharide chain of *Pseudomonas pseudomallei* lipopolysaccharides. *Carbohyd. Res.* **233** 185-193.

Kolter, R., M. Inuzuka and D. R. Helinski. 1978. Transcomplementation dependent replication of low molecular weight origin fragment from plasmid RK6. *Cell* **15** 119-128.

Kondo, E., P. Wangroongsaub and K. Kanai. 1994. Effects of tunicamycin on the pH-activity pattern of acid phosphatase in *pseudomonas pseudomallei. Southeast Asian J. Trop. Med. Public Health* **25** 144-151.

Kondo, E., S. Dejsirilert, N. Wejprasit, D. Chiewsilp and K. Kanai. 1991a. Heat-stable and heat-labile components of nonspecific acid phosphatase detected in *Pseudomonas pseudomallei. Jpn. J. Med. Sci. Biol.* **44** 51-62.

Kondo, E., T. Kurata, P. Naigowit and K. Kanai. 1996. Evolution of cell-surface acid phosphatase of *Burkholderia pseudomallei. Southeast Asian J. Trop. Med. Public Health* **27** 592-599.

Kondo, E., V. Petkanchanapong, P. Naigowit, T. Kurata and K. Kanai. 1991b. Demonstration of acid phosphatase activity in antigenic glycoprotein fractions obtained from the culture filtrate of *Pseudomonas pseudomallei. Jpn. J. Med. Sci. Biol.* **44** 213-224.

Korbsrisate, S., N. Suwanasai, A. Leelaporn, T. Ezaki, Y. Kawamura and S. Asrasombath. 1999. Cloning and characterization of a nonhemolytic phospholipase C gene from *Burkholderia pseudomallei. J. Clin. Microbiol.* **37** 3742-3745.

Krivan, H. C., D. D. Roberts and V. Ginsburg. 1988. Many pulmonary pathogenic bacteria bind specifically to the carbohydrate sequence Galanc-beta-1-4gal found in some glycolipids. *Proc. Natl. Acad. Sci. USA* **16** 6157-6161.

Kuehn, M., K. Lent, J. Hass, J. Hagenzieker, M. Cervin and A. L. Smith. 1992. Fimbriation of *Pseudomonas cepacia. Infect. Immun.* **60** 2002-2007.

Kuehn, M. J. 1997. Establishing communication via Gram-negative bacterial pili. *Trends Microbiol.* **5** 130-132.

Kunakorn, M. and R. B. Markham. 1995. Clinically practical seminested PCR for *Burkholderia pseudomallei* quantitated by enzyme immunoassay with and without solution hybridization. *J. Clin. Microbiol.* **33** 2132-2135.

Kunakorn, M., B. Petchclai, K. Khupulsup and P. Naigowit. 1991. Gold blot for detection of immunoglobulin M (IgM)- and IgG- specific antibodies for rapid serodiagnosis of Melioidosis. *J. Clin. Microbiol.* **29** 2065-2067.

Kunakorn, M., K. Raksakait, C. Sethaudom, R. W. Sermswan and T. Dharakul. 2000. Comparison of three PCR primer sets for diagnosis of septicemic melioidosis. *Acta Tropica* **74** 247-251.

Kunakorn, M., P. Boonma, K. Khupulsup and B. Petchclai. 1990. Enzyme-linked immunosorbent assay for immunoglobulin M specific antibody for the diagnosis of Melioidosis. *J. Clin. Microbiol.* **28** 1249-1253.

Kunakorn, M., P. Jayanetra and D. Tanphaichitra. 1991. Man-to-man transmission of melioidosis. *Lancet* **337** 1290-1291.

Kyte, J. and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157** 105-132.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. *Nature* **222** 680-685.

Lajudie, P. and E. R. Brygoo. 1953. Contribution a l'etude du pouvoir pathogene du bacille de Whitmore. *Ann. Instit. Pasteur* **85** 99-106.

Lath, R., V. Rajshekhar and V. George. 1998. Brain abscess as the presenting feature of melioidosis. *Br. J. Neurosurgery* **12** 170-172.

Laver, W. G., G. M. Air, R. G. Webster and S. J. Smith-Gill. 1990. Epitopes on protein antigens- misconceptions and realities. *Cell* **61** 553-556.

Laville, J., C. Voisard, C. Keel, M. Maurhofer, G. Defago and D. Hass. 1992. Global control in *Pseudomonas fluorescens* mediating antibiotic synthesis and suppression of black root rot of tobacco. *Microbiology.* **89** 1562-1566.

Laws, L. W. and T. K. Hall. 1963. Melioidosis in animals in north Queensland. 1. Incidence and pathology, with special reference to central nervous system lesions. *Qld. J. Agric. Sci.* **20** 499-513.

Leakey, A. K., G. C. Ulett and R. G. Hirst. 1998. BALB / c and C57BI / 6 mice infected with virulent *Burkholderia pseudomallei* provide contrasting animal models for the acute and chronic forms of human melioidosis. *Microb. Pathog.* **24** 269-275.

Lee, C. A. 1997. Type III secretion systems: machines to deliver bacterial proteins into eukaryotic cells. *Trends Microbiol.* **5** 148-156.

Lee, M. A. and Y. Liu. 2000. Sequencing and characterization of a novel serine metalloprotease from *Burkholderia pseudomallei. FEMS Microbiol. Lett.* **192** 67-72.

Lee, S. H., C. Kim, W. S. Lee and S. M. Byun. 1996. RcsC-mediated induction of colanic acid by secretion of streptokinase in *Escherichia coli* K-12. *FEMS Microbiol. Lett.* **139** 189-193.

Leelarasamee, A. 1998. *Burkholderia pseudomallei*: the unbeatable foe? *Southeast Asian J. Trop. Med. Public Health* **29** 410-415.

Leelarasamee, A. 2000. Melioidosis in south-east Asia. *Acta Tropica* **74** 129-132.

Leelarasamee, A. and S. Bovornkitti. 1989. Melioidosis : review and update. *Rev. Infect. Dis.* **11** 413-425.

Leimeisterwachter, M., E. Domann and T. Chakraborty. 1992. The expression of virulence genes in *Listeria monocytogenes* is thermoregulated. *J. Bacteriol* **174** 947-952.

Letoffe, S., P. Delepelaire and C. Wandersman. 1991. Cloning and expression in *Escherichia coli* of the *Serratia marcescens* metalloprotease gene- secretion of the protease from *Escherichia coli* in the presence of the *Erwinia chrysanthemi* protease secretion functions. *J. Bacteriol* **173** 2160-2166.

Levine, H. B. and H. Wolochow. 1960. Occurrence of poly-β-hydroxybutyrate in *Pseudomonas pseudomallei*. *J. Bacteriol.* **79** 305-306.

Lew, A. E. and P. M. Desmarchelier. 1994. Detection of *Pseudomonas pseudomallei* by PCR and hybridization. *J. Clin. Microbiol.* **32** 1326-1332.

Li, L. and H. You-win. 1992. *Pseudomonas pseudomallei* and Melioidosis in China. *Chinese Med. J.* **105** 775-779.

Lim, T. K. 1997. Emerging pathogens for pneumonia in Singapore. *FAMS. M. Med (Int med)* **26** 651-658.

Lin, M. Y., M. C. Cheng, K. J. Huang and W. C. Tsai. 1993. Classification, pathogenicity and drug susceptibility of haemolytic gram-negative bacteria isolated from sick or dead chickens. *Avian Dis.* **37** 6-9.

Lindler, L. E. and B. D. Tall. 1993. *Yersinia pestis* pH 6 antigen forms fimbriae and is induced by intracellular association with macrophages. *Mol. Microbiol.* **8** 311-324.

Lingnau, A., E. Domanne and M. Hudel. 1995. Expression of the *Listeria monocytogenes* InlA and InlB genes, whose products mediate bacterial entry into tissue-culture cell-lines, by PrfA-dependent and PrfA-independent mechanisms. *Infect. Immun.* **63** 3896-3903.

Liu, J. 1990. Molecular genetics of the chemotactic signalling pathway in *Escherichia coli*. Thesis. University of Utah, Salt Lake City, Utah.

Liu, J. T. 1981. Histopathology studies in horses experimentally infected with *Pseudomonas pseudomallei*. *People Military Surgeon* 12-16.

Liu, P. V. 1957. Survey of hemolysin production among species of Pseudomonads. *J. Bacteriol* **74** 718-727.

Loomis, W. F., G. Shaulsky and N. Wang. 1997. Histidine kinases in signal transduction pathways of eukaryotes. *J. Cell Science* **110** 1141-1145.

Loprasert, S., R. Sallabhan, W. Whangsuk and S. Mongkolsuk. 2000. Characterization and mutagenesis of *fur* gene from *Burkholderia pseudomallei*. *Gene* **245** 129-137.

Lory, S. 1992. Determinants of extracellular protein secretion in Gram-negative bacteria. *J. Bacteriol.* **174** 3423-3428.

Lukat, G. S., B. H. Lee, J. M. Mottonen, A. M. Stock and J. B. Stock. 1991. Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. *J. Biol.Chem.* **226** 8348-8354.

Lukat, G. S., W. R. McCleary, A. M. Stock and J. B. Stock. 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc. Natl. Acad. Sci. USA* **89** 718-722.

Machielag, M. J., J. P. Demers, S. A. Fraga-Spano, D. J. Hlasta, S. G. Johnson, R. M. Kanojia, R. K. Russell, Z. H. Sui, M. A. Weidner-Wells, H. Werblood, B. D. Foleno, R. M. Goldschmidt, M. J. Loeloff, G. C. Webb and J. F. Barnett. 1998. Substituted salicylanilides as inhibitors of two-component regulatory systems in bacteria. *J. Med. Chem.* **41** 2939-2945.

Mack, K. and R. W. Titball. 1996. Transformation of *Burkholderia pseudomallei* by electroporation. *Anal. Biochem.* **242** 73-76.

Mack, K. and R. W. Titball. 1998. The detection of insertion sequences within the human pathogen *Burkholderia pseudomallei* which have been identified previously in *Burkholderia cepacia*. *FEMS Microbiol. Lett.* **162** 69-74.

Madigan, M. T., J. M. Martinko and J. Parker. 2000. Brock Biology of Microorganisms. P.F.Corey (Ed). Ninth Edition. USA, Prentice-Hall International.

Mahdy, H. M. 1997. Pseudomonads role in the transformation of certain heavy metals in dead lake. *Ecologia (Bratislava)* **16** 323-330.

Makino, K., M. Amemura, S. K. Kim, A. Nakata and H. Shinagawa. 1994. Mechanism of transcriptional activation of the phosphate regulon in *Escherichia coli*. In: Phosphate in microorganisms: Cellular and molecular biology. A.Torriani-Gorini, E.Yagil & S.Silver (Eds), pp.5-12.Washington, DC: ASM Press.

Markovitz, A. 1977. Genetics and regulation of bacterial capsular polysaccharide biosynthesis and radiation sensitivity. I.W.Sutherland (Ed). Surface carbohydrates of the prokaryotic cell. pp 415-462. New York, Academic Press.

Masoud, H., M. Ho, T. Schollaardt and M. B. Perry. 1997. Characterization of the capsular polysaccharide of *Burkholderia* (*Pseudomonas*) *pseudomallei* 304b. *J. Bacteriol.* **179** 5663-5669.

Matsuura, M., K. Kawahara, T. Ezaki and M. Nakano. 1996. Biological activities of lipopolysaccharide of *Burkholderia* (*Pseudomonas*) *pseudomallei*. *FEMS Microbiol. Lett.* **137** 79-83.

McCallum, K. L. and C. Whitfield. 1991. The rcsA gene of *Klebsiella pneumoniae* O1:K20 is involved in expression of the serotype-specific K (capsular) antigen. *Infect. Immun.* **59** 494-502.

McCallum, K. L., D. H. Laakso and C. Whitfield. 1989. Use of a bacteriophage-encoded glycanase enzyme in the generation of lipopolysaccharide O side chain deficient mutants of *Escherichia coli* 09:K30 and *Klebsiella* 01:K20 of O and K antigens in resistance to complement-mediated serum killing. *Can. J. Microbiol.* **35** 994-999.

McCleary, W. R. 1996. The activation of PhoB by acetylphosphate. *Mol. Microbiol.* **20** 1155-1163.

McCormick, J. B., D. L. Sexton, J. G. McMurray, E. Carey, A. B. Peggy Hayes and R. A. Feldman. 1975. Human-to-human transmission of *Pseudomonas pseudomallei*. *Ann. Int. Med.* **83** 512-513.

McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady and M. Rosenberg. 1981. A system to study promoter and terminator signals recognised by *Escherichia coli* RNA polymerase. In: J.G.Chirikjian and T.S.Papas (Eds). Gene Amplification and Analysis: Analysis of Nucleic Acids by Enzymatic Methods., pp 383-415. New York, Elsevier-North Holland.

McKevitt, A. I. and D. E. Woods. 1984. Characterization of *Pseudomonas cepacia* isolates from patients with cystic fibrosis. *J. Clin. Microbiol.* **19** 291-293.

McKevitt, A. I., S. Bajaksouzian, J. D. Klinger and D. E. Woods. 1989. Purification and characterization of an extracellular protease from *Pseudomonas cepacia*. *Infect. Immun.* **57** 771-778.

McIver, K. S., E. Kessler, J. C. Olson and D. E. Ohman. 1995. The elastase propeptide functions as an intramolecular chaperon required for elastase activity and secretion in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **18** 877-889.

McIver, K. S., J. C. Olson and D. E. Ohman. 1993. *Pseudomonas aeruginosa lasB*1 mutants produce an elastase, substituted at active-site His-223 that is defective in activity, processing, and secretion. *J. Bacteriol.* **175** 4008-4015.

Medvedkin, V. N., E. A. Permyakov, L. Klimenko, Y. V. Mitin, N. Matsushima, S. Nakayama and R. H. Kretsinger. 1995. Interactions of (ala-asterisk-ala-asterisk-lys-asterisk-pro) (n) and (lys-asterisk-lys-asterisk-ser-asterisk-pro) (n) with DNA-proposed coiled-coil structure of AlgR3 and AlgP from *Pseudomonas aeruginosa. Protein Eng.* **8** 63-70.

Mehra, N. K. 1990. Role of HLA linked factors in governing susceptibility to leprosy and tuberculosis. *Trop. Med. Parasitol.* **41** 352-354.

Mekalanos, J. J. 1992. Environmental signals controlling expression of virulence determinants in bacteria. *J. Bacteriol.* **174** 1-7.

Merianos, A., M. Patel and J. M. Lane. 1993. The 1990-1991 outbreak of melioidosis in the Northern Territory of Australia: epidemiology and environmental studies. *Southeast Asian J. Trop. Med. Public Health* **24** 425-435.

Merkel, T. J., C. Barros and S. Stibitz. 1998. Characterization of the *bvgR* locus of *Bordetella pertussis. J. Bacteriol.* **180** 1682-1690.

Meyer, J. M., D. Hohnadel and F. Halle. 1989. Cepabactin from *Pseudomonas cepacia*, a new type of siderophore. *J. Gen. Microbiol.* **135** 1479-1487.

Michiels, T. and G. R. Cornelis. 1991. Secretion of hybrid proteins by the *Yersinia* Yop export system. *J. Bacteriol.* **173** 1677-1685.

Michiels, T., P. Wattiau, R. Brasseur, J. M. Ruysschaert and G. Cornelis. 1990. Secretion of Yop proteins by *Yersiniae. Infect. Immun.* **58** 2840-2849.

Migula, W. 1984. Arbeiten aus dem Bakteriologischen Institute der Technischen Hochschule zu Karlsruhe. 1, 235-238.

Miller, J. F., J. J. Mekalanos and S. Falkow. 1989. Coordinate regulation and sensory transduction in the control of bacterial virulence. *Science* **243** 916-922.

Miller, S. I., A. M. Kukral and J. J. MeKalanos. 1989. A two-component regulatory system (*phoP phoQ*) controls *Salmonella typhimurium* virulence. *Proc. Natl. Acad. Sci. USA* **86** 5054-5058.

Miller, S. I. and J. J. MeKalanos. 1990. Constitutive expression of the phoP regulon attenuates *Salmonella* virulence and survival within macrophages. *J. Bacteriol.* **172** 2485-2490.

Miller, S. I., W. P. Loomis, C. A. Aranda, I. Behlau and E. Hohmann. 1993. The PhoP virulence regulon and live oral *Salmonella* vaccines. *Vaccine* **11** 122-125.

Miller, V. L. and J. J. Mekalanos. 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J. Bacteriol.* **170** 2575-2583.

Miller, W. R., L. Pannell, L. Cravity, W. A. Tanner and T. Rosebury. 1948. Studies on certain biological characteristics of *Malleomyces mallei* and *Malleomyces pseudomallei*. II. Virulence and infectivity for animals. *J. Bacteriol.* **55** 127-135.

Milton, D. L., R. O'Toole, P. Horstedt and H. Wolfwatz. 1996. Flagellin A is essential for the virulence of *Vibrio anguillarum*. *J. Bacteriol.* **178** 1310-1319.

Mims, C., N. Dimmock, A. Nash and J. Stephen. 1997. Mim's Pathogenesis of Infectious Disease. 4th Edn. San Diego, Academic Press Inc.

Mims, C. A., N. J. Dimmock, A. Nash and J. Stephen. 1995. Mechanisms of cell and tissue damage. Mims' Pathogenesis of Infectious Diseases. 4th, 207-208. London, Academic Press.

Mitchell, T. J. 1998. Molecular basis of virulence. *Arch. Dis. Child.* **78** 197-199.

Mitsuchi, M., P. Wichaidlit and S Jeungnijnirund. 1986. Outlines of soils of the Northeastern Plateau, Thailand: Their characteristics and constraints. Japan International Cooperation Agency Report. Thailand, Siriphan Offset Pub. Co. Ltd.

Miyagi, K., K. Kawakami and A. Saito. 1997. Role of reactive nitrogen and oxygen intermediates in gamma interferon-stimulated murine macrophage bactericidal activity against *Burkholderia pseudomallei*. *Infect. Immun.* **65** 4108-4113.

Mizuno, T. 1997. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Res.* **4** 161-168.

Mizuno, T. 1998. His-Asp phosphotransfer signal transduction. *J. Biochem.* **123** 555-563.

Mizuno, T., T. Kaneko and S. Tabata. 1996. Compilation of all genes encoding bacterial two-component signal transducers in the genome of the cyanobacterium, *Synechocystis* sp. strain PCC 6803. *DNA Res.* **3** 407-414.

Moayeri, M. and R. A. Welch. 1998. Bacterial exotoxins. In: P.Williams, J. Ketley and G. Salmond (Eds). Methods in Microbiology: Bacterial Pathogenesis. pp 287-300. London, Academic Press.

Moens, S. and J. Vanderleyden. 1996. Functions of bacterial flagella. *Crit. Rev. Microbiol.* **22** 67-100.

Mohamed, R., S. Nathan, N. Embi, N. Razak and G. Ismail. 1989. Inhibition of macromolecular synthesis in cultured macrophages by *Pseudomonas pseudomallei* exotoxin. *Microbiol. Immunol.* **33** 811-820.

Moseley, S. L., I. Huq, A. R. M. Alim, M. SO M. Samadpour-Motalebi and S. Falkow. 1980. Detection of enterotoxigenic *Escherichia coli* by DNA colony hybridization. *J. Infec. Dis.* **142** 892-898.

Moulder, J. W. 1985. Comparative biology of intracellular parasitism. *Microbiol. Rev.* **49** 298-337.

Msadek, T. 1999. When the going gets tough: survival strategies and environmental signalling networks in *Bacillus subtilis*. *Trends Microbiol.* **7** 201-207.

Neilands, J. B. 1993. Siderophores. *Arch. Biochem. Biophys.* **302** 1-3.

Nelson, J. W., S. L. Butler, D. Krieg and J. R. W. Govan. 1994. Virulence factors of *Burkholderia cepacia*. *FEMS Immunol. Med. Microbiol.* **8** 89-98.

Nicholls, L. 1930. Melioidosis, with special reference to the dissociation of *Bacillus whitmori*. *Br. J. Exp. Pathol.* **11** 393-399.

Nigg, C. 1963. Serological studies on subclinical melioidosis. *J. Immunol.* **91** 18-28.

Nigg, C., R. J. Heckly and M. Colling. 1955. Toxin produced by *Malleomyces pseudomallei*. *Proc. Soc. Exp. Biol. Med.* **89** 17-20.

Nimtz, M., V. Wray, T. Domke, B. Brenneke, S. Haussler and I. Steinmetz. 1997. Structure of an acidic exopolysaccharide of *Burkholderia pseudomallei*. *Eur. J. Biochem.* **250** 608-616.

Ninfa, A. J. and B. Magasanik. 1986. Covalent modification of the *glnG* product, NRI, by the *glnL* product, NRII, regulates the transcription of the *glnALG* operon in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **83** 5909-5913.

Ninfa, A. J., E. Brodsky and B. Magasanik. 1989. DNA-Protein Interactions in Transcription. J. Gralla (Ed). pp 43-52. New York, Liss.

Nixon, B. T., C. W. Ronson and F. M. Ausubel. 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. *Proc. Natl. Acad. Sci. USA* **83** 7850-7854.

Nomura, S., M. Yanabu, H. Kido, X. Lan, H. Ichiyoshi and K. Katsura. 1995. Significance of cytokines and CD68-positive microparticles in immune thrombocytopenic purpura. *Eur. J. Haem.* **55** 49-56.

Norton, R., B. Roberts, M. Freeman, M. Wilson, C. Ashhurst-Smith, W. Lock, D. Brookes and J. La Brooy. 1998. Characterization and molecular typing of *Burkholderia pseudomallei*: are disease presentations of melioidosis clonally related? *FEMS Immunol. Med. Microbiol.* **20** 37-44.

Obrig, T. G. 1994. Toxins that inhibit host protein synthesis. *Methods Enzymol.* **235** 647-656.

Oosawa, K., J. F. Hess and M. I. Simon. 1988. Mutants defective in bacterial chemotaxis show modified protein phosphorylation. *Cell* **53** 89-96.

Ophier, T. and D. L. Gutnick. 1994. A role for exopolysaccharides in the protection of microorganisms from desiccation. *Appl. Environ. Microbiol.* **60** 740-745.

Oppenheim, D. S. and C. Yanofsky. 1980. Translational coupling during expression of the tryptophan operon of *E. coli*. *Genetics* **95** 785-795.

Orskov, F., I. Orskov, B. Jann and K. Jann. 1971. Immuno-electrophoresis patterns of extracts from all *Escherichia coli* O and K antigen test strains: correlation with pathogenicity. *Acta Pathol. Microbiol. Scand.* **79B** 142-152.

Osbourn, A. E., B. R. Clarke, B. J. H. Stevens and M. J. Daniels. 1990. Use of oligonucleotide probes to identify members of two-component regulatory system in *Xanthomonas campestris* pathovar *campestris*. *Mol. Gen. Genet.* **222** 145-151.

Owen-Hughes, T. A., G. D. Pavitt, D. S. Santos, M. J. Sidebotham, C. S. Hulton, J. C. Hinton and C. F. Higgins. 1992. The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* **71** 255-265.

Pabo, C. O. and R. T. Sauer. 1984. Protein-DNA recognition. *Ann. Rev. Biochem.* **53** 293-321.

Pabo, C. O. and R. T. Sauer. 1992. Transcriptional factors: structural families and principles of DNA recognition. *Ann. Rev. Biochem.* **61** 1053-1095.

Pabo, C. O. and M. Lewis. 1982. The oprator-binding domain of λ repressor: structure and DNA recognition. *Nature* **298** 443-447.

Palleroni, N. J. and B. Holmes. 1981. *Pseudomonas cepacia* sp. *Int. J. Syst. Bacteriol.* **31** 479-481.

Pao, G. M. and M. H. Jr. Saier. 1995. Response regulators of bacterial signal transduction systems: selective domain shuffling during evolution. *J. Mol. Evolution* **40** 136-154.

Parkinson, J. S. 1993. Signal transduction schemes of bacteria. *Cell* **73** 857-871.

Parkinson, J. S. 1995. Genetic approaches for signalling pathways and proteins. In: J.A.Hoch & T.J.Silhavy (Eds). Two-Component Signal Transduction. pp 9-24. Washington, DC., American Society for Microbiology.

Parkinson, J. S. and E. C. Kofoid. 1992. Communication modules in bacterial signalling proteins. *Ann. Rev. Gen.* **26** 71-112.

Patamasucon, P., U. B. Schaad and J. D. Nelson. 1982. Melioidosis. *J. Pediatr.* **100** 175-182.

Pearson, W. R. and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85** 2444-2448.

Pegues, D. A., M. J. Hantman, I. Behlau and S. I. Miller. 1995. PhoP/PhoQ transcriptional repression of *Salmonella typhimurium* invasion genes: evidence for a role in protein secretion. *Mol. Microbiol.* **17** 169-181.

Penn, C. W. 1992. Chronic infections, latency and the carrier state. In: C.E.Hormaeche, C. W. Penn and C. J. Smyth (Eds). Molecular Biology of Bacterial Infection: Current Status and Future Prespective. pp 107-126. Wiltshire, Cambridge University Press.

Percheron, G., F. Thibault, J. C. Paucod and D. Vidal. 1995. *Burkholderia pseudomallei* requires Zn^{2+} for optimal exoprotease production in chemically defined media. *Appl. Environ. Microbiol.* **61** 3151-3153.

Perona, J. J. and C. S. Craik. 1997. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **48** 29987-29990.

Perry, M. B., L. L. Maclean, T. Schollaardt, L. E. Bryan and M. Ho. 1995. Structural characterization of the lipopolysaccharide O antigens of *Burkholderia pseudomallei. Infect. Immun.* **63** 3348-3352.

Petkanjanapong, V., P. Naigowit, E. Kondo and K. Kanai. 1992. Use of endotoxin antigens in enzyme-linked immunosorbent assay for the diagnosis of *P. pseudomallei* infections (melioidosis). *Asian Pacific J. Allergy Immunol.* **10** 145-150.

Pettijohn, D. E. 1988. Histone-like proteins and bacterial chromosome structure. *J. Biol. Chem.* **263** 12793-12796.

Pitt, T. L., H. Aucken and D. A. B. Dance. 1992. Homogeneity of lipopolysaccharide antigens in *Pseudomonas pseudomallei. J. Infect. Dis.* **25** 139-146.

Pitt, T. L., S. Trakulsomboon and D. A. B. Dance. 2000. Molecular phylogeny of *Burkholderia pseudomallei. Acta Tropica* **74** 181-185.

Poetter, K. and D. L. Coplin. 1991. Structural and functional analysis of the *rcsA* gene from *Erwinia stewartii. Mol. Gen. Genet.* **229** 155-160.

Popov, S. F., N. G. Tikhonov and V. Y. Kurilov. 1994. Influence of capsule formation in *Pseudomonas pseudomallei* and *Pseudomonas mallei* on their relations with host cells. *Immunol. Infect. Dis.* **4** 142-145.

Pruckler, J. M., R. F. Benson, M. Moyenuddin, W. T. Martin and B. S. Fields. 1995. Association of flagellum expression and intracellular growth of *Legionella pneumophila. Infect. Immun.* **63** 4928-4932..

Pruksachartvuthi, S., N. Aswapokee, and K. Thankerngpol. 1990. Survival of *Pseudomonas pseudomallei* in human phagocytes. *J. Med. Microbiol* **31** 109-114.

Puente, J. L., A. Verdugo-Rodriguez and E. Calva. 1991. Expression of *Salmonella typhi* and *Escherichia coli* OmpC is influenced differently by medium osmolarity; dependence on *Escherichia coli* OmpR. *Mol. Microbiol.* **5** 1205-1210.

Pugsley, A. P. 1993. The complete general secretory pathway in Gram-negative bacteria. *Microbiol. Rev.* **57** 50-108.

Punyagupta, S. 1989. Review of 686 cases and presentation of a new clinical classification. In: S.Punyagupta, T. Sirisanthana and B. Stapatayavong (Eds). Melioidosis. pp 217-229. Bangkok, Bangkok Medical Publisher.

Puthucheary, S. D., J. Vadivelu, C. Cecile, W. K. Thong and G. Ismail. 1996. Short report: electron microscopic demonstration of extracellular of *Burkholderia pseudomallei. Am. J. Trop. Med. Hyg.* **54** 313-314.

Puthucheary, S. D., N. Parasakthi and M. K. Lee. 1992. Septicemic melioidosis: a review of 50 cases from Malaysia. *Trans. R. Soc. Trop. Med. Hyg.* **86** 683-685.

Rani, R., M. A. Fernandez-Vina, S. A. Zaheer, K. R. Beena and P. Stastny. 1993. Study of HLA class II alleles by PCR oligotyping in leprosy patients from North India. *Tissue Antigens* **42** 133-137.

Rapaport, F. T., J. W. Millar and J. Ruch. 1961. Endotoxic properties of *Pseudomonas pseudomallei*. *Arch. Pathol.* **71** 89-96.

Rattanathongkom, A., R. W. Sermswan and s. Wongratanacheewin. 1997. Detection of *Burkholderia pseudomallei* in blood samples using polymerase chain reaction. *Mol. Cell. Probes* **11** 25-31.

Razak, C., G. Ismail, N. Embi and O. Omar. 1986. Protection studies using whole cells and partially purified toxic material (PPTM) of *Pseudomaonas pseudomallei*. *Malays. Appl. Biol.* **15** 105-111.

Redfearn, M. S. 1964. Toxic lipid: isolation from *Pseudomonas pseudomallei*. *Science* **146** 648-649.

Redfearn, M. S., N. J. Palleroni and R. Y. Stanier. 1966. A comparative study of *Pseudomonas pseudomallei* and *Bacillus mallei*. *J. Gen. Microbiol.* **43** 293-313.

Reubi, J. C., J. A. Laissue, B. Waser, D. L. Steffen, R. W. Hipkin and A. Schonbrunn. 1999. Immunohistochemical detection of somatostatin sst2a receptors in the lymphatic, smooth muscular, and peripheral nervous systems of human gastrointestinal tract: facts and artifacts. *J. Clin. Endoc. Metabol.* **84** 2942-2950.

Rietschel, E. T., H. Brade, O. Holst, L. Brade, S. Muller-Loennies and U. Mamat. 1994. Bacterial endotoxin: chemical constitution, biological recognition, host response, and immunological detoxification. *Curr. Top. Microbiol. Immunol.* **216** 39-81.

Roberts, I. S. 1996. The biochemistry and genetics of capsular polysaccharide production in bacteria. *Ann. Rev. Microbiol.* **50** 285-315.

Roberts, J. A. 1996. Tropism in bacterial infections: urinary tract infections. *J. Urol.* **156** 1552-1559.

Rode, J. W. and D. A. Webling. 1981. Melioidosis in the northern territory of Australia. *Med. J. Aust* **1** 181-184.

Rodriguez, A., Y. Quentin, A. Lazdunski, V. Mejean and M. Foglino. 2000. Two-component systems in *Pseudomonas aeruginosa*: why so many? *Trends Microbiol.* **8** 498-504.

Roychoudhury, S. N., A. Zielinski, A. J. Ninfa, N. E. Allen, L. N. Jungheim, T. I. Nicas and A. M. Chakrabarty. 1993. Inhibitors of 2-component signal transduction systems- inhibition of alginate gene activation in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **90** 965-969.

Rozalski, A., Z. Sidorezyk and K. Kotelko. 1997. Potential virulence factors of *Proteus* bacilli. *Microbiol. Mol. Biol. Rev.* **61** 65-74.

Rubin, F. A., D. J. Kopecko, K. F. Noon and L. S. Baron. 1985. Development of a DNA probe to detect *Salmonella typhi. J. Clin. Microbiol.* **22** 600-605.

Rugdech, P., N. Anuntagool and S. Sirisinha. 1995. Monoclonal antibodies to *Pseudomonas pseudomallei* and their potential for diagnosis of melioidosis. *Am. J. Trop. Med. Hyg.* **52** 231-235.

Ruiz-Lozano, J. M. and P. Bonfante. 1999. Identification of a putative P-transporter operon in the genome of a *Burkholderia* strain living inside the arbuscular mycorrhizal fungus *Gigaspora margarita. J. Bacteriol.* **181** 4106-4109.

Russell, P. UK Melioidosis Meeting. 2000. Personal Communication.

Russo, T. A., S. Wenderoth, U. B. Carlino, J. M. Merrick and A. J. Lesse. 1998. Identification, genomic organization, and analysis of the group III capsular polysaccharide genes *kpsD, kpsM, kpsT, and kpsE* from an extraintestinal isolate of *Escherichia coli* (CP9, O4/K54/H5). *J. Bacteriol.* **180** 338-349.

Sajjan, U. S. and J. F. Forstner. 1992. Identification of the mucin-binding adhesin of *Pseudomonas capacia* isolated from patients with cystic fibrosis. *Infect. Immun.* **60** 1434-1440.

Salisbury, W. A. and J. J. Likos. 1970. *Pseudomonas pseudomallei* in a case of chronic melioidosis. *Am. J. Clin. Path.* **54** 602-606.

Salmond, G. P. C. and P. J. Reeves. 1993. Membrane traffic wardens and protein secretion in Gram-negative bacteria. *Trends Biochem.Sci.* **18** 7-12.

Salyers, A. A. and D. D. Whitt. 1994. Bacterial pathogenesis: A molecular approach. Washington, American Society for Microbiology.

Sambrook, J., E. F. Fritsch and T. Maniatis. 1989. Molecular cloning: A laboratory manual. Cold Spring Habor Laboratory Press. 2nd Edn. New York.

Sanford, J. P. 1985. *Pseudomonas* species (including melioidosis and glanders). In: G.L.Mandell, R. G. Douglas Jr J. E. Bennett (Eds). Principles and Practice of Infectious Disease. 2nd, pp 2003-2009. New York, Wiley & Sons.

Sanford, J. P. 1990. Principles and Practice of Infectious Diseases. G.L.Mandell, R. G. Douglas Jr J. E. Bennett (Eds). pp 1692-1696. New York, Churchill Livingstone.

Sanford, J. P. 1995. Principles and Practice of Infectious Diseases. G.L.Mandell, R. G. Douglas Jr J. E. Bennett (Eds). 4th Edn, pp 2003-2009. New York, Churchill Livingstone.

Sanford, J. P. and W. L. Jr. Moore. 1971. Recrudescent melioidosis: a south east Asian legacy. *Am. Rev. Respir. Dis.* **104** 452-457.

Sangiambut, S., M. Mahfouz, M. Gilpin, K. A. Brown and N. Fairweather. 2000. Identification of genes involved in two-component signal transduction in *Burkholderia pseudomallei*. SGM/SfAM Millennium Meeting 10-14 April, 33. Abstract.

Santanirand, P., V. S. Harley, D. A. B. Dance B. S. Drasar and G. J. Bancroft. 1999. Obligatory role of gamma interferon for host survival in a murine model of infection with *Burkholderia pseudomallei*. *Infect. Immun.* **67** 3593-3600.

Scarlato, V., A. Prugnola, B. Arico and R. Rappuoli. 1990. Positive transcriptional feedback at the *bvg* locus controls expression of virulence factors in *Bordetella pertussis*. *Proc. Natl. Acad. Sci. USA* **87** 6753-6757.

Scarlato, V., B. Arico, M. Domenighini and R. Rappuoli. 1993. Environmental regulation of virulence factors in *Bordetella* species. *BioEssays* **15** 99-105.

Schaechter, M. 1998. Damage by microbial agents. In: M.Schaechter, N. C. Engleberg B. I. Eisenstein and G. Medoff (Eds). Mechanism of Microbial Disease. 3rd Edn, pp 115-125. Maryland, Williams & Wilkins.

Schippers, B., A. W. Bakker and P. A. Bakker. 1987. Interactions of deleterious and beneficial rhizosphere microorganisms and the effect of cropping practices. *Annu. Rev. Phytopathol.* **25** 339-358.

Schmid, M. B. 1990. More than just "histone-like" proteins. *Cell* **63** 451-453.

Schmitt, M. P. and R. K. Holmes. 1994. Cloning, sequence, and footprint analysis of two promoter/operators from *Corynebacterium diphtheriae* that are regulated by diphtheria repressor (DtxR) and iron. *J. Bacteriol.* **176** 1141-1149.

Schnaitman, C. A. and J. D. Klena. 1993. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol. Rev.* **57** 655-682.

Schroder, I., C. D. Wolin, R. Cavicchioli and R. P. Gunsalus. 1994. Phosphorylation and dephosphorylation of the NarQ, NarX, and NarL proteins of the nitrate-dependent two-component regulatory system of *Escherichia coli*. *J. Bacteriol.* **176** 4985-4992.

Schumacher, D. A. and K. E. Klose. 1999. Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. *J. Bacteriol.* **181** 1508-1514.

Schuster, S. C., R. V. Swanson, L. A. Alex, R. B. Bourret and M. I. Simon. 1993. Assembly and function of a quaternary signal transduction complex monitored by surface plasmon resonance. *Nature* **365** 343-347.

Sermswan, R. W., S. Wongratanacheewin, U. Tattawasart and S. Wongwajana. 1994. Construction of a specific DNA probe for diagnosis of melioidosis and use as an epidemological marker of *Pseudomonas pseudomallei*. *Mol. Cell. Probes* **8** 1-9.

Sexton, M. M., A. L. Jones, W. Chaowagul and D. E. Woods. 1994. Purification and characterization of a protease from *pseudomonas pseudomallei*. *Can. J. Microbiol.* **40** 903-910.

Shinde, U. and M. Inouye. 1993. Intramolecular chaperones and protein folding. *Trends Biochem. Sci.* **18** 442-446.

Shine, J. and L. Dalgarno. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosomal binding sites. *Proc. Natl. Acad. Sci. USA* **71** 1342-1346.

Simon, R., B. Hote, B. Klauke and B. Kosier. 1991. Isolation and characterization of insertion sequence elements from Gram-negative bacteria by using a broad host range positive selection vector. *J. Bacteriol* **173** 1502-1508.

Simon, R., U. Priefer and A. Puhler. 1983. A broad host range mobilisation system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. *BioTechnol.* **1** 784-971.

Simpson, A. J. H. and V. Wuthiekanun. 2000. Interaction of insulin with *Burkholderia pseudomallei* may be caused by a preservative. *J. Clin. Pathol.* **53** 159-160.

Sirikulchayanonta, V. and T. Subhadrabandhu. 1994. Melioidosis: another aetiology of granulomatous osteomyelitis. *Clin. Orthop.* **308** 183-186.

Sirisinha, S. 1991. Diagnostic value of serological tests for melioidosis in an endemic area. *Asian Pacific J. Allergy Immunol.* **9** 1-3.

Sirisinha, S., N. Anuntagool, P. Intachote, V. Wuthiekanun, S. D. Puthucheary, J. Vadivelu and N. J. White. 1998. Antigenic differences between clinical and environmental isolates of *Burkholderia pseudomallei*. *Microbiol. Immunol.* **42** 731-737.

Sirisinha, S., N. Anuntagool, T. Dharakul, P. Ekpo, S. Wongratanacheewin, P. Naigowit, B. Petchclai, V. Thamliktikul and Y. Suputtamongkol. 2000. Recent developments in laboratory diagnosis of melioidosis. *Acta Tropica* **74** 235-245.

Slauch, J. M., A. A. Lee and M. J. Mahan. 1996. Molecular characterisation of the *oafA* locus responsible for acetylation of *Salmonella typhimurium* O-antigen. OafA is a member of a family of integral membrane trans-acylases. *J. Bacteriol* **178** 5904-5909.

Sledjeski, D. D. and S. Gottesman. 1996. Osmotic shock induction of capsule synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **178** 1204-1206.

Smith, C. J., J. C. Allen, M. N. Embi, O. Othman, N. Razak and G. Ismail. 1987. Human Melioidosis: an emerging medical problem. *MIRCEN Journal* **3** 343-366.

Smith, G. A., H. Marquis, S. Jones, N. C. Johnston, D. A. Portony and H. Goldfine. 1995. The two distinct phospholipases C of *Listeria monocytogenes* have overlapping roles in escape from a vacuole and cell-to-cell spread. *Infect. Immun.* **63** 4231-4237.

Smith, H. 1990. Pathogenicity and the microbe *in vivo*. *J. Gen. Microbiol.* **136** 377-383.

Smith, H. 2000. Host factors that influence the behaviour of bacterial pathogens *in vivo*. *Int. J. Med. Microbiol.* **290** 207-213.

Smith, M. D., B. J. Angus, V. Wuthiekanun and N. J. White. 1997. Arabinose assimilation defines a nonvirulent biotype of *Burkholderia pseudomallei*. *Infect. Immun.* **65** 4319-4321.

Smith, M. D., V. Wuthiekanun, A. L. Walsh, N. Teerawattanasook, V. Desakorn, Y. Suputtamongkol, T. L. Pitt and N. J. White. 1995. Latex agglutination for rapid detection of *Pseudomonas pseudomallei* antigen in urine of patients with melioidosis. *J. Clin. Pathol.* **48** 174-176.

Smith, M. D., V. Wuthiekanun, A. L. Walsh and N. J. White. 1995. Quantitative recovery of *Burkholderia pseudomallei* from soil in Thailand. *Trans. Roy. Soc. Trop. Med. Hyg.* **89** 488-490.

Sokol, P. A. 1986. Production and utilization of pyochelin by clinical isolates of *Pseudomonas cepacia*. *J. Clin. Microbiol.* **23** 560-562.

Sokol, P. A. and D. E. Woods. 1988. Effect of pyochelin on *Pseudomonas cepacia* respiratory infections. *Microb. Pathog.* **5** 197-205.

Songsivilai, S. and T. Dharakul. 2000. Multiple replicons constitute the 6.5-megabase genome of *Burkholderia pseudomallei*. *Acta Tropica* **74** 169-179.

Sookpranee, M., P. Boonma, K. Nuntirooj, P. Lumbiganon, V. Cowsuwon, A. Pauapairoji, P. Puapermpoonsiri, S. Paupermpoonsiri and V. Laopaiboon. 1989. Melioidosis at Srinagarind hospital. In: S.Punyagupta, T. Sirisanthana and B. Stapatayavong (Eds). Melioidosis. pp 34-47. Bangkok, Thailand, Bangkok Med. Pub.

Sookpranee, M., P. Boonma, W. Susaengrat, K. Bhuripanyo and S. Punyagupta. 1992. Multicenter prospective randomized trial comparing ceftazimide plus co-trimoxazole with chloramphenicol plus doxycycline and co-trimoxazole for treatment of severe melioidosis. *Antimicrob. Agents Chemother.* **36** 158-162.

Soubrier, F., S. Levy-Schil, J. Mayaux, D. Petre and A. Arnaud. 1992. Cloning and primary structure of the wide-spectrum amides from *Brevibacterium* sp. R312: high homology to the amiE product from *Pseudomonas aeruginosa*. *Gene* **116** 99-104.

Southern, E. M. 1975. Detection of specific sequence among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98** 503-517.

Soy, S. Y., P. Y. Chang, Y. K. Leung, W. K. Lam and D. Y. C. Yu. 1983. Successful treatment of melioidosis caused by a multiresistant strain in an immunocompromised host with third generation cepharosporins. *Am. Rev. Resp. Dis.* **127** 650-654.

Spiller, O. B., C. L. Harris and B. P. Morgan. 1999. Efficient generation of monoclonal antibodies against surface-expressed proteins by hyperexpression in rodent cells. *J. Immunol. Metho.* **224** 51-60.

Spitznagel, J. K., F. G. Dalldorf, M. S. Leffell, J. D. Folds, R. H. Welsh, M. H. Cooney and L. E. Martin. 1974. Characterization of azurophil and specific granules purified from human polymorphonuclear leukocytes. *Lab. Invest.* **30** 774-785.

Sprich, M. G. 1956. Phagocytosis of *M. tuberculosis* and *M. smegmatis* stained with indicator dyes. *Am. Rev. Tuberc.* **74** 552-565.

Stanton, A. T. and W. Fletcher. 1921. Melioidosis, a new disease of the tropics. *Trans. Fourth Cong. Far Eastern Associ. Trop. Med.* **2** 196-198.

Stanton, A. T. and W. Fletcher. 1932. Melioidosis studies from the Institute of Medical Research, Federated Malay States Vol. 21. John Balc & Sons and Danielson Ltd. London.

Steinmetz, I., A. Reganzerowski, B. Brenneke, S. Haussler, A. Simpson and N. J. White. 1999. Rapid identification of *Burkholderia pseudomallei* by latex agglutination based on an exopolysaccharide-specific monoclonal antibody. *J. Clin. Microbiol.* **37** 225-228.

Steinmetz, I., M. Nimtz, V. Wary, S. Haubler, A. Reganzerowski and B. Brenneke. 2000. Exopolysaccharides of *Burkholderia pseudomallei*. *Acta Tropica* **74** 211-214.

Steinmetz, I., M. Rohde and B. Brenneke. 1995. Purification and characterization of an exopolysaccharide of *Burkholderia (Pseudomonas) pseudomallei*. *Infect. Immun.* **63** 3959-3965.

Stewart, R. C. 1993. Activating and inhibitory mutations in the regulatory domain of CheB, the methylesterase in bacterial chemotaxis. *J. Biol. Chem.* **268** 1921-1930.

Stock, A. M., V. L. Robinson and P. N. Goudreau. 2000. Two-component signal transduction. *Annu. Rev. Biochem.* **69** 183-215.

Stock, J. B., A. J. Ninfa and A. M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* **53** 450-490.

Stock, J. B., A. M. Stock and M. Mottonen. 1990. Signal transduction in bacteria. *Nature* **334** 857-871.

Stock, J. B., M. G. Surette, M. Levit and A. M. Stock. 1995. Two-component signal transduction system: Structure-function relationships and mechanisms of catalysis. In: J.A.Hoch & T.J.Silhavy (Eds). Two-component signal transduction. pp 25-32. Washington, DC, American Society for Microbiology.

Stout, V. 1994. Regulation of capsule synthesis includes interactions of the RcsC/RcsB regulatory pair. *Res. Microbiol* **145** 389-392.

Stout, V. 1996. Identification of the promoter region for the colanic acid polysaccharide biosynthetic genes in *Escherichia coli* K-12. *J. Bacteriol.* **178** 4273-4280.

Stout, V., A. T. Cabassa, M. R. Maurizi, D. Gutnick and S. Gottesman. 1991. RcsA, an unstable positive regulator of capsular polysaccharide synthesis. *J. Bacteriol.* **173** 1738-1747.

Stout, V. and S. Gottesman. 1990. RcsB and RcsC: a two-component regulator of capsule synthesis in *Escherichia coli*. *J. Bacteriol.* **172** 659-669.

Straley, S. C. and R. D. Perry. 1995. Environmental modulation of gene expression and pathogenesis in *Yersinia*. *Trends Microbiol.* **3** 310-317.

Strauch, M. A., D. deMendoza and J. A. Hoch. 1992. Cis-unsaturated fatty-acids specificity inhibit a signal-transducing protein-kinase required for initiation of sporulation in *Bacillus subtilis*. *Mol. Microbiol.* **6** 2909-2917.

Straus, D. C., M. K. Lonon, D. E. Woods and C. W. Garner. 1989. Production of an extracellular toxic complex by various strains of *Pseudomonas cepacia*. *J. Med. Microbiol* **30** 17-22.

Strauss, J. M., A. D. Alexander, G. Rapmund, E. Gan and A. E. Dorsey. 1969. Melioidosis in Malaysia. III: Antibodies to *Pseudomonas pseudomallei* in the human population. *Am. J. Trop. Med. Hyg.* **18** 703-707.

Strauss, J. M., D. W. Ellison, E. Gan, S. Jason, J. L. Marcaselli and G. Papmund. 1969. Melioidosis in Malaysia. IV. Intensive ecological study of Carey Island, Selangor, for *Pseudomonas pseudomallei*. *Med. J. Malaya.* **24** 94-100.

Subhadrabandhu, T., S. Prichasuk and B. Sathapatayavongs. 1995. Localised melioidotic osteomyelitis. *J.Bone Joint Surg.* **77** 445-449.

Supputtamongkol, Y., D. Kwiatkowski, D. A. B. Dance, W. Chaowagul and N. J. White. 1992. Tumor necrosis factor in septicaemic melioidosis. *J. Infect. Dis.* **165** 561-564.

Suputtamongkol, Y., A. J. Hall, D. A. B. Dance, W. Chaowagul, A. Rajchanuvong, M. D. Smith and N. J. White. 1994. The epidemiology of melioidosis in Ubon Ratchatani, northeast Thailand. *Int. J. Epidemiol.* **23** 1082-1090.

Sura, T., M. D. Smith, G. M. Cowan, A. L. Walsh, N. J. White and S. Krishna. 1997. Polymerase chain reaction for the detection of *Burkholderia pseudomallei*. *Diagn. Microbiol. Infect. Dis.* **29** 121-127.

Takasuka, N., T. Tokunaga and K. S. Akagawa. 1991. Preexposure of macrophages to low-doses of lipopolysaccharide inhibits the expression of the tumor necrosis factor alpha messenger RNA but not of IL-Beta messenger RNA. *J. Immunol.* **146** 3824-3830.

Tan, A. P., M. H. Pui and L. K. Tan. 1995. Imaging patterns in melioidosis. *Australas. Radiol.* **39** 260-264.

Tanphaichitra, D. 1989. Tropical disease in the immunocompromised host: melioidosis and pythiosis. *Rev. Infect. Dis.* **11 (Suppl)** S1629-S1643.

Thin, R. N., F. M. Grovers, G. Rapmund and M. Mariappan. 1971. *Pseudomonas pseudomallei* in surface water of Singapore. *Singapore Med. J.* **12** 181-185.

Thomas, A. D. 1981. Prevalence of melioidosis in animals in Northern Queensland. *Aust.Vet. J.* **57** 146-148.

Thomas, A. D., J. C. Forbes-Faulkner, J. H. Norton and K. F. Trueman. 1988. Clinical and pathological observation on goats experimentally infected with *Pseudomonas pseudomallei*. *Aust.Vet. J.* **65** 43-46.

Thomas, A. D., J. H. Norton, J. C. Forbes-Faulkner and G. Woodland. 1981. Melioidosis in an intensive piggery. *Aust. Vet. J.* **57** 144-145.

Toder, D. S. 1998. *Pseudomonas aeruginosa*: a ubiquitous pathogen. In: M.Schaechter, N. C. Engleberg B. I. Eisenstein and G. Medoff (Eds). Mechanisms of Microbial Diseases. 3rd Edn, pp 199-204.

Tong, S., S. Yang, Z. Lu and W. He. 1996. Laboratory investigation of ecological factors influencing the environmental presence of *Burkholderia pseudomallei*. *Microbiol. Immunol.* **40** 451-453.

Torres-Cabassa, A., S. Gottesman, R. D. Fraederick, P. J. Dolph and D. L. Coplin. 1987. Control of extracellular polysaccharide synthesis in *Erwinia stewartii* and *Escherichia coli* K-12: a common regulatory function. *J. Bacteriol.* **169** 4525-4531.

Torres-Cabassa, A. S. and S. Gottesman. 1987. Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. *J. Bacteriol.* **169** 981-989.

Torres, M., R. Cano and J. C. Palomares. 1991. Evaluation of a DNA probe of plasmid origin for the detection of *Neisseria gonorrhoeae* in cultures and clinical specimens. *Mol. Cell. Probes* **5** 49-54.

Towbin, H. M., T. Staehelin and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76** 4350-4354.

Trakulsomboon, S., D. A. B. Dance, M. D. Smith, N. J. White and T. L. Pitt. 1997. Ribotype differences between clinical and environmental isolates of *Burkholderia pseudomallei*. *J. Med. Microbiol* **46** 565-570.

Trakulsomboon, S., T. L. Pitt and D. A. B. Dance. 1994. Molecular typing of *Pseudomonas pseudomallei* from imported primates in Britain. *Vet. Rec.* **135** 65-66.

Ulett, G. C., N. Ketkunam and R. G. Hirst. 1998. Macrophage-lymphocyte interactions mediate anti-*Burkholderia pseudomallei* activity. *FEMS Immunol. Med. Microbiol.* **21** 283-286.

Ulett, G. C., N. Ketkunam and R. G. Hirst. 2000. Proinflammatory cytokine mRNA responses in experimental *Burkholderia pseudomallei* infection in mice. *Acta Tropica* **74** 229-234.

Ullrich, M., A. Penaloza-Vazquez, A. Bailey and C. L. Bender. 1995. A modified two-component regulatory system is involved in temperature-dependent biosynthesis of the *Pseudomonas syringae* phytotoxin coronatine. *J. Bacteriol.* **177** 6160-6169.

Utusumi, R., S. Katayama, M. Taniguchi, T. Horie, M. Ikeda, S. Igaki, H. Nakagawa, A. Miwa, H. Tanabe and M. Noda. 1994. Newly identified genes involved in the signal transduction of *E. coli* K-12. *Gene* **140** 73-77.

Van Alphen, W. and B. Lugtenberg. 1977. Influence of osmolarity of the growth medium on the outer membrane protein pattern of *Escherichia coli*. *J. Bacteriol.* **131** 623-630.

Van Regenmortel, M. H. 1998. From absolute to exquisite specificity. Reflections on the fuzzy nature of species, specificity and antigenic sites. *J. Immunol. Methods.* **216** 37-48.

Vedros, N. A., D. Chow and E. Liong. 1988. Experimental vaccine against *Pseudomonas pseudomallei* infections in captive cetaceans. *Dis. Aquatic Org.* **5** 157-161.

Vedros, N. A., E. Liong and D. D. Hammond. 1983. A potential vaccine for melioidosis in marine mammals. *Proc.14th Annual Conf. Intl. Assoc. Aquatic Anim. Med.* 17.

Vesselinova, A., H. Najdenski, S. Nikolova and V. Kussovski. 1996. Experimental Melioidosis in hens. *J. Vet. Med.* **B 43** 371-378.

Virlogeux, I., H. Waxin, C. Ecobichon, J. O. Lee and M. Y. Popoff. 1996. Characterization of the *rcsA* and *rcsB* genes from *Salmonella typhi: rscB* through *tviA* is involved in regulation of Vi antigen synthesis. *J. Bacteriol.* **178** 1691-1698.

Vogel, K. and A. Hinnen. 1990. The yeast phosphatase system. *Mol. Microbiol.* **4** 2013-2017.

Volz, K. 1993. Structural conservation in the CheY superfamily. *Biochemistry* **32** 11741-11753.

Volz, K. 1995. Structural and functional conservation in response regulators. In: J.A.Hoch & T.J.Silhavy (Eds). Two-component signal transduction. pp 53-64. Washington, DC, ASM Press.

Volz, K. and P. Matsumura. 1991. Crystal structure of *Escherichia coli* CheY refined at 1.7°A resolution. *J. Biol. Chem.* **266** 15511-15519.

Von Heijne, G. 1992. Membrane protein structure prediction, hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225** 487-494.

Vorachit, M., K. Lam, P. Jayanetra and J. W. Costerton. 1993. Resistance of *Pseudomonas pseudomallei* growing as a biofilm on silastic discs to ceftazidime and co-trimoxazole. *Antimicrob. Agents Chemother.* **37** 2000-2002.

Vorachit, M., K. Lam, P. Jayanetra and J. W. Costerton. 1995. Electron microscopy study of the mode of growth of *Pseudomonas pseudomallei in vitro* and *in vivo*. *J. Trop. Med. Hyg.* **98** 379-391.

Vorachit, M., P. Chongtrakool, S. arkomsean and S. Boonsong. 2000. Antimicrobial resistance in *Burkholderia pseudomallei*. *Acta Tropica* **74** 139-144.

Vriesma, A. J., M. J. Dankert and S. A. J. Zaat. 2000. A shift from oral to blood pH is a stimulus for adaptive gene expression of *Streptococcus gordonii* CH1 and induces protection against oxidative stress and enhanced bacterial growth by expression of *msrA*. *Infect. Immun.* **68** 1061-1068.

Wacharotayankun, R., Y. Arakawa, M. Ohta, T. Hasegawa, M. Mori, T. Horii and N. Kato. 1992. Involvement of *rcsB* in *Klebsiella* K2 capsule synthesis in *Escherichia coli* K-12. *J. Bacteriol.* **174** 1063-1067.

Walsh, A. L. and V. Wuthiekanun. 1996. The laboratory diagnosis of melioidosis. *Br. J. Biom. Sci.* **53** 249-253.

Walsh, A. L., M. D. Smith, V. Wuthiekanun, Y. Suputtamongkol, V. Desakorn, W. Chaowagul and N. J. White. 1994. Immunofluorescence microscopy for the rapid diagnosis of Melioidosis. *J. Clin. Pathol.* **47** 377-379.

Walther, P., M. Muller and M. E. Schweingruber. 1984. The ultrastructure of the cell-surface and plasma-membrane of exponential and stationary phase cells of *Schizosaccharomyces pombe*, growth in different media. *Arch. Microbiol* **137** 128-134.

Walton, K. M. and J. E. Dixon. 1993. Protein-tyrosine phosphatase. *Ann. Rev. Biochem.* **62** 101-120.

Wang, C. Y., B. H. Yap and A. E. Delilkan. 1993. Melioidosis pneumonia and blast injury. *Chest* **103** 1897-1899.

Wang, W. C., L. M. Goldman, D. M. Schleider, M. M. Appenheimer, J. R. Subjeck, E. A. Repasky and S. S. Evans. 1998. Fever-range hyperthermia enhances L-selection-dependent adhesion of lymphocytes to vascular endothelium. *J. Immunol.* **160** 961-969.

Watson, M. E. E. 1984. Compilation of published signal sequences. *Nucleic Acids Res.* **12** 5145-5164.

Watson, N., D. S. Dunyak, E. L. Rosey, J. L. Slonczewski and E. R. Olson. 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J. Bacteriol.* **174** 530-540.

Watson, R. R., E. E. Galyov, S. M. Paulin, P. W. Jones and T. S. Wallis. 1998. Mutation of *invH*, but not *stn*, reduces Salmonella-induced enteritis in cattle. *Infect. Immun.* **66** 1432-1438.

Watt, W. B. and A. M. Dean. 2000. Molecular-functional studies of adaptive genetic variation in prokaryotes and eukaryotes. *Ann. Rev. Gen.* **34** 593-622.

Wedemayer, G. J., P. A. Patten, L. H. Wang, P. G. Schultz and R. C. Stevens. 1997. Structural insights into the evolution of an antibody combinding site. *Science* **276** 1665-1669.

Wehland, M. and F. Bernhard. 2000. The RcsAB Box: Characterization of a new operator essential for the regulation of exopolysaccharide biosynthesis in enteric bacteria. *J. Biol. Chem.* **275** 7013-7020.

Weinberg, A. N. and H. M. Heller. 1997. Infectious Disease of the Lung. pp 2413-2430. New York,Thieme-Stratton.

Weinberg, E. D. and G. A. Weinberg. 1995. The role of iron in infection. *Curr. Opin. Infect. Dis.* **8** 164-169.

Wendrich, T. M. and M. A. Marafiel. 1997. Cloning and characterization of a *relA/spoT* homologue from *Bacillus subtilis*. *Mol. Microbiol.* **26** 65-79.

Werner, A. S., C. G. Cobbs, D. Kays and E. W. Hook. 1967. Studies on the bacteremia of bacterial endocarditis. *J. Am. Med. Assoc.* **202** 199-203.

Wernery, R., J. Kinne, J. Haydn-Evans and A. Ul-Haq. 1998. Melioidosis in a seven year old camel, a new disease in the United Arab Emirates (UAE). *J. Camel Prac. Res.* **4** 141-143.

Wetmore, P. W. and W. S. Gochenour. 1956. Comparative studies of the genus *Malleomyces* and selected *Pseudomonas* species. *J. Bacteriol.* **72** 79-89.

White, N. J. 1994. Melioidosis. *Zbl. Bakt.* **280** 439-443.

White, N. J., W. Chaowagul, V. Wuthiekanun, D. A. B. Dance, Y. Watanagoon and N. Pitakwatchara. 1989. Halving of mortality of severe melioidosis by ceftazidime. *Lancet* **ii** 697-701.

Whitmore, A. 1913. An account of a glanders-like disease occurring in Rangoon. *J. Hyg.* **13** 135-148.

Whitmore, A. and C. S. Krishnaswami. 1912. An account of the discovery of a hitherto undescribed infective disease occurring among the population of Rangoon. *Indian Med. Gaz.* **47** 262-267.

Wilkinson, S. G. and T. L. Pitt. 1995. *Burkholderia* (*Pseudomonas*) *cepacia*: pathogenicity and resistance. *Rev. Med. Microbiol.* **6** 10-17.

Winstanley, C. and C. A. Hart. 2000. Presence of Type III secretion genes in *Burkholderia pseudomallei* correlates with ara⁻ phenotypes. *J. Clin. Microbiol.* **38** 883-885.

Winstanley, C., B. A. Hales, J. E. Corkill, M. J. Gallagher and C. A. Hart. 1998. Flagellin gene variation between clinical and environmental isolates of *Burkholderia pseudomallei* contrasts with the invariance among clinical isolates. *J. Med. Microbiol.* **47** 689-694.

Wohlfarth, S., C. Hoesche, C. Strunk and U. K. Winkler. 1992. Molecular genetics of the extracellular lipase of *Pseudomonas aeruginosa* PAO1. *J. Gen. Microbiol.* **138** 1325-1335.

Wong, K. C., B. S. W. Ho, S. I. Egglestone and W. H. P. Lewis. 1995. Duplex PCR system for simultaneous detection of *Neisseria gonorrhoeae* and *Chlamydia trachomatis* in clinical specimens. *J. Clin. Pathol.* **48** 101-104.

Wong, K. T., S. D. Puthcheary and J. Vadivelu. 1995. The histopathology of human Melioidosis. *Histopathol.* **26** 51-55.

Wong, S. M., P. A. Carroli, L. G. Rahme, F. M. Ausubel and S. B. Calderwood. 1998. Modulation of expression of the ToxR regulation in *Vibrio cholerae* by a member of the two-component family of response regulators. *Infect. Immun.* **66** 5854-5861.

Wongratanacheewin, S., K. Komutrin and R. W. Sermswan. 2000. Use of multiplex PCR patterns as genetic markers for *Burkholderia pseudomallei. Acta Tropica* **74** 193-199.

Wongratancheewin, S., U. Tattawasart and V. Lulitanond. 1990. An avidin-biotin enzyme-linked immunosorbent assay for the detection of *Pseudomonas pseudomallei* antigens. *Trans. R. Soc.Trop. Med. Hyg.* **84** 429-430.

Wongwanich, S., P. Chotanachan, E. Kondo and K. Kanai. 1996. Multifactorial pathogenic mechanisms of *Burkholderia pseudomallei* as suggested from comparison with *Burkholderia cepacia. Southeast Asian J. Trop. Med. Public Health* **27** 111-118.

Woodgett, J. R., K. L. Gould and T. Hunter. 1986. Substrate-specificity of protein-kinase-c - use of synthetic peptides corresponding to physiological sites as probes for substrate recognition requirements. *Eur. J. Biochem.* **161** 177-184.

Woods, D. E., A. L. Jones and P. A. Hill. 1993. Interaction of insulin with *Pseudomonas pseudomallei. Infect. Immun.* **61** 4045-4050.

Woods, M. L., B. J. Currie and D. M. Howard. 1992. Neurological melioidosis: seven cases from the Northern Territory of Australia. *Clin. Infect. Dis.* **15** 163-169.

Wren, B. W., S. M. Colby, R. R. Cuberley and M. J. Pallen. 1992. Degenerate PCR primers for the amplification of fragments from genes encoding response regulators from a range of pathogenic bacteria. *FEMS Microbiol. Lett.* **99** 287-292.

Wurgler-Murphy, S. M. and H. Saito. 1997. Two-component signal transducers and MAPK cascades. *TIBS* **22** 172-176.

Wuthiekanun, V., D. A. B. Dance, Y. Wattanagoon, Y. Supputtamongkol, W. Chaowagul and N. J. White. 1990. The use of selective media for the isolation of *Pseudomonas pseudomallei* in clinical practice. *J. Med. Microbiol.* **33** 121-126.

Wuthiekanun, V., M. D. Smith, D. A. B. Dance and N. J. White. 1995. Isolation of *Pseudomonas pseudomallei* from soil in north-eastern Thailand. *Trans. Roy. Soc. Trop. Med. Hyg.* **89** 41-43.

Wuthiekanun, V., M. D. Smith and N. J. White. 1995. Survival of *Burkholderia pseudomallei* in the absence of nutrients. *Trans. R. Soc.Trop. Med. Hyg.* **89** 491-491.

Wuthiekanun, V., M. D. Smith, D. A. B. Dance, A. L. Walsh, T. L. Pitt and N. J. White. 1996. Biochemical characteristics of clinical and environmental isolates of *Burkholderia pseudomallei. J. Med. Microbiol* **45** 408-412.

Wyman, C., I. Rombel, A. K. North, C. Bustamente and S. Kustu. 1997. Unusual oligomerisation required for activity of NtrC, a bacterial enhancer-binding protein. *Science* **275** 1658-1661.

Xiao, Y., S. Heu, J. Yi, Y. Lu and S. W. Hutcheson. 1994. Identification of a putative alternate sigma factor and characterisation of a multicomponent regulatory cascade controlling the expression of *Pseudomonas syringae* pv. syringae Pss61 *hrp* and *hrmA* genes. *J. Bacteriol.* **176** 1025-1036.

Yabuuchi, E. and M. Arakawa. 1993. *Burkholderia pseudomallei* and Melioidosis: be aware in temperate area. *Microbiol. Immunol.* **37** 823-836.

Yabuuchi, E., L. Wang, M. Arakawa and I. Yano. 1993. Survival of *Pseudomonas pseudomallei* strains at 5°C. *J. Jpn. Assoc. Infect. Dis.* **67** 331-335.

Yabuuchi, E., N. Miyajima, H. Hotta, A. Ohyama and N. Tanaka. 1970. *Pseudomonas cepacia* from blood burn patient. *Med. J. Osaka Univ.* **21** 1-6.

Yabuuchi, E., Y. Kosako, M. Arakawa, H. Hotta and I. Yano. 1992. Identification of Oklahoma isolate as a strain of *Pseudomonas pseudomallei. Microbiol. Immunol.* **36** 1239-1249.

Yahr, T. L., J. Goranson and W. Frank. 1996. Exoenzyme S of *Pseudomonas aeruginosa* is secreted by a type III pathway. *Mol. Microbiol.* **22** 991-1003.

Yang, H., C. D. Kool and P. A. Sokol. 1993. Ability of *Pseudomonas pseudomallei* malleobactin to acquire transferrin-bond, lactoferrin-bond, and cell-derived iron. *Infect. Immun.* **61** 656-662.

Yang, H., W. Chaowagul and P. A. Sokol. 1991. Siderophore production by *Pseudomonas pseudomallei. Infect. Immun.* **59** 776-780.

Yang, S., S. Tong, C. Mo, Z. Jiang, S. Yang, Y. Ma and Z. Lu. 1998. Prevalence of human Melioidosis on Hainan Island in China. *Microbiol. Immunol.* **42** 651-654.

Yang, Y. and M. Inouye. 1991. Intermolecular complementation between two defective mutants of *Escherichia coli* signal-transducing receptors. *Proc. Natl. Acad. Sci. USA* **88** 11057-11061.

Yeh, K. S., L. S. Hancox and S. Clegg. 1995. Construction and characterisation of a *fimZ* mutant of *Salmonella typhimurium. J. Bacteriol.* **177** 6861-6865.

Yh, N., H. Hy, P. J. Chen and M. H. Chang. 1999. Protamine enhances the efficiency of liposome-mediated gene transfer in a cultured human hepatoma cell line. *J. Formosan Med. Assoc.* **98** 562-566.

Zanetti, F., G. De Luca and S. Stampi. 2000. Recovery of *Burkholderia pseudomallei* and *B. cepacia* from drinking water. *Int. J. Food Microbiol.* **59** 67-72.

Zhang, Y., D. J. Isaacman, R. M. Wadowsky, J. Rydquist-White, J. C. Post and G. D. Ehrlich. 1995. Detection of *Streptococcus pneumoniae* in whole blood by PCR. *Mol. Cell. Probes* **11** 25-31.

Zhou, J. and B. G. Spratt. 1992. Sequence diversity within the *argF, fbp* and *recA* genes of natural isolates of *Neisseria meningitidis*: interspecies recombination within the argF gene. *Mol. Microbiol.* **6** 2135-2146.

I dedicate this work to my wife Heba, my son Ahmed and my daughter Nahla