The identification of genes important to the growth of *Staphylococcus aureus* in *in vitro* models mimicking infection

A thesis submitted for the degree of Doctor of Philosophy

by

Michael David Wiltshire, B.Sc.

(University of Warwick)

July 2000

Department of Molecular Biology and Biotechnology,

University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN.

SUMMARY

Staphylococcus aureus is a major pathogen, which causes a wide range of infections. Despite its obvious clinical importance, little is known about the mechanisms of pathogenesis. An *in vitro* model mimicking infection was developed in order to identify putative virulence determinants. The model involves the growth of *S. aureus* in serum under microaerobic conditions. All known virulence factors tested were shown not to be required for growth, or preferentially expressed, in serum.

Tn*917* transposon libraries of *S. aureus* were screened to identify genes preferentially expressed in serum, compared to a nutrient-rich growth medium. 73 clones were identified and the transposon insertion site was characterised for 23 of these clones. Analysis of sequence flanking the transposon insertion revealed the identity of the mutated loci. 10 out of 23 sequenced clones, contained transposons inserted within genes involved in the biosynthesis of the aspartate family of amino acids (lysine, threonine, methionine and isoleucine). These were: the two common pathway enzymes; aspartokinase (*lysC*), and aspartate semialdehyde dehydrogenase (*asd*), along with; dihydrodipicolinate dehydrogenase (*dapA*), and cystathionine γ-synthase (*yjcI*), involved in the biosynthesis of lysine and methionine respectively.

Analysis of methionine biosynthesis indicated that *S. aureus* possesses only a single pathway, which proceeds via cystathionine. Several genes encoding methionine biosynthetic enzymes were found clustered on the *S. aureus* chromosome.

The genes *lysC, asd* and *dapA* were found to be encoded by the first three genes of an eight gene operon, which also contains three other genes involved in lysine biosynthesis. This operon named the *dap* operon, is the major lysine biosynthetic operon of *S. aureus*. *lysC, asd* and *dapA* were all found to be repressed at the transcriptional level primarily by lysine, although factors other than the availability of lysine may be responsible for the regulation of lysine biosynthetic gene expression in serum. *lysC, asd* and *dapA* were all found to be expressed *in vivo,* in a murine pyelonephritis model using both RT-PCR and TaqMan techniques. However, these genes were not found to be important in three murine pathogenicity models.

Finally, in addition to the development of a model mimicking infection, and the identification of genes with a potentially important role *in vivo*, this thesis has enhanced our understanding of both methionine and lysine biosynthesis in *S. aureus*.

This thesis is dedicated to my parents
for all their love and support

Acknowledgments

There are many people that I would like to thank, for their help, advice, and drunken nights out, during the last three and a half wonderful years in Sheffield.

Most importantly I would like to thank my supervisor Professor Simon Foster for his invaluable advice and guidance throughout the project, but most annoyingly of all, for more often than not knowing more about my project than myself. Cheers!

To the many fantastic Post Doc's of E54 especially Mark, Pan, and Mal, thank you for all your assistance in matters scientific. Also, thanks to Jodi for putting up with me next to her for 2 years and helping me to keep things in perspective.

To the people of E54, E58, E46, and further a field throughout MBB, for their help, sport playing, and friendship, thank you and long may it continue.

At SB in Philadelphia, I would like to thank Dr. Martin Burnham and everyone who helped me during my time in Philadelphia, especially Pan for housing me and showing me around.

Lastly, but by no means least importantly, thank you to all my 'Sheffield' friends both past and present who are too numerous to mention, although I would like to single out Laurissa, Steve, Kevin, Kate and Claire, words really cannot describe how much your friendship has meant over the years.....

This work was funded by a BBSRC Case Studentship in collaboration with SmithKline Beecham (well at least that was the company name 3 $^{1}/_{2}$ years ago).

Abbreviations

A_{XYZ}	Absorbance (at XYZ nm)
AAA	Aspartate amino acid
AAAB	Aspartate amino acid biosynthetic
AIP	Autoinducing peptide
AK	Aspartokinase
AK HDH	Aspartokinase homoserine dehydrogenase
Amp	Ampicillin
ASADH	Aspartate semialdehyde dehydrogenase
BHI	Brain heart infusion
β-gal	β-galactosidase
CDM	Chemically defined medium
CF	Cystic fibrosis
Cm	Chloramphenicol
CNS	Central nervous system
CSF	Cerebrospinal fluid
Ct	Threshold cycle
CTBL	Cystathionine β-lyase
CTGL	Cystathionine γ-lyase
CTGS	Cystathionine γ-synthase
D-Ala	D-Alanine
DAP	Diaminopimelate
DEPC	Diethyl-pyrocarbonate
DFI	Differential fluorescence induction
DHPS	Dihydrodipicolinate synthase
DHPR	Dihydrodipicolinate reductase
DIG	Digoxigenin
DMF	N,N-dimethylformide
dNTPs	Deoxyribonucleoside-5'-triphosphate
2DGE	Two-dimensional gel electrophoresis
Ery	Erythromycin
FACS	Fluorescence-activated cell sorting
FSB	Frozen storage buffer
G	Genomic DNA
GAMBIT	Genomic analysis and mapping by *in vitro* transposition
GFP	Green fluorescent protein
HDH	Homoserine dehydrogenase
HUVEC	Human umbilical vein endothelial cells
IVET	*In vivo* expression technology
ivi	*In vitro* induced
Kan	Kanamycin
LA	L agar
LB	L broth

LBE	Lysine biosynthetic enzymes
Lin	Lincomycin
MBE	Methionine biosynthetic enzymes
MRSA	Methicillin-resistant *Staphylococcus aureus*
MU	MUG units
4-MU	4-Methyl umbelliferone
MUG	4-Methyl umbelliferone-β-D-galactopyranoside
n	Sample/population size
NBT	Nitroblue tetrazoium
NTC	No template control
PMNS	Polymorphonuclear leukocytes
ΔRn	Fluorescence of the reporter divided by the fluorescence of the passive reference dye (total fluorescence)
SB	SmithKline Beecham
sCSF	Synthetic cerebrospinal fluid
seg	Sera expressed gene(s)
SDS	Sodium dodecyl sulphate
SOD	Superoxide dismutase
sp.	Species
SSC	Saline sodium citrate (buffer)
STM	Signature-tagged mutagenesis
TAE	Tris-acetate EDTA (buffer)
TE	Tris-EDTA (buffer)
Tet	Tetracycline
THPA	Tetrahydrodipicolinate acetyltransferase
THPS	Tetrahydrodipicolinate succinylase
Tn	Transposon
TSB	Tryptone soy broth
TSS	Toxic shock syndrome
TSST-1	Toxic shock syndrome toxin-1
Tris	Tris(hydroxymethyl)aminomethane
UNG	Uracil-DNA-glycosylase
v/v	Volume for volume
VRSA	Vancomycin-resistant *Staphylococcus aureus*
WT	Wild type
w/v	Weight for volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
σ	Sigma factor
~	Approximately
∇	Tn*917* insertion site

Table of Contents

List of Figures

Chapter 4 Figures:

Chapter 5 Figures:

List of Tables

CHAPTER 1

INTRODUCTION

1.1 The staphylococci

Staphylococci are Gram-positive cocci of 0.5-1.5 μm diameter that characteristically divide in more than one plane to form irregular or grape-like clusters. It was due to this clustering that Ogston in the late nineteenth century, named a group of micrococci, causing inflammation and suppuration *Staphylococcus*, from staphyle' meaning bunch of grapes (Kloos and Lambe, 1991). Currently the genus *Staphylococcus* is comprised of 32 species and eight subspecies, many of whose natural hosts include humans (Kloos and Bannerman, 1994; Kloos and Lambe, 1991). The genus *Staphylococcus* is well defined and phylogenetically belongs to the *Bacillus-Lactobacillus-Streptococcus* cluster of Gram-positive bacteria, with a low G + C content (Woese, 1987).

The staphylococci are non-motile non-spore forming bacteria that are not generally encapsulated. The cell wall peptidoglycan is unique owing to multiple glycine residues in the interpeptide bridge, rendering them susceptible to lysostaphin. Members of the genus are commonly catalase-positive but oxidase-negative, and are generally facultative anaerobes, capable of generating energy by respiration or fermentative pathways. Most species have relatively complex nutritional requirements, and it is not possible to specify a defined medium for all staphylococcal strains, as nutritional requirements vary between strains. However, in general, staphylococci require 5-12 essential amino acids and 2 or 3, B-group vitamins (Wilkinson, 1997).

Staphylococci are typically found on the skin, skin glands and mucus membranes of warm blooded animals. They are sometimes found at other sites including, the mouth, blood, mammary glands and upper respiratory tracts of hosts. The staphylococci are also wide-spread in the environment for example, soil, air and water. They are some of the hardiest of the non-spore-forming bacteria, having the ability to survive for long periods on dry inanimate objects. Staphylococci also demonstrate resistance to heat (withstanding 50 °C for 30 minutes) and tolerance to high concentrations of salt, with most species capable of growth in the presence of 1.7 M (10 % w/v) NaCl (Kloos and Lambe, 1991; Kloos and Bannerman, 1994).

The identification of pathogenic staphylococci is most commonly by the ability to produce coagulase and thus clot plasma (Kloos and Lambe, 1991). The vast majority of

staphylococci are coagulase-negative and do not cause disease. A few coagulase-negative strains do however cause disease, but are generally considered to be opportunistic pathogens, causing infection following trauma, pre-disposing disease, or due to indwelling foreign bodies. To date only three species of staphylococci have been found to be coagulase-positive; *S. intermedius* and *S. hyicus,* which are found in animals (Deveriese *et al.*, 1978; Hajek, 1976) and *S. aureus,* which causes infections in humans.

1.2 *Staphylococcus aureus*

Staphylococcus aureus is the second most common species in the genus *Staphylococcus*, but the best known and by far the most studied staphylococcal species, on account of its frequency and highly versatile pathogenicity. The accolade of the most common staphylococci goes to *S. epidermidis,* which lives on the skin of many people but rarely causes disease. The species name *aureus* ('golden') refers to the fact that on agar, its colonies are pigmented with a bronze colour, while other species form white colonies. *S. aureus* however is probably best distinguished from other species of the genus using the coagulase test (Chapter 1.1). Strains of *S. aureus* are alkaline phosphatase-positive and usefully show no β-galactosidase activity. The acidic amino acids, aspartate and glutamate; and K^+, represent the major amino acids and cation in terms of concentration in the cytoplasmic pool (Wilkinson, 1997). The chromosome of *S. aureus* (strain 8325) has been mapped by a variety of methods including *Sma*I digestion (Pattee *et al.*, 1990) and has been shown to be approximately 2.78 megabase pairs in size. Using the assumption of a 1 kb average gene size, this suggests that *S. aureus* encodes approximately 3,000 genes (Coulter *et al.*, 1998).

1.3 Diseases and infections caused by *S. aureus*

S. aureus is a virulent and frequent pathogen that is probably the most common cause of bacterial infection in humans, although the majority are minor and superficial skin infections. *S. aureus* has been reported as the fourth most frequently isolated pathogen in hospital environments (Kloos and Bannerman, 1994), but is also isolated from community acquired infections. Infections caused by *S. aureus* range from the insignificant, to the life threatening and can be viewed as three general types: superficial lesions such as small skin abscesses and wound infections; systemic and life threatening conditions such as bacteraemia, endocarditis, osteomyelitis, pneumonia and

sepsis; and toxinoses such as toxic shock, toxic epidermal necrolysis, and food poisoning.

S. aureus is also a particular problem following invasive hospital techniques. For example, 20 % of infections involving long-term intravascular canulae (Bluhm, 1985) and prosthetic joint infections, in up to 1 % of patients undergoing hip replacement (Fitzgerald, 1989) are caused by *S. aureus*.

S. aureus produces a range of heat-stable extracellular toxins (Table 1.1) that on ingestion causes food-poisoning. Food poisoning is characterised by nausea, vomiting, abdominal pains and diarrhoea, with symptoms generally presenting within 2-6 hours of consuming contaminated food (Bergdoll, 1989). The foods most regularly implicated include; cooked meat, fish, poultry, and dairy produce. Importantly, death from staphylococcal induced food poisoning is rare.

A more important toxin related disease associated with infection or colonization with *S. aureus* is toxic shock syndrome (TSS). However this disease is only attributed to strains of *S. aureus* containing the gene for toxic shock syndrome toxin-1 (TSST-1). Symptoms of this syndrome include shock, high fever, vomiting, diarrhoea, renal and hepatic dysfunction and a scarlatiniform rash. Menstrual tampons have been shown to be an important accessory factor in the development of TSS (Arbuthnott *et al.*, 1990).

Serious *S. aureus* infections generally occur in association with a predisposing condition. Entry into the body is enhanced by skin damage, as *S. aureus* does not usually penetrate into deep tissues unless the skin or mucosal membranes are damaged or actually cut. This may come about for example by burns, accidental wounds, surgical intervention or associated skin disease. Susceptibility is also caused by chronic debilitating disorders such as acquired immune deficiency or cancer. Thus it is not surprising that *S. aureus* is a pathogen commonly associated with nosocomial infections (Kloos and Bannerman, 1994). Life-threatening diseases of *S. aureus* include bacteraemia, septicaemia, endocarditis, pneumonia and meningitis. Staphylococcal bacteraemia is generally as a result of a local infection with dissemination into the blood stream, but also due to direct inoculation as a result of intravenous drug abuse. Septicaemia differs from bacteraemia, in that it is the result of the multiplication of bacteria in the blood. This often results in secondary site infections for example endocarditis.

Problematically, for both food preparation and hospital workers, is that 25-50 % of healthy individuals are carriers of *S. aureus,* with the mucous membrane of the nose being the main site of carriage. *S. aureus* however, is also found on the skin and in faeces (Novick, 1990; Kloos and Bannerman, 1994).

1.4 Treatment of *S. aureus* infections

The introduction of penicillin in the 1940's, for the treatment of severe *S. aureus* infections allowed control of these infections, and had a great effect on morbidity and mortality. However, resistance due to the presence of a penicillinase emerged almost immediately and by the late 1940's approximately 60 % of hospital-acquired strains were resistant to penicillin (Barber and Rozwadow-Dowzenko, 1948). Furthermore the staphylococci have proved to be extremely capable in developing resistance to the commonly used antimicrobial agents with resistant strains to streptomycin, tetracycline and erythromycin emerging shortly after their introduction into clinical practice (Demerec, 1948; Haight and Finland, 1952). The resistance to antibiotics in most cases is coded for by genes carried on plasmids, this probably accounts for the rapid spread of resistant organisms. The early widespread arrival of antibiotic resistant strains by the 1950's caused grave therapeutic problems and sent a shock wave through the medical community. Although it seems entirely in character, that among the earliest organisms to accommodate to these powerful drugs should have been the adaptable *Staphylococcus*. In the 1960's β-lactamase resistant antibiotics, such as methicillin were introduced, but methicillin-resistant *S. aureus* (MRSA) were reported soon after (Jevons, 1961) and they currently account for a large proportion of clinical isolates of *S. aureus* worldwide. Methicillin resistance is due primarily to a novel form of a penicillin binding protein (PBP2a (MecA)) that has reduced affinity for methicillin (Hartman and Tomasz, 1984). Vancomycin is often the last line of defence antibiotic, in the treatment of multiple antibiotic resistant strains, in particular MRSA. Unfortunately intermediate level vancomycin-resistant *S. aureus* (VRSA) have now been detected. These resistant strains were first reported in Japan (Kremery *et al.*, 1996), but have since been reported in both the United States and France. Concerningly it has been reported that up to 20 % of MRSA isolated from hospitals throughout Japan are heterogeneously resistant to vancomycin, and that this resistance develops into VRSA upon exposure to vancomycin (Hiramatsu *et al.*, 1997). The global spread of resistance to this antibiotic is now almost inevitable and could potentially lead to a return to the pre-antibiotic age.

1.5 Virulence factors of *S. aureus*

S. aureus elaborates a great variety of extracellular proteins, many of which are potential virulence factors, although only a few have actually been implicated in virulence (Arvidson, 1983). Several definitions of a virulence factor exist, however for the purpose of this study a virulence factor is described as a factor, for which there is at least a hypothetical involvement in the causation of disease. A list of putative and proven virulence factors is shown in Table 1.1, although it is not comprehensive. Furthermore, many of the genes for these factors have yet to be identified. The virulence factors listed in Table 1.1 have been divided into four broad categories: 1) Attachment; factors involved in attaching the bacteria to cells or extracellular matrices. 2) Evasion of host defence; factors that prevent or reduce phagocytosis, or that interfere with the function of specific antistaphylococcal antibodies or other specific host-defence mechanisms. 3) Invasion/tissue penetration; factors that specifically attack host cells and factors that specifically degrade components of the extracellular matrices. 4) Antibiotic inactivation; factors that destroy/inactivate antibiotics administered against *S. aureus* or other infections (Projan and Novick, 1997).

Importantly, the purpose of a given virulence factor is thought not to be to cause disease but to enhance survival of the bacterium. This is because *S. aureus* are not inherently pathogenic organisms (unlike for example the influenzae virus which must cause disease to propagate). Thus for every illness caused by *S. aureus,* many more individuals are benignly colonized with no sign of disease. In these terms, the infected host should really be viewed as the environment to which the bacterium must adapt in order to survive. Within the context of the ability to grow and divide, few if any of the virulence factors are essential, and thus these factors are best viewed as a subset of accessory gene products, which are advantageous in particular environments. Furthermore, with the exception of toxinoses, no single factor is responsible for staphylococcal infections but is instead due to the coordinate action of multiple different factors (Projan and Novick, 1997).

1.6 Control of production of virulence determinants

1.6.1 Coordinate regulation of virulence

Bacterial growth and survival is competitive and is largely dependent on a cell's ability to adapt to environmental changes. Infections caused by *S. aureus* are at least in

Factor	Role	Gene	Expressed?
Clumping factor	Attachment	clfA	Log phase
Fibrinogen-binding protein	Attachment	*fbpA*	Log phase
Fibronectin-binding protein A	Attachment	*fnbA*	Log phase
Fibronectin-binding protein B	Attachment	*fnbB*	Log phase
Collagen-binding protein	Attachment	*cna*	
Coagulase	Attachment	*cga / coa*	Log phase
Polysaccharide / adhesin (PS/A)	Attachment		
Polysaccharide intracellular adhesin	Attachment		
Enterotoxin A	Evasion of host defences	*sea*	Constitutive
Enterotoxins B, C1-3, D, E, H	Evasion of host defences	*seb-h*	Post-exponential
Toxic shock syndrome toxin-1	Evasion of host defences	*tst*	Post-exponential
Exfoliative toxins A, B	Evasion of host defences	*eta, etb*	Post-exponential
Protein A	Evasion of host defences	*spa*	Log phase
Lipase	Evasion of host defences	*geh*	Post-exponential
V8 protease	Evasion of host defences	*ssP*	Post-exponential
Fatty acid modifying enzyme (FAME)	Evasion of host defences		
Panton-Valentine leukocidin	Evasion of host defences	*lukPV, lukS-PV*	Post-exponential
Leukocidin R	Evasion of host defences	*lukF-R, lukS-R*	
Capsular polysaccharide type 1	Evasion of host defences	*cap1* locus	Post-exponential
Capsular polysaccharide type 5	Evasion of host defences		Post-exponential
Capsular polysaccharide type 8	Evasion of host defences	*cap8* locus	Post-exponential
Staphylokinase	Evasion of host defences	*sak*	
Catalase	Evasion of host defences	*katA*	
α-Toxin	Invasion/tissue penetration	*hla*	Post-exponential
β- Haemolysin	Invasion/tissue penetration	*hlb*	Post-exponential
γ- Haemolysin	Invasion/tissue penetration	*hlgA, B, C*	Post-exponential
δ- Haemolysin	Invasion/tissue penetration	*hld*	Post-exponential
Phospholipase C	Invasion/tissue penetration	*plc*	
Metalloprotease (elastase)	Invasion/tissue penetration	*sepA*	
Hyaluronidase, hyaluronate lyase	Invasion/tissue penetration	*hysA*	Post-exponential
β-lactamase	Antibiotic inactivation	*bla*	

Table 1.1

Putative and proven virulence factors of *S. aureus*

Adapted from Projan and Novick (1997).

part related to the organisms' striking ability to react to changing environments during the infection process. All bacteria express their genetic information in a regulated manner, with relatively few genes being continuously expressed ('housekeeping' genes). The remainder of the genes, are deemed non-essential for growth and are often referred to as accessory genes (Chapter 1.5). The constitutive production of accessory factors rather than on a needs only basis would be energetically unfavourable. As a result bacteria have evolved control systems that coordinately regulate groups of genes and operons. This modulation occurs by the expression of appropriate genes via signal transduction pathways, so that accessory factors are only produced when needed (Miller *et al.*, 1989).

Several characteristics of a host environment are known to induce the production of virulence factors and toxins, such as temperature and nutrient availability (Bliska *et al.*, 1993). Sensing of these signals leads to a change in the control of the transcription of a wide variety of unlinked genes, many of which are involved in pathogenesis.

1.6.2 Bacterial signal transduction

Adaptive gene expression necessitates that the cell be able to detect and respond to environmental changes. As the majority of stimuli to which cells are required to respond, do not generally penetrate the cytoplasm, signalling systems have evolved in all cells. These signal transduction systems in bacteria often share a significant degree of structural similarity with one an other, despite the signals detected being quite diverse (Stock *et al.*, 1989). A common mechanism for the transduction of environmental stimuli to bring about changes in gene expression are the two-component sensor regulators, these have been identified in eukaryotes as well as in many bacteria (Parkinson and Kofoid, 1992; Kavanaugh and Williams, 1996). Two-component systems generally consist of two proteins, which are often encoded in a single operon. Commonly one of the proteins is a transmembrane sensor, and the other a transcription factor known as a response regulator. In general the transmembrane protein (sensor) binds to a specific ligand, the signal, and autophosphorylates. The phosphorylated sensor then relays the phosphate to the response regulator. The regulator is now in its active conformation and can stimulate or repress target genes at the trancriptional level.

A further method of global regulation of gene expression involves a process of intercellular communication that enables cells to sense population levels. This signalling mechanism is known as quorum sensing and has been described in many

bacteria, including several Gram-positive species. This mechanism of communication uses a diffusible signalling molecule (pheromone) to regulate physiological processes in a cell density dependent manner (Kleerebezem *et al.*, 1997).

1.6.3 Expression of virulence factors in *S. aureus*

The majority of information on this important subject has been gained from the study of bacterial cells in culture. A classical growth curve of *S. aureus,* grown in nutrient medium is divided into four phases; lag, log, post-exponential and stationary, and is represented diagrammatically in Figure 1.1. During exponential growth, bacterial metabolism is rapid and efficient and can be regarded as a steady state in which cells are geared to ensure constancy of growth. As a culture ages and growth stops (post-exponential), there is a reorganization of cellular metabolism for long-term survival under unfavourable conditions. A striking feature of this reorganization is the rapid synthesis of exoproteins, of which many are virulence factors. In fact it has been observed that almost all the exoproteins produced by *S. aureus* are produced post exponentially (Novick and Projan, 1997). There are two exceptions to this, enterotoxin A, which is produced constitutively (Tremaine *et al.*, 1993), and coagulase, which is produced during logarithmic growth (Duthie, 1954). At the same time the production of many proteins essential for growth and cell division, along with a wide range of cell wall associated proteins including protein A, clumping factor and the fibronectin-binding proteins, ceases. Therefore it is clear that there is a complex regulatory system governing this change over, which can be regarded as a metabolic 'toggle switch' that is set at the end of exponential phase for exoprotein synthesis. Table 1.1 shows a list of virulence factors, and indicates where possible, during which growth phase they are expressed.

1.7 Regulators of virulence determinant production of *S. aureus*

Several regulatory loci have been discovered in *S. aureus* by transposon mutagenesis. Each mutation results in pleiotropic differences in the expression of exoproteins and/or virulence factors. These loci have been termed global regulators and include: *agr* (Recsei *et al.*, 1986; Morfeldt *et al.*, 1988), *sar* (Cheung *et al.*, 1992), and *sae* (Girauldo *et al.*, 1994). A further global regulator, *xpr* (extracellular protein regulator) was thought to have been identified (Smeltzer *et al.*, 1992), but this was later shown to contain an additional mutation within the *agr* locus (McNamara and Iandolo, 1998). It

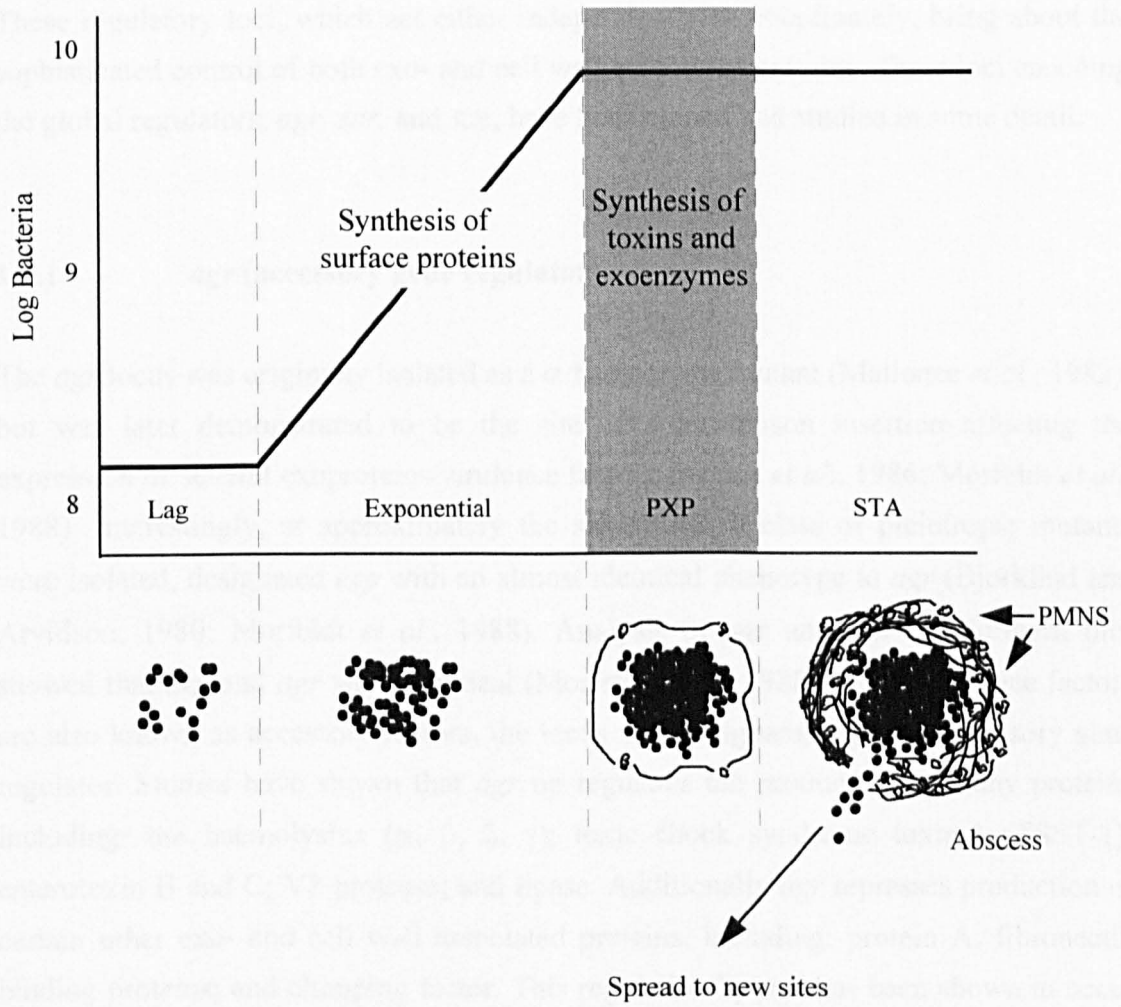


Figure 1.1
Hypothetical role of virulence factors in staphylococcal infection

In this model, bacteria in lag phase initiate the infection and would enter exponential phase, where multiplication would begin, accompanied by synthesis of surface proteins. Crowding during post-exponential phase (PXP) would activate a density sensing mechanism that would stimulate toxic exoprotein production. This enables the organism to escape from the localised infection (abscess) during stationary phase (STA) and spread to new sites, where the cycle would be repeated.

Reproduced from Projan and Novick (1997).

is via these global regulators, along with many other unknown regulatory components that *S. aureus* coordinates production of a wide range of virulence factors (Table 1.1). These regulatory loci, which act either independently or coordinately, bring about the sophisticated control of both exo- and cell wall associated proteins. Three loci encoding the global regulators; *agr, sar,* and *sae*, have been cloned and studied in some detail.

1.7.1 *agr* (accessory gene regulator)

The *agr* locus was originally isolated as a α-haemolysin mutant (Mallonee *et al.*, 1982), but was later demonstrated to be the site of a transposon insertion affecting the expression of several exoproteins/virulence factors (Recsei *et al.*, 1986; Morfeldt *et al.*, 1988). Interestingly, at approximately the same time, a class of pleiotropic mutants were isolated, designated *exp* with an almost identical phenotype to *agr* (Bjorklind and Arvidson, 1980; Morfeldt *et al.*, 1988). Analysis of *agr* and *exp* by Northern blot showed that *exp* and *agr* were identical (Morfeldt *et al.*, 1988). Since virulence factors are also known as accessory factors, the locus was designated *agr*, for accessory gene regulator. Studies have shown that *agr* up regulates the production of many proteins including: the haemolysins (α, β, δ, γ); toxic shock syndrome toxin-1 (TSST-1); enterotoxin B and C; V8 protease; and lipase. Additionally *agr* represses production of certain other exo- and cell wall associated proteins, including: protein A; fibronectin binding proteins; and clumping factor. This regulation by *agr* has been shown to occur in a growth phase dependent manner (Bjorklind and Arvidson 1980; Recsei *et al.*, 1986). The *agr* locus was cloned and the complete sequence determined (Peng *et al.*, 1988; Kornblum *et al* 1990; Arvidson *et al.*, 1990). The *agr* operon is shown diagrammatically in Figure 1.2.

The *agr* locus consists of two divergent transcriptional units, whose transcripts RNAII and RNAIII are transcribed from promoters P2 and P3 respectively (Kornblum *et al.*, 1990; Novick *et al.*, 1995). A third transcript RNAI has also been detected and is transcribed from promoter P1. The role of P1 is unclear, and has been shown to be weak and constitutive, occurring throughout the growth cycle and independently of *agr* function (Peng *et al.*, 1988). The P2 operon produces a transcript, RNAII of approximately 3.5 kb long which encompasses four genes, designated *agrA, agrB, agrC* and *agrD* (Novick *et al.*, 1995). The final *agr* transcript is RNAIII, which is a transcript of approximately 0.5 kb and is transcribed in the opposite direction to RNAII (P2) (Kornblum *et al.*, 1990) (Figure 1.2).

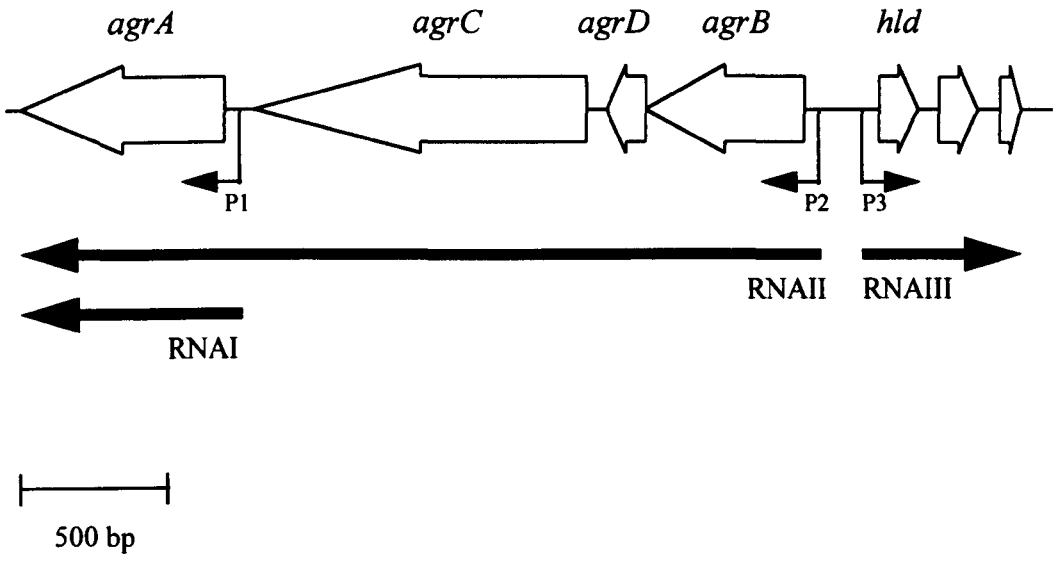


Figure 1.2
The *agr* locus of *S. aureus*

Map of the *agr* locus showing the positions of the three *agr* promoters, P1, P2 and P3, together with the extent of their transcripts. The *agrC and agrA* genes encode a sensor and response regulator, respectively of a two-component signal transduction system. The *agrD* gene product is modified by the product of the *agrB* gene to give an octapeptide know as AIP (autoinducing peptide). The *hld* gene encodes δ-haemolysin.

Adapted from Kornblum *et al.* (1990).

1.7.1.1 RNAIII

The P3 operon produces a 512 base RNAIII transcript, containing three putative open reading frames (ORFs). The first encodes the 26 amino acid, δ-haemolysin protein (*hld*). The second and third ORFs, potentially encode 21 and 19 amino acid polypeptides respectively, however only the second ORF has a putative ribosome binding site.

Therefore it would appear that the 3' portion of RNAIII is not translated. However studies have shown that the 3' portion of RNAIII is required for *agr* activity. Mutagenesis studies have shown than none of the ORFs can be responsible for RNAIII activity, thus indicating that RNAIII itself is the effector molecule of the positive and negative regulation by *agr*, of genes at the end of exponential growth (Novick *et al.*, 1993). The role of RNAIII as the molecule regulating exoprotein gene expression was also supported by the finding that an insertion in the RNAIII 3' region, did not impair δ-haemolysin production, but resulted in a typical *agr* $^{-}$ phenotype (Arvidson *et al.*, 1990).

The manner in which RNAIII regulates target genes has been the subject of much study and is not fully understood. Transcriptional fusions to *hla, spa* and *tst* have determined that RNAIII acts at the transcriptional level (Novick *et al.*, 1993). However a series of experiments using an exogenous inducer to regulate RNAIII expression showed that while induction of RNAIII at virtually any point in the growth of *S. aureus* resulted in an immediate inhibition of *spa* transcription, transcription of *hla* while dependent on RNAIII, did not always immediately follow induction of RNAIII, but instead occurred at a later point in growth (Vandenesch *et al.*, 1991). From this result it was concluded that another 'temporal factor' must exist in addition to *agr,* that is required for the expression of α-haemolysin and thus probably other exoproteins produced by *S. aureus*.

There is also evidence demonstrating that RNAIII acts as a translational activator, stimulating translation of α-haemolysin mRNA (Novick *et al.*, 1993). The mechanism for this translational control, was determined by Morfeldt *et al.* (1995), who demonstrated that, RNAIII functions as an antisense RNA to complex with sequences on the α-haemolysin mRNA upstream from the translational start site. This prevents a stem-loop from forming and frees the ribosome binding site and start codon, thus allowing translation of the *hla* mRNA. Therefore RNAIII has two likely regulatory roles, functioning at both the transcriptional and translational levels.

1.7.1.2 RNAII

Studies established that there are four ORFs within the P2 transcript, all of which have been shown to be essential for complete *agr* function. These ORFs are designated *agrA, agrB, agrC* and *agrD* (Novick *et al.*, 1995). Studies have demonstrated that AgrA shows N-terminal similarity to the response regulatory component, and that AgrC has C-terminal amino acid sequence homology to the sensory component of a two-component sensor regulator system (Novick, 1990). The primary role of the AgrA/AgrC signalling pathway is to up regulate both of the major *agr* promoters, P2 and P3. It has not however been demonstrated that AgrA binds directly to DNA. Instead AgrA is thought to work in conjunction with a second transcription factor SarA (Chapter 1.7.2) (Morfeldt *et al.*, 1996a/b) (Figure 1.3). The remaining two ORFs, AgrB and AgrD are involved in the generation of a small autoinducing peptide (AIP), (Ji *et al* 1995) which is involved in cell density signalling.

1.7.3 Cell density control of virulence determinants by AIP

Transcription of the P2 operon of *agr,* has been found to be auto-regulated in a density dependent manner by a signalling pheromone system (Ji *et al.*, 1995). Two genes of the P2 operon, *agrB* and *agrD* are responsible for this auto-regulation and combine to produce the autoinducer. The autoinducer is a small peptide, which is the ligand that binds to and activates AgrC. *agrD* encodes the autoinducer as a 46 amino acid propeptide, which is subsequently processed to giving an octapeptide known as AIP (autoinducing peptide). *agrB* encodes a protein that is required for this processing and possibly for the secretion of AIP (Ji *et al.*, 1997). AIP functions as an extracellular signal that stimulates the coordinated virulence response in *S. aureus*. This autocrine regulation suggests that the *S. aureus agr* operon is a quorum sensing regulon, detecting the presence of a critical cell density (1.6.2), analogous to density-sensing regulatory systems in other bacteria that utilize homoserine lactones as autoinducers. Achievement of the threshold concentration of AIP causes activation of both the P2 and P3 promoters of *agr* (Ji *et al.*, 1995).

A survey of some 400 strains of *S. aureus* and other staphylococci revealed that the vast majority synthesize and secrete a soluble autoinducer of RNAIII transcription. Interestingly, this study also revealed that the peptides of some strains inhibit the expression of *agr* in other strains (Ji *et al.*, 1997). In total four different groups of *S. aureus* have been identified with AIP from one group inhibiting *agr* expression of the

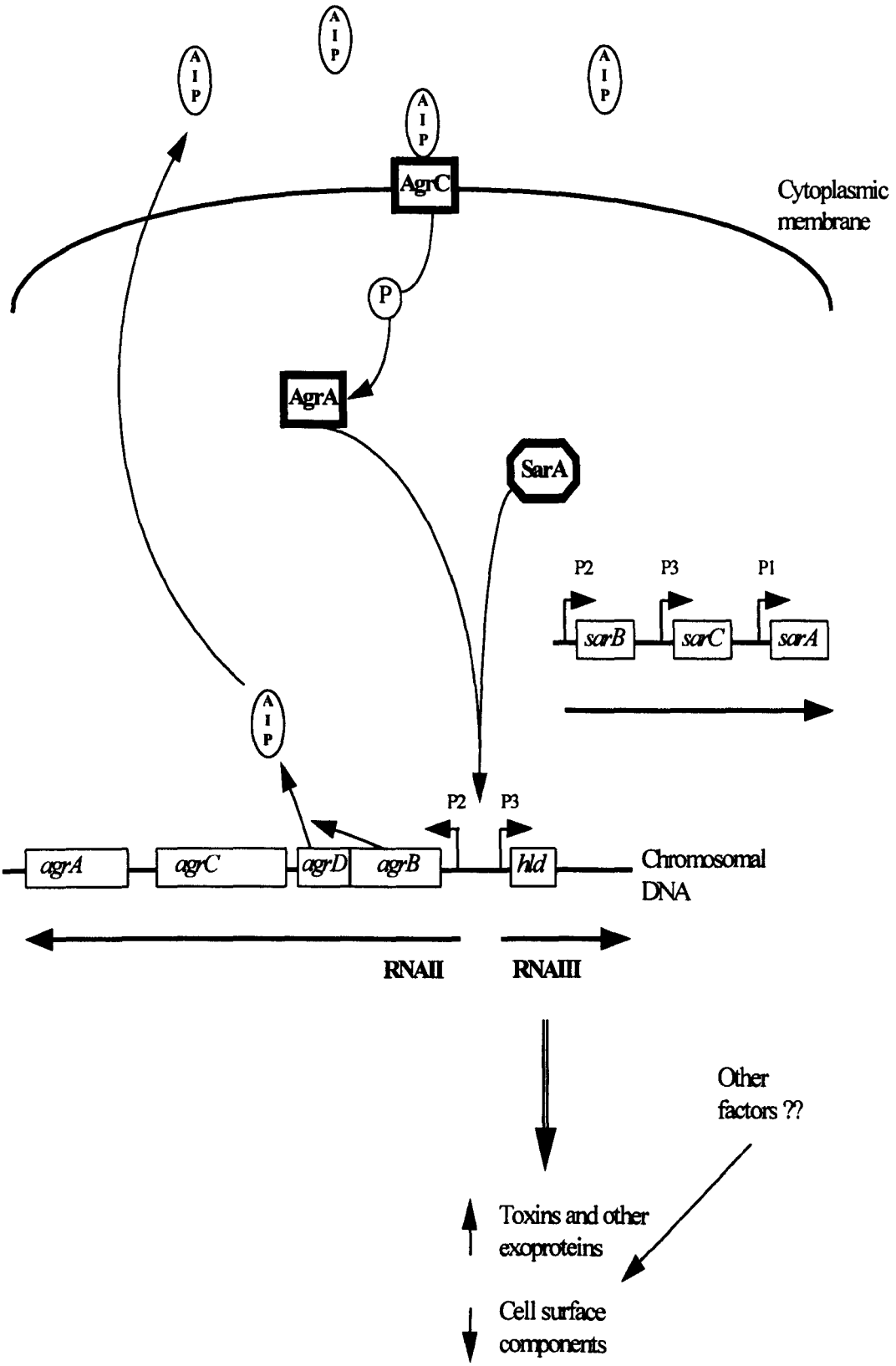

AIP
AIP
AIP
AIP
AgrC
Cytoplasmic membrane
P
AgrA
SarA
P2
P3
P1
sarB
sarC
sarA
AIP
P2
P3
agrA
agrC
agrD
agrB
hld
Chromosomal DNA
RNAII
RNAIII
Other factors ??
Toxins and other exoproteins
Cell surface components

Figure 1.3

Model of the interaction between the *agr* and *sar* gene products (depicted in bold type) which mediate changes in expression from the P2 and P3 promoters of *agr*

AgrC and AgrA represent the sensor and regulator, respectively of a two-component signal transduction system. AIP (Autoinducing peptide) is encoded within the *agrD* gene, and is modified from a 46 amino acid propeptide to an octapeptide, by the gene product of the *agrB* gene. AIP has been proposed as a signalling pheromone that, probably via AgrC, leads to increased transcription from the P2 and P3 promoters of *agr,* in a density dependent manner. The RNAIII transcript has been shown to both up and down regulate production of some exoproteins, as indicated (not to scale).

Adapted from Fairhead (1998).

other groups. This finding represents a novel type of bacterial interference, which could be correlated with the ability of one strain to exclude others from infection and/or colonization sites (Ji *et al.*, 1997). The sequence of *agrD,* which encodes AIP, and the two flanking genes *agrB* (AIP processor) and *agrC* (sensor), is highly variable among different strains of *S. aureus*. Thus these three genes must have co-evolved in order to retain the specificity of autoinducer processing and receptor interaction (Ji *et al.*, 1997).

1.7.2 *sar* (staphylococcal accessory regulator)

A second regulatory locus, was identified in *S. aureus* by Cheung and colleagues (1992), once again using transposon mutagenesis. This mutation resulted in a pleiotropic effect on the expression of several virulence factors. This locus was designated *sar* (staphylococcal accessory regulator), and was shown to be distinct from *agr* (Cheung *et al.*, 1992). A mutation in the *sar* locus was found to result in decreased production of several extracellular enzymes including; α, β and δ-haemolysin, along with an increase in serine protease (V8), and metalloprotease (Cheung and Ying, 1994; Chan and Foster, 1998a). Finally studies have shown that the *sar* gene products are necessary for the optimal transcription of RNAIII, suggesting that *sar* is essential for *agr*-dependent regulation (Heinrichs *et al* 1996). The *sar* locus has been cloned and mapped (Bayer *et al.*, 1996), and is shown diagrammatically in Figure 1.4.

The entire *sar* locus encompasses 1.2 kb and studies have demonstrated there to be three overlapping transcripts in a parallel array, each transcribed from a distinct promoter (Bayer *et al.*, 1996; Heinrichs *et al* 1996). The three *sar* transcripts designated *sarA, sarC* and *sarB,* have common 3' ends and are initiated from a triple-promoter system, designated P1, P3 and P2, respectively (Figure 1.4). Interestingly, multiple promoters directing the transcription of environmentally regulated genes are a common architectural feature in the control of transcription (Projan and Novick, 1997). All three *sar* transcripts encode an ORF (*sarA*) of 372 bp which in complementation experiments has been show to be necessary for α-haemolysin production (Cheung and Ying, 1994). However all three transcripts have been shown to be required for wildtype levels of α-haemolysin. This indicates a role for the two small upstream ORFs (ORF3 and ORF4) of 39 and 18, amino acids respectively (Bayer *et al.*, 1996), which together are only present on the *sarB* (P2) transcript. All three promoters, have been shown to be regulated in a growth phase dependent manner, with expression of *sarA* (P1) and *sarB* (P2) greatest during mid-exponential phase, and *sarC* (P3) expression maximal at the end of stationary phase (Bayer *et al.*, 1996). This links well with the fact that promoters

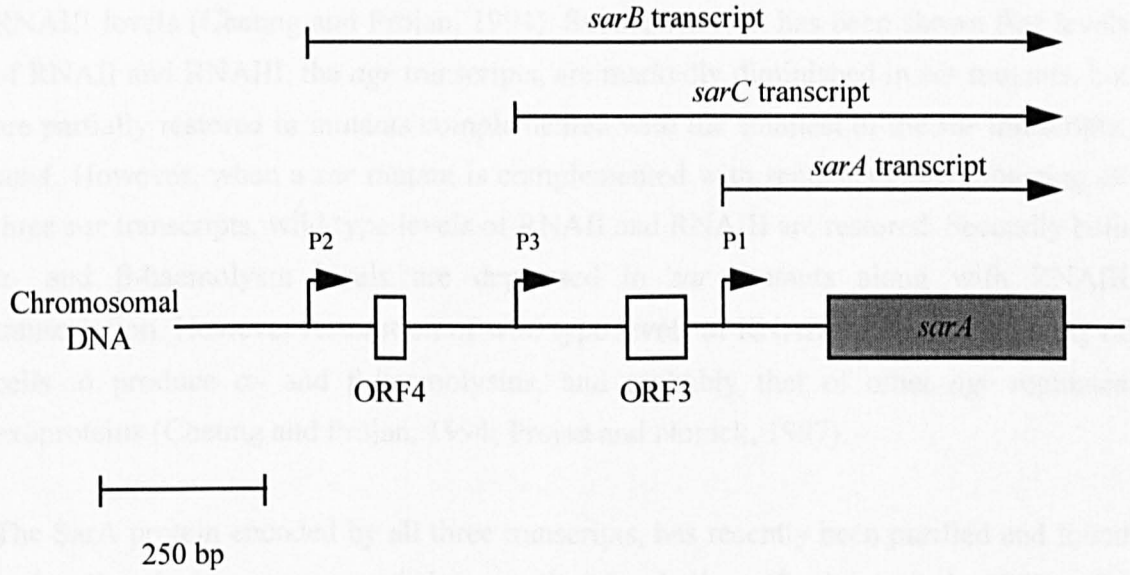


Figure 1.4

The *sar* locus of *S. aureus*

The location of the three *sar* promoters and the approximate extent of their transcripts are marked.

Adapted from Bayer *et al.* (1996).

P1 and P2 show sequence similarity to sigma^{70} dependent promoters, whilst P3 showed a high degree of similarity to a *B. subtilis* sigma^{B} (σ^{B}) dependent promoter (Deora *et al.*, 1997), which are known to be induced by σ^{B} (Chapter 1.7.4.1) during stationary phase (Boyland *et al.*, 1993). However, further analysis of these promoters has revealed that transcription from P1 is stronger than that from P2 or P3 (Cheung *et al.*, 1998).

From comparison of *sar* with the *virF* gene of *Shigella flexneri,* it was hypothesized that *sar* may control the expression of a range of exoproteins by positively controlling RNAIII levels (Cheung and Projan, 1994). Subsequently it has been shown that levels of RNAII and RNAIII, the *agr* transcripts, are markedly diminished in *sar* mutants, but are partially restored in mutants complemented with the smallest of the *sar* transcripts, *sarA*. However, when a *sar* mutant is complemented with sequences encompassing all three *sar* transcripts, wild type levels of RNAII and RNAIII are restored. Secondly both α- and β-haemolysin levels are depressed in *sar* mutants along with RNAIII transcription. However restoration of wild type levels of RNAIII, restores the ability of cells to produce α- and β-haemolysins, and probably that of other *agr* regulated exoproteins (Cheung and Projan, 1994; Projan and Novick, 1997).

The SarA protein encoded by all three transcripts, has recently been purified and found to be dimeric in nature, consisting mostly of α-helices (Rechtin *et al.*, 1999). The function of the SarA protein has been studied and shown to be a DNA-binding protein. Two recent reports (Morfeldt *et al.*, 1996a; Chein and Cheung, 1998) have shown SarA binds to *cis*-regulating elements upstream of both the P2 and P3 promoters of *agr* activating the expression of both *agr* operons. However there are significant discrepancies between the two studies regarding the binding sites for SarA in the *agr* locus. The situation however has recently been clarified by Rechtin *et al.* (1999), who showed that neither of the earlier reports were wholly accurate. DNase I footprint analysis revealed three high-affinity binding sites for SarA in the putative regulatory region *cis* to the *agr* promoters. Each of these binding sites consists of two half sites, each of approximately 18 bp. Thus the current model of SarA interactions with the *agr* region involves a stable SarA dimer binding to three specific sites, each containing two half sites, upstream of the regulated P2 and P3 promoters (Rechtin *et al.*, 1999). Therefore it appears that *sar* and *agr* do interact at some level and increased SarA levels have been shown to correlate with greater *agr* activation (Cheung *et al.*, 1998). However, the fact that a *agr, sar* double mutant is more impaired in virulence factor production, than either *agr* or *sar* single mutants implies that *sar* also has a role in expression of virulence factors, independent of *agr* (Cheung *et al.*, 1994; Booth *et al.*, 1997).

SarA is also known to directly regulate collagen adhesin (*cna*), although definitive SarA binding sites upstream of this gene have yet to be identified. Furthermore significant homology has been observed among sequences *cis* to genes with altered expression in SarA mutants, for example *tst, spa, hla* and *hlb,* to the SarA binding sites of *agr* (Chan and Foster 1998a). However, none of these genes have been shown to be regulated directly by SarA.

Several models exist for the mechanism by which SarA affects gene expression (reviewed by Rechtin *et al.*, 1999). The most interesting model was originally proposed by Morfeldt *et al.* (1996a). This model centres around SarA inducing changes in the superhelicity of DNA. Several pieces of evidence support this model including: i) the spacing between the -10 and -35 regions of the P2 and P3 promoters of *agr* is approximately 3 bp too far (Morfeldt *et al.*, 1996a), thus one function of SarA may be to overcome this spacing; and ii) the gene encoding the collagen adhesin (*cna*) is repressed by SarA in late-stage cultures, at the same time as SarA activates P2 and P3 (Gillaspy *et al.*, 1997; Blevins *et al.*, 1999). The spacing of the -10 and -35 regions is near optimal and thus SarA binding may serve to overwind regulatory promoters, thus acting as both a repressor and an activator (Rechtin *et al.*, 1999).

1.7.3 *sae* (*S. aureus* exoprotein expression)

A third global regulator of virulence determinant production, designated *sae* (*S. aureus* exoprotein expression), has been identified (Giraudo *et al.*, 1994), and subsequently shown to be different from both *agr* and *sar* (Giraudo *et al.*, 1999). Inactivation of *sae* leads to the diminished production of α- and β-haemolysins, DNase, coagulase and protein A. However, production of proteases, staphylokinase and lipase remained constant (Giraudo *et al.*, 1994). Northern blot analysis revealed that *sae* affects exoprotein expression at the transcriptional level (Giraudo *et al.*, 1997).

The *sae* locus has recently been cloned, and sequencing revealed the presence of two genes that are potentially co-transcribed, designated *saeR* and *saeS*, encoding a response regulator and a histidine protein kinase, respectively (Figure 1.5). These genes show high homology to other bacterial two-component regulatory systems, including the PhoP/PhoR system of *B. subtilis* (Giraudo *et al.*, 1999). Interestingly, while *sae* is required for the transcription of *hla, hlb* and *coa* genes, its inactivation appears to have no effect on *agr* or *sar* transcription (Giraudo *et al.*, 1997). Therefore it is believed that

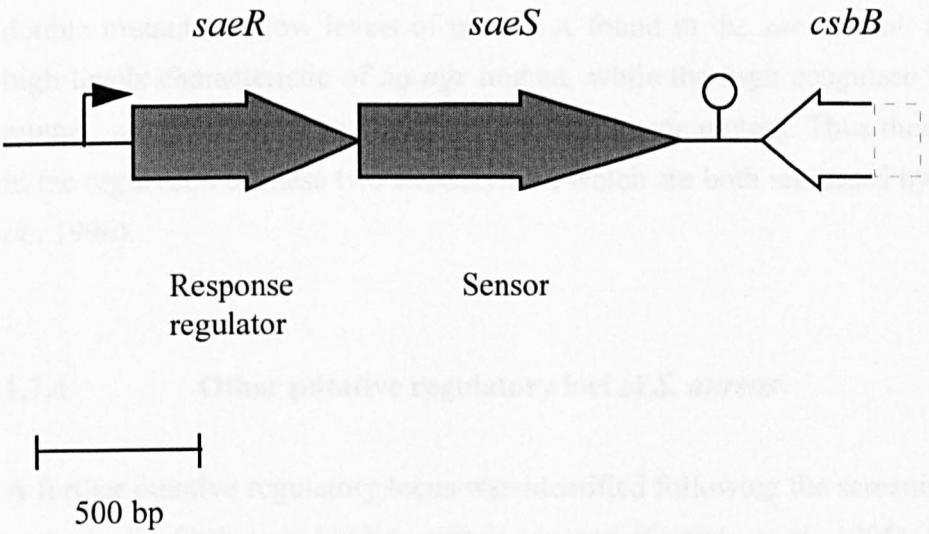


Figure 1.5

The *sae* locus of *S. aureus*

Schematic drawing indicating the location and orientation of the *saeR-saeS* operon and part of an incomplete ORF, represented by dotted lines, with homology to *cbsB* of *B. subtilis*.

Adapted from Giraudo *et al.* (1999).

SaeR exerts its effects by direct interaction with target genes, which are also subject to regulation by *agr* (Giraudo *et al.*, 1999).

However, a direct interaction between *sae* and *agr* presently cannot be ruled out as reduced levels of exoproteins are found in a *sae, agr* double mutant, as compared to their already diminished levels in either single mutant. This suggests that there may be a synergistic interaction between the two loci. Furthermore the *sae agr* double mutant showed further decreased virulence (Giraudo *et al.*, 1996). Importantly in the same double mutant, the low levels of protein A found in the *sae* mutant, are raised to the high levels characteristic of an *agr* mutant, while the high coagulase levels of an *agr* mutant, are reduced to the very low levels of an *sae* mutant. Thus there is a difference in the regulation of these two exoenzymes, which are both repressed by *agr* (Giraudo *et al.*, 1996).

1.7.4 Other putative regulatory loci of *S. aureus*

A further putative regulatory locus was identified following the screening of transposon mutants, for fibrinogen binding protein mutants (Cheung *et al.*, 1995). One such mutant displayed nil levels of surface protein A, and increased haemolysin production, in addition to a decreased fibrinogen binding capacity. Importantly, this locus has been shown to be distinct from other regulatory loci, including *agr* and *sar* (Cheung *et al.*, 1995).

Several other two-component regulatory systems, in addition to *agrA/agrC* and *saeS/ saeR* have been identified in *S. aureus*. These include: *lytS/lytR* which effect the expression of murein hydrolases and thus influencing control over the rate of autolysis (Brunskill and Bayles, 1996); and *phoR/phoP* which in *E. coli* up regulates the expression of at least 30 genes (Lammers *et al.*, 2000; Wanner *et al.*, 1993). However it is predicted that many more two-component regulatory systems will be discovered in *S. aureus,* some of which will potentially be involved in pathogenesis and in the regulation of virulence determinants.

Finally sigma factors (σ), which are proteins that bind to core RNA polymerase to form the holoenzyme that binds to, specific promoters (Moran, 1993), may have a role in the regulation of virulence determinant production. There are two broad groups of sigma factors: primary sigma factors that are responsible for the expression of housekeeping genes, whose products are necessary for exponential growth (Deora *et al.*, 1997); and

the alternative sigma factors, that control the expression of a wide range of genes involved in many cellular functions. It is this latter class of sigma factors that have been proposed to play a role in virulence determinant production (Cheung *et al.*, 1999; Nicholas *et al.*, 1999).

1.7.4.1 The role of sigma factor σ^B in the regulation of virulence determinant production

σ^B encoded by *sigB*, is an example of an alternative sigma factor (Haldenwang and Losick, 1980) and has been proposed to control the expression of genes involved in the environmental stress response of *S. aureus* (Chan *et al.*, 1998). The *sar* P3 promoter was identified as having a σ^B-like consensus sequence and studies confirmed that σ^B initiated transcription from the *sar* P3 promoter (Deora *et al.*, 1997), which like most σ^B dependent promoters, is most active during stationary phase (Cheung *et al.*, 1997). A *sigB* mutant has been found to be deficient in clumping factor, coagulase and a pigment (Nicholas *et al.*, 1999). However σ^B has been shown not to have a major role in virulence (Chan *et al.*, 1998; Nicholas *et al.*, 1999), although the σ^B mutant has been reported to lead to an alteration in *hla* transcription (Cheung *et al.*, 1999).

1.8 The role of environmental stimuli in the expression of virulence determinants

S. aureus is a successful pathogen that is capable of colonizing a diverse range of environments in the human host. This is substantially aided by the fact that *S. aureus* has evolved a range of sophisticated mechanisms to rapidly respond and adapt to environmental changes. This is important as the interaction between *S. aureus* and the human host during infection, involves both survival and multiplication of the pathogen in the presence of the host immune system. Thus it is essential for the bacterium to be able to sense from environmental signals, its association with a host. These signals probably include temperature, osmolarity, oxygen concentration, pH and the concentration of metal ions such as iron, calcium and magnesium (Mekalanos, 1992).

Some of these environmental factors have been found to modulate virulence determinant production, in both *agr/sar* dependent and independent ways. For example, it was discovered that α-haemolysin is specifically induced under microaerobic conditions. This induction was found to be independent of *agr*, but dependent on *sarA*,

and to occur at the transcriptional level (Chan and Foster, 1998a). Whereas high NaCl levels have been shown to repress transcription of *spa, hla* and *tst* expression. This appears to be independent of *agr* and *sarA*. Thus additional regulatory mechanisms are also involved in the control of the transcription of these *agr*-dependent genes (Chan and Foster, 1998b).

Further studies have also revealed both pH and glucose to have an effect on *sec, hla* and *agr* expression. It was observed that the *agr* transcripts, RNAII and RNAIII, were poorly expressed when the pH of the culture was maintained at 6.5 or higher (Regassa *et al.*, 1992). Therefore *agr* expression is to some extent dependent on a drop in pH that is usually observed as a culture grows. Finally the presence of glucose in the growth medium has also been shown to repress transcription of RNAII and RNAIII, in addition to *sec* and *hla* (Regassa *et al.*, 1992). The fact that the expression of *agr,* and thus that of virulence factors of *S. aureus* are subjected to catabolite repression is interesting, as it may imply that *agr* plays a role in nutrition, as well as virulence and thus, these two roles may be at least in part linked.

Therefore it can be seen that *S. aureus* is capable of detecting a wide range of environmental stimuli and modulating their production of virulence factors accordingly. The majority of these mechanisms involved here however remain to be elucidated.

1.9 The expression of virulence factors and their regulators *in vivo*

The vast majority of work conducted on the expression of virulence factors and their regulators has been carried out *in vitro,* primarily in nutrient-rich growth media. Recently however, several studies have been conducted, aimed at analyzing the expression of virulence factors (for example, protein A and α-haemolysin) and their regulators (*agr* and *sar*) *in vivo.*

1.9.1 *In vivo* expression of *agr, spa* and *hla*

In order to gain insight into the regulatory response of *S. aureus* during infection, a system has recently been described allowing bacterial gene expression to be monitored in the host (Goerke *et al.*, 2000). This study developed a method for the direct analysis of gene expression during bacterial infections, and allowed for the first time the activity of the global regulator *agr* (via RNAIII expression), controlling the expression of major

virulence factors to be analysed during infection. Additionally, the transcription of *spa* (encoding protein A, a gene repressed by *agr*), and of *hla* (encoding α-haemolysin, a gene activated by *agr*), was monitored. This method enabled direct transcriptional analysis from clinical specimens of *S. aureus* and quantification of specific transcripts was performed by competitive reverse transcription-PCR (RT-PCR). Analysis was performed on *ex vivo* material, namely sputum from cystic fibrosis (CF) patients, infected with *S. aureus*. Interestingly, the bacteria expressed RNAIII poorly, indicating an inactive *agr, in vivo*. This is thought, to be linked to the fact that in sputum samples, the bacterial density is below the threshold for *agr* activation *in vitro* (10^9 cfu ml^{-1} (Balaban and Novick, 1995)). Furthermore, this study revealed that *agr*-deficient strains were able to infect CF airways (Goerke *et al.*, 2000), therefore showing that *agr* is not activated or essential for chronic lung infection in CF patients. This result however, is in contrast to other animal models were *agr* mutants have been shown to be less virulent than the corresponding wild type (Abdelnour *et al.*, 1993). Therefore it is possible that the *agr* operon is only activated during particular types, and/or at certain stages, of a given infection.

Despite the observed low levels of RNAIII expression, *spa* was greatly diminished compared to its transcription *in vitro*. This was even true of *agr*-deficient strains, thus indicates *agr*-independent regulatory mechanisms of *spa in vivo*. Finally the transcription of α-haemolysin was detectable in sputum samples, but its levels were found not to correlate with RNAIII expression, Once more indicating the presence of signals and additional regulators not linked to *agr* (Goerke *et al.*, 2000).

1.9.2 *In vivo* expression analysis of *sar*

Studies on the three transcripts of *sar* showed these to be expressed temporally (Bayer *et al.*, 1996; Blevins *et al.*, 1999) (Chapter 1.7.2). In an attempt to determine if the individual *sar* promoters were differentially expressed *in vitro* and *in vivo,* green fluorescent protein (*gfp*) transcriptional fusions were constructed (Cheung *et al.*, 1998). *In vitro* findings showed P1 (*sarA*) to be the only active promoter as determined, by fluorescence microscopy. *In vivo* activity was determined in the rabbit endocarditis model 24 h after infection. This once more showed P1 to be active and P3 to be inactive. However, P2 (*sarB*) which was inactive *in vitro,* became highly activated *in vivo*. Thus *in vivo* promoter activation of *sar,* differed from that *in vitro*. Finally *gfp* studies showed that, the *sar* P2 promoter was activated on the surface, but not in the centre of the lesions. Therefore the individual *sar* promoters may be differentially

expressed within the host, presumably reflecting the physiological response to distinct host micro-environments (Cheung *et al.*, 1998).

1.10 Virulence factors and disease

The important question remains as to how do the concerted action of the virulence determinants and their global regulators lead to disease? Results suggest that attachment and growth are precursors to the production of most exotoxins. The fact that the attachment proteins require on going cell wall assembly to become anchored to the cell wall, is good reason for cells to only produce these under conditions which support cell division, and not post-exponentially. Likewise to have an optimum effect in the immediate environment of the bacterium, extracellular proteins would be best produced when the cells are localised and in sufficient quantity to produce enough material to have a 'beneficial effect' for the organism. This schaem suggests that sepsis due to staphylococcal infection is not the result of a generally disseminated contamination of the bloodstream. Rather it is the focal infection in which growing bacteria reach a cell density sufficient to lower the local pH, deplete available nutrients, and then produce degrading enzymes, pore-forming invasins, and immunomodulators that produce the symptoms associated with the disease (Projan and Novick, 1997).

Finally, despite the obvious central role played by the global regulators, *agr, sar* and *sae*, mutations at these loci, while demonstrating reduced virulence in virtually all models used (Abdelnour *et al.*, 1993; Cheung *et al.*, 1994; Booth *et al.*, 1995), did not render the bacteria avirulent. This is an important consideration for the design of antimicrobial substances, as a therapeutic approach designed against these targets might well fail.

1.11 *In vitro* and *in vivo* screens for bacterial genes potentially important in pathogenicity

Many techniques have been developed over the years, that have increased our knowledge and understanding of the mechanisms employed by bacteria to enable them to survive and proliferate inside the host, causing disease. Our current level of knowledge is however by no means complete, but has been greatly increased by the development of a succession of ingenious techniques, designed to identify key

components and determine their role in pathogenesis. Some of these techniques are described in the following sections.

1.11.1 *In vitro* analysis of pathogenicity

A variety of *in vitro* systems have been developed that recreate certain aspects of infection, enabling the study of bacterial gene expression, along with the behaviour of mutant strains in physiological conditions that reflect the situation *in vivo* (DiRita and Mekalanos, 1989). These include the use of specific culture conditions to mimic the host environment, as well as tissue culture assays for adhesion, invasion or cytotoxicity studies. These *in vitro* assays mimicking the infectious process have been immensely useful and continue to provide much information into the mechanisms of bacterial pathogenesis.

1.11.1.1 Plate-based screens

1.11.1.1.1 Screens for genes important in starvation-survival

Nutrient limitation of *S. aureus* has been shown to induce a starvation-survival state, which enables the bacterium to survive until sufficient nutrients become available to support growth (Watson *et al.*, 1998a). A library of Tn*917* mutants of *S. aureus* was screened for strains defective in their starvation-survival response, using a simple technique based on their ability to survive on nutrient limiting agar plates. 20 mutants were isolated displaying approximately 10- to 100-fold lowered survival potential under amino acid limiting conditions (Watson, 1997). These included *sodA,* encoding the major superoxide dismutase of *S. aureus,* and a gene with homology to *ctaA* of *B. subtilis*, which encodes a haem A synthase. Interestingly the *ctaA* mutant demonstrated a decreased production of both α-haemolysin and toxic shock syndrome toxin-1 (Clements *et al.*, 1999a). Despite the apparent importance of these two genes *in vitro,* later studies showed that neither of these mutants were defective in the ability to form lesions in a subcutaneous mouse infection model (Clements *et al.*, 1999a/b).

Interestingly a further screen for starvation-survival mutants under glucose limiting conditions (as described by Watson, 1997) identified the gene in *S. aureus* encoding catalase (*katA*). A defective *katA* was found not only to be impaired in its ability to enter the starvation-survival state, but importantly was also found to be attenuated in a

mouse subcutaneous lesion model (Horsburgh and Foster, unpublished data). This demonstrates that simple *in vitro* screens can identify genes that are both expressed and important *in vivo*.

1.11.1.1.2 Screens for genes expressed during growth in milk

Recently a study was undertaken to identify *S. aureus* genes expressed during growth in milk, in an attempt to gain an insight into the genes important for bovine mastitis, of which *S. aureus* is the major cause (Lammers *et al.*, 2000). In this study a Tn*917* transposon library was screened during growth on bovine milk and mutants were selected which expressed β-galactosidase in milk, but not in a nutrient-rich growth medium. This screen identified 28 genes, including those involved in cell wall synthesis, carbohydrate metabolism and transcriptional regulation (Lammers *et al.*, 2000). Furthermore many of the genes identified in this approach were also identified from *S. aureus* by *in vivo* approaches in various murine models of infection (Mei *et al.*, 1997; Coulter *et al.*, 1998; Lowe *et al.*, 1998) (Chapter 1.11.2).

1.11.1.2 *In vitro* studies using tissue culture for the identifications of *in vivo* induced genes

Several studies have been undertaken aimed at looking at the expression of genes on interaction with a variety of mammalian cell lines *in vitro*.

1.11.1.2.1 Screening for *L. monocytogenes* genes preferentially expressed in mammalian cells

A method for the identification of genes preferentially expressed in mammalian cells has been used with some success on the facultative intracellular pathogen *Listeria monocytogenes*. *Listeria* genes, with preferentially intracellular expression (*pic* genes), were identified using a library of Tn*917*::*lacZ* insertional mutants, by screening for transcriptional fusions with greater *lacZ* expression inside a macrophage like cell line, than in a nutrient-rich growth medium (Klarsfeld *et al.*, 1994). This screen identified five genes with as much as 100-fold induction within macrophages, including the isolation of a gene encoding a phosphatidylinositol-phospholipase C (*plcA*), which has been identified as a virulence factor (Camilli *et al.*, 1993). Importantly, the *plcA*

promoter drives the majority of transcription of *prfA,* the major virulence gene regulator of *L. monocytogenes* (Mengaud *et al.*, 1991). Furthermore a mutation within *plcA* led to a 25-fold increase in LD_{50} (Mengaud *et al.*, 1991).

A more recent screen has been conducted on *L. monocytogenes* in an attempt to further study the response of the interaction of bacterial cells with a variety of mammalian cells. This work showed that PrfA is induced upon adherence to mammalian cells. Additionally, it was also discovered that PrfA was up regulated by mammalian cell extracts, (Renzoni *et al.*, 1999) in the same manner as by whole cells, via the *plc* promoter (Klarsfeld *et al.*, 1994). This result strongly argues the role of host cellular component(s) in the activation of PrfA, either directly or indirectly (Renzoni *et al.*, 1999). However, the identity of the host signal and the sensory mechanism leading to increased *prfA* transcription remains unknown.

1.11.1.2.2 Screening for *S. aureus* genes preferentially expressed in mammalian cells

A similar study to the one described above (Chapter 1.11.1.2.1) has recently been conducted in *S. aureus*, in an attempt to assess its response to the cellular environment. In this study an expression library of *S. aureus* was used to identify genes whose expression was induced after exposure to human umbilical vein endothelial cells (HUVEC) (Vriesema *et al.*, 2000). This screen identified 33 different clones, containing transposons that had integrated into a gene showing preferential expression under the selective conditions. 19 of these clones showed homology at the protein level to entries in the database, while the remainder, were novel genes. The genes identified could be divided into different categories based on function of the encoded proteins: transport, catabolism, biosynthesis, and DNA repair. Finally a subset of five of these clones were analysed further, and showed that human serum, in addition to HUVEC were capable of triggering gene expression in *S. aureus* (Vriesema *et al.*, 2000).

Importantly these *in vitro* studies show that cell extracts/human serum induce the same genes respectively as whole cells, which are presumed to be the same as those induced *in vivo*. Therefore simple *in vitro* assays can produce credible results for the analysis of *in vivo* expressed genes.

1.11.1.3 Other *in vitro* screens

1.11.1.3.1 Two-dimensional gel electrophoresis (2DGE)

An alternative to studying differential gene expression is the analysis of differential protein production. Two-dimensional gel electrophoresis (2DGE) is a powerful method with high resolution and sensitivity and has been shown to be useful for the comparative analysis and detection of proteins from bacterial cells. The 2DGE system has been used to compare proteins from *S. typhimurium,* grown in laboratory broth and macrophages. This demonstrated that a number of proteins were, induced by host tissues, including the stress proteins GroEL and DnaK (Buchmeier and Heffron, 1990).

1.11.1.3.2 Subtractive and differential hybridization

Several strategies for the identification of genes preferentially induced by host tissues have been described, which involve the study of either mRNA or cDNA for differences in gene expression under different conditions. Differential subtractive hybridization systems have recently been developed for use in prokaryotes and are particularly useful for organisms that do not have a well-defined background. These techniques allow the identification of genes, transcriptionally induced under different conditions, for example in response to different external stimuli (reviewed by Handfield and Levesque, 1999). Differential and subtractive hybridization techniques are variations on a thaem, with both systems based on mRNA expression levels. However, rather than relying on specific probes, as other mRNA based studies do (for example, Northern blotting), differential subtractive hybridization systems rely on random isolations of unknown and differently expressed genes. An example of the application of this technique was in the analysis of gene expression of *Mycobacterium avium* following phagocytosis by human macrophages (Plum and Clarck-Curtiss, 1994). In this study total RNA was isolated from *M. avium* grown in parallel in human-derived macrophages and nutrient-rich broth. Bacterial mRNA was then converted to cDNA, and biotin-labeled cDNA prepared from broth grown *M. avium* was used to subtract housekeeping genes. After each round of subtraction, a sample of the unsubtracted cDNA was amplified, labeled and hybridized to a cosmid library under stringent conditions. This study identified one gene as specifically induced in *M. avium* grown in cell culture, although this gene was shown to have no significant homologue in the database (Plum and Clarck-Curtiss, 1994). Interestingly all other cell culture induced genes were also expressed in broth grown bacteria.

1.11.1.3.3 Oligonucleotide arrays

A panoramic view of bacterial transcription can also be obtained with the use of high-density oligonucleotide probe arrays to monitor gene expression. Bacterial transcript imaging was studied on a chip containing probes for 106 *H. influenzae* and 100 *S. pneumoniae* genes which demonstrated the applicability of oligo arrays for monitoring global gene expression (de Saizieu *et al.*, 1998). As well as allowing simultaneous quantification of the transcript level, the technique was found to be very sensitive (range of 1-5 transcripts per cell). It has been stated that with the increasing number of fully sequenced bacterial genomes, parallel methods for monitoring gene expression should provide information for understanding the bacterial response in an infection and studying virulence gene induction *in vivo* (Handfield and Levesque, 1999).

1.11.2 *In vivo* genetic analysis of bacterial virulence

Upon entering the host, pathogenic organisms find themselves in a situation that must differ significantly from any encountered in the environmental reservoir. Bacteria are able to respond to this change in circumstance by modulating their pattern of gene expression accordingly, down regulating genes that are no longer necessary, and up regulating the expression of those that are specifically required for survival within the host (Chapter 1.6.3). It is therefore postulated that the majority of these up regulated components will be virulence factors and that at least some will play a critical role in pathogenesis.

Many potential virulence factors of *S. aureus* have been studied in great detail *in vitro* over the last few decades, and these *in vitro* assays continue to contribute a great deal to the understanding of bacterial pathogenesis. However, generally they cannot replicate the complex environment encountered by pathogens during infection. For example, a pathogen may encounter several radically different environments within the host, and it may therefore have very different requirements at various points during infection, particularly in terms of a developing immune response. Thus the information gained from such *in vitro* studies is of limited use. Therefore, *in vivo* experimental models are highly desirable as they permit direct assessment of a pathogen's ability to colonize, survive and cause disease, or damage, in a living host.

In vivo models allow the analysis of bacterial pathogenesis *in situ,* however despite their obvious advantages they have their own limitations being generally expensive, and

often laborious. This has to some extent, reduced the inclination to conduct studies in living animals especially on a large-scale. As a result until recently, very little has been determined about the role of virulence factors and their expression *in vivo*. This is true for almost all pathogenic bacteria, although some screening of individual mutant strains for altered virulence has been carried out on a limited scale with both plant and animal infection models, using either randomly selected transposon mutants (Bowe *et al.*, 1998; Osbourne *et al.*, 1987) or strains affected in cell surface or extracellular protein production (Miller *et al.*, 1989).

Recently however, steps have been taken to address the abyss in our knowledge, with the development of several methods that greatly simplify *in vivo* analysis, allowing the simultaneous analysis of large numbers of strains. This is important as the discovery and characterisation of genes preferentially induced *in vivo* during infection is important, as it will comprise the next phase in the study of bacterial virulence at the molecular level.

Most of these *in vivo* methods fall into two broad groups: *in vivo* expression technology, and signature-tagged mutagenesis. *In vivo* expression technology (IVET) is a promoter trap strategy designed to identify promoters that are specifically activated within the host. Signature-tagged mutagenesis (STM) is the alternative method and uses comparative hybridisation to isolate mutants unable to survive specified environmental conditions, and has been used to identify genes critical for survival in the host. These methods and their major findings are outlined in the following sections.

1.11.2.1 *In vivo* expression technology (IVET)

In vivo expression technology (IVET) was initially designed to identify putative virulence genes in *Salmonella typhimurium* (Mahan *et al.*, 1993), but has since been adapted for use in many bacterial species. This method for the first time allowed the study of the bacterial response to the host environment using a gene expression system within the animal host itself, for the selection of genes that are specifically expressed during infection. The IVET system functions as a promoter trap for genes that are specifically induced in host tissues. Four variants of IVET have been designed to date, all enabling the isolation of *in vivo* induced (*ivi*) genes. These use: auxotrophic selection, antibiotic selection, genetic recombination and green fluorescent protein, as the reporter to detect *in vivo* promoter activity.

1.11.2.1.1 Auxotrophic complementation as a reporter for *in vivo* gene expression

The original *S. typhimurium* IVET system required the availability of a bacterial auxotroph (Mahan *et al.*, 1993). This nutritional deficiency needed to be supplemented *in vitro* and complemented *in vivo* by expression from the *ivi* promoter fused to the gene used for selection. Transcriptional fusions were created by inserting fragments of chromosomal DNA into a suicide delivery plasmid, upstream of a promoterless operon fusion of the *lacZY* genes, fused to *purA* (purine metabolism). This library was then introduced via homologous recombination into *S. typhimurium* (*purA*$^-$ strain). Importantly the integration of fusions by homologous recombination into the host bacterium chromosome generates a duplication of the cloned DNA, leaving a functional copy of the wild type gene, in addition to the partial copy fused to the reporter operon. This is important as it allows the analysis of genes essential for growth *in vivo* and assured that no polar effects were present (Mahan *et al.*, 1993).

Bacteria representing the pool of chromosomal fusions were injected intraperitoneally into mice, and the surviving bacteria (those able to overcome the purine deficiency), were recovered following three days infection. These were screened on laboratory medium for clones with low promoter activity, via β-galactosidase activity. Clones containing fusions to genes specifically induced in the animal model showed little *lacZ* expression in synthetic laboratory media, although clearly these clones had sufficient transcription levels to survive within the mouse (Mahan *et al.*, 1993). This led to the identification of 15 strains carrying promoters specifically induced *in vivo*. Analysis of mutations of some of these *ivi* genes, were found to have attenuated virulence, thereby validating the ability of IVET to isolate virulence genes.

A slight variation on the auxotrophy thaem has been described (Wang *et al.*, 1996a), in which the products of *purEK* (purine metabolism) were used as selection in an adenine auxotrophic background. This system relied solely on colony size, rather than expression from a reporter gene, to measure *in vitro* activities of gene fusions. The rational here was that a lower activity would result in smaller colony size on minimal medium containing limiting amounts of adenine. Application of this strategy identified *ivi* genes in *Pseudomonas aeruginosa* induced in both mice (Wang *et al.*, 1996a) and in respiratory mucus collected from cystic fibrosis (CF) patients (Wang *et al.*, 1996b). These studies identified novel loci specifically induced under their respective conditions. Additionally two loci were identified by both methods one of which lead to an approximately 100-fold increase in LD_{50}.

This IVET method, although successful in some bacteria, requires the existence of an attenuating and complementary auxotrophy, which is not available in all bacterial systems.

1.11.2.1.2 Antibiotic selection as a reporter for *in vivo* gene expression

An alternative to the above auxotrophic complementation method, has been developed that uses a selection based on antibiotic resistance in the host, rather than alleviation of nutritional deficiencies. This was originally carried out using the gene encoding chloramphenicol acetyltransferase (*cat*) as a selection marker fused to *lacZY*. The antibiotic based IVET was developed for *S. typhimurium* (Mahan *et al.*, 1995), and was used for the identification of *ivi* genes in both cultured macrophages as well as in mice. In this system, *ivi* genes were identified, by host induced chloramphenicol resistance *in vivo* and low promoter activity *in vitro* (β-galactosidase). Screens of this nature in *S. typhimurium* have lead to the identification of several genes that were known virulence determinants. For example, *phoP* was identified which induces the expression of multiple virulence genes following invasion of macrophages (Heithoff *et al.*, 1997). However a large proportion had either no known homologues or none with known function (reviewed by Chiang *et al.*, 1999). Antibiotic based IVET has also been used successfully to identify host responsive elements (*hre*) in *Yersinia entrocolitica*. Disruption of four of these *hre* genes lead to increased LD_{50}, decreased persistence in host tissues, or both (Young and Miller, 1997).

Significantly, adjustment of the antibiotic dosage or timing of its administration may permit the isolation of *ivi* promoters with different levels of activity, or expressed differentially during the infection cycle. This is a major advantage of this system since gene activation and expression is not an all or nothing phenomenon. This system could however present problems in systemic distribution of chloramphenicol and limitations for selection of chloramphenicol resistance in certain animal models.

1.11.2.1.3 Genetic recombination as a reporter for *in vivo* gene expression

The previous IVET variations (auxotrophic complementation or antibiotic selection) both favour the identification of genes that are highly expressed throughout infection, as fusions to genes that are transiently or lowly expressed during infection, generally would be unable to survive. To allow the identification of promoters that are only lowly

or transiently expressed during infection, an alternative IVET system was designed that relies on the use of *tnpR* operon fusions encoding a site-specific resolvase (γ δ resolvase), as the selection gene and *lacZY* as reporter genes. Resolvase is an enzyme that catalyses the irreversible recombination between specific DNA sequences, termed *res* sites (Camilli *et al.*, 1994).

In this system resolvase activity results in the permanent excision from the chromosome of a tetracycline resistant cassette flanked by *res* sites. This *in vivo* activity can be detected by screening, following the recovery of bacteria from the animal (Camilli *et al.*, 1994). As the action of resolvase is permanent this IVET system permits the detection of promoter activity even if the promoter is only briefly active *in vivo*. Unfortunately this method does not have the benefits of positive selection, as auxotrophic complementation or antibiotic selection both do, although this technique should be significantly more sensitive.

An early application of resolvase IVET was for the identification of *ivi* genes in *Vibrio cholerae*. Credibility of this technique was confirmed on the determination that three of the *ivi* loci showed colonization defects in infant mouse competition assays (Camilli and Mekalanos, 1995).

Resolvase IVET technology has also been conducted without homologous recombination, in *S. aureus,* using the murine renal abscess model to identify *ivi* genes (Lowe *et al.*, 1998). A collection of 45 *ivi* genes, were isolated of which, only 17 had homology to known genes. Interestingly, one of these included *agrA,* part of the accessory gene regulator of *S. aureus* (Chapter 1.7.1). However, the majority of the *ivi* genes identified were not classical virulence factors but were involved in adaptation to the *in vivo* environment. Notably, a selection of 12 *ivi* genes, were insertionally inactivated, of which 7, were found to be attenuated in virulence, in a murine renal abscess model (Lowe *et al.*, 1998).

1.11.2.1.4 Differential fluorescence induction using green fluorescent protein as a reporter for *in vivo* gene expression

The final and most sophisticated IVET technique for the identification of *ivi* is based on differential fluorescence induction (DFI) (Valdivia and Falkow, 1996), which allows the rapid identification of bacterial genes induced upon association with host cells. This assay uses green fluorescent protein (GFP) (Cormack *et al.*, 1996), as a selectable

marker, in conjunction with fluorescence-activated cell sorting (FACS). DFI was originally developed to identify *S. typhimurium* genes that are induced *in vitro* by phagocytic cells, but was also found to be capable of identifying genes induced within murine cells *in vivo* (Valdivia and Falkow, 1997). Following infection, host cells infected with a bacterium bearing a transcriptionally active *gfp* gene fusion, were separated by FACS and lysed. The recovered bacteria were then grown under *ex vivo* conditions, and bacteria, demonstrating low, or no fluorescence in the absence of host cells, were isolated by FACS (Valdivia and Falkow, 1997).

This methodology led to the identification of 14 *S. typhimurium ivi* genes. Of these eight had bacterial homologues of known function, some of which had previously been shown to have a role in virulence. For example, *ssaH*, a gene encoding part of the type III secretion system, located within the *Salmonella* pathogenicity island 2 (SPI-2) which has been shown to be necessary for *S. typhimurium* survival in mice (Shea *et al.*, 1996). Five of the novel *ivi* genes were disrupted in order to establish their role in virulence. Results showed that two of the novel loci demonstrated decreased spleen colonization in a murine competition assay (Valdivia and Falkow, 1997). The advantage of this system over others is that it allows real-time measurement of gene activity *in vivo* and should be widely applicable to the identification of genes from pathogenic organisms that are induced upon association with host cells or tissues.

The collective success of these IVET approaches is demonstrated by the identification of more than 100 *S. typhimurium* genes that are induced during infection (Mahan *et al.*, 1993; Mahan *et al.*, 1995). More recently the discovery by Heithoff and colleagues (1999), that a *S. typhimurium* strain deficient in DNA adenine methylase (Dam) was totally avirulent, along with the discovery that this locus represses the expression of more than 20 *in vivo* induced genes, further highlighted the importance of IVET and its continued use in the future.

1.11.2.2 Signature-tagged mutagenesis (STM)

The second technique enabling the study of bacterial pathogenesis *in vivo* is signature-tagged mutagenesis (STM) (Hensel *et al.*, 1995). This approach uniquely circumvents the previous problems associated with exploiting transposon mutagenesis, due not only to the inability to identify mutants with attenuated virulence within pools of mutants, but also to the impracticalities of separately assessing the virulence of individual

mutants. STM is a comparative hybridization technique that uses a collection of uniquely tagged transposons, each one modified by the incorporation of a different DNA sequence (tag) (Hensel *et al.*, 1995). This system allowed the identification of bacteria recovered from hosts infected with a mixed population of mutants, in addition to the selection of mutants with attenuated virulence. Furthermore, because STM is a random process, a large enough study would represent a whole genome scan for *ivi* genes.

The tags used to label transposons are DNA segments that contain a 40 bp variable central region, flanked by 20 bp invariant 'arms' that facilitate the co-amplification and labeling of the central portion by PCR (Hensel *et al.*, 1995). When tagged transposons are used to mutagenize a bacterium, each individual mutant can in theory be distinguished from all the other mutants based on the different tag carried by its integrated transposon. Tagged mutants are assembled separately in microtitre dishes, and combined prior to infection to form the inoculum pool. Pools of up to 96 differently tagged mutants (pool complexity is discussed by Chiang *et al.* (1999)) are subjected to a selective process, such as infection of an animal. Following a set period of infection, bacteria are recovered, and DNA extracted from the infection pools. The tags present in these DNA samples are amplified and labeled by PCR. Colony blots are then probed and compared with the hybridization patterns obtained with the use of tags amplified from the inoculum pool as a probe (Hensel *et al.*, 1995). The technique of STM relies on the ability of the pathogen in question to replicate *in vivo* as part of a mixed population, and thus it can only be expected to identify virulence genes whose mutant phenotype cannot be complemented by other clones present in the same inoculum (Chiang *et al.*, 1999).

The original application of STM was for the identification of *ivi* in *S. typhimurium* (Hensel *et al.*, 1995), and resulted in the identification of 28 mutants with attenuated virulence by the identification of tags that were present in the inoculum pool, but not in the pool recovered from the infected mice. Significantly this study resulted in the identification of the second pathogenicity island, SPI2 of *S. typhimurium*. This was later determined to encode the type III secretion system (Shea *et al.*, 1996) that is essential for infection in two animal models (Hensel *et al.*, 1997).

The virulence of *S. aureus* has been assayed on two separate occasions using STM. In the first study (Mei *et al.*, 1997), pools of Tn*917* mutants were tested in a murine model of bacteraemia. This screen resulted in the identification of 50 mutants demonstrating attenuated virulence (Mei *et al.*, 1997). DNA sequence analysis of regions flanking the

transposon insertions revealed, approximately half of these genes have no known function. The majority of the remaining loci identified could be categorized into three broad groups according to their predicted identity from sequence analysis. These groups are those involved in: cell surface metabolism (for example, peptidoglycan cross-linking and transport functions); nutrient biosynthesis; and cellular repair processes.

The second *S. aureus* STM assay (Coulter *et al.*, 1998) was a more complex study involving the analysis of pools of *S. aureus* Tn*917* mutants in three murine infection models: abscess; bacteraemia; and wound infection, for growth attenuation. This enabled the identification of various genes affecting growth and virulence in specific disease states. In total 237 mutants were isolated from all three *in vivo* screens. Interestingly a subset of 23 mutants showed significant and reproducible attenuation in all three *in vivo* environments. Notably, 18 of the genes identified in this latter assay (Coulter *et al.*, 1998), were also isolated by Mei *et al.* (1997).

1.11.2.3 Advantages and disadvantages of the IVET and STM systems

Chapter 1.11.2.1 described how IVET has been used for the identification of both novel and known *ivi* virulence genes. A significant advantage of this technique over others (including STM) is that with the exception of recombinant IVET (Chapter 1.11.2.1.3), this methodology employs a positive selection approach for the identification of *ivi*. However not surprisingly only a proportion of the genes identified as specifically expressed *in vivo*, were found to have a significant role in virulence, as determined by pathogenicity assays. These non-essential *ivi* genes however could be responsible for damage to the host, which is not assessed in most pathogenicity studies. Furthermore it may be that many *ivi* loci make small individual contributions to virulence, and thus their effects may be additive or synergistic (Chiang *et al.*, 1999).

Ideally, IVET should be capable of studying both increases and decreases in promoter activity. However to date, there has been no report of an IVET strategy designed to identify genes whose expression must be down regulated during infection. Additionally it would also be desirable to be able to analyse changes in expression level quantitatively. This type of analysis is potentially possible with the antibiotic-based IVET system, which theoretically permits the identification of *ivi* promoters with different levels of activity through variation of antibiotic levels (Chiang *et al.*, 1999). However no such studies have been carried out and thus the logistics of such an approach are not fully understood.

However, the most significant disadvantage of IVET is that in most of its present guises, it selects against genes that are expressed *in vitro*. This is however in spite of the fact that there is no evidence suggesting that *in vitro*-expressed genes are not important for either survival *in vivo* or in the causation of host damage.

Results from STM experiments differ from those of IVET, as by definition all genes identified via STM are attenuated in virulence, regardless of the expression pattern of the mutated gene in that strain. Therefore STM is a more direct method for isolating genes necessary for *in vivo* survival, as all genes identified by IVET, must subsequently be mutated to enable their role in virulence to be determined. Unfortunately however, STM is a technique, based on negative selection whereby mutants bearing the desired phenotype (attenuated virulence) are selected against by the host (Hensel *et al.*, 1995). These mutants of interest must then be identified following infection by hybridization screening. Therefore STM is not as simple a selection process as IVET, although the vast majority of mutant strains identified, in *S. typhimurium, S. aureus* and *V. cholerae* by STM have subsequently been shown by LD_{50} tests or competition analysis to be important for growth *in vivo* (reviewed by Chiang *et al.*, 1999).

A further problem with STM is that this technique may not be able to identify mutations in genes leading to weakly or even moderately attenuated virulence, depending on the sensitivity of the strategy used for detecting such changes in tagged populations. This is particularly important as this subset includes mutations in genes that are essential for the causation of disease, but do not affect the survival of the bacterium in the host. Cholera toxin is an example of such a gene, as while its action is primarily responsible for the lethality of cholera, its inactivation does not affect the bacterium's ability to colonize the host (Taylor *et al.*, 1987).

It is potentially of interest, that while many of the genes identified by IVET and STM in *S. typhimurium* fall within the definition of a classical virulence factor, the majority of *S. aureus* genes identified by these techniques appear to have more fundamental roles in bacterial metabolism (Mahan *et al.*, 1993; Mahan *et al.*, 1995; Hensel *et al.*, 1995; Chiang *et al.*, 1999; Lowe *et al.*, 1998; Mei *et al.*, 1997; Coulter *et al.*, 1998). Several plausible explanations exist to explain this phenomenon. Firstly it may be due to the fact that classical virulence factors are expressed at too high a level *in vitro* to qualify as *in vivo*-induced. This theory is supported by work carried out by Chan and Foster (1998b), who showed that the toxins; α-haemolysin and toxic shock syndrome toxin-1 along with the global regulator Sar are expressed to a significant level (as determined by β-galactosidase production), in laboratory medium. Secondly a transposon insertion

within a gene encoding a toxin may result in a phenotype that can be complemented by other mutant strains within the infection pool. Finally, no STM strategy of *S. aureus* to date, has examined more than approximately 1,500 mutants, thus potentially these studies have not been of sufficient size to allow saturation of the *S. aureus* chromosome.

In conclusion these two systems appear to be widely applicable with both IVET and STM having been used for the identification of *ivi* genes in a range of bacterial pathogens including; *S. typhimurium, S. aureus,* and *V. cholerae*. Of significant importance is the fact IVET and STM appear to be complementary approaches in respect to the genes which they identify. Therefore these techniques can be used in combination, enabling a comprehensive search for *ivi* genes. Of great concern however, is the fact that neither IVET nor STM can be expected to identify genes that are essential for bacterial growth. This has major implications, as essential genes provide prime targets for antimicrobial programs.

1.11.2.4 Identifications of essential genes using GAMBIT

Essential genes are by definition vital for growth and proliferation *in vitro*. As a result it is predicted that these genes would be expressed under all conditions, and thus unless expressed only at extremely low levels *in vitro*, they would not be identified by IVET. Additionally mutations within genes essential for pathogenesis would invariably be lethal, thus STM would equally be unable to assess this class of genes. For these reasons a systematic and effective method for studying essential genes was developed.

A technique termed GAMBIT for 'genomic analysis and mapping by *in vitro* transposition' allows the identification of essential genes through the application of extended-length PCR, *in vitro* transposon mutagenesis, transformation, and genetic footprinting (Akerley *et al.*, 1998). This methodology like STM constitutes a negative selection, in which certain mutants are eliminated by the selection conditions. These mutants are recognized by the loss of PCR products, corresponding to insertions in the essential genes. A particular advantage of GAMBIT is its ability to target specific genes or regions of the chromosome for example pathogenicity islands (Akerley *et al.*, 1998). However, its application to an entire genome would entail a substantial amount of work (Chiang *et al.*, 1999).

The technique of GAMBIT has already been applied to both *Haemophilus influenzae* and *Streptococcus pneumoniae* (Akerley *et al.*, 1998). This study detected known essential genes and assigned essentiallity to several genes of unknown function from both organisms, thus confirming its potential for the analysis of pathogenesis (Akerley *et al.*, 1998). The GAMBIT approach is predicted to prove equally useful for the identification of genes essential for pathogenic organisms to grow and survive *in vivo*. Unfortunately this methodology is severely limited as it is only available for naturally transformable organisms whose genomes have been sequenced. However, the completion of genome sequencing projects and the development of efficient linear DNA transformation techniques, should allow the use of GAMBIT in a range of bacterial pathogens (Akerley *et al.*, 1998; Chiang *et al.*, 1999).

1.11.3 Summary of strategies for the *in vitro* and *in vivo* analysis of bacterial virulence

As described in Chapter 1.11.1 and 1.11.2, many different techniques exist for the analysis of bacterial virulence under both *in vitro* and *in vivo* conditions. Studies of this nature have been greatly advanced by the recent sequencing of bacterial genomes. Furthermore with the adoption of techniques such as IVET, STM it is now possible to study genes expressed or preferentially induced *in vivo*, a proportion of which will be critical in bacteria-host interactions. Importantly with the introduction of DFI, gene expression can now be studied *in situ,* in real time. Finally, with the development of GAMBIT, there now exists a method for the identification of genes with an essential role.

With the introduction of this multitude of *in vivo* assays there has been a dramatic shift away from the use of classical *in vitro* assays for the analysis of bacterial components. However one problem arising from these new technologies is the fact that they rely on animal models, with the limitations that such models may impose, along with the financial and ethical implications. Therefore there is still a place for the development of new *in vitro* models for the analysis of bacterial virulence.

This thesis aims to describe the development of such an *in vitro* model and describes its use for the identification and subsequent analysis of genes potentially involved in virulence.

1.12 Rationale and objectives for this project

Over many years bacteria have evolved highly sophisticated regulatory mechanisms to modulate their gene expression in response to disparate environments. Our understanding of such processes during infection is limited by our ability to recreate host conditions in an experimental setting. It has been shown that different approaches have been undertaken in an attempt to gain a better understanding of the mechanisms by which *S. aureus* causes disease, the majority of these are both time consuming and technically demanding. Therefore the main objective of this project was to design a simple *in vitro* assay system, whose conditions mimic those *in vivo*. This model would then be used initially for the analysis and characterisation of the role of known virulence determinants in an attempt to gain further information into their role/expression *in vivo*. Secondly the model system will be used for the identification and analysis of novel components potentially involved *in vivo*.

CHAPTER 2

MATERIALS AND METHODS

2.1 Media and antibiotics

2.1.1 Media

Unless otherwise stated all media was prepared using distilled water (dH_2O) and was sterilised by autoclaving for 20 min at 121 °C (15 pounds per square inch).

2.1.1.1 B2

Casein acid hydrolysate	10 g l^{-1}
Difco Bacto yeast extract	25 g l^{-1}
K_2HPO_4	1 g l^{-1}
NaCl	25 g l^{-1}

The pH was adjusted to 7.5. Once autoclaved and cooled, sterile glucose was added to 1 mM (final concentration).

2.1.1.2 Brain heart infusion (BHI) (Oxoid)

Brain heart infusion (BHI)	37 g l^{-1}

Oxoid Agar No. 1 (1 % (w/v)) was used for BHI agar

Details of the composition of BHI are shown in Table A.1.1.

2.1.1.3 Chemically defined medium (Hussain *et al.*, 1991)

A chemically defined medium (CDM) as devised by Hussain *et al.* (1991), is composed of 5 solutions (1-5) which were mixed in a ratio of 7 (solution 1) : 1 (solution 2) : 0.5 (solution 3) : 0.1 (solution 4) : 1 (solution 5).

Details of the composition of solutions 1-5 are given in Table A.1.2.

Solutions 1, 3 and 4 were mixed and autoclaved together, whereas solution 2 was filter sterilised (0.2 μm pore size) and added after cooling to 50 °C. Solution 5 was autoclaved separately and added after cooling. Difco Bacto agar (1.5 % (w/v)) was used for CDM agar.

2.1.1.3.1 Aspartate amino acid varying CDM

CDM 1-16 was identical to that described in Chapter 2.1.1.3, except for the presence or absence of the aspartate amino acids (AAA): lysine; threonine; methionine; and isoleucine. In CDM 1-16 these four amino acids were either included in the same amounts as shown in Table A.1.2, or they were totally absent. Details of the aspartate amino acid composition of CDM 1-16 are shown in Table 5.2. All CDM (1-16) were made up lacking all four AAA (CDM 15). The AAA were then added from stock solutions (Chapter 2.2.2), as required.

2.1.1.4 Luria-Bertani (LB) (Miller, 1972)

Difco Bacto tryptone	10 g l^{-1}
Difco Bacto yeast extract	5 g l^{-1}
NaCl	5 g l^{-1}

The pH was adjusted to 7.2 using NaOH. Oxoid Agar No. 1 (1.0 % (w/v)) was added for L agar (LA).

2.1.1.5 LK

Difco Bacto tryptone	10 g l^{-1}
Difco Bacto yeast extract	5 g l^{-1}
KCl	7 g l^{-1}
Sodium Citrate	0.5 g l^{-1}

Oxoid Agar No. 1 (1.0 % (w/v)) was added for LK bottom agar.
Oxoid Agar No. 1 (0.7 % (w/v)) was added for LK top agar.

2.1.1.6 Phage agar

Casamino acids	3 g l^{-1}
Difco Bacto yeast extract	3 g l^{-1}
NaCl	5.9 g l^{-1}

Oxoid Agar No. 1 (1.0 % (w/v)) was added for phage bottom agar.
Oxoid Agar No. 1 (0.33 % (w/v)) was added for phage top agar.

2.1.1.7 Sera (Sigma)

Sera were purchased from Sigma as frozen liquid. This serum had been pooled from normal animals/humans and had been filter sterilised (0.1 μm pore size) and aseptically filled. All sera were collected from clotted whole blood to preclude additives. Unless otherwise stated pig serum was used in all serum experiments.

Pig serum	(Sigma, S7398)
Human serum	(Sigma, S7023)

Details of the composition of serum are shown in Table A.1.5-A.1.7.

Liquid - undiluted serum (100 % (v/v)).

Plates - 1 volume of serum pre-warmed to 50 °C, was added to 1 volume of 50 °C agar (Difco Bacto agar (2.0 % (w/v)) giving a final concentration of agar (1.0 % (w/v)). Agar plates contained 50 % serum (v/v).

2.1.1.7.1 Serum plus iron

A 50 mM stock solution of iron sulphate ($FeSO_4 \cdot 7H_2O$) dissolved in distilled H_2O was filter sterilised (0.2 μm pore size). This stock solution was added to serum to give a range of final concentrations of iron: 0 μM; 2 μM; 20 μM; 200 μM; 500 μM; and 2000 μM.

2.1.1.7.2 Serum plus catalase

A 10 mg ml^{-1} stock solution of catalase (Sigma) dissolved in 10 mM Tris pH 7.5, was filter sterilised (0.45 μm pore size) and stored at -20 °C. This stock solution was added to serum to give a range of concentrations of catalase: 0 μg ml^{-1}; 10 μg ml^{-1}; 100 μg ml^{-1}; 500 μg ml^{-1}; 1000 μg ml^{-1}.

2.1.1.8 Sheep blood agar

Blood base agar (Difco)	39.5 g l^{-1}

Agar was allowed to cool to 50°C, and 5 % (v/v) sterile sheep blood was added (TCS Microbiology).

2.1.1.9 SOB (Sambrook *et al.*, 1989)

Difco Bacto tryptone	2 % (w/v)
Difco Bacto yeast extract	0.5 % (w/v)
NaCl	10 mM
KCl	2.5 mM

Once autoclaved and cooled, sterile supplements of $MgCl_2$ and $MgSO_4$ were added, each to 10 mM (final concentration).

2.1.1.10 SOC (Sambrook *et al.*, 1989)

This was prepared by the addition of sterile glucose (to 20 mM) to SOB.

2.1.1.11 Synthetic cerebrospinal fluid I (van Dijk *et al.*, 1994)

NaCl	127.64 mM
KCl	2.55 mM
$CaCl_2$	1.26 mM
$MgCl_2$	0.93 mM

2.1.1.12 Synthetic cerebrospinal fluid II (Smith and Johanson, 1983)

NaCl	130 mM
$NaHCO_3$	18 mM
KH_2PO_4	3.0 mM
$CaSO_4$	1.4 mM
$MgSO_4$	0.8 mM
Urea	2.0 mM
Glucose	12.1 mM

2.1.1.13 Tryptic Soy Broth (TSB) (Difco)

Tryptic soy broth	30 g l^{-1}

Oxoid Agar No. 1 (1.0 % (w/v)) was used for TSB agar.

2.1.2 Antibiotics

All antibiotics used in this study are listed in Table 2.1. The stock solutions were filter sterilised (0.2 μm pore size) and stored at -20 °C. Once the media had cooled to below 55 °C the antibiotic stock solutions were added to the media as appropriate. Concentrations of antibiotics used for selection were as in Table 2.1, unless otherwise stated in the text.

2.2 Buffers and stock solutions

All buffers were prepared in dH_2O and stored at room temperature. Solutions for use in microbiological work and *in vitro* DNA manipulations were sterilised by autoclaving. All of the methods in this chapter are as in Sambrook *et al.* (1989), unless otherwise stated. Solutions for RNA work were treated with 0.1 % (v/v) DEPC overnight at room temperature to remove RNases, before sterilisation by autoclaving.

Antibiotic		* Stock concentration (mg ml^{-1})	Working concentration (µg ml^{-1})	Inducing concentration (µg ml^{-1})
Erythromycin	(Ery)	5[a]	5	0.15
Tetracycline	(Tet)	5[a]	5	-
Lincomycin	(Lin)	25[b]	25	-
Kanamycin	(Kan)	50[b]	50	-
Ampicillin	(Amp)	50[b]	50	-
Chloramphenicol	(Cm)	30[c]	5	-

Table 2.1

Concentrations of antibiotics used during this study

* The antibiotics were dissolved in: 50 % (v/v) ethanol[a], water[b], ethanol[c], filter sterilised (0.2 µm pore size) and stored at -20 °C.

2.2.1 ABT buffer

NaCl	100 mM
K_2HPO_4	60 mM
KH_2PO_4	40 mM
Triton® X-100	0.1 % (v/v)

2.2.2 Aspartate amino acid stock solutions

Stock solutions of the aspartate amino acids (lysine, threonine, methionine and isoleucine) were made up (100x working concentration (Table A.1.2)) and sterilised.

Lysine	14.29 g l^{-1}
Threonine	21.43 g l^{-1}
Methionine	14.29 g l^{-1}
Isoleucine	21.43 g l^{-1}

These amino acids were added to the CDM lacking all four AAA (CDM 15), as required (Chapter 2.1.1.3.1).

2.2.3 DNA loading buffer (6x)

Bromophenol blue	25 mg
Xylene cyanol FF	25 mg
Ficoll (MW 400,000)	1.5 g
dH_2O	to 10 ml

2.2.4 Frozen storage buffer (FBS)

KCl	7.4 g
$MnCl_2 \cdot 4H_2O$	8.9 g
$CaCl_2 \cdot 2H_2O$	1.5 g
$Co(NH_3)_6Cl_3$	0.8 g
Potassium acetate (1 M pH 7.5)	10 ml
Glycerol	100 g

The above were mixed with sterile dH_2O to 1 l and the pH adjusted to 6.4 with 0.1 M HCl. The solution was divided into aliquots and stored at -20 °C. Aliquots were defrosted when needed and filter sterilised (0.45 μm pore size) prior to use.

2.2.5 HPLC buffers

2.2.5.1 HPLC buffer A

NaOH	40 mM
Methanol	20 % v/v

Buffer A was adjusted to pH 4.23 with phosphoric acid, and a small amount of sodium azide (145 μl from 1 % (w/v) solution) was added to 1 l of buffer A to equalise its A_{202} with that of buffer B (Chapter 2.2.5.2).

2.2.5.2 HPLC buffer B

NaOH	40 mM

Buffer B was adjusted to pH 4.1 with phosphoric acid.

2.2.6 Phage buffer

$MgSO_4$	1 mM
$CaCl_2$	4 mM
Tris-HCl pH 7.8	50 mM
NaCl	0.59 % (w/v)
Gelatin	0.1 % (w/v)

2.2.7 Phosphate buffered saline (PBS)

NaCl	8 g l^{-1}
Na_2HPO_4	1.4 g l^{-1}
KCl	0.2 g l^{-1}
KH_2PO_4	0.2 g l^{-1}

The pH was adjusted to 7.4 using NaOH.

2.2.8 QIAGEN buffers

2.2.8.1 Buffer P1

Tris-HCl, pH 8	50 mM
EDTA	10 mM
RNase A	100 µg ml^{-1}

2.2.8.2 Buffer P2

NaOH	200 mM
SDS	1 % (w/v)

2.2.8.3 Buffer P3

Potassium acetate, pH 5.5	3.0 M

2.2.8.4 Buffer B1

Tris-HCl, pH 8	50 mM
EDTA, pH 8	50 mM
Tween® 20	0.5 % (v/v)
Triton® X-100	0.5 % (v/v)

2.2.8.5 Buffer B2

Guanidine HCl	3.0 M
Tween® 20	20 % (v/v)

2.2.8.6 Buffer QBT

NaCl	750 mM
MOPS, pH 7.0	50 mM
Ethanol	15 % (v/v)
Triton® X-100	0.15 % (v/v)

2.2.8.7 Buffer QC

NaCl	1.0 M
MOPS, pH 7.0	50 mM
Ethanol	15 % (v/v)

2.2.8.8 Buffer QF

NaCl	1.25 M
MOPS, pH 8.5	50 mM
Ethanol	15 % (v/v)

2.2.8.9 Buffer QG

Supplied in the QIAquick kit, details not provided.

2.2.8.10 Buffer PB

Supplied in the QIAquick kit, details not provided.

2.2.8.11 Buffer PE

Supplied in the QIAquick kit, details not provided.

2.2.8.12 Buffer EB

Tris-HCl, pH 8.5	10 mM

2.2.9 RNA extraction buffers and solutions (BIO101)

2.2.9.1 CRSR-BLUE

Chaotrophic RNA stabilising reagent

Supplied in the FastRNA™ Kit, BLUE, details not provided.

2.2.9.2 PAR

Phenol acid reagent

Supplied in the FastRNA™ Kit, BLUE, details not provided.

2.2.9.3 CIA

Chloroform isoamyl alcohol (24:1), supplied in the FastRNA™ Kit, BLUE.

2.2.9.4 DIPS

DEPC-treated/isopropanol precipitation solution

Supplied in the FastRNA™ Kit, BLUE, details not provided.

2.2.9.5 SEWS

Salt/ethanol wash solution, RNase-free

Supplied in the FastRNA™ Kit, BLUE, details not provided.

2.2.9.6 SAFE

DEPC-treated water for elution, supplied in the FastRNA™ Kit, BLUE.

2.2.10 Southern blot buffers and solutions

2.2.10.1 Southern depurination solution

HCl	250 mM

2.2.10.2 Southern denaturing buffer

NaOH	0.5 M
NaCl	1.5 M

2.2.10.3 Southern neutralising buffer

Tris-HCl (pH 7.5)	0.5 M
NaCl	3 M

2.2.10.4 SSC (20x)

NaCl	3 M
Tri-sodium citrate·$2H_2O$	300mM

The pH was adjusted 7.0 with 1 M NaOH.

2.2.10.5 Pre-hybridisation solution

SSC	5x
Blocking reagent (Roche)	1 % (w/v)
N-lauroylsarcosine, Na salt,	0.1 % (w/v)
SDS	0.02 % (w/v)

2.2.10.6 Hybridisation solution

The probe was diluted in prehybridisation buffer to 5-25 ng ml^{-1}

2.2.10.7 2x wash solution

SSC	2x
SDS	0.1 % (w/v)

2.2.10.8 0.5x wash solution

SSC	0.5x
SDS	0.1 % (w/v)

2.2.10.9 Maleic acid buffer

Maleic acid	0.1 M
NaCl	0.15 M

The pH was adjusted to 7.5 with solid NaOH.

2.2.10.10 Washing buffer

Maleic acid	0.1 M
Tween® 20	3 % (v/v)

2.2.10.11 Blocking solution

Maleic acid buffer containing 1 % (w/v) blocking reagent (Roche)

Blocking reagent was dissolved in a microwave and store at -20 °C.

2.2.10.12 Antibody solution

Blocking solution, containing 75 Um ml^{-1} Anti-DIG-AP conjugate (Roche)

2.2.10.13 Detection buffer

Tris-HCl (pH 9.5)	100 mM
NaCl	100 mM
$MgCl_2 \cdot 6H_2O$	50 mM

The pH was adjusted to 7.5

2.2.10.14 Colour-substrate solution

Detection buffer containing 2 % (v/v) NBT/BCIP (Roche).

2.2.11 TAE (50x)

Trisma base	242 g
Glacial acetic acid	57.1 ml
Na_2EDTA (0.5 M pH 8.0)	100 ml
dH_2O	to 1 l

Before use the buffer was diluted 1:50 to produce TAE.

2.2.12 TE (10x)

Trisma base	100 mM
Na_2EDTA	10 mM

The pH was adjusted to 7.5 using 1 M HCl. This solution was diluted 1:10 with sterile dH_2O before use, to give TE.

2.3 Enzymes and chemicals

All chemicals and enzymes were of analytical grade and purchased from Sigma, Fisons or Merck (BDH), unless otherwise stated. All restriction enzymes, RNase A/H, DNase, reverse transcriptase, T4 ligase, polymerases, TaqMan components, and buffers, for the modification of DNA, were purchased from, ABI Perkin-Elmer, Life Technologies (formally Gibco BRL), Northumbria Biologicals Limited (NBL), Promega, or Roche (formally Boehringer Mannheim).

α-Amylase (Sigma) was dissolved in dH_2O to 25 mg ml^{-1} and stored at -20 °C. Cellosyl (gift from R. Marquardt, Hoechst AG, Frankfurt, Germany) was dissolved in 25 mM sodium phosphate buffer (pH 5.6), to 2 mg ml^{-1}, and stored at -20 °C. Lysostaphin (Sigma) was dissolved in 20 mM sodium acetate to 5 mg ml^{-1} and stored at -20 °C. MUG (4-methylumbelliferyl-β-D-galactopyranoside) (Sigma) was dissolved in DMSO to 4 mg ml^{-1} and stored at -20 °C. 4-MU (4-methylumbelliferone) (Sigma) was dissolved in DMSO to 1 mM and stored at -20 °C. Pronase E (Sigma) was dissolved in TES to 10 mg ml^{-1}, pre-digested for 30 min at 37 °C, and stored at -20 °C. RNase A (DNase-free) (Sigma), dissolved in dH_2O to 10 mg ml^{-1}, and stored at -20 °C. X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (Sigma) was dissolved in DMF to 20 mg ml^{-1} and stored at -20 °C. Phenol was purchased pre-equilibrated from NBL. TRISOL™ LS (ready to use monophasic solution of phenol and guanidine isothiocyanale) was purchased from Life Technologies.

2.4 Bacterial strains, plasmids and phages

2.4.1 Bacterial maintenance, culture and storage conditions

Bacterial strains used in this study are listed in Table 2.2-2.3.

2.4.1.1 *Staphylococcus aureus* strains

S. aureus NCTC 8325, the propagating strain for typing phage 47, has been the organism of choice for studies of the *S. aureus* genome. This strain is known to carry the prophages ϕ11, ϕ12 and ϕ13; an ultraviolet (u.v.) cured/prophage-minus derivative, 8325-4, has been used for most genetic studies in *S. aureus*. 8325-4 was the background of choice for all work described in this thesis unless otherwise stated.

Staphylococcus aureus strains (Table 2.2) were grown on BHI, serum, and CDM plates, containing antibiotics where appropriate to maintain selection of resistance markers. Strains were grown at 37 °C under aerobic, microaerobic or anaerobic conditions (Chapter 2.9). Plate cultures were stored for up to one month at 4 °C, before re-streaking onto fresh agar plates. For long-term storage, strains were grown overnight in BHI medium. The cells were harvested by centrifugation (5,000 rpm (Centaur 2), for 10 min) and then resuspended in sterile BHI containing 20 % (w/v) glycerol and stored at -20 °C and -70 °C. Liquid cultures were normally prepared by inoculation of culture medium with a single isolated colony. These cultures were grown overnight in sterile plastic universals containing 1:6 their volume of medium, and were aerated on a rotary shaker at 250 rpm. These cultures were then used to inoculate conical flasks containing one-fourth their volume (for example 25 ml; 100 ml flask) of medium, and were aerated on a rotary shaker at 250 rpm. All liquid cultures were grown at 37 °C and under aerobic conditions unless specified. These conditions are referred to as standard conditions.

2.4.1.2 *Escherichia coli* strains

Escherichia coli strains were cultured using LB or LA medium (Chapter 2.1.1.4), containing antibiotics where necessary to ensure the selection of plasmid-bearing strains. Plate cultures were stored at 4 °C for up to one month before re-streaking onto fresh agar plates. Liquid cultures were prepared as for *S. aureus* (Chapter 2.4.1.1) strains except strains were grown using LB or LA medium. All *E. coli* cultures were incubated at 37 °C. For long-term storage, *E. coli* strains and their plasmid-bearing derivatives were treated as for *S. aureus* strains (Chapter 2.4.1.1) except LB media was used rather than BHI, and stored at -20 °C and/or -70 °C. *E. coli* strains used in this study are shown in Table 2.3.

Strain	Relevant genotype/markers	Source/reference
8325-4	8325 cured of prophages ϕ11, ϕ12 and ϕ13	Novick, 1963
WCUH29		SmithKline Beecham
RN4220	Restriction minus, modification plus	Kreiswirth *et al.*, 1983
PC6911	*agr*::*tetM* (Tet^r)	Chan and Foster, 1998
PC1839	*sarA*::*km* (Kan^r)	Chan and Foster, 1998
PC18391	*sarA*::*km agr*::*tetM* (Kan^r, Tet^r)	Chan and Foster, 1998
PC400	*sigB*::*tetM* (Tet^r)	Chan *et al.*, 1998a
ST1	*katA*::Tn*917*	Horsburgh, unpublished
SPW1	*sodA*::Tn*917*	Watson *et al.*, 1998b
SPW2	*lysA*::Tn*917*	Watson *et al.*, 1998b
SPW3	*ctaA*::Tn*917*	Watson *et al.*, 1998b
SH109	*agr*$^+$, *hld*$^+$ *hld*::*lacZ* Ery^r	Fairhead, 1998
PC161	*sarA*$^+$ *sarA*::*lacZ* Ery^r	Chan and Foster, 1998
MC100	*sigB*$^+$ *sigB*::*lacZ* Ery^r	Chan *et al.*, 1998a
MJH005	*fur*$^+$ *fur*::*lacZ* Ery^r	Horsburgh, unpublished
PC203	*spa*$^+$ *spa*::*lacZ* Ery^r	Chan and Foster, 1998
PC322	*hla*$^+$ *hla*::*lacZ* Ery^r	Chan and Foster, 1998
MJH006	*katA*$^+$ *katA*::*lacZ* Ery^r	Horsburgh, unpublished
MDW1941	8325-4/pLTV1 Tet^r, Cm^r, Ery^r, pLTV1	This study
*seg*1	*seg*1::Tn*917* Ery^r	This study
*seg*3	*seg*3::Tn*917* Ery^r	This study
*seg*4	*seg*4::Tn*917* Ery^r	This study
*seg*5	*seg*5::Tn*917* Ery^r	This study
*seg*7	*seg*7::Tn*917* Ery^r	This study
*seg*8	*seg*8::Tn*917* Ery^r	This study
*seg*10	*seg*10::Tn*917* Ery^r	This study
*seg*13	*seg*13::Tn*917* Ery^r	This study
*seg*14	*seg*14::Tn*917* Ery^r	This study
*seg*21	*seg*21::Tn*917* Ery^r	This study
*seg*22	*seg*22::Tn*917* Ery^r	This study
*seg*24	*seg*24::Tn*917* Ery^r	This study

Strain Cont.	Relevant genotype/markers	Source/reference
*seg*25	*seg*25::Tn*917* Ery^r	This study
*seg*26	*seg*26::Tn*917* Ery^r	This study
*seg*29	*seg*29::Tn*917* Ery^r	This study
*seg*30	*seg*30::Tn*917* Ery^r	This study
*seg*33	*seg*33::Tn*917* Ery^r	This study
*seg*35	*seg*35::Tn*917* Ery^r	This study
*seg*37	*seg*37::Tn*917* Ery^r	This study
*seg*39	*seg*39::Tn*917* Ery^r	This study
*seg*40	*seg*40::Tn*917* Ery^r	This study
*seg*43	*seg*43::Tn*917* Ery^r	This study
*seg*45	*seg*45::Tn*917* Ery^r	This study
*seg*48	*seg*48::Tn*917* Ery^r	This study
*seg*50	*seg*50::Tn*917* Ery^r	This study
MDW41	*lysC*$^+$ *lysC*::*lacZ* Ery^r	This study
MDW42	*asd*$^+$ *asd*::*lacZ* Ery^r	This study
MDW43	*dapA*$^+$ *dapA*::*lacZ* Ery^r	This study
SWC1	WCUH29 *seg*1::Tn*917*	This study
SWC24	WCUH29 *seg*24::Tn*917*	This study
SWC10	WCUH29 *seg*10::Tn*917*	This study

Table 2.2

List of *S. aureus* strains used during this study

Tet^r, tetracycline resistance; Kan^r, kanamycin resistance; Ery^r erythromycin resistance; Cm^r, Chloramphenicol resistance;

Strain of *E. coli*	Relevant genotype/markers	Source/reference
DH5α	*supE44Δ lacU169* (F80 *lacZ* ΔM15) *hsdR17 rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1	Hanahan, 1983
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) F80 *lacZ* ΔM15 Δ*lac*X74 *rec*A1 *deo*R *ara*D139 Δ(*ara-leu*)7697 *gal*K *rps*L (Str^r) *end*A1 *nup*G	Invitrogen

Table 2.3

List of *E. coli* strains used during this study

2.4.2 Plasmids

Plasmid DNA was purified using QIAGEN plasmid kits in accordance with the manufacturer's instructions (Chapter 2.17.1.2-2.17.1.3). Purified plasmid DNA was stored in TE buffer at 4 °C (short-term storage), or -20 °C (long-term storage).

2.4.3 Phages

Bacteriophage ϕ11 (Table 2.4) was used in this study for phage transduction (Chapter 2.14). This phage is a *S. aureus* specific, temperate transducing phage of serological group B, requiring Ca^{2+} ions for maintenance. ϕ11 has an approximate genome size of 45 kb (Novick, 1991).

2.5 Centrifugation

A number of different centrifuges were used for the pelleting of cells and precipitated material. These were: i) Eppendorf microfuge S415c; maximum volume - 2 ml, maximum speed 14,000 rpm (11,000 **g**); ii) Centaur 2 centrifuge (Sanyo); maximum volume - 50 ml, maximum speed 5,000 rpm; iii) Avanti™ J25I (Beckman), maximum volumes and speeds dependent on the rotor used: JA-20; maximum volume - 50 ml, maximum speed 20,000 rpm (48,384 **g**); JA-14; maximum volume - 250 ml, maximum speed 14,000 rpm (30,074 **g**); JA-10.5; maximum volume - 500 ml, maximum speed 10,000 rpm (18,480 **g**).

All centrifugations were carried out at room temperature unless stated otherwise.

2.6 Starvation-survival

2.6.1 Starvation-survival assays (Watson, 1998b)

A 50 ml culture of CDM was inoculated with *S. aureus* from an overnight plate and incubated with shaking (250 rpm) for 18 hours at 37 °C. At this stage the approximate cell density was 1-5 $x10^8$ cfu ml^{-1}. Starvation in synthetic cerebrospinal fluids (sCSF) (Chapters 2.1.1.11-2.1.1.12) or PBS (Chapter 2.2.7) was examined by harvesting the CDM cultures by centrifugation (4000 rpm; 10 min), and twice washing with

Plasmid/Phage	Relevant genotype/markers	Source or reference
***S. aureus* plasmids**		
pLTV1	Tet^r, Cm^r, Ery^r	Camilli *et al.*, 1990
pMDW41	Ery^r	This study
pMDW42	Ery^r	This study
pMDW43	Ery^r	This study
***S. aureus* phages**		
ϕ11		Mani *et al.*, 1993
***E. coli* plasmids**		
pSEG1	Amp^r	This study
pSEG3	Amp^r	This study
pSEG4	Amp^r	This study
pSEG5	Amp^r	This study
pSEG7	Amp^r	This study
pSEG8	Amp^r	This study
pSEG10	Amp^r	This study
pSEG13	Amp^r	This study
pSEG14	Amp^r	This study
pSEG21	Amp^r	This study
pSEG22	Amp^r	This study
pSEG24	Amp^r	This study
pSEG25	Amp^r	This study
pSEG26	Amp^r	This study
pSEG29	Amp^r	This study
pSEG30	Amp^r	This study
pSEG33	Amp^r	This study
pSEG35	Amp^r	This study
pSEG37	Amp^r	This study
pSEG39	Amp^r	This study
pSEG40	Amp^r	This study
pSEG43	Amp^r	This study
pSEG45	Amp^r	This study
pSEG48	Amp^r	This study
pSEG50	Amp^r	This study
pAZ106	Amp^r	Kemp *et al.*, 1991
pMDW41	Amp^r	This study
pMDW42	Amp^r	This study
pMDW43	Amp^r	This study
pCR®2.1-TOPO	Amp^r, Kan^r	Invitrogen
pMTOPO9	Amp^r, Kan^r	This study

Table 2.4

List of plasmids and phages used in this project

sCSF I, II or PBS by centrifugation and resuspension. Cells were subsequently resuspended in 50 ml sCSF I, II or PBS and the cell suspensions were incubated with or without shaking at 37 °C. For starvation-survival in amino acid limiting CDM (16), cells were grown as described above and incubated without shaking at 37 °C.

2.6.2 Determination of *S. aureus* viability

Viability (starvation-survival) was determined by serial dilution of samples in PBS and dotting 10 µl of each dilution onto a CDM agar plate. After overnight incubation at 37 °C the number of colony forming units (cfu) was determined.

2.7 Determination of bacterial cell density

2.7.1 Spectrophotometrically (A_{600})

To quantify the bacterial yield in a culture, spectrophotometric measurements at 600 nm (A_{600}), were carried out. These measurements were taken using a Jenway 6100 spectrophotometer.

2.7.2 Direct cell counts (cfu)

An alternative method for the quantification of cell numbers is the method of direct cell counts. The number of cells was determined by serial dilution of samples in PBS and spreading 100 µl of each dilution onto a BHI agar plate. After overnight incubation at 37 °C the number of colony forming units (cfu) was determined.

2.8 Serum model

S. aureus strains to be grown in the serum model were grown overnight on BHI plates containing the relevant antibiotics where necessary, at 37 °C. For analysis on solid serum media a single colony was inoculated onto serum agar (Chapter 2.1.1.7) and grown for two days at 37 °C under aerobic, microaerobic or anaerobic conditions (Chapter 2.9). Growth was determined after this period by colony diameter measurements. Values were taken as an average of at least 3 colonies. For analysis in

liquid serum, a single colony was resuspended in 5 ml serum in plastic universals. These cultures were also grown overnight at 37 °C under aerobic, microaerobic or anaerobic conditions. Growth was determined after this period by A_{600} (Chapter 2.7.1) and the appropriate dilution made to allow a starting A_{600} of 0.2 in 25 ml of pre-warmed serum in 100 ml flasks. These cultures were grown at 37 °C for 24 h under aerobic, microaerobic or anaerobic conditions (Chapter 2.9). Growth was monitored through out this period by either A_{600} or direct counts (Chapter 2.7.2). Alternatively for analysis in 5 ml cultures, the overnight culture was used to inoculate 5 ml of fresh serum, with a starting A_{600} of 0.2. These cultures were grown for 15 h.

2.9 Gaseous composition of environmental conditions used for growth

S. aureus cultures and plates were grown under aerobic, microaerobic or anaerobic conditions:

Aerobic:	O_2 - 20.95 %	159.2 mmHg
	CO_2 - 0.03 %	0.228 mmHg
	N_2 - 78.08 %	593.4 mmHg

Unless other wise stated all cultures were incubated under aerobic conditions.

Microaerobic:	O_2 - 10 %	76 mmHg
	CO_2 - 5 %	38 mmHg
	N_2 - 85 %	646 mmHg

A Don Whitley Scientific Limited Variable Atmosphere incubator was used for the maintenance of microaerobic conditions as described.

Anaerobic:	O_2 - 0 %	0 mmHg
	CO_2 - 10 %	76 mmHg
	H_2 - 5 %	38 mmHg
	N_2 - 85 %	646 mmHg

A Don Whitley Scientific Limited MK3 Anaerobic Work Station was used for the maintenance of anaerobic conditions as described.

2.10 Culture aeration

Culture aeration was varied, by adjusting the volume of media in the flask:

Standard	25 ml	100 ml flask
High aeration	20 ml	250 ml flask
Low aeration	100 ml	250 ml flask

Unless other wise stated all cultures were grown using standard aeration conditions.

2.11 β-galactosidase assays

2.11.1 Using X-Gal as a substrate

β-galactosidase activity was detected directly on solid media using the substrate X-Gal. X-Gal is hydrolysed by β-galactosidase resulting in β-D-galactopyranoside and 5-bromo-4-chloro-3-indolyl. 5-bromo-4-chloro-3-indolyl has a blue colouration and is thus a visible indicator (blue) of β-galactosidase activity. 80 µl X-Gal solution (Chapter 2.3) was added to 5 ml of BHI; CDM; or serum agar cooled to 50 °C. This molten overlay was mixed briefly by inverting and poured on to warmed (~ 60 °C) agar plates (BHI; CDM; or serum respectively). These plates were stored at - 4 °C for up to a week.

2.11.2 Using MUG as a substrate

100 µl samples of cell culture were removed, pelleted (11,000 **g**; 3 min) and snap frozen (-195.8 °C). Samples were stored at -20 °C for later analysis. The frozen cells were thawed and resuspended in an equal volume of ABT (Chapter 2.2.1). 50 µl of cells were then diluted (appropriate dilutions generally 1:10-1:10000) with ABT in microtitre plates giving a total volume of 45 µl. These plates were warmed to 25 °C for approximately 5 min, prior to the addition of 5 µl MUG (4 mg ml^{-1} (Chapter 2.3)). These reactions were mixed gently with a pipette and incubated at 25 °C for exactly 60 min. MUG is hydrolysed to β-D-galactopyranoside and 4-methylumbelliferyl (4MU) by the action of β-galactosidase. 4MU is a fluorescent compound and so is a quantifiable indicator of β-galactosidase activity. The reaction between β-galactosidase and MUG was stopped with the addition of 0.2 M Na_2CO_3 (200 µl). A

fluorimeter (Victor^2$_w$, Wallac) was used to measure the fluorescence of each sample (355/460nm, 0.1 sec). The concentration of 4MU and thus promoter activity was calculated from a calibration curve (Chapter 2.11.3; Figure 2.1). For each sample the concentration of 4MU was related to β-galactosidase activity and expressed in MUG units of β-galactosidase activity. Importantly the background level of β-galactosidase activity measured from the 8325-4 samples (control samples) was deducted from the *lacZ* fusion strains β-galactosidase activity.

β-galactosidase activity = U / (D x t x A_{600})
(MUG units)

Where: U = pM of MU produced; A_{600} = A_{600} at a given time point; t = time of incubation at 25 °C (min); and D = dilution factor.

Thus, 1 MUG unit of β-galactosidase activity is defined as the activity that hydrolyses: 1 pM of MUG min^{-1} ml^{-1} A_{600}^{-1}

2.11.3 β-galactosidase assay calibration curve

A calibration curve was prepared each time β-galactosidase assays were performed. These were created with the fluorescent product (4-methyl umbelliferone) with final concentrations of: 1250; 625; 312.5, 156.25; 78.13; 39.06; 19.531; 9.77 pM. An example of such a calibration curve is shown in Figure 2.1.

2.12 Transposon mutagenesis

2.12.1 Transposon mutagenesis using Tn*917*-pLTV1

A single colony of *S. aureus* strain MDW1941, from a overnight BHI plate (Tet and Ery 5 μg ml^{-1}) was inoculated into 10 ml BHI broth (Tet and Ery 5 μg ml^{-1}). This culture was grown overnight, at 37 °C with shaking (250 rpm). The A_{600} was determined the following morning and the appropriate dilution was made to allow a starting A_{600} of 0.02 in 50 ml pre-warmed BHI (Tet and Ery 5 μg ml^{-1}). This culture was grown aerobically at 37 °C with shaking (250 rpm) until an A_{600} of ~ 1.0 had been reached. A 3 ml sample was removed after this period, and pelleted (11,000g, 2 min).

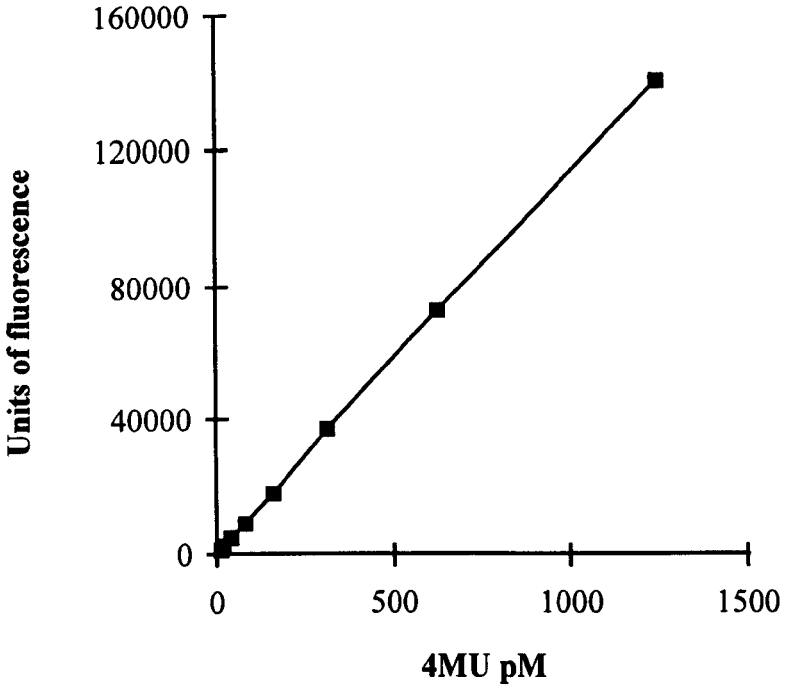


Figure 2.1

Calibration curve of concentration of 4MU against units of fluorescence

The pellet was subsequently resuspended in 100 ml BHI (Ery 5 μg ml^{-1}) pre-warmed to 41-41.8 °C, and grown at this temperature with shaking (250 rpm). Once the A_{600} had reached 0.3-0.4, 3 ml of culture was transferred to a further 100 ml BHI (Ery 5 μg ml^{-1}), again pre-warmed to 41-41.8 °C. This culture was once more grown at the increased temperature with shaking (250 rpm) and was allowed to grow to stationary phase (18 h). After this time, cells were pelleted (10,000 **g**; 10 min; 4 °C) and resuspended in 4 ml BHI containing 10 % (v/v) glycerol, giving approximately $5x10^9$ cfu ml^{-1}. Cells were then snap frozen (-195.8 °C) and stored at -20 °C.

2.12.2 Tn*917* transposon library evaluation and enumeration

The efficiency of transposon insertion was determined by serial dilutions of cells in PBS (Chapter 2.2.7) and plating a range of cell dilutions (10^4-10^8) onto BHI agar plates containing Ery, or Ery and Tet (5 μg ml^{-1}). The transposon insertion frequency was calculated by comparing the number of colonies that grow on plates containing Ery (bacteria containing the transposon (Tn*917*) inserted into the chromosome along with bacteria containing the intact plasmid (Tn*917*-LTV1)) against the number of colonies that grow on plates containing Ery and Tet (bacteria containing the intact plasmid (Tn*917*-LTV1)).

2.13 Screening for genes important for growth in sera

2.13.1 Isolation of putative sera essential mutants of *S. aureus* from a Tn*917* library

An aliquot of a Tn*917* library was suitably diluted in PBS (Chapter 2.2.7) to allow ~100 colonies per BHI plate (Ery 5 μg ml^{-1}) and the plates incubated overnight at 37 °C. The following day the colonies were replica plated onto BHI (control plate) and pig sera agar plates (Chapter 2.1.1.7) using sterile velvets. BHI plates were incubated aerobically while sera plates were incubated microaerobically, both at 37 °C for 24 h. Putative sera-essential mutants were isolated by comparing the serum plate with the control plate for absent colonies. These colonies were selected from the control plate and re-tested under the above conditions. Strains absent on serum were selected from the control plate and glycerol stocks made (Chapter 2.4.1.1).

2.13.2 Identification of putative sera expressed genes of *S. aureus* from a Tn*917* library

An aliquot of a Tn*917* library was suitably diluted in PBS (Chapter 2.2.7), to allow ~100 colonies per BHI plate (Ery 5 μg ml^{-1}) and the plates incubated overnight at 37 °C. The following day the colonies were replica plated onto BHI (control plate) and pig sera agar plates, containing X-Gal, (Chapter 2.3) using sterile velvets and were grown for 24 h at 37 °C. BHI plates were incubated aerobically while sera plates were incubated microaerobically. Putative sera expressed mutants were isolated by comparing the serum plate with the control plate for colonies that were blue on serum but white on BHI. These colonies were selected from the control plate and re-screened under the same conditions. Colonies blue on serum and white on BHI were selected and glycerol stocks made (Chapter 2.4.1.1).

2.14 Phage techniques

2.14.1 Preparation of phage lysates

S.aureus strains were inoculated into BHI containing the relevant antibiotic, and grown at 37 °C with shaking (250 rpm) until in log phase (A_{600} ~ 0.5). 2-5 ml of cells were pelleted (11,000 **g**; 3 min) and resuspended in 5 ml BHI in a plastic universal. 5 ml phage buffer (Chapter 2.2.6) and 100 μl stock lysate (ϕ11 propagated on 8325-4 cell : phage ratio of approx. 20:1) were added to the cells and the mixture was allowed to incubate for 10 min at room temperature. The tube was then placed at 30 °C shaking slowly (50 rpm) for 2-4 hours, until clear. If cells failed to lyse in this time, they were left overnight (static) at 25 °C. Lysates were filter sterilised (0.2 μm pore size) and stored at 4 °C. The resulting lysates were between 10^9 and 10^{10} pfu ml^{-1}.

2.14.2 Determination of phage titres

S. aureus 8325-4 was inoculated into BHI and incubated overnight at 37 °C with shaking (250 rpm). This was used to inoculate (A_{600} 0.05) 25 ml of fresh BHI and grown until log phase (~ 0.5) at 37 °C with shaking (250 rpm). Phage lysates were diluted in phage buffer (Chapter 2.2.6) to 10^{-7}. 100 μl of diluted phage was mixed with 50 μl of 1M $CaCl_2$, and 400 μl of 8325-4 culture, and incubated for 10 min at room temperature. 5 ml of phage top agar (Chapter 2.1.1.6) cooled to 50 °C, was added to

the phage mixture and used to overlay a pre-warmed (~ 60 °C) phage bottom agar plate (Chapter 2.1.1.6). After setting (30 min), plates were incubated overnight at 37 °C, after which time the number of pfu ml^{-1} was determined.

2.14.3 Phage transduction

S. aureus 8325-4 was inoculated into LK broth (Chapter 2.1.1.5) and incubated overnight at 37 °C with shaking (250 rpm). This was used to inoculate (A_{600} 0.05) 25 ml of fresh LK, which was grown until log phase (A_{600} ~ 0.5) at 37 °C with shaking (250 rpm). 500 µl of cell culture was transferred to a plastic universal containing 500 µl of phage lysate and 1 ml LK broth. After gentle mixing the lysate/cell mixture was incubated at 37 °C without shaking, for 25 min, followed by 15 min with shaking (250 rpm). 1 ml ice-cold 0.02 M sodium citrate was added at this time and cells were pelleted (5000 rpm (Centaur 2); 10 min) and resuspended in 1 ml 0.02 M sodium citrate and left on ice for 2 h. The lysate/cell mixture was spread on LK bottom agar plates (Chapter 2.1.1.5) containing 0.05 % (v/v) sodium citrate and inducing levels of Ery, 0.15µg ml^{-1}. These plates were incubated at 37 °C for 90 min, after which time they were overlaid with 5 ml LK top agar (Chapter 2.1.1.5) containing 25 mg ml^{-1} Ery. After setting (30 min), plates were incubated for 18-72 h at 37 °C. Any resulting colonies were putative phage transductants.

2.15 β-Haemolysin activity

Blood agar plates were prepared as described in Chapter 2.1.1.8, containing 5 % (v/v) sheep blood. *S. aureus* strains were inoculated by picking colonies onto the test plate. After overnight incubation at 37 °C β-haemolysin activity was visualised by a zone of clearing (lysis) around the bacterial colony.

2.16 Transformation techniques

2.16.1 Transformation of *E. coli*

2.16.1.1 Preparation of *E. coli* competent cells (Hanahan, 1983)

A generous scoop of *E. coli* DH5α cells from an overnight plate was inoculated into 1 ml of SOB (Chapter 2.1.1.9) and the cells dispersed by vortexing. This was inoculated into a 1 l flask containing 100 ml of pre-warmed SOB and incubated at 37 °C, with shaking (250 rpm), until A_{600} 0.5-0.6. The culture was transferred to 50 ml centrifuge tubes and chilled on ice for 15 min. The cells were then pelleted by centrifugation (11,000 **g**; 15 min; 4 °C) and the pellet drained and resuspended in 33 ml of frozen storage buffer (FSB) (Chapter 2.2.4). The cells were incubated on ice for 15 min and then pelleted and drained as before. This pellet was resuspended in 8 ml FSB. 280 µl of dimethyl sulphoxide (DMSO) was added, mixed by swirling and the tube incubated on ice for 5 min. An equal volume of DMSO was added as before and the tube incubated on ice for a further 15 min. Aliquots of 250 µl were pipetted into microfuge tubes, used immediately or snapped frozen at -195.8 °C and stored at -70 °C.

2.16.1.2 Transformation

To transform cells of *E. coli* DH5α, 200 µl of a competent cell suspension was transferred into a pre-chilled microfuge tube, mixed with the required DNA (~ 50 ng) in a 10-20 µl volume and incubated on ice for 40 min. The tube was then transferred to 42 °C for 90 sec and immediately returned to the ice for a further 2 min. 800 µl SOC (Chapter 2.1.1.10) was added to the cells, which were then incubated for 45 min at 37 °C to allow the expression of plasmid encoded antibiotic resistance markers. The transformed cells were plated onto LB agar containing the appropriate antibiotics and then incubated overnight at 37 °C or until colonies were visible.

2.16.2 Transformation of *S. aureus*

2.16.2.1 Preparation of *S. aureus* electrocompetent cells

S. aureus RN4220 were inoculated into B2 media (Chapter 2.1.1.1) and incubated overnight at 37 °C with shaking (250 rpm). This was used to inoculate (A_{600} 0.02)

500 ml of fresh B2 and grown until late log phase (A_{600} 0.5-1.0). Cells were pelleted (5000 rpm; 10 min; 4°C) and washed 3x with 150 ml sterile water at 4 °C. Cells were then concentrated in a series of pelleting (5000 rpm (Centaur 2); 10 min; 4 °C), and resuspension steps: i) 1/10 volume of 10 % (v/v) ice-cold glycerol (50 ml); ii) 1/20 volume of 10 % (v/v) ice-cold glycerol (25 ml); and iii) 1/40 volume of 10 % (v/v) ice-cold glycerol (12.5 ml). All the supernatant was removed after each spin. The pellet was finally resuspended in 625 µl 10 % (v/v) ice-cold glycerol. Cells were snap frozen (-195.8 °C) and stored at -70 °C. The cfu ml^{-1} of the competent cells was determined prior to use.

2.16.2.2 Transformation of *S. aureus* (Electroporation)

5-20 µg of plasmid DNA was precipitated by isopropanol precipitation (Chapter 2.17.1.7), and washed twice with 70 % (v/v) ice-cold (-20 °C) ethanol, ensuring the complete removal of salt. The DNA was then resuspended in 100 µl of sterile dH_2O and transferred to a 0.1 cm gap cuvette. 2 x 10^{10}-2 x 10^{11} competent cells, which had been allowed to defrost for 5 min at room temperature, were then added to the cuvette. The plasmid/cell mix was electroporated at room temperature at 100 ohms, 25 µF and 1.25 kV using a BioRad Gene Pulser. Cells were recovered by adding 1 ml of pre-warmed B2 containing an inducing concentration of Ery (0.15 µg ml^{-1}) to the cuvette. Cells were then transferred to sterile 2 ml microfuge tubes and incubated at 37 °C with shaking (250 rpm) for 2 h. Recovered cells were plated directly onto BHI Ery/Lin plates (5 µg ml^{-1}/25 µg ml^{-1}) and incubated at 37 °C for 24-36 h.

2.17 Nucleic acid purification techniques

2.17.1 DNA purification techniques

2.17.1.1 Chromosomal DNA preparation

Chromosomal DNA was isolated and purified using a QIAGEN 100/G kit according to the manufacturer's instructions, as follows. A single colony of the strain of interest was used to inoculate 5 ml of BHI and grown overnight at 37 °C. The cells were pelleted (5,000 **g**; 10 min) and resuspended in 5 ml of Buffer B1 containing 50 µg ml^{-1} RNase A (Chapter 2.3). After rigorous vortexing, 25 µl lysostaphin (5 mg ml^{-1}) (Chapter 2.3), and 200 µl Pronase (20 mg ml^{-1}) (Chapter 2.3) were added and

incubated for at least 30 min at 37 °C. 1.2 ml of Buffer B2 was added and vortexed briefly. The tube was then incubated at 50 °C for at least 30 min.

The clear lysate was applied to a QIAGEN 100/G tip, which had been equilibrated with 5 ml of Buffer QBT. Once the lysate had passed through, the tip was washed 2x with 10 ml of Buffer QC. The genomic DNA was then eluted into a sterile 50 ml tube by the addition of 5 ml buffer QF.

Genomic DNA was precipitated by adding 3.5 ml of isopropanol and inverting the tube 10-20 times. The genomic DNA was centrifuged (5000 rpm (Centaur 2); 15 min) to pellet the DNA. The pellet was washed with 2 ml cold (-20 °C) 70 % (v/v) ethanol and air dried for 10 min. The DNA was dissolved in 1 ml TE overnight at 4 °C, and stored at this temperature.

All QIAGEN solution details are given in Chapter 2.2.8.

2.17.1.2 Small scale plasmid preparation from *E. coli*

Plasmid DNA was isolated and purified using a QIAprep Spin Miniprep kit DNA purification system (QIAGEN) according to the manufacturer's instructions, as follows. 1-5 ml LB containing appropriate antibiotic was inoculated with a single colony of plasmid-bearing cells of *E. coli* DH5α and incubated overnight at 37 °C. The cells were pelleted (10,000 **g**; 5 min), resuspended in 250 µl of Buffer P1 and transferred to a microfuge tube. 250 µl of Buffer P2 was added and mixed by inverting the tube 4-6 times. 350 µl of Buffer P3 was then added and again mixed by inversion. The lysate was centrifuged (11,000 **g**; 10 min) and the supernatant was applied to a QIAprep column (QIAGEN). The plasmid DNA was bound to the column by centrifuging for 1 min (11,000 **g**) and the flow-through discarded. The column was then washed, by the addition of 750 µl of Buffer PE and centrifuging (11,000 **g**; 1 min) and the flow-through was once more discarded before centrifuging for an additional 1 min (11,000 **g**). The plasmid DNA was then eluted into a clean microfuge tube with 50 µl of Buffer EB after standing for 1 min by centrifugation (11,000 **g**; 1 min). Plasmid DNA was stored at -20 °C.

All QIAGEN solution details are given in Chapter 2.2.8.

2.17.1.3 Large scale plasmid preparation from *E. coli*

Plasmid-containing cells of *E. coli* DH5α from 500 ml overnight (37 °C) LB cultures (containing appropriate antibiotics) were harvested by centrifugation (15,000 **g**; 3 min; 4 °C). The plasmid was then purified from these cells using a Plasmid Maxi kit DNA purification system (QIAGEN), according to the manufacturer's guidelines, as follows. The pellet was resuspended completely in 10 ml Buffer P1. 10 ml of Buffer P2 was then added and the sample, mixed gently by inversion (4-6 times) before incubating at room temperature for 5 min. 10 ml of chilled Buffer P3 was then added and the sample was immediately mixed (gently) and incubated on ice for 20 min. The resulting white precipitate was then pelleted by centrifugation (20,000 **g**; 30 min; 4 °C) and the supernatant was promptly removed to a clean tube and filtered using 4 layers of cheesecloth. The supernatant was then applied to a QIAGEN tip 500 (equilibrated with 10 ml Buffer QBT) and allowed to enter the resin by gravity flow. The tip was then washed with 2x 30 ml Buffer QC and the DNA eluted with 15 ml Buffer QF. Plasmid DNA was precipitated with 10.5 ml isopropanol and pelleted by centrifugation (27,000 **g**; 30 min; 4 °C). The DNA pellet was washed with 5 ml 70 % (v/v) ethanol, re-centrifuged (27,000 **g**; 20 min; 4 °C) and air dried for 10 min. Purified plasmid DNA was then resuspended in an appropriate volume of TE and stored at -20 °C.

All QIAGEN solution details are given in Chapter 2.2.8.

2.17.1.4 Phenol/chloroform purification of DNA

To DNA in a volume of 200 µl TE, 100 µl phenol and 100µl chloroform/isoamyl alcohol (24:1) was added. After mixing, the resultant emulsion was centrifuged (11,000 **g**; 10 min). The mixture separates into an upper layer of aqueous DNA, and a lower layer of phenol/chloroform. The upper layer was removed and transferred to a fresh microfuge tube containing 200 µl chloroform/isoamyl alcohol (24:1) and mixed gently. After centrifugation (11,000 **g**; 2 min), the top layer was again removed to a fresh microfuge tube. The DNA was recovered by ethanol precipitation (Chapter 2.17.1.5).

2.17.1.5 Ethanol precipitation of DNA

DNA was precipitated by the addition of 0.1 volume of 3 M sodium acetate (adjusted to pH 5.2 with acetic acid) and 2.5 volumes of 100 % (v/v) ethanol followed by incubation at -20 °C for 10 min. The precipitated DNA was pelleted by centrifugation (11,000 **g**; 10 min) and the pellet washed with 70 % (v/v) ethanol. The pellet was air dried and dissolved in an appropriate volume of TE. DNA was stored at 4 °C (short-term) and -20 °C (long-term).

2.17.1.7 Ethanol/sodium acetate precipitation of DNA for sequencing

2 µl 3 M sodium acetate (adjusted to pH 5.2 with acetic acid) and 50 µl of 95 % (v/v) ethanol were added to each completed PCR sequencing reaction (Chapter 2.21.3.1). The tubes were vortexed and then placed on ice for 10 min. The DNA was pelleted by centrifugation (11,000 **g**; 30 min). The pellet was rinsed in 250 µl of 70 % (v/v) ethanol and then centrifuged for a further 10 min. The DNA pellet was air dried for 10 min. DNA was then sequenced (Chapter 2.21.3.1).

2.17.1.7 Isopropanol precipitation of DNA

DNA was precipitated by the addition of 0.1 volume of 3 M sodium acetate (adjusted to pH 5.2 with acetic acid) and an equal volume of isopropanol, followed by thorough mixing and incubation at room temperature for 2 min. The precipitated DNA was pelleted by centrifugation (11,000 **g**; 10 min) and the pellet washed with cold (-20 °C) 70 % (v/v) ethanol. The precipitated DNA was pelleted by centrifugation (11,000 **g**; 5 min). The 70 % (v/v) ethanol wash step was then repeated before the pellet was air dried and dissolved in an appropriate volume of TE. DNA was stored at 4 °C (short-term) and -20 °C (long-term).

2.17.1.8 Gel extraction purification of DNA

Double stranded DNA fragments were purified from agarose gels (Chapter 2.19.1.3) with the QIAquick gel extraction kit (QIAGEN) according to the manufacturer's instructions, as follows. Agarose blocks containing the DNA fragment were dissolved in 3 volumes of Buffer QG at 50 °C for 10 min. One volume of isopropanol was added

to the mix, which was then applied to a QIAquick spin column (QIAGEN). The DNA was bound to the column by centrifuging the sample for 1 min (11,000 **g**) and the flow through was discarded. The column was then washed with 750 µl Buffer PE, left to stand for 2-5 min and once more centrifuged for 1 min (11,000 **g**). The flow through was discarded and the column centrifuged again for 1 min (11,000 **g**). The DNA was then eluted into a clean microfuge tube with 50 µl of Buffer EB by centrifugation (11,000 **g**; 1 min). Purified DNA was stored at 4 °C (short-term) and -20 °C (long-term).

All QIAGEN solution details are given in Chapter 2.2.8.

2.17.2 RNA purification techniques

2.17.2.1 RNA extraction

S. aureus WCUH29 was inoculated into either, *in vitro* media (BHI or TSB) or into a murine pyelonephritis model (*in vivo*). The *in vitro* cultures were grown under standard conditions (Chapter 2.4.1.1) and allowed to reach mid exponential phase (A_{600} 0.5-1.0), before 10 ml of cells were harvested by centrifugation (11,000 **g**; 3 min; 4 °C) and resuspended in 100 µl sterile dH_2O. The infected kidneys were removed after 7 days and contained ~ 10^7-10^8 cfu ml^{-1}. Total RNA was isolated from these *S. aureus* cells using the FastRNA™ kit BLUE (BIO101) according to the manufacturer's instructions, as follows. On ice 500 µl CRSR-BLUE, 500 µl PAR and 100 µl CIA were added to a FastRNA™ BLUE tube (BIO101) along with the *in vitro* and *in vivo* grown cells (10 ml cells concentrated/1 infected kidney). The tube was then placed in the FastPrep™ instrument (BIO101) and processed for 2x 45 sec at speed rating 6, with 1 min cooling on ice between each processing. The tube was then removed from the machine and cooled on ice for 5 min before centrifuging (11,000 **g**; 15 min; 4 °C). The top phase was removed to a clean microfuge tube without disturbing the interphase. 500 µl of CIA was then added and the sample was vortexed for 10 sec and the layers separated by centrifugation (11,000 **g**; 2 min). The top layer was once more removed to a fresh tube, to which 500 µl of DIPS was added. The sample was mixed and incubated at room temperature for 10 min to precipitate RNA. The RNA was then pelleted by centrifugation (11,000 **g**; 5 min) and washed with 500 µl SEWS, without resuspension. The fluid was removed with a small-bore pipette tip and the pellet air dried for 10 min. The pellet was then dissolved in 50 µl SAFE, at 55 °C for 10 min and then stored at -70 °C. The yield and integrity of the RNA was

then assessed both spectrophotometrically (Chapter 2.18.2) and through a 1.2 % (w/v) agarose gel (Chapter 2.19.1.3).

All BIO101 solution details are given in Chapter 2.2.9.

2.18 Quantification of nucleic acids (Sambrook *et al.*, 1989)

2.18.1 Quantification of DNA

To quantify the DNA concentration in a solution, spectrophotometric measurements at 260 nm were carried out. An absorbance of 1 corresponds to approximately 50 μg ml^{-1} for double stranded DNA, and approximately 20 μg ml^{-1} for single stranded oligonucleotides. These measurements were taken using a Jenway 6105 spectrophotometer. The concentration of DNA could also be estimated independently by agarose gel electrophoresis (Chapter 2.19.1.3), comparing the intensity of the ethidium bromide stained bands, to bands containing known amounts of DNA (Table 2.5).

2.18.2 Quantification of RNA

To quantify the RNA concentration in a solution, spectrophotometric measurements at 260 nm were carried out. An absorbance of 1 corresponds to approximately 40 μg ml^{-1} RNA. These measurements were taken using a Jenway 6105 spectrophotometer.

2.19 *In vitro* nucleic acid manipulation techniques

2.19.1 *In vitro* DNA manipulation techniques

2.19.1.1 DNA restriction

Restriction digests were performed with up to 17.5 μl of DNA (genomic, plasmid or PCR product). This solution was combined with 2 μl of appropriate 10x restriction enzyme buffer and 0.5 μl of the appropriate restriction enzyme. The volume was made up to 20 μl with dH_2O and incubated at 37 °C for up to 15 h. Before the DNA was

manipulated it was cleaned up using the QIAquick PCR purification kit (QIAGEN) (Chapter 2.21.5) according to the manufacturer's instructions.

2.19.1.2 DNA ligation

2.19.1.2.1 Standard

The following were added on ice:

DNA in TE buffer	15 µl*
10x ligase buffer	2 µl
dH_2O	2 µl
T4 DNA ligase	1 µl (3 Weiss units)

The mixture was incubated overnight at 15 °C or two days at 4 °C. This ligation mixture could be used directly in the transformation of *E. coli*.

* 15 µl of digested DNA (vector to insert ratio 1:3, 0.2-1 µg of vector DNA)

2.19.1.2.2 Rescue of transposon insertions

The following were added on ice:

DNA in TE buffer	250 µl
10x ligase buffer	30 µl
dH_2O	19 µl
T4 DNA ligase	1 µl (3 Weiss units)

Ligations were performed in a total of 300 µl to minimise the possibility of multimers. A vector to insert ratio 1:3, 0.2-1 µg of vector DNA in TE buffer was used. The mixture was incubated overnight at 15 °C or for two days at 4 °C. Following ethanol precipitation (Chapter 2.17.1.5) and resuspension in 20 µl TE, this ligation mixture could be used directly in the transformation of *E. coli*.

2.19.1.3 Agarose gel electrophoresis

DNA fragments were separated by horizontal gel electrophoresis using various horizontal electrophoresis tanks (Life Technologies). Appropriate volumes of agarose gel (0.8-1.5 % (w/v) dissolved in TAE (Chapter 2.2.11) were submerged in suitable volumes of TAE electrophoresis buffer. The gel contained 0.2 μg ml^{-1} ethidium bromide. DNA samples were mixed with a fifth their volume of 6x DNA loading buffer (Chapter 2.2.3) and loaded into slots in the gel. The gel was run at 80-100 V for, 1-2 h and visualised by means of a u.v. transilluminator at 260 nm. DNA fragment sizes were estimated using 1 μg of λ DNA digested with *Hin*dIII or double digested with *Hin*dIII/*Eco*RI, 1 μg of a 100 bp ladder (Life Technologies) as size standards (Table 2.5).

2.19.1.3.1 Gel photography

A permanent record of agarose gels was obtained by photographing the ethidium bromide-stained gels illuminated with u.v. light at 260 nm. A Kodak 203 red-orange filter, and Polaroid 667 (ASA 3000) film were used.

2.19.2 *In vitro* manipulation of RNA

2.19.2.1 DNase treatment of RNA

The following were added on ice:

RNA	50 μg
DNase	1 μl
RNase inhibitor	2 μl
10x buffer	6 μl
dH_2O	up to 60 μl

These reactions were then incubated at 37 °C for 30 min.

RNA samples were then mixed with 500 μl TRIZOL™ LS (Chapter 2.3) and incubated at room temperature for 5 min. 100 μl of chloroform was then added and the tubes were shaken for 20 sec, before incubating at room temperature for 5 min. The

DNA size markers		
*Hind*III (λ)	*Hind*III/*Eco*RI (λ)	100 bp ladder
23.15	21.23	2072
9.24	5.15	1500
6.56	4.97	1400
4.38	4.27	1300
2.32	3.53	1200
2.02	2.02	1100
0.56	1.90	1000
0.13	1.58	900
	1.38	800
	0.95	700
	0.83	*600
	0.56	500
	0.13	400
		300
		200
		100

Table 2.5

The size of fragments used as size markers for agarose gel electrophoresis

*600 bp fragment is 2-3 x brighter than the other bands for easy reference.

layers were then separated by centrifugation (11,000 **g**; 15 min; 4 °C), after which the aqueous layer was removed to a clean tube. The RNA was then precipitated overnight at -20 °C with an equal volume of isopropanol, before pelleting the RNA by centrifugation (11,000 **g**; 10 min; 4 °C). The pellet was finally washed with DEPC treated (Chapter 2.2) 75 % (v/v) ethanol before air-drying the pellet for 10 min. The pellet was then dissolved in 50 µl DEPC treated dH_2O at 55 °C for 10 min. RNA samples were stored at -70 °C. The purity of the RNA prep was tested at this point using a pair of DNA designed PCR primers (sequence not supplied by SmithKline Beecham), in a PCR reaction as described in Chapter 2.21.2.1, except that RNA was used as the template rather than DNA, and the number of cycles was increased from 35 to 40. A DNA free RNA prep was shown by the absence of bands, for the RNA samples.

2.19.2.2 Production of cDNA from RNA

The following were added on ice:

RNA	6 µg
Random hexamers	6 µl (300 ng)
DEPC treated dH_2O	up to 24 µl

These reactions were incubated at 70 °C for 10 min and then cooled to 4 °C in a Perkin-Elmer GeneAmp PCR System 9600. The reaction mix was then split into two: A) + reverse transcriptase; B) - reverse transcriptase.

The following were added at room temperature:

	A	B
RNA mix from the above step	12 µl	12 µl
10x buffer	2 µl	2 µl
$MgCl_2$	2 µl	2 µl
dNTPs	1 µl	1 µl
DTT	2 µl	2 µl
RNase inhibitor	1 µl	1 µl

These reactions were then incubated at room temperature for 5 min, after which the following were added to each reaction.

Reverse transcriptase	1 µl	-
DEPC treated dH_2O	-	1 µl

Incubations of these reactions were carried out using a Perkin-Elmer GeneAmp PCR System 9600.

Room temperature	10 min
42°C	50 min
70°C	15 min

Once complete samples were cooled to 4 °C, before the addition of 1 µl RNase H to all samples. Tubes were then incubated at 37 °C for 20 min. cDNA samples were stored at -70 °C.

2.20 DNA Hybridisation techniques

2.20.1 Labelling of DNA probes with Digoxigenin

DNA fragments were labelled by a random priming method using a commercially available digoxigenin (DIG) DNA labelling and detection kit (Roche). The DNA template (approximately 3 µg in 5 µl) was heat denatured in a boiling water bath for 10 min and chilled on ice for another 10 min. To the denatured DNA, 2 µl of each of 10x hexanucleotide mixture and 10x dNTP labelling mixture were added. The volume of the reaction was made up to 20 µl with dH_2O and 1 µl of Klenow enzyme. The reaction was incubated overnight at 37 °C, and was terminated by the addition of 2 µl of 200 mM EDTA pH 8.0. The labelled DNA was precipitated by the addition of 0.1 volume of 4 M LiCl, and 2.5 volumes of 100 % (v/v) ice-cold ethanol. After thorough mixing the tube was incubated at -70 °C for 30 min, and the DNA pelleted by centrifugation (11,000 **g**; 15 min; 4 °C). The ethanol was decanted and the pellet washed in 70 % (v/v) ethanol. The pellet was air dried for 10 min, and resuspended in 50 µl of TE buffer. The labelled DNA probe was either used immediately, or stored at -20 °C. The amount of labelled DNA present was quantified by comparison to labelled control DNA of known concentration as follows. The control labelled DNA and the probe were diluted and 1 µl of various dilutions from 1:20 to 1:2,000,000 were spotted onto Hybond-N+ Extra nylon membrane. The DNA was fixed to the membrane by u.v. fixation (Chapter 2.20.3), and the DIG labelled DNA was then detected immunologically, using alkaline phosphatase linked anti-digoxigenin

antibody (Chapter 2.2.10.12). The spot intensities of the control and probe dilutions were compared to estimate the concentration of the probe.

Prior to use, the probe was denatured by placing the tube in a boiling water bath for 10 min followed by immediate chilling on ice for 10 min.

2.20.2 Southern blotting

Agarose gel electrophoresis was performed and photographed as described in Chapters 2.19.1.3 and 2.19.1.3.1. The gel was soaked in Southern depurination solution (Chapter 2.2.10.1) for 10 min to depurinate the DNA. This acts to nick large DNA molecules aiding their transfer onto the membrane during the blotting procedure. The gel was washed in dH_2O and soaked in Southern denaturing buffer (Chapter 2.2.10.2) for 45 min. The gel was neutralised by soaking in several volumes of Southern neutralisation buffer (Chapter 2.2.10.3) for 45 min. The DNA was then transferred overnight from the gel to an Hybond-N+ extra membrane, by capillary transfer overnight using 10x SSC as the transfer buffer (Figure 2.2).

2.20.3 Fixing the DNA to the membrane

For all hybridisation experiments Amersham Life Sciences Hybond-N+ Extra (positively charged nylon membrane) was used. DNA was fixed to the membrane with the use of a u.v. crosslinker (Amersham Life Sciences RPN 2500), 70 mJ / cm^2; 15 sec.

2.20.4 Prehybridisation and hybridisation

Membranes to be probed with the DIG-labelled DNA were prehybridised for 1 h at 68 °C in pre-hybridisation solution (20 ml per 100 cm^2 of membrane) (Chapter 2.2.10.5). The membrane was then hybridised with the labelled probe (Chapter 2.11.1) overnight at 68 °C (Chapter 2.2.10.6). The concentration of probe used was 5-25 ng ml^{-1} diluted in pre-hybridisation solution. Unbound probe was removed by washing the membrane twice in 2x wash (Chapter 2.2.10.7) for 5 min at room temperature. The membrane was then washed twice in 0.5x wash (Chapter 2.2.10.8) for 15 min at 68 °C.

2.20.5 Detection of DIG-labelled DNA

The hybridised and washed membranes were equilibrated with washing buffer (Chapter 2.2.10.10) for 1 min and then blocked for 30 min with gentle rocking in blocking solution (Chapter 2.2.10.11). The membrane was then transferred to antibody solution (Chapter 2.2.10.12) containing a 1:5000 dilution of the stock solution of anti-DIG-alkaline phosphatase antibody (Roche). After 30 min incubation with gentle rocking, the membrane was rinsed and then washed twice (15 min) with washing buffer with gentle rocking. The membrane was then equilibrated for 2 min with detection buffer (Chapter 2.2.10.13) and transferred to 10 ml colour substrate solution which contains: 66 µl nitro-blue tetrazolium (NBT); (50 mg ml^{-1} in 70 % (v/v) DMF); and 33 µl BCIP (50 mg ml^{-1}) (Chapter 2.2.10.14). This was then incubated in the dark to develop the membrane. After the colour development the membrane was washed in 1x TE to stop the reaction and stored dry in the dark.

2.21 Polymerase chain reaction

2.21.1 Primer design

The primers used for PCR amplification were short synthetic oligonucleotides (19-28 bp) that were based on DNA sequences obtained from the *S. aureus* databases and sequences produced as part of this study. Suitable restriction sites were introduced where necessary at the 5' ends to enable subsequent cloning. Note this is unnecessary for cloning into pCR®2.1-TOPO (Chapter 2.22.2).

2.21.2 PCR procedure

2.21.2.1 DNA amplification

2.21.2.1.1 Taq polymerase

Standard PCR amplification reactions were performed using Taq polymerase.

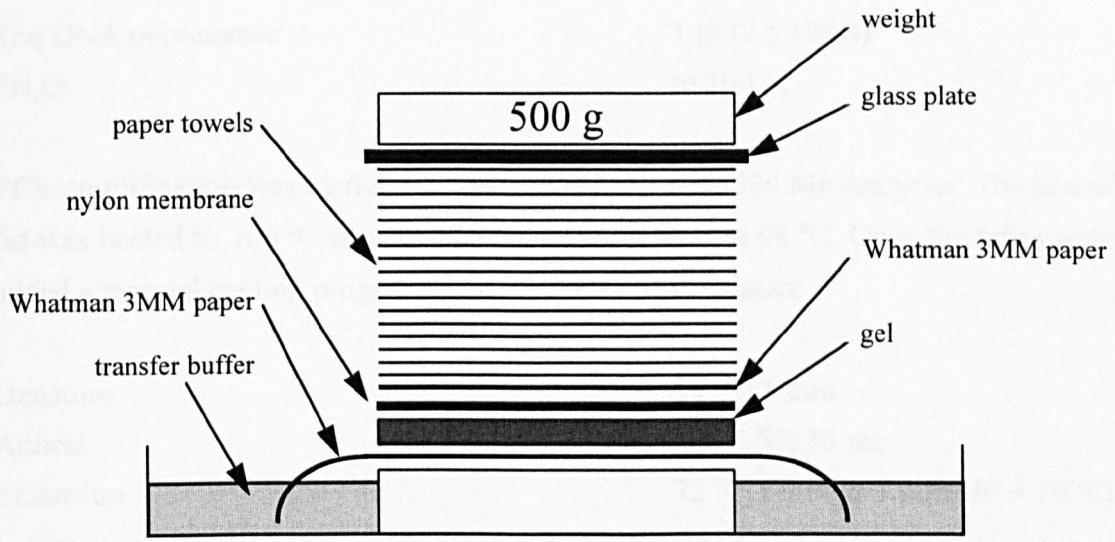


Figure 2.2
Capillary transfer of nucleic acids from gels

Buffer is drawn from a reservoir and passes through the gel into a stack of paper towels. The nucleic acids (DNA/RNA) is eluted from the gel by the moving stream of buffer and is deposited on the nylon membrane. A weight is applied to the top of the paper towel stack to ensure a tight connection between the layers of material used in the transfer system.

Reproduced from Sambrook *et al.* (1989)

The following were added on ice:

Template DNA	100-500 ng
Forward primer	100 pM
Reverse primer	100 pM
10x Mg-free buffer	10 µl
$MgCl_2$	1.5-4.5 mM (final concentration)
dNTPs	0.1 mM (final concentration)
Taq DNA polymerase	1 µl (2.5 Units)
dH_2O	to 100 µl

PCR amplification was carried out using an Eppendorf 5330 Mastercycler. The heated lid was heated to 106 °C and the block was pre-heated to 96 °C. Once the tubes were added a thermal cycling programme of 35 cycles was initiated:

Denature	96 °C; 1 min
Anneal	50-55 °C; 30 sec
Extension	72 °C; t min (t=1 min kb^{-1} + 10 %)

Once the 35 cycles were complete a final step of 72 °C for 10 min was added to allow complete extension of the primers. The PCR products were stored at 4 °C.

2.21.2.1.2 Pwo polymerase

DNA amplifications of < 3 kb that required 3'-5' proof-reading activity were performed with Pwo polymerase (Roche).

The following were added on ice:

Master mix 1:

Template DNA	0.1-0.75 µg
dNTPs	0.2 mM (final concentration)
Forward primer	300 nM (final concentration)
Reverse primer	300 nM (final concentration)
dH_2O	to 50 µl

Master mix 2:

10x PCR buffer (20 mM $MgSO_4$)	10 µl
Pwo polymerase	0.5 µl (2.5 Units)
dH_2O	to 50 µl

Master mix 1 and 2 were combined on ice in a PCR tube and cycling immediately commenced.

PCR amplification was carried out using an Eppendorf 5330 Mastercycler. The heated lid was heated to 106 °C and the block was pre-heated to 94 °C. Once the tubes were added a thermal cycling programme of 10 cycles (A), followed by a second one of 20 cycles (B) was initiated:

Cycle A - 10 cycles	
Denature	94 °C; 15 sec
Anneal	50-55 °C; 30 sec
Extension	72 °C; t min (t= ~1 min 1.5 kb^{-1})
Cycle B - 20 cycles	
Denature	94 °C; 15 sec
Anneal	50-55 °C; 30 sec
Extension	72 °C; t min (t= ~1 min 1.5 kb^{-1})
	+ 20 sec for each cycle

Once the all cycles were complete a final step of 72 °C for 7 min was added to allow complete extension of the primers. The PCR products were stored at 4 °C.

2.21.2.1.3 Expand™ polymerase

DNA amplifications of >3 kb that required 3'-5' proof-reading activity were performed with the Expand™ Long Template PCR system (Roche). This system is composed of a unique enzyme mix containing the thermostable Taq and Pwo polymerases, and is optimised for amplifying fragments of >3-27 kb in length.

The following were added on ice:

Master mix 1:

Template DNA	0.1-0.5 µg
dNTPs	0.35 mM (final concentration)
Forward primer	300 nM (final concentration)
Reverse primer	300 nM (final concentration)
dH_2O	to 50 µl

Master mix 2:

10x PCR buffer (1.75 mM $MgCl_2$)	10 µl
Expand™ polymerase	0.75 µl (2.6 Units)
dH_2O	to 50 µl

Master mix 1 and 2 were combined on ice in a PCR tube and cycling immediately commenced.

PCR amplification was carried out using an Eppendorf 5330 Mastercycler. The heated lid was heated to 106 °C and the block was pre-heated to 94 °C. Once the tubes were added a thermal cycling programme of 10 cycles (A), followed by a second one of 20 cycles (B) was initiated:

Cycle A - 10 cycles	
Denature	94 °C; 10 sec
Anneal	50-55 °C; 30 sec
Extension	68 °C; t min*
Cycle B - 20 cycles	
Denature	94 °C; 10 sec
Anneal	50-55 °C; 30 sec
Extension	68 °C; t min*
	+ 20 sec for each cycle

t min*	2	4	8	15	20	30
kb	3	6	10	20	30	40

Once all cycles were complete a final step of 68 °C for 7 min was added to allow complete extension of the primers. The PCR products were stored at 4 °C.

2.21.2.2 RT-PCR

The following were added on ice:

PCR master mix (Life Technologies)	43 µl
DNA template*	5 µl (700 ng)
Forward primer	1 µl (25 pM)
Reverse primer	1 µl (25 pM)

* cDNA +/- RT, prepared from *in vitro* (BHI/TSB), and *in vivo* (kidney) grown *S. aureus*; genomic DNA control.

Note: Primers were designed to a central portion of each gene by SmithKline Beecham, sequences not supplied.

PCR amplification was carried out using a Perkin-Elmer GeneAmp PCR System 9600. The heated lid was heated to 106 °C and the block was pre-heated to 94 °C. Once the tubes were added a thermal cycling programme of 40 cycles was initiated:

Denature	94 °C; 30 sec
Anneal	50 °C; 30 sec
Extension	72 °C; 30 sec

Once the 40 cycles were complete a final step of 72 °C for 7 min was added to allow complete extension of the primers. The PCR products were visualised by agarose gel electrophoresis (Chapter 2.19.1.3) and stored at 4 °C.

2.21.2.3 TaqMan PCR

The following were added to 96 well microtitre plates:

DNA template*	4 µl (~ 500 ng)
10x buffer	5 µl
$MgCl_2$	11 µl (2.75 mM)
dNTP	4 µl
UNG	0.5 µl
Probe	0.1 µl

Forward primer	1.8 µl (45 pM)
Reverse primer	1.8 µl (45 pM)
TaqGOLD	0.5 µl
dH_2O	to 50 µl

* cDNA +/- RT, prepared from *in vitro* (BHI/TSB), and *in vivo* (kidney) grown *S. aureus*; dilutions of *S. aureus* chromosomal DNA (calibration curve); and a no template control (dH_2O).

S. aureus chromosomal DNA dilutions (ng): 100, 25, 6.25, 1.5625, 0.3906, 0.0977, 0.0244, 0.0061.

Note: Primers and probes were designed to a central portion of each gene by SmithKline Beecham, sequences not supplied.

Reactions were carried out with enzymes and reagents from ABI Perkin-Elmer. All reactions were performed in at least duplicate in adjacent wells on the microtitre plate. PCR amplification was carried out using an ABI PRISM 7700 Sequence Detection System (Perkin-Elmer Corporation). The microtitre plate was added to the machine and was heated to 50 °C for 2 min and then 95 °C for 10 min. A thermal cycling programme of 40 cycles was then initiated:

Denature	95 °C; 15 sec
Anneal/extension	60 °C; 1 min

Once the 40 cycles were complete the reaction was over and results were analysed using the TaqMan computer software (Perkin-Elmer)

2.21.3 Automated DNA sequencing

2.21.3.1 Short single read sequencing reactions

Short (300-700 bp) single read sequencing reactions were carried out on the rescued DNA flanking the transposon insertion (Chapter 4.2.6.3), and to confirm correct construct construction (Chapter 2.21.3.1) using the following protocol.

The following were added on ice:

Terminator ready reaction mix	8 µl
Template DNA	0.2-0.5 µg
Primer	3.2 pM
dH_2O	to 20 µl

PCR amplification was carried out using an Eppendorf 5330 Mastercycler. The heated lid was heated to 106 °C and the block was pre-heated to 96 °C. Once the tubes were added a thermal cycling programme of 25 cycles was initiated:

96 °C for 30 sec
50 °C for 15 sec
60 °C for 4 min

The PCR products were cleaned by ethanol/sodium acetate precipitation (Chapter 2.17.1.6). The sequencing reactions were carried out on an ABI PRISM dye terminator cycler (Perkin-Elmer) according to the manufacturer's instructions.

2.21.3.1 Single read sequencing reactions for both strands

PCR products requiring accurate sequencing, for example those necessary for the design of TaqMan primers and probes (Chapter 5.2.5.1.4), were sequenced by MWG-BIOTECH. For these reactions, both strands of the PCR product were sequenced, with an accuracy of >99 %. This sequencing was carried out using the LI-COR 4200 system.

2.21.3.3 DNA sequencing of the *dap* operon

2.21.3.3.1 Publication quality sequencing

The plasmid pSEG10 containing the 5' end of the *dap* operon, rescued from *seg*10 (Chapter 5.2.2.1), required sequencing to confirm the structure of the *dap* operon. A large prep of this plasmid was prepared from *E. coli* (Chapter 2.17.1.3) and the relevant portion of the plasmid was sequenced by MWG-BIOTECH. For this reaction,

both strands of the plasmid were completely sequenced, with an accuracy of >99.95 %. This sequencing was carried out using the LI-COR 4200 system.

2.21.3.3.2 Primer walking quality sequencing

The plasmid pMTOPO9 containing the 3' end of the *dap* operon (Chapter 5.2.2.2.1), constructed as described in Chapter 2.22.2, required sequencing to confirm the structure of the 3' portion of the *dap* operon. A large prep of this plasmid was produced from *E. coli* (Chapter 2.17.1.3) and the plasmid insert was sequenced by MWG-BIOTECH. For this sequencing reaction only one strand of the insert was sequenced in a plasmid vector (pCR®2.1-TOPO (Invitrogen)). This sequencing was carried out using the LI-COR 4200 system.

2.21.4 PCR screening

2.21.4.1 *E. coli* cells

Colonies from *E. coli* transformation plates were patched using sterile toothpicks onto LB agar plates containing the appropriate antibiotic and grown overnight at 37 °C. A small amount of the patched positive clone was transferred to a PCR tube containing the correct PCR mix (as in Chapter 2.21.2.1.1 minus the DNA) and the primers appropriate for the fragment to be screened. The PCR programme followed for screening was as in Chapter 2.21.2.1.1 with the tubes heated to 95 °C initially for 5 min to ensure lysis of the cells and release of DNA. Transformants containing the plasmid and correct insertion, were isolated by subjecting the PCR products to gel electrophoresis (Chapter 2.19.1.3) and then identifying the associated colonies from the LB agar plate.

2.21.4.2 *S. aureus* cells

A small amount of patched clone was resuspended in 30 µl PBS (Chapter 2.2.7). These cells were then lysed with 3 µl of 5 mg ml^{-1} lysostaphin (Chapter 2.3) at 37 °C for 10 min. 5 µl of the lysed cells were then used in a PCR reaction to determine if pAZ106 had integrated into the gene of interest. PCR was conducted as described in Chapter 2.21.2.1.2, using primers OLP3 and OL 2-4 (Table A.1.2.1).

2.21.5 Purification of PCR fragments

Following PCR, the DNA fragments were cleaned using the QIAquick PCR purification kit (QIAGEN) according to the manufacturer's instructions, as follows. 5 volumes of Buffer PB were added to the PCR reaction mix. The DNA was bound to the column by centrifuging the sample for 1 min (11,000 **g**) and the flow through discarded. The column was then washed with 750 µl Buffer PE, and once more centrifuged for 1 min (11,000 **g**). The flow through was discarded and the column was again centrifuged for 1 min (11,000 **g**). The DNA was then eluted into a clean microfuge tube with 50 µl of Buffer EB after standing for 1 min by centrifugation (11,000 **g**; 1 min). DNA was stored at 4 °C (short-term) and -20 °C (long-term).

All QIAGEN solution details are given in Chapter 2.2.8.

2.22 DNA cloning

2.21.1 Cloning of Tn*917* insertion sites

2.22.1.1 Rescue of chromosomal DNA flanking Tn*917* insertion sites

Chromosomal DNA was prepared for each *seg* mutant (Chapter 2.17.1.1). 10 µg of chromosomal DNA from each *seg* strain was digested overnight at 37 °C (Chapter 2.19.1.1) with *Eco*RI or *Xba*I. Digests were cleaned by phenol/chloroform extraction (Chapter 2.17.1.4), and a fraction of the DNA was analysed by agarose gel electrophoresis (Chapter 2.19.1.3). 0.2 µg of the clean digested DNA was then ligated (Chapter 2.19.1.2.2) and transformed in to *E. coli* DH5α (Chapter 2.16.1). *E. coli* containing a plasmid comprising part of the transposon (the *col*E1 replicon and the ampicillin resistance gene (*bla*)) plus flanking chromosomal DNA out to the nearest relevant restriction site in the chromosomal DNA were recovered followed selection on ampicillin (50 µg ml^{-1}).

2.22.1.2 Recovery of DNA flanking Tn*917* insertion sites

Plasmid DNA was recovered from two independent transformants for each *seg* mutant by a small scale plasmid preparation (Chapter 2.17.1.2). This plasmid DNA was analysed by digesting with *Eco*RI, and *Eco*RI + *Sal*I (or *Xba*I, and *Xba*I + *Sal*I) to

confirm the presence of a plasmid derived from Tn*917* (pSEGX). Confirmed plasmid preparations were stored at -20 °C.

2.22.1.3 Determination of the Tn*917* insertion sites

The transposon insertion site, was identified by sequencing (Chapter 2.21.3.1) plasmids recovered in Chapter 2.22.1.2. These plasmids were sequenced using primer TnS1 (Table A.1.2.1), which is complementary to a region at the *Sal*I end of the transposon.

2.22.2 Cloning the 3' end of the *dap* operon

To allow the complete sequencing of the *dap* operon in 8325-4, the 3' end of the operon (3' end of *dapA* - 3' of *lysA*) was cloned into a suitable vector (pCR®2.1-TOPO (Invitrogen)) using the TOPO™ TA Cloning® kit (Invitrogen). PCR primers (MW41 and MW44 (Table A.1.2.1)) were designed from sequence in the *dap* database. These primers were used to PCR amplify a 5.7 kb fragment using Expand™ DNA polymerase mix (Chapter 2.21.2.1.3). The PCR product was gel purified (Chapter 2.17.1.8) and then cloned into pCR®2.1-TOPO according to the manufacturer's guidelines as follows.

The following were added at room temperature:

PCR product	5-20 ng
pCR®2.1-TOPO	1 µl
Sterile water	to 5 µl

The reaction was mixed gently and incubated at room temperature for 5 min. The mix was then briefly centrifuged (15 sec; 11.000 **g**) and place on ice. 2 µl of 0.5 M β-mercaptoethanol was added to each vial of TOP10 One Shot™ cells (Invitrogen), followed by 2 µl of the TOPO cloning reaction. This reaction was gently mixed and incubated on ice for 30 min, after which time the cells were heat shocked at 42 °C for 30 sec, before returning back to ice for a further 2 min. 250 µl of room temperature SOC medium (Chapter 2.1.1.10) was then added to the reaction and the tube was shaken horizontally at 37 °C for 30 min and then placed on ice. 10-50 µl of the transformation mix was then plated onto pre-warmed (37 °C) LB agar plates

(Chapter 2.1.1.4), containing 50 μg ml^{-1} Amp, spread with 80 μl 20 mg ml^{-1} X-Gal. The plates were then incubated overnight at 37 °C. 10 white or light blue colonies were selected for analysis. These clones were cultured overnight in LB medium containing 50 μg ml^{-1} Amp and the plasmids isolated using a small scale *E. coli* plasmid preparation (Chapter 2.17.1.2). Positive clones were identified by restriction analysis using *Eco*RV, *Hind*III, *Nsi*I or *Sac*I (Chapter 2.19.1.1). Confirmation of a positive result was achieved by automated sequence analysis (Chapter 2.21.3.1) using M13 forward and reverse primers. A positive clone pMTOPO9 was then sequenced (Chapter 2.21.3.3.2).

2.22.3 *lacZ* reporter gene construction

lacZ reporter gene fusions were created without gene inactivation using pAZ106. PCR primers (OL 1-4) (Table A.1.2.1) were designed with unique restriction sites (*Eco*RI and *Bam*HI) to allow cloning into pAZ106. The PCR primers were designed (Chapter 2.21.1) as shown in Figure 5.18. PCR amplification was performed using Pwo polymerase (Chapter 2.21.2.1.2) and the resulting PCR products were gel purified (Chapter 2.17.1.8). Both the vector (pAZ106) and the PCR fragments were digested with *Eco*RI and *Bam*HI (Chapter 2.19.1.1) and then cleaned up using the QIAquick PCR purification kit (QIAGEN) (Chapter 2.21.5). The PCR fragments and the cut vector were then ligated (Chapter 2.19.1.2.1) and transformed into *E. coli* (Chapter 2.16.1.2). Transformants were selected on LB containing 50 μg ml^{-1} Amp. 5 transformants for each construct were picked and analysed. These transformants were cultured overnight in LB medium containing 50 μg ml^{-1} Amp. The plasmids were then isolated using a small scale *E. coli* plasmid prep (Chapter 2.17.1.2). Positive transductants were identified by restriction analysis using *Eco*RI and *Bam*HI (Chapter 2.19.1.1). Following confirmation of positive clones, a large scale *E. coli* plasmid prep (Chapter 2.17.1.3) was performed for each construct. The purified plasmids were then electroporated into *S. aureus* RN4220 and transformants selected on BHI Ery/Lin plates (5 μg ml^{-1}/25 μg ml^{-1}) (Chapter 2.16.2). Positive transformants were patched onto BHI and sera plates with an X-Gal overlay (Chapter 2.11.1) and grown overnight at 37 °C under aerobic and microaerobic conditions respectively (Chapter 2.9). Clones found to be blue on serum but white on BHI were presumed to contain the correct *lacZ* construct. Confirmation of correct *lacZ* construction was afforded by PCR screening (Chapter 2.21.4.1) using primers OLP3 and OL 2-4 (Table A.1.2.1). Correct *lacZ* fusions were determined by PCR products 95 bp larger than the original cloned products. Positive constructs were then transduced into 8325-4 by phage transduction

(Chapter 2.14). Colonies were once again screened using blue white selection and PCR screening as described above. Final confirmation for the construction of *lacZ* transcriptional fusion strains was gained by Southern blot analysis using a DIG labelled probe representing the cloned region (Chapter 2.20).

2.23 Sequence and database analysis

2.23.1 DNA and amino acid sequence analysis

Automated sequence data from the ABI PRISM dye terminator cycler was edited and analysed using the Applied Biosystems 373A DNA sequencer data analysis program. For all mutants putative open reading frames (ORFs) were identified using the Gene Jockey II program (BIOSOFT). This package includes tools for determining potential ORFs, translations of nucleic acid sequence to protein sequence, and restriction enzyme analysis. The amino acid sequence of these putative ORFs were analysed for homology with proteins from *S. aureus* and other bacteria using the BLASTP programme via the BLAST 2.0.4 database on the Internet (http://www.ncib.nlm.nih.gov/cgi-bin/BLAST/). DNA sequence was also analysed using the BLASTX programme via the BLAST 2.0.4 database. This same technique was used for the analysis of DNA sequence produced by MWG BIOTECH (Chapters 2.21.3.3.1-2.21.3.3.1). DNA sequences were also analysed using the Staden based NIP program (Staden, 1990). This program was also used for determining potential ORFs and translations of nucleic acid sequence to protein sequence, but also additionally for the analysis of stem-loop structures.

2.23.2 Database searching

Genome databases of *B. subtilis* (SubtiList; http://www.pasteur.fr/Bio/SubtiList.html) and *E. coli* (Colibri; http://www.pasteur.fr/Bio/Colibri.html) were searched on the Internet using: gene; key word; and BLAST search functions, within the databases. These searches allow the identification of: genes and their enzymes; potential isozymes; homologues; genetic locations; the identity of surrounding genes (and their function); and both the DNA and protein sequences of specified genes. The incomplete publicly available (Internet) *S. aureus* databases (TIGR, strain Col; http://www.TIGR.org/cgi-bin/BlastSearch/blast.cgi? and NCTC 8325; http://www.genome.ou.edu/staph.html) were searched using BLAST analysis. Contigs of interest were then

extracted and analysed in Gene Jockey II for the presence of ORFs. Putative ORFs were then screened for similarity to protein sequences in the NCBI protein sequence database library using the BLAST programs (Chapter 2.23.1). The SmithKline Beecham, *S. aureus* databases (strains: WCUH29; and Buttle) were searched in an identical manner to that described above for the publicly available *S. aureus* databases (TIGR and NCTC 8325). However extracted contigs were analysed using in-house programs and databases (details not supplied). Finally sequences of known proteins that could not be obtained from another source were gained from the SWISS-PROT and TrEMBL, protein databases via ExPASy (http://www.expasy.ch/) using SRS (Sequence Retrieval System). These databases and programs were all accessed through the Internet. These sequences were then used to search for homologues in other databases.

2.23.3 Database construction

Sequence data, both sequenced during this project, and retrieved from the *S. aureus* databases (TIGR and NCTC 8325), were assembled into a database using the Staden SAP program (Staden, 1990). The database was then analysed using both Gene Jockey II and the Staden NIP package (Staden, 1990).

2.24 Auxotrophy experiments

2.24.1 Solid media auxotrophy experiments

16 different types of chemically defined media were prepared (Chapter 2.1.1.3.1). *S. aureus* strains were inoculated in triplicate onto the different types of CDM agar from fresh standard CDM (16) agar plates. These plates were then incubated aerobically at 37 °C for 48 hours. After this time plates were scored for growth (+/-).

2.24.2 Liquid media auxotrophy experiments

5 ml CDM (16) cultures were grown overnight aerobically at 37 °C. Cells were pelleted by centrifugation (11,000 **g**; 3 min) and washed twice with 5 ml CDM 15 (-all 4 AAA). The pellet was then resuspended in 5 ml CDM 15 and used as an inoculum. 25 ml CDM cultures (CDM 1-4 and 11-16 (Table 5.2)) were inoculated to

A_{600} 0.07, a value previously determined to be a suitable inoculum for CDM +/- the aspartate family of amino acids. Strains were grown aerobically at 37 °C with shaking (250 rpm) for up to 48 h. Growth was monitored throughout this time by A_{600} (Chapter 2.7.1).

2.25 Murine pathogenicity models

2.25.1 Murine subcutaneous lesion model (Chan *et al.*, 1998)

S. aureus strains were grown to stationary phase (t = 15 h) in BHI containing no antibiotics as described in Chapter 2.4.1.1. The bacteria were harvested by centrifugation (4,000 rpm (Centaur 2); 5 min) and resuspended in phosphate-buffered saline (PBS) (Chapter 2.2.7) and adjusted to 5 x 10^8 cfu ml^{-1}. Male and female bald mice were inoculated subcutaneously (n = 7 (*seg*1 (*lysC*), *seg*24 (*asd*)) and n = 8 (*seg*10 (*dapA*))) with 0.2 ml of bacterial suspension (~ 1 x 10^8 cfu ml^{-1}). After 7 days the mice were sacrificed and skin lesions were aseptically removed and stored at -70 °C. The lesions were weighed and homogenised in 5 ml PBS at 4 °C for approximately 60 min. The total number of bacteria recovered from the lesions (cfu $lesion^{-1}$) was determined by direct cell counts (Chapter 2.7.2) on BHI agar. Statistical evaluation of the data was determined by the Mann-Whitney U test.

2.25.2 Murine pyelonephritis model (Nicholas *et al.*, 1999)

Mice were inoculated (n = 5 (WCUH29, SWC24 (*asd*)) n = 4 SWC1 (*lysC*), and n = 3 (SWC10 (*dapA*))) with ~ 2 x 10^7 cfu ml^{-1} WCUH29 cells. After 7 days the mice were sacrificed and their kidneys were aseptically removed and stored at -70 °C. The kidneys were homogenised at 4 °C to extract bacteria. The total number of bacteria recovered from the kidney (cfu $kidney^{-1}$) was determined by direct cell counts (Chapter 2.7.2) on TSB agar. Statistical evaluation of the data was determined by the Mann-Whitney U test.

2.25.3 Murine wound infection model (Nicholas *et al.*, 1999)

Mice were inoculated (n = 5) with 7 x10^6-2 x 10^7 cfu ml^{-1} WCUH29 cells. After 7 days the mice were sacrificed and wounds were aseptically removed and stored at -70 °C.

The wounds were homogenised at 4 °C to extract bacteria. The total number of bacteria recovered from the wound (cfu wound^{-1}) was determined by direct cell counts (Chapter 2.7.2) on TSB agar. Statistical evaluation of the data was determined by the Mann-Whitney U test.

2.26 Light microscopy

S. aureus strains were grown under standard conditions (Chapter 2.4.1.1) in both BHI and serum. Samples were taken throughout growth and into stationary phase, and were viewed under a light microscope using oil immersion at 1000x magnification.

2.27 HPLC analysis

2.27.1 HPLC peptidoglycan analysis

2.27.1.1 Preparation of cell wall peptidoglycan (Atrih *et al.*, 1996)

S. aureus strains were grown overnight as described in Chapter 2.4.1.1 and used to inoculate (A_{600} 0.05) 100 ml of fresh media (BHI/serum). These cultures were grown using standard conditions (Chapter 2.4.1.1) until the end of exponential phase (A_{600} dependent on media used). These cells were then pelleted by centrifugation (10,000 **g**; 10 min; 4 °C), resuspended in 25 ml of hot (80 °C) 5 % (v/v) SDS and boiled for 30 min. Insoluble material was recovered by centrifugation (10,000 **g**; 10 min) and the pellet resuspended in 25 ml warm (50 °C) 4 % (v/v) SDS and boiled for 15 min. The samples were then centrifuged (10,000 **g**; 10 min) and the pellet washed with warm (50 °C) dH_2O 4-6 times to remove all trace of SDS. These samples were then centrifuged (11,000 **g**; 5 min) and resuspended in 50 mM Tris-HCl pH 7.5 and treated with amylase (200 μg ml^{-1} final concentration) (Chapter 2.3) for 2 h at 37 °C. After this treatment the samples were treated with pronase (1 mg ml^{-1} final concentration) for 2 h at 60 °C. The walls were then recovered by centrifugation (11,000 **g**; 5 min) washed with dH_2O and pelleted again (11,000 **g**; 5 min). The pellet was then resuspended in hydrofluoric acid (300 μl of 48 % v/v solution) and incubated on ice for 36 h. The sample was then pelleted (11,000 **g**; 5 min) once more, and washed repeatedly with 100 mM Tris-HCl pH 7.5 until the pH was neutral. Peptidoglycan was stored at -20 °C.

2.27.1.2 Enzymatic hydrolysis of peptidoglycan and RP-HPLC separation of soluble muropeptides (Atrih *et al.*, 1996)

Peptidoglycan samples (Chapter 2.27.1.1) were digested overnight at 37 °C with 250 μg ml^{-1} of Cellosyl (Chapter 2.3). The reaction was stopped by boiling the reaction for 3 min. Insoluble material was removed by centrifugation (14,000 **g**; 8 min), and the resulting peptidoglycan hydrolysate stored at -20 °C. Muropeptide-containing samples were reduced at room temperature using sodium borohydride at a final concentration of 8 mg ml^{-1}. The reaction was stopped after 13 min by lowering the pH to 4 with phosphoric acid. The reduced muropeptides were then separated by reverse-phase HPLC (RP-HPLC). A Hypersil octadecylsilane column purchased from Sigma (4.6 by 250 mm; particle size 5 μm) was used without a guard column. All samples were filtered (0.45 μm pore size) prior to analysis, to remove insoluble material. The column was equilibrated at 52 °C with buffer A (Chapter 2.2.5.1) at a flow rate of 0.5 ml min^{-1} for 17 min at 40 °C with a Waters 600E delivery system and column heater. Soluble-reduced muropeptides (60 μl), were then loaded by injection (Waters 717 plus autosampler). A linear gradient of 0-100 % buffer B (Chapter 2.2.5.2) over a period of 270 min was started 5 min after sample injection. The flow rate was constant at 0.5 ml min^{-1} over the course of the gradient and the eluted compounds were detected by monitoring A_{202} (Waters 486 absorbance detector).

2.27.2 Amino acid analysis of sera by HPLC (Atrih *et al.*, 1996)

The pico-Tag amino acid analysis system (Waters) was used for the analysis of both pig and human sera. Prior to amino acid analysis, desalted, HPLC-purified muropeptides were hydrolysed with 6 M HCl for 14 h at 100 °C. The hydrolysates were dried and derivatised with phenylisothiocynate as recommended by the manufacturer. The phenylthiocarbamyl amino acid derivatives were separated on an HPLC Pico-Tag column (Waters; 3.9 by 150 mm) maintained at 38 °C. Peaks were detected at 254 nm and integrated by use of a Waters 746 data module integrator. The results of this analysis are shown in Table A.1.7.

CHAPTER 3

DEVELOPMENT OF *IN VITRO* MODELS MIMICKING INFECTION

3.1 Introduction

3.1.1 Background to the *in vitro* models mimicking infection

Two different *in vitro* models mimicking infection were selected for this study. A brief introduction to each model along with the reasoning behind its selection follows.

3.1.1.1 Synthetic cerebrospinal fluid model

Cerebrospinal fluid (CSF), is the fluid filling the cavity in the brain and the spinal chord, and is the fluid in which they are suspended. CSF is secreted primarily by the choroid plexus epithelium in the lateral 3rd and 4th ventricles of the brain, from where it circulates to the entire surface of the brain. The fluid has a vital role, supporting and cushioning the brain within the skull. The CSF filled subarachnoid space lies protected behind the blood brain barrier, which severely limits the passage of even small molecules from the blood, into the brain or CSF (Tuomanen, 1996). Thus the composition of the CSF is not determined directly by the composition of the plasma, but is a specially controlled privileged environment. The ionic composition of CSF however is broadly similar to that of plasma, mainly sodium with chloride and bicarbonate. Further details of the composition of CSF are given in Tables, A.1.3. Importantly, several types of pathogenic bacteria including *S. aureus* are capable of entering the central nervous system (CNS), causing infections.

S. aureus along with other staphylococci have been shown to be important causative agents of a wide range of infections of the CNS. Examples include CSF shunt infections, acute purulent meningitis, brain abscess, and epidural abscess. The importance of the staphylococci in these infections is highlighted by the fact that staphylococci are the etiologic organisms in 79 % of CSF shunt infections, with *S. aureus* causing approximately 20 % of all infections (reviewed by Roos and Scheld, 1997). These serious infections have varying prognoses for infected individuals, ranging from high rates of mortality to permanent neurologic sequelae and postoperative epilepsy. It was because of the important role of *S. aureus* in infections of

the CNS, that synthetic cerebrospinal fluid (sCSF) was chosen as the first of the models mimicking infection.

3.1.1.2 Serum model

Blood consists of a cellular portion (formed elements), which are suspended and carried in the plasma. These formed elements and their major functions include: erythrocytes (O_2 transport); leukocytes (immune defence); and platelets (blood clotting). Plasma is a straw coloured liquid consisting of water and dissolved solutes. The major solute of the plasma in terms of its concentration is Na^+. In addition to Na^+, plasma contains many other ions, as well as organic molecules such as metabolites, hormones, and enzymes. Plasma proteins constitute 7-9 % of the plasma (Scully, *et al.*, 1986). The three types of proteins are albumins, globulins and fibrinogen. These are involved in: maintenance of blood volume, and pressure; transport of lipids, and fat soluble vitamins; as well as functioning as antibodies, and clotting factors. The blood plasma normally has a pH of between 7.35 and 7.45, with an average value of 7.40. The pH is maintained at this constant value through the function of the lungs and kidneys. The lungs regulate the CO_2 concentration of the blood, and the kidneys regulate the bicarbonate concentration. Further details of the composition of blood are given in Tables A.1.4-A.1.7.

During the process of clot formation, fibrinogen is converted into insoluble threads of fibrin. The fluid from clotted blood, which is called serum, thus does not contain fibrinogen but is otherwise identical to plasma. Pathogenic bacteria, including *S. aureus* are capable of growth in blood causing both septicaemia and bacteraemia.

The incidence of both community and hospital acquired staphylococcal bacteraemia has risen substantially over the past two decades, resulting in its recognition as a primary nosocomial infection. This increase is related mainly to the increasing use of indwelling vascular devices (Ing *et al.*, 1997). Of particular relevance is the fact that by the early nineties approximately one fifth of all cases of bacteraemia were caused by *S. aureus* (Ing *et al.*, 1997). Overall mortality from *S. aureus* bacteraemia remains high, ranging from 16 to 48 %, and is greatly dependent on the presence or absence of, a removable primary focus, underlying respiratory or cardiac disease, associated infective endocarditis or septic shock. Interestingly however, although methicillin-resistant *S. aureus* (MRSA) has emerged as a more prevalent problem in *S. aureus* bacteraemia during the past 10-15 years, this has apparently not affected mortality rates (reviewed by Ing *et al.*, 1997).

Thus *S. aureus* bacteraemia is an extremely important infection that is definitely on the increase, and with the further increased use of catheters and other indwelling vascular devices in combination with the increased prevalence of antibiotic resistant strains, much work needs to be done to investigate the role of virulence factors in the induction and propagation of *S. aureus* in bacteraemia. It was for these reasons along with the ease and availability of serum that serum was chosen as the second of the models mimicking infection.

3.1.2 Reporter gene fusions for the analysis of gene expression

Reporter gene fusions are important tools for the analysis of gene expression and have been constructed in numerous bacterial species including *S. aureus*. These function by inserting promoterless genes, encoding a measurable gene product, downstream of a promoter of interest. Therefore, when the promoter is transcriptionally active, the reporter gene is transcribed and the resulting gene product can be assayed to quantitatively determine activity of the promoter of interest (Camilli *et al.*, 1990). There are several different types of reporter gene systems in use in bacteria, including the bioluminescence (*lux*), and β-lactamase (*lac*) systems. For the purpose of this study, gene fusions using the *lacZ* gene, encoding β-galactosidase were employed. The *lac* system is a commonly used assay system and the genetic and biochemical aspects of the system are well understood (Silhavy and Beckwith, 1985). An advantage of this system over others is that several substrates are available for this system including: X-Gal, and MUG. Fusions used in this assay were constructed in such a way that they were transcriptionally fused to the gene of interest. This therefore provides a method for distinguishing between transcriptional and post-transcriptional control.

3.2 Results

3.2.1 Development of *in vitro* models mimicking infection

3.2.1.1 Synthetic cerebrospinal fluid

Two synthetic cerebrospinal fluids (sCSF) were selected for this study. The first synthetic cerebrospinal fluid (sCSF I) is a very simple medium developed by van Dijk *et al.* (1994). sCSF I is comprised of only a few chlorides and completely lacks a carbon source (Chapter 2.1.1.11). The second synthetic cerebrospinal fluid (sCSF II)

was developed by Smith and Johanson (1985), and is once more a simple medium. Notably however, sCSF II contains both a source of phosphorus and carbon (glucose) (Chapter 2.1.1.12). These two types of sCSF form the first of the *in vitro* models mimicking infection. Details of the composition of CSF are given in Table A.1.3. Results revealed that the *S. aureus* wild type strain (8325-4) would not grow in either sCSF I, or sCSF II. Therefore, further studies involved the use of sCSF as a starvation-survival medium.

3.2.1.1.1 Kinetics of starvation-survival in sCSF I and II

Nutrient limitation of *S. aureus* induces a starvation-survival state, which enables cells to survive for long periods of time until sufficient nutrients become available to support growth. Starvation results in decreased cell size (Watson *et al.*, 1998a) and a reduced metabolic state (Clements and Foster, 1998). Survival was measured as the ability to form a colony on nutrient media (Chapter 2.6.2). This is an accurate measure of viability in *S. aureus* as previous studies have shown there to be no apparent 'viable but nonculturable state' in this organism (Watson *et al.*, 1998a).

The kinetics of the starvation-survival response of *S. aureus* were examined in both sCSF I, and II. Cultures were grown in amino acid limiting CDM until approximately 6 h post-exponential. Cells were subsequently washed and resuspended in sCSF I, or II at the same cell density as the original culture (~ 5 x10^8 - 1 x 10^9 cfu ml^{-1}) (Chapter 2.6.1). Cultures were incubated at 37 °C without shaking and viability was determined over 35 days (Chapter 2.6.2). Results showed that cells starved in sCSF I retained a higher degree of viability than those starved in sCSF II (Figure 3.1). After 35 days a drop in viability of 10^3 was observed in sCSF I, and a 10^4 decrease in viability in sCSF II. Comparable survival kinetics, were observed if cultures were shaken (250 rpm) during starvation (results not shown).

3.2.1.2 Serum model

Two types of sera were selected for this study, pig serum and human serum (Chapter 2.1.1.7). Serum is a medium of complex composition (Tables A.1.5-A.1.7) and although different batches of serum were used during the study, this did not affect reproducibility. The majority of experimentation was carried out using pig serum due to

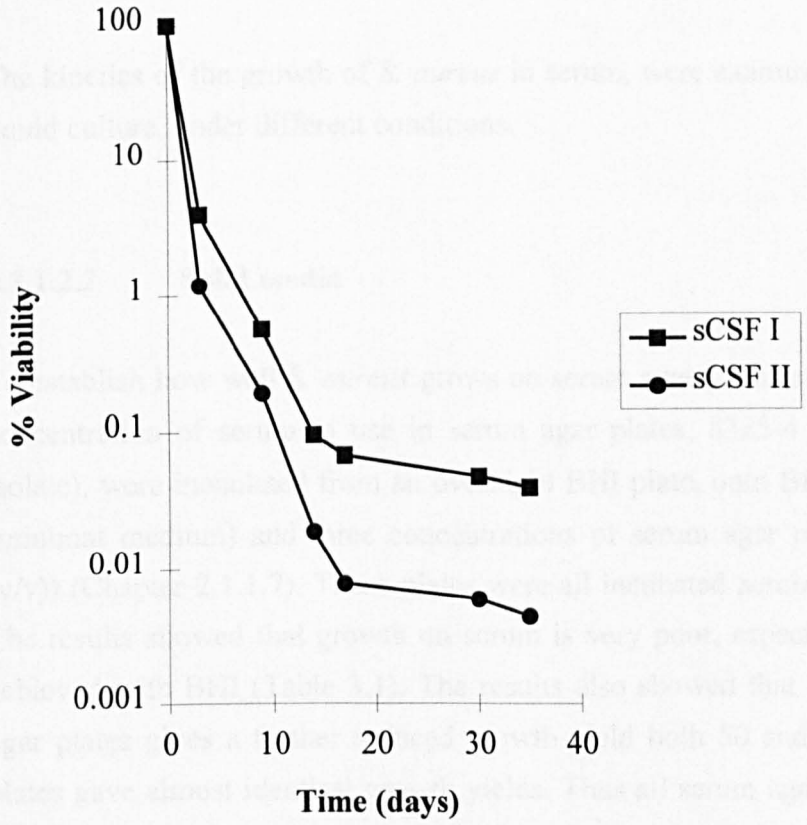


Figure 3.1

Starvation-survival kinetics of *S. aureus* 8325-4 in sCSF I and II at 37 °C

Approximately 6h post-exponential phase cultures of *S. aureus* 8325-4, grown in amino acid limiting CDM, were washed and resuspended in sCSF I, or II and incubated at 37 °C without shaking. Viability (cfu ml^{-1}) was determined by serial dilution on CDM agar plates. 100 % viability is ~ 1 $x10^9$ for both sCSF I, and II. The results are representative of two separate experiments.

cost constraints, although human serum was used in certain situations to confirm results obtained using pig serum. However, unless otherwise stated pig serum was used.

3.2.1.2.1 Development of the serum model

The kinetics of the growth of *S. aureus* in serum, were examined on agar plates and in liquid culture, under different conditions.

3.2.1.2.2 Solid media

To establish how well *S. aureus* grows on serum agar plates and to determine the best concentration of serum to use in serum agar plates, 8325-4 and WCUH29 (clinical isolate), were inoculated from an overnight BHI plate, onto BHI (rich medium), CDM (minimal medium) and three concentrations of serum agar plates (80, 50 and 25 % (v/v)) (Chapter 2.1.1.7). These plates were all incubated aerobically at 37 °C for 48 h. The results showed that growth on serum is very poor, especially if compared to that achieved with BHI (Table 3.1). The results also showed that while 25 % (v/v) serum agar plates gives a further reduced growth yield both 50 and 80 % (v/v) serum agar plates gave almost identical growth yields. Thus all serum agar plates used were 50 % (v/v) serum. This was the case for both pig and human sera (Chapter 2.1.1.7). Finally results showed that both 8325-4 and WCUH29 grew comparably (Table 3.1).

To investigate the role of environmental oxygen and carbon dioxide concentration on the growth of *S. aureus* on serum, serum agar plates were inoculated from an overnight BHI agar plate and grown at 37 °C, under both aerobic and microaerobic conditions (Chapter 2.9). 8325-4 was shown to grow under both of these conditions, although the growth yield was greatly increased by incubating plates microaerobically (Table 3.2).

3.2.1.2.3 Liquid culture

Serum cultures (25 ml in 100 ml flasks) were grown at 37 °C with shaking (250 rpm) under both aerobic and microaerobic conditions (Chapter 2.9). Cultures were grown for 24 h, and growth was monitored by A_{600} or by direct counts on BHI agar plates (Chapter 2.7). Results revealed 8325-4 and WCUH29 reached a final A_{600} of 4.0-6.0 in pig serum (Figure 3.2). 8325-4 was shown to grow to a higher A_{600} under microaerobic than

Media	Average colony diameter (mm)	
	8325-4	WCUH29
BHI	3.8	3.6
CDM	1.3	1.3
Sera 80 % (v/v)	< 0.1	< 0.1
Sera 50 % (v/v)	< 0.1	< 0.1
Sera 25 % (v/v)	<< 0.1	<< 0.1

Table 3.1
The growth of *S. aureus* 8325-4 and WCUH29 on different solid media

BHI, CDM and three concentrations of pig serum agar plates (80, 50 and 25 % (v/v)) were inoculated with a single colony of either 8325-4 or WCUH29. Plates were incubated aerobically at 37 °C for 48 h. Growth was measured according to colony size (diameter). Average diameters were calculated after measuring at least three colonies from two separate experiments.

Condition	Average colony diameter (mm)	
	24 h	48 h
Aerobic	<< 0.1	< 0.1
Microaerobic	0.15	1.25

Table 3.2

The effect of environmental oxygen and carbon dioxide on the growth of 8325-4 on serum agar plates

S. aureus 8325-4 was grown on pig serum agar plates, under aerobic and microaerobic conditions at 37 °C. Growth yield was scored according to colony size (diameter) after both 24 h and 48 h incubation. Average diameters were calculated after measuring at least three colonies from two separate experiments.

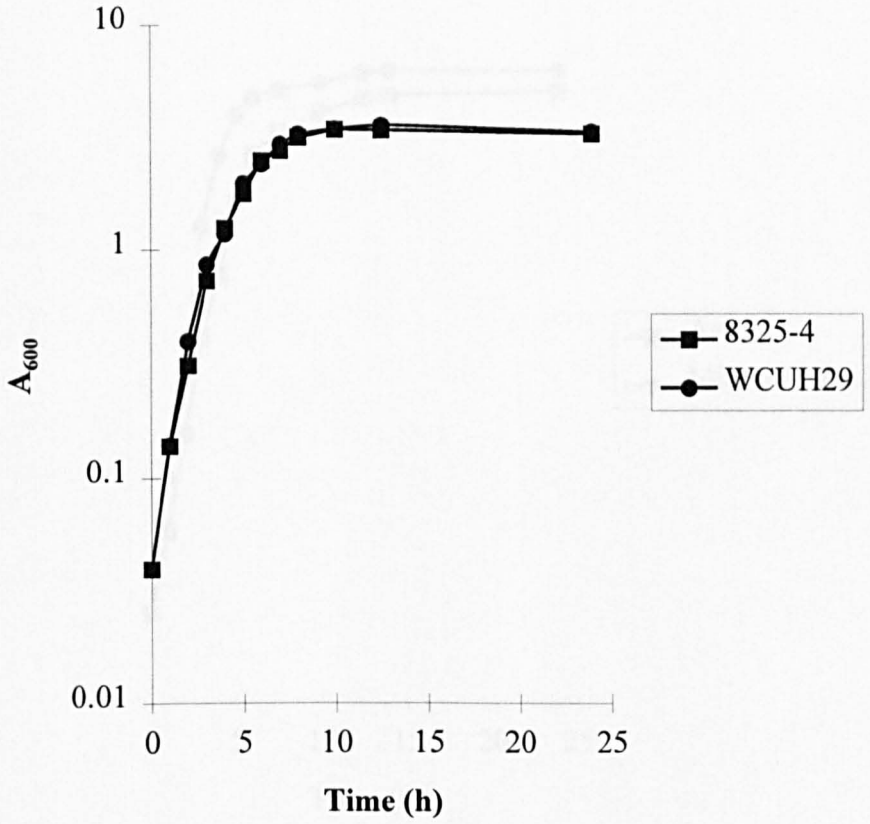


Figure 3.2

The growth of *S. aureus* 8325-4 and WCUH29 in pig serum

S. aureus 8325-4 and WCUH29 were grown at 37 °C in pig serum under aerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

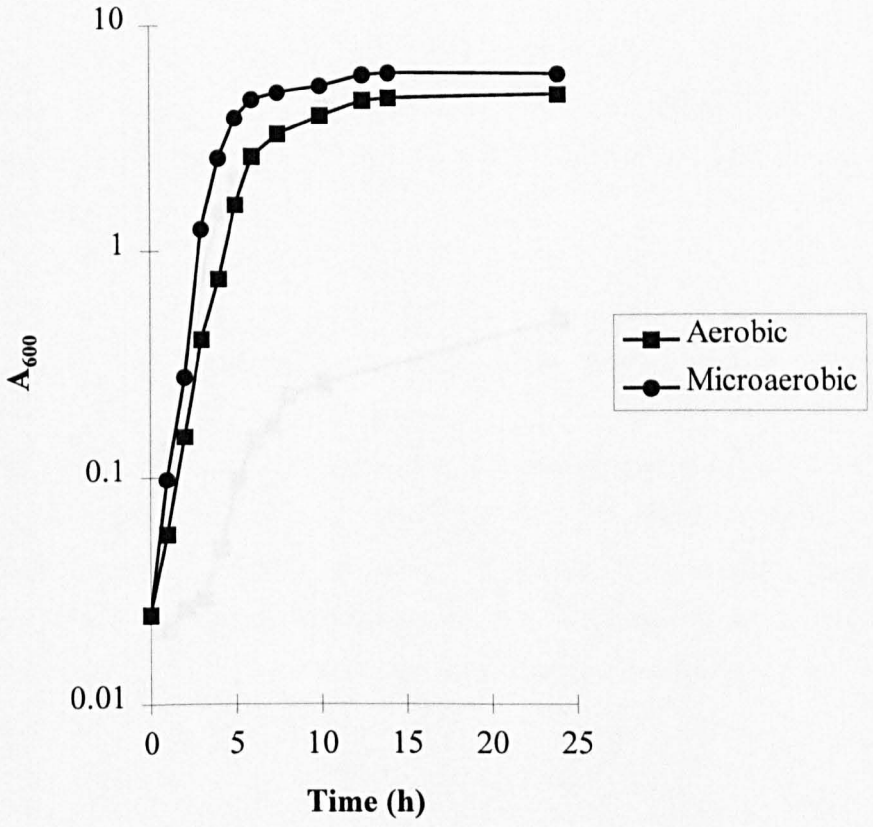


Figure 3.3

The growth of *S. aureus* 8325-4 in pig serum under aerobic and microaerobic conditions

S. aureus 8325-4 was grown at 37 °C in pig serum under both aerobic and microaerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

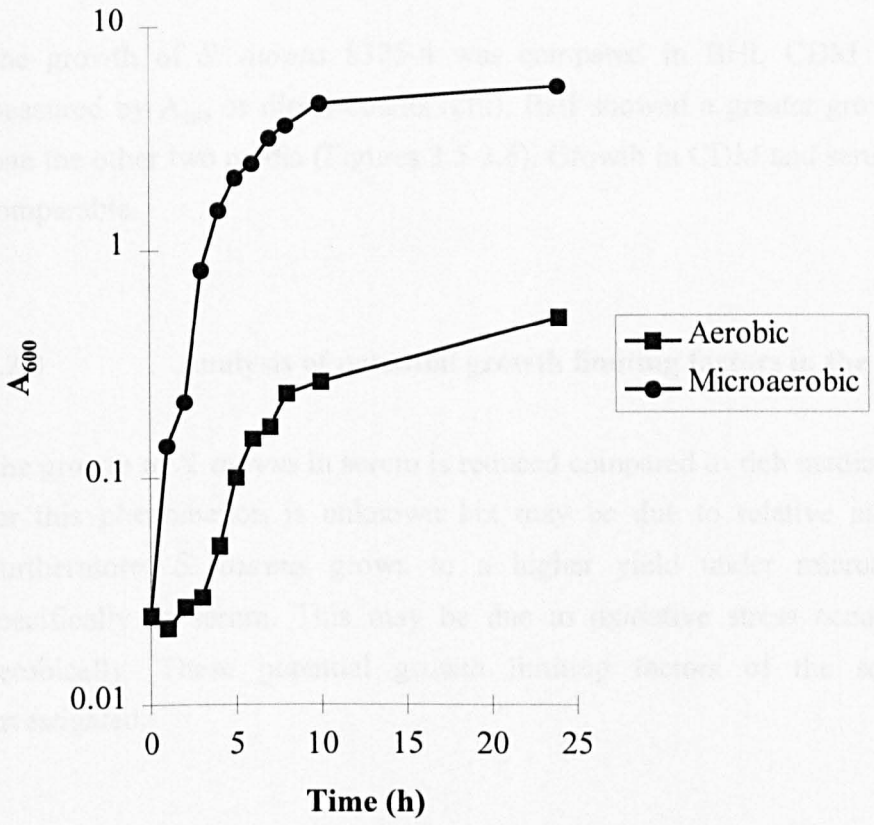


Figure 3.4

The growth of *S. aureus* 8325-4 in human serum under aerobic and microaerobic conditions

S. aureus 8325-4 was grown at 37 °C in human serum under both aerobic and microaerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

aerobic conditions (A_{600} of 4.8 Vs 6.0) (Figure 3.3). The same trend is true for growth in human serum, with 8325-4 growing to a higher A_{600} under microaerobic than aerobic conditions (Figure 3.4). However, the difference in human serum (A_{600} of 0.5-5.4) is much greater compared with that seen in pig serum.

3.2.2 Comparison of the growth of *S. aureus* 8325-4 in different media

The growth of *S. aureus* 8325-4 was compared in BHI, CDM and serum. When measured by A_{600} or direct counts (cfu), BHI showed a greater growth rate and yield, than the other two media (Figures 3.5-3.6). Growth in CDM and serum was found to be comparable.

3.2.3 Analysis of potential growth limiting factors in the serum model

The growth of *S. aureus* in serum is reduced compared to rich media (BHI). The reason for this phenomenon is unknown but may be due to relative nutrient deprivation. Furthermore, *S. aureus* grows to a higher yield under microaerobic conditions specifically in serum. This may be due to oxidative stress occurring when grown aerobically. These potential growth limiting factors of the serum model were investigated.

3.2.3.1 The effect of the addition of iron on the growth of 8325-4 in serum

Iron is an important element for the growth of almost all living cells (Neilands, 1972) and iron availability is of major importance in bacterial pathogenesis (Courcol *et al.*, 1997). *In vivo*, intracellular iron is plentiful however the great majority of extracellular iron is bound to specific binding proteins such as transferrin and lactoferrin. Therefore, for pathogenic bacteria such as *S. aureus* to cause disease in host tissues, they must possess at least one mechanism for scavenging iron, for example by the production of siderophores (Neilands, 1981). Importantly, *S. aureus* has been shown to have a decreased growth rate and yield under iron restricted conditions (Trivier *et al.*, 1995). In serum it is predicted that a large percentage of the iron will be bound to proteins such as transferrin and haemoglobin. In order to establish whether iron is limiting in serum, iron in the form of iron sulphate (Chapter 2.3), was added to serum at a range of concentrations (0-2000 μM). Experiments were carried out using both solid and liquid

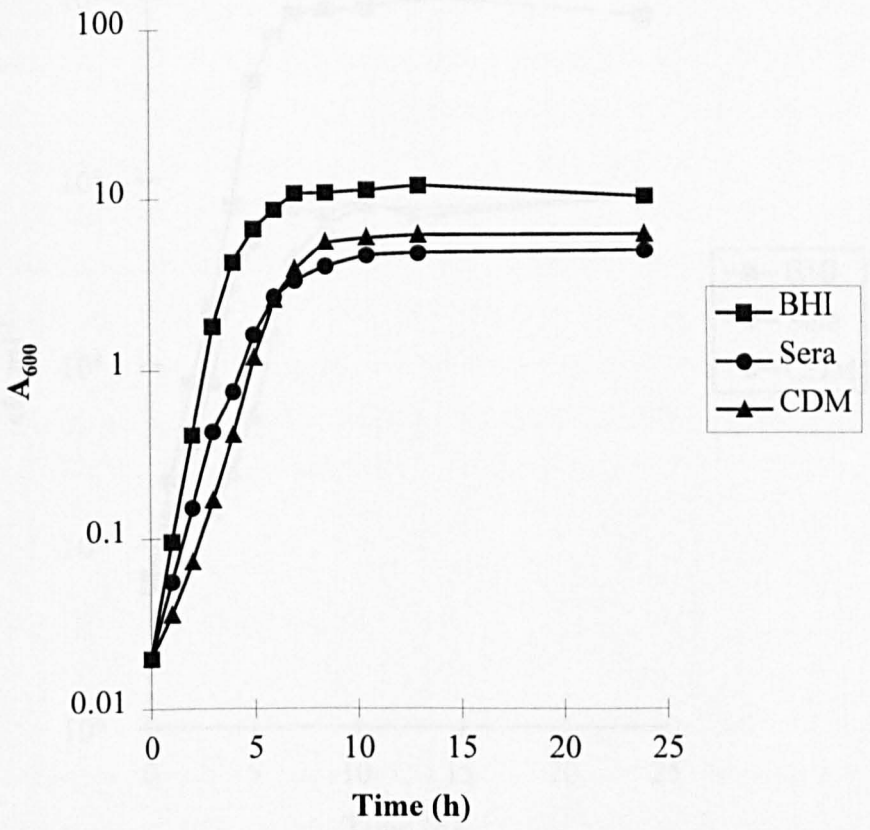


Figure 3.5

The growth of *S. aureus* 8325-4 in BHI, CDM and serum measured by A_{600}

S. aureus 8325-4 was grown at 37 °C in BHI, CDM and serum under aerobic conditions with shaking (250 rpm). Growth was monitored for 24 h by A_{600}. The results are representative of at least two separate experiments.

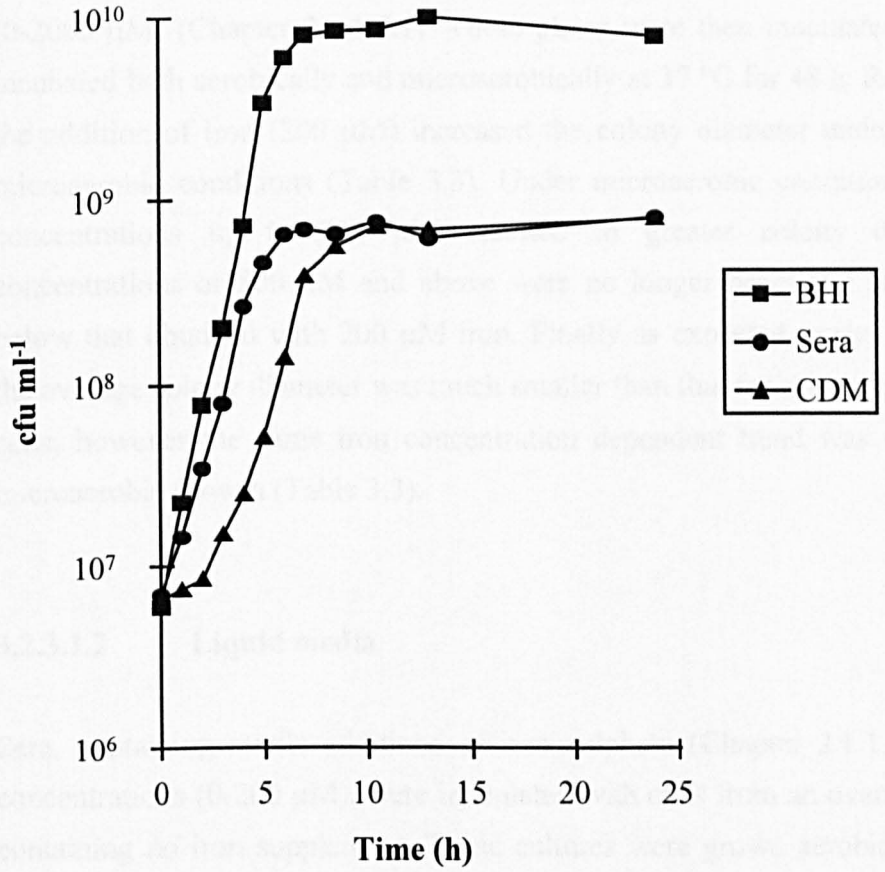


Figure 3.6

The growth of *S. aureus* 8325-4 in BHI, CDM and serum measured by direct counts

S. aureus 8325-4 was grown at 37 °C in BHI, CDM and serum, under aerobic conditions with shaking (250 rpm). Growth was measured by dilution and plate counts on BHI agar plates for 24 h. The results are representative of at least two separate experiments.

media, and growth was determined by colony diameter (solid media) and A_{600} (liquid cultures).

3.2.3.1.1 Solid media

Serum agar plates were prepared containing appropriate concentrations of iron sulphate (0-2000 μM) (Chapter 2.1.1.7.1). These plates were then inoculated with 8325-4 and incubated both aerobically and microaerobically at 37 °C for 48 h. Results revealed that the addition of iron (200 μM) increased the colony diameter under both aerobic and microaerobic conditions (Table 3.3). Under microaerobic conditions, increasing iron concentrations up to 200 μM resulted in greater colony diameter. Whereas concentrations of 500 μM and above were no longer beneficial and reduced growth below that obtained with 200 μM iron. Finally as expected, under aerobic conditions the average colony diameter was much smaller than that for microaerobically incubated cells, however the same iron concentration dependent trend was visible as seen for microaerobic growth (Table 3.3).

3.2.3.1.2 Liquid media

Sera, containing sterile additions of iron sulphate (Chapter 2.1.1.7.1) at a range of concentrations (0-200 μM), were inoculated with cells from an overnight serum culture containing no iron supplement. These cultures were grown aerobically at 37 °C with shaking (250 rpm) for 24 h. Results revealed 8325-4 showed no significant increased growth yield, as determined by A_{600}, when iron was added to serum (Figure 3.7).

3.2.3.2 Examination of the role of oxygen and carbon dioxide levels on the growth of 8325-4 in serum

Early observations revealed that *S. aureus* grew to a higher growth yield in serum under microaerobic than aerobic conditions (Table 3.2). This could be due to a reduced O_2 (10 %), or an increased CO_2 (5 %) tension aiding growth, or a combination both.

Concentration of iron ($FeSO_4$)	Average Colony Diameter (mm)			
	Aerobic 24 h	Aerobic 48 h	Microaerobic 24 h	Microaerobic 48 h
0 μM	<< 0.1	< 0.1	0.15	1.25
2 μM	<< 0.1	< 0.1	0.5	1.65
20 μM	<< 0.1	< 0.1	1.0	2.0
200 μM	0.25	0.5	1.2	2.45
500 μM	0.2	0.25	0.9	1.8
2000 μM	0.2	0.3	0.8	1.5

Table 3.3

The effect of iron on the growth of 8325-4 on serum agar

S. aureus 8325-4 was inoculated onto serum agar plates containing a range of concentrations of iron sulphate (0-2000 μM). Plates were incubated aerobically and microaerobically at 37 °C. Growth yield was scored according to colony size (diameter), after both 24 h and 48 h incubation. Average colony diameters were calculated after measuring at least three colonies from two separate experiments.

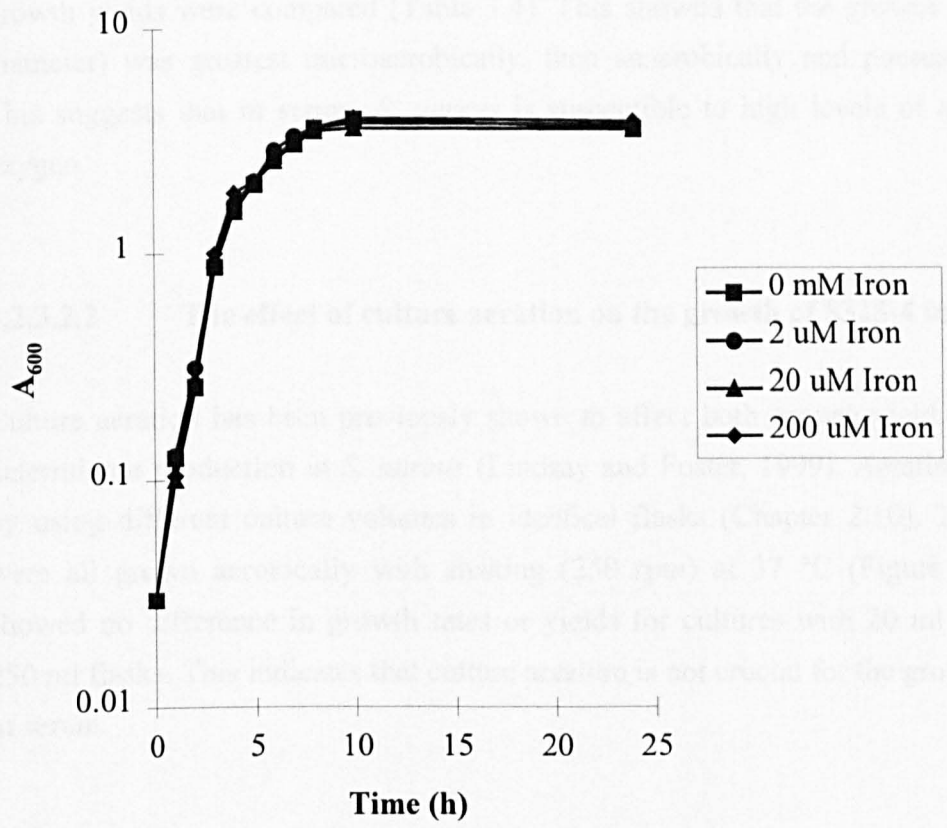


Figure 3.7

The effect of iron on the growth of 8325-4 in liquid serum cultures

S. aureus 8325-4 was grown at 37 °C in serum containing a range of concentrations of iron sulphate (0-200 μM). Cultures were incubated aerobically with shaking (250 rpm) and growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

3.2.3.2.1 The effect of oxygen availability on the growth of 8325-4 in serum

The simplest way to look at oxygen sensitivity of *S. aureus* grown in serum was to grow plate cultures under different environmental oxygen concentration conditions. As *S. aureus* is a facultative aerobe it can be grown both aerobically and anaerobically. To investigate oxygen sensitivity in serum, serum agar plates inoculated with 8325-4, were incubated aerobically, microaerobically and anaerobically (Chapter 2.9) at 37 °C, and growth yields were compared (Table 3.4). This showed that the growth yield (colony diameter) was greatest microaerobically, then anaerobically and poorest aerobically. This suggests that in serum, *S. aureus* is susceptible to high levels of environmental oxygen.

3.2.3.2.2 The effect of culture aeration on the growth of 8325-4 in serum

Culture aeration has been previously shown to affect both growth yield and virulence determinant production in *S. aureus* (Lindsay and Foster, 1999). Aeration, was varied by using different culture volumes in identical flasks (Chapter 2.10). These cultures were all grown aerobically with shaking (250 rpm) at 37 °C (Figure 3.8). Results showed no difference in growth rates or yields for cultures with 20 ml or 100 ml in 250 ml flasks. This indicates that culture aeration is not crucial for the growth of 8325-4 in serum.

3.2.3.3 The effect of the addition of catalase on the growth of 8325-4 in serum

Certain supplements added to liquid or solid media have been reported to increase the aerotolerance of bacteria (reviewed by Krieg and Hoffman, 1986). The only common link among these diverse agents is their ability to quench toxic forms of oxygen. These include catalase, which destroys hydrogen peroxide (H_2O_2). Sterile additions of catalase (Chapter 2.3) were added to serum at a range of concentrations (0-1000 $\mu g\ ml^{-1}$). Experiments were carried out using both solid and liquid media and growth yield was determined by colony diameter (solid media) and A_{600} (liquid cultures).

Condition		Average colony diameter (mm)
Aerobic	(21 % O_2)	< 0.1
Microaerobic	(10 % O_2)	1.25
Anaerobic	(0 % O_2)	0.75

Table 3.4

The effect oxygen availability on the growth of 8325-4 on serum

S. aureus 8325-4 was grown on serum agar plates under aerobic, microaerobic and anaerobic conditions at 37 °C. Growth yield was scored according to colony size (diameter) after 48 h. Average diameters were calculated after measuring at least three colonies from two separate experiments.

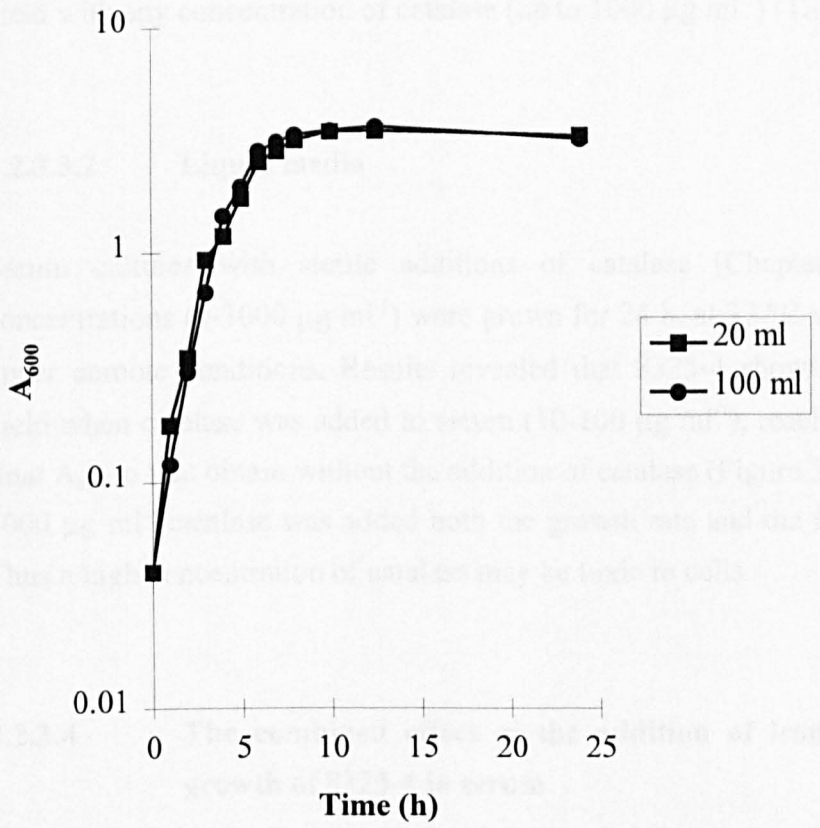


Figure 3.8
The effect of culture aeration on the growth of 8325-4 in serum

S. aureus 8325-4 was grown at 37 °C in serum in 20, or 100 ml cultures in 250 ml flasks. Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

3.2.3.3.1 Solid media

Serum plates containing the appropriate concentration of catalase (0-1000 μg ml^{-1}) (Chapter 2.1.1.7.2), were inoculated with 8325-4 and incubated both aerobically and microaerobically at 37 °C for 48 h. Results revealed that the addition of catalase (500 and 1000 μg ml^{-1}) increased the growth yield of 8325-4 under aerobic conditions (Table 3.5). Under microaerobic conditions there appeared to be no increase in growth yield with any concentration of catalase (up to 1000 μg ml^{-1}) (Table 3.5).

3.2.3.3.2 Liquid media

Serum cultures with sterile additions of catalase (Chapter 2.3) at a range of concentrations (0-1000 μg ml^{-1}) were grown for 24 h, at 37 °C with shaking (250 rpm), under aerobic conditions. Results revealed that 8325-4 showed no increased growth yield when catalase was added to serum (10-100 μg ml^{-1}), reaching an almost identical final A_{600} to that obtain without the addition of catalase (Figure 3.9). Interestingly, when 1000 μg ml^{-1} catalase was added both the growth rate and the final A_{600} were reduced. Thus a high concentration of catalase may be toxic to cells.

3.2.3.4 The combined effect of the addition of iron and catalase on the growth of 8325-4 in serum

The addition of both iron and catalase to serum plates improved the growth of *S. aureus*, but not beyond the level achieved by the addition of iron alone (results not shown).

3.2.4 Analysis of the role and expression of known components in the serum model

Within our laboratory there exists a large number of insertionally inactivated and transcriptional fusion strains in the 8325-4 background. These are within genes with known function, and are all putative virulence factors, or regulators of putative virulence factors. These strains were grown in the serum model (and in BHI as a control), to ascertain their importance, if any (mutant strains), and any difference in transcriptional regulation (fusion strains).

Concentration of catalase	Average colony diameter (mm)			
	Aerobic 24 h	Aerobic 48 h	Microaerobic 24 h	Microaerobic 48 h
0 μg ml^{-1}	<< 0.1	< 0.1	0.15	1.25
10 μg ml^{-1}	<< 0.1	< 0.1	0.15	1.1
100 μg ml^{-1}	<< 0.1	< 0.1	0.15	1.2
500 μg ml^{-1}	< 0.1	0.1	0.15	1.25
1000 μg ml^{-1}	0.1	0.2	0.25	1.15

Table 3.5

The effect of catalase on the growth of *S. aureus* 8325-4 on serum agar plates

S. aureus 8325-4 was grown at 37 °C, on serum agar plates containing a range of concentrations of catalase (0-1000 μg ml^{-1}). Plates were incubated under aerobic or microaerobic conditions. Growth yield was scored according to colony diameter after both 24 h and 48 h incubation. Average colony diameters were calculated after measuring at least three colonies from two separate experiments.

3.2.4.1 The role of the major global regulators *agr* and *sarA* in the serum model

The accessory gene regulator (*agr*), and the staphylococcal accessory regulator (*sarA*), are important regulators of virulence factor production in *S. aureus* (Chapter 1.7.1-1.7.2). Previous studies revealed these regulators to be important for infection *in vivo* (Abdelnour *et al.*, 1993; Booth *et al.*, 1995), and thus it was postulated that these regulators might be important for growth in the serum model. Insertionally inactivated mutants, of both *agr*, *sarA* and a double *agr*, *sarA* mutant constructed in the isogenic parental background 8325-4 (Chan and Foster, 1998a), were grown in the serum model.

Growth experiments revealed that in both BHI and serum under all conditions, PC6911 (*agr*) and PC18391 (*agr sarA*) strains consistently reach a slightly higher A_{600} than the parental strain (8325-4). PC1839 (*sarA*) however grew at approximately the same rate and gave the same yield as 8325-4 (results not shown). Thus *agr* and *sarA* do not play an important role during growth in the serum model.

3.2.4.2 The role of sigma factor σ^B in the serum model

Under deleterious conditions, such as a lack of nutrients or other environmental stresses, *S. aureus* induce sophisticated response mechanisms, in order to protect against environmental stress and thus allow continued growth. The alternative sigma factor σ^B (*sigB*) has been shown to control the environmental stress response, in *S. aureus*. Therefore it is possible that sigma factor σ^B, may be important for growth in the serum model.

To test this theory, a *sigB* insertionally inactivated mutant (PC400) (Chan *et al.*, 1998b) was grown in serum to determine whether the loss of this stress sigma factor was deleterious to growth. Cultures of 8325-4, along with PC400 (*sigB*) were grown in BHI and serum at 37 °C with shaking (250 rpm), under both aerobic and microaerobic conditions. Cultures were grown for 24 h and growth was monitored by A_{600}. Results revealed that the growth rate and final growth yield of PC400 (*sigB*) and 8325-4 were identical to each other under all four conditions (results not shown). Thus σ^B has no role in the ability of *S. aureus* to grow in the serum model.

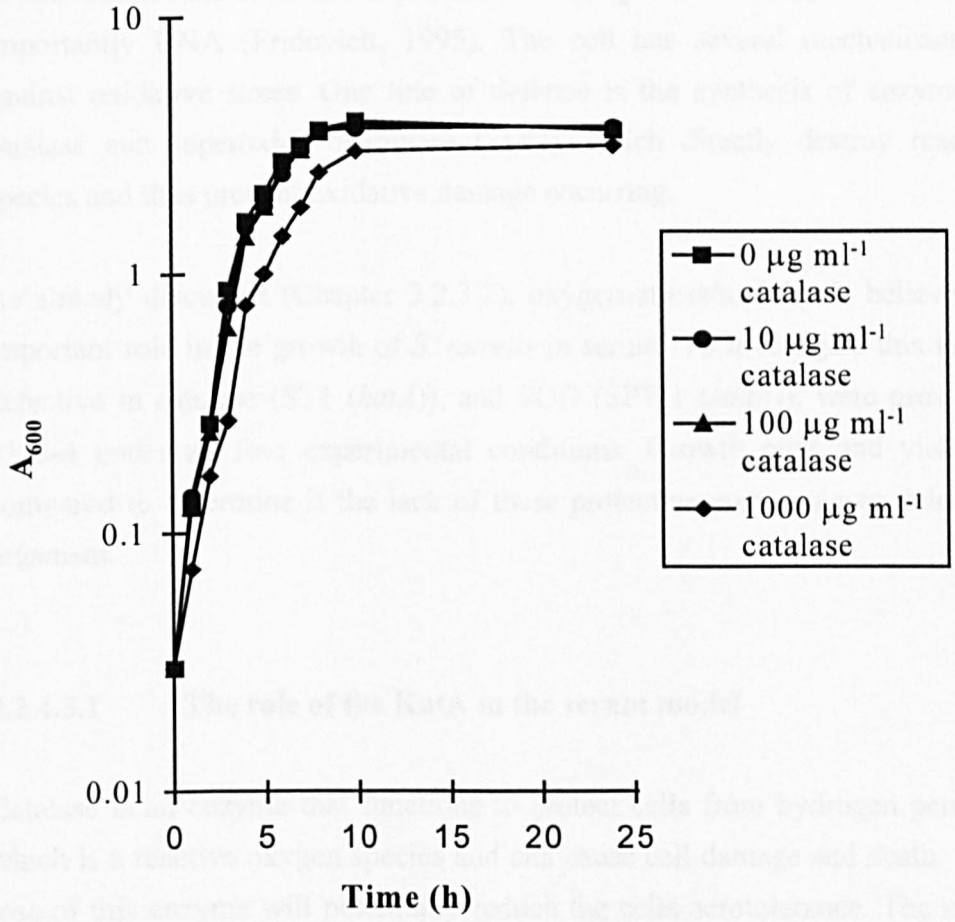


Figure 3.9

The effect of catalase on the growth of 8325-4 in liquid serum cultures

S. aureus 8325-4 was grown at 37 °C in serum containing a range of concentrations of catalase (0-1000 μg ml^{-1}). Cultures were incubated aerobically with shaking (250 rpm) and growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

3.2.4.3 The role of enzymes protecting against toxic forms of oxygen in the serum model

Resistance against oxidative stress is important for the growth and survival of *S. aureus* in all conditions were oxygen is present. Reactive oxygen species cause severe damage to macromolecular cellular components, including membranes, proteins, RNA and most importantly DNA (Fridovich, 1995). The cell has several mechanisms of defence against oxidative stress. One line of defence is the synthesis of enzymes, including catalase and superoxide dismutase (SOD), which directly destroy reactive oxygen species and thus prevent oxidative damage occurring.

As already discussed (Chapter 3.2.3.2), oxygen stress/toxicity is believed to play an important role in the growth of *S. aureus* in serum. To investigate this further, strains defective in catalase (ST1 (*katA*)), and SOD (SPW1 (*sodA*)), were grown along with 8325-4 under all four experimental conditions. Growth rates and yields were then compared to determine if the lack of these protective enzymes was deleterious to the organism.

3.2.4.3.1 The role of the KatA in the serum model

Catalase is an enzyme that functions to protect cells from hydrogen peroxide (H_2O_2), which is a reactive oxygen species and can cause cell damage and death. Therefore the loss of this enzyme will potentially reduce the cells aerotolerance. The major catalase (KatA) of *S. aureus* has recently been identified as a starvation-survival mutant and furthermore has been shown to be important in a mouse abscess model of pathogenicity (Horsburgh and Foster, unpublished data). To investigate the role of KatA in the ability of *S. aureus* to grow in serum, the growth kinetics of ST1 (*katA*) were compared to the parent 8325-4. Strains were grown in BHI or serum at 37 °C with shaking (250 rpm), under both aerobic and microaerobic conditions.

The growth rate and final yield of ST1 (*katA*) and 8325-4 were identical to each other in BHI under all conditions (Figure 3.10). In contrast, the results in serum showed that while the growth yield is not significantly altered, there is a clear defect in the growth kinetics of ST1 (*katA*) compared to 8325-4 (Figure 3.11a and 3.11b). This is true for both aerobic (Figure 3.11a), and microaerobic conditions (Figure 3.11b). The growth defect is greater under aerobic than microaerobic conditions. This result is therefore in

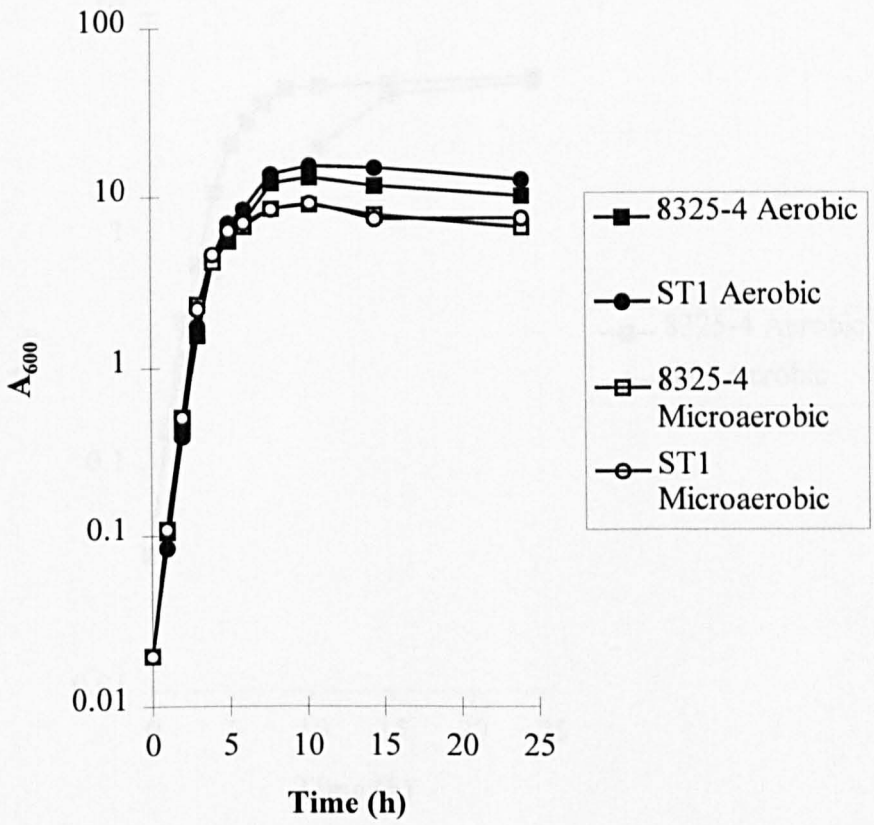


Figure 3.10

The role of the KatA in the growth of *S. aureus* in BHI

S. aureus 8325-4 and ST1 (*katA*) were grown at 37 °C in BHI under aerobic and microaerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

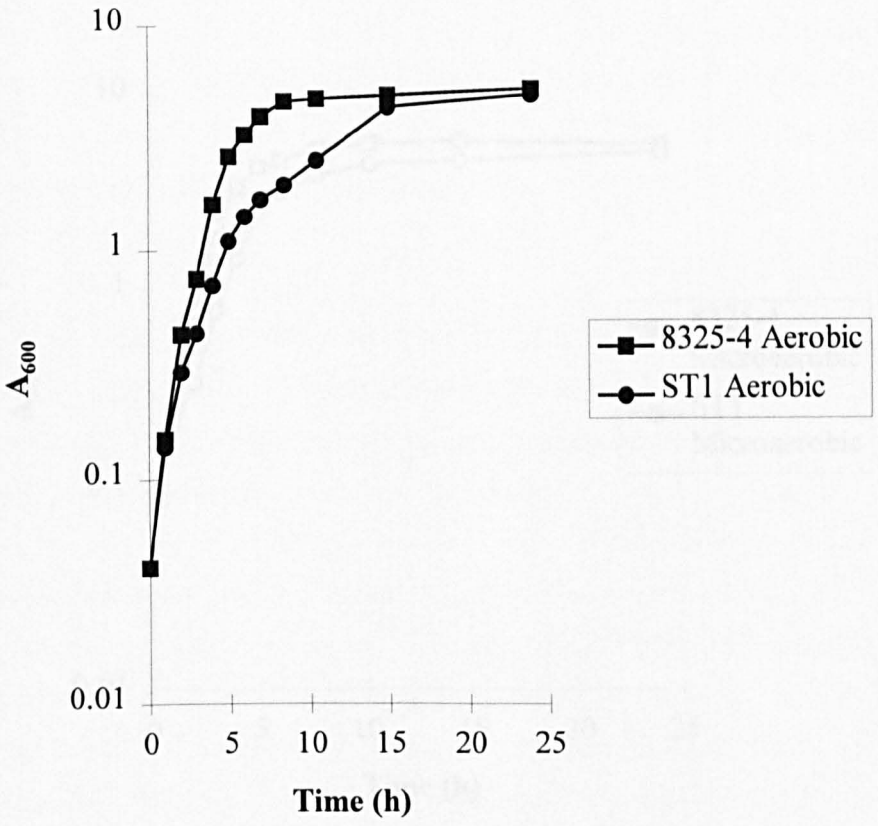


Figure 3.11a

The role of the KatA in the growth of *S. aureus* in serum under aerobic conditions

S. aureus 8325-4 and ST1 (*katA*) were grown at 37 °C in serum under aerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

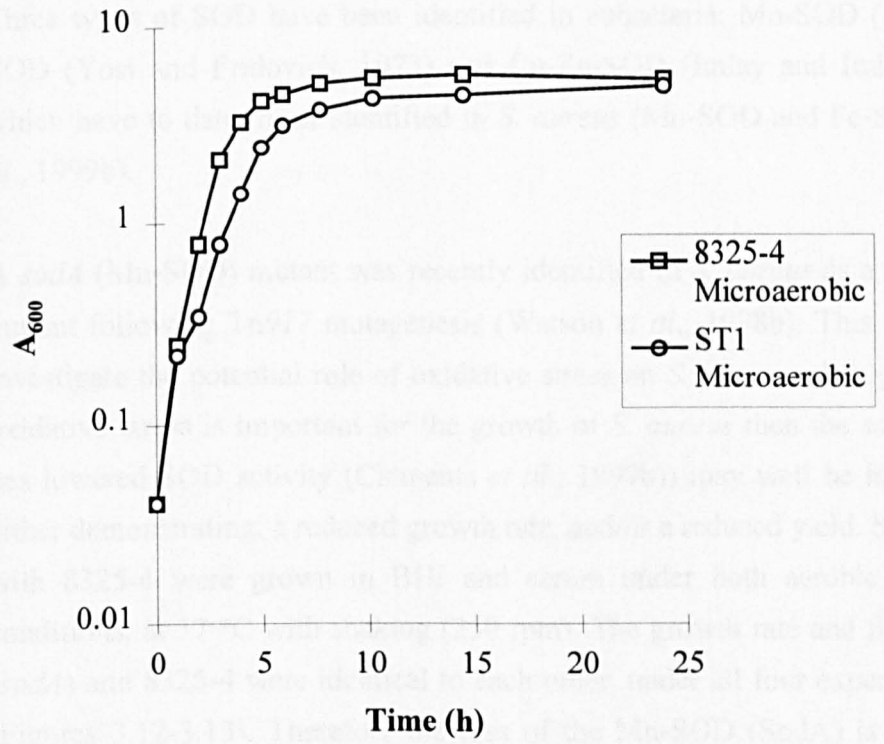


Figure 3.11b

The role of the KatA in the growth of *S. aureus* in serum under microaerobic conditions

S. aureus 8325-4 and ST1 (*katA*) were grown at 37 °C in serum under microaerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

agreement with the postulation that oxidative stress is lower under reduced O_2 levels (microaerobic conditions).

3.2.4.3.2 The role of SodA in the serum model

SODs are a class of enzyme that function to protect cells from superoxide radicals (O_2^-). Thus loss of these enzymes will potentially reduce the aerotolerance of bacterial cells. Three types of SOD have been identified in eubacteria: Mn-SOD (Touati, 1988), Fe-SOD (Yost and Fridovich, 1973) and Cu-Zn-SOD (Imlay and Imlay, 1996), two of which have to date, been identified in *S. aureus* (Mn-SOD and Fe-SOD) (Clements *et al.*, 1999b).

A *sodA* (Mn-SOD) mutant was recently identified in *S. aureus* as a starvation-survival mutant following Tn*917* mutagenesis (Watson *et al.*, 1998b). This strain was used to investigate the potential role of oxidative stress on *S. aureus* when grown in serum. If oxidative stress is important for the growth of *S. aureus* then the *sodA* mutant (which has lowered SOD activity (Clements *et al.*, 1999b)) may well be impaired in growth, either demonstrating, a reduced growth rate, and/or a reduced yield. SPW1 (*sodA*) along with 8325-4 were grown in BHI and serum under both aerobic and microaerobic conditions, at 37 °C with shaking (250 rpm). The growth rate and final yield of SPW1 (*sodA*) and 8325-4 were identical to each other, under all four experimental conditions (Figures 3.12-3.13). Therefore the loss of the Mn-SOD (SodA) is not detrimental to growth in serum.

3.2.4.4 The role of CtaA in the serum model

A Tn*917* insertion mutant with a defect in starvation-survival and recovery was recently identified (Watson *et al.*, 1998b). Sequencing and analysis of chromosomal DNA flanking the Tn*917* insert revealed that the transposon had inserted in a gene with 39.7 % identity to *ctaA* of *B. subtilis* (Clements *et al.*, 1999a). *ctaA* encodes a haem O monoxygenase which converts haem O to haem A (Svensson *et al.*, 1993). SPW3 (*ctaA*) had a major defect in its ability to recover following starvation. However, this defect was partially alleviated, by the addition of catalase to the recovery medium, indicating a probable involvement of oxidative stress (Clements *et al.*, 1999a). To investigate the role of CtaA in the ability of *S. aureus* to grow in serum, the growth kinetics of SPW3 (*ctaA*) were compared to the parental strain 8325-4. Strains were

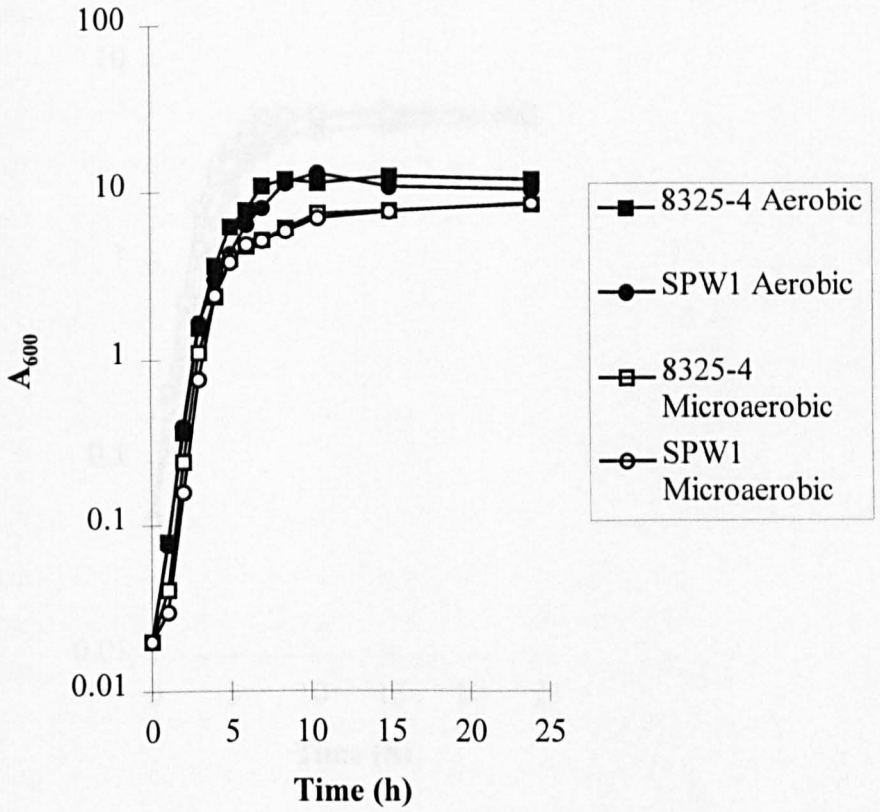


Figure 3.12

The role of SodA in the growth of *S. aureus* in BHI

S. aureus 8325-4 and SPW1 (*sodA*) were grown at 37 °C in BHI under aerobic and microaerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

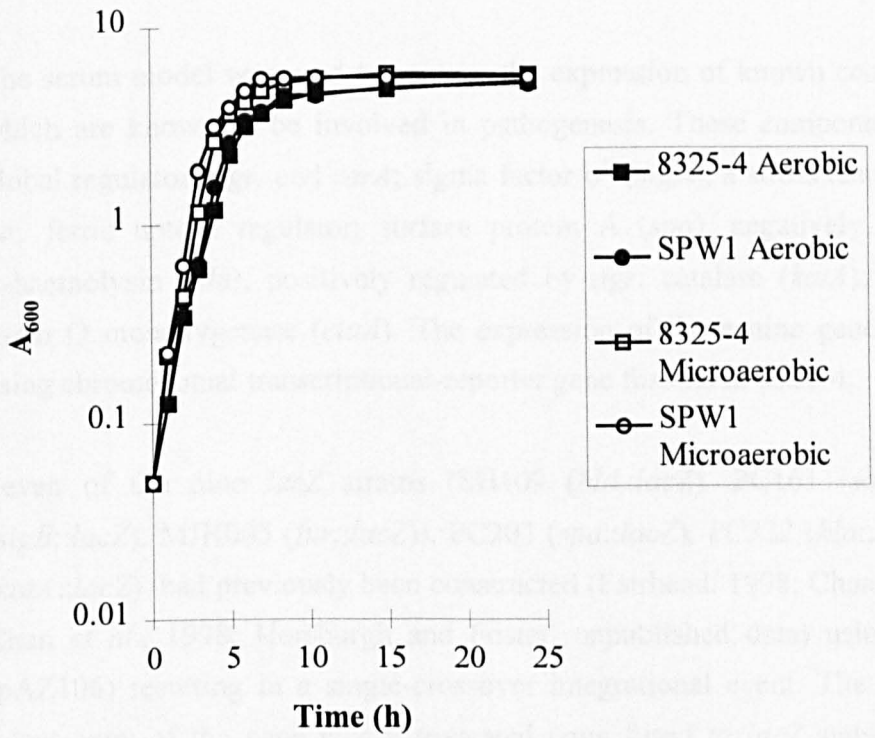


Figure 3.13

The role of SodA in the growth of *S. aureus* in serum

S. aureus 8325-4 and SPW1 (*sodA*) were grown at 37 °C in serum under aerobic and microaerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 24 h. The results are representative of at least two separate experiments.

grown in BHI and serum under both aerobic and microaerobic conditions, at 37 °C with shaking (250 rpm). The growth rate and final yield of SPW3 (*ctaA*) and 8325-4 were identical to each other under all four experimental conditions (results not shown). Therefore the loss of haem A synthase is not detrimental to growth in serum. However, when SPW3 (*ctaA*) is plated onto serum agar plates, small colony variants are observed, as previously reported for other types of solid media (Clements *et al.*, 1999a).

3.2.5 Expression of known components in the serum model

The serum model was used to analyze the expression of known components, many of which are known to be involved in pathogenesis. These components are: the major global regulators *agr,* and *sarA*; sigma factor σ^B (*sigB*), a stress response sigma factor; *fur,* ferric uptake regulator; surface protein A (*spa*), negatively regulated by *agr*; α-haemolysin (*hla*), positively regulated by *agr*; catalase (*katA*); SOD (*sodA*); and haem O monoxygenase (*ctaA*). The expression of these nine genes was investigated using chromosomal transcriptional-reporter gene fusions in 8325-4.

Seven of the nine *lacZ* strains (SH109 (*hld*::*lacZ*), PC161 (*sarA*::*lacZ*), MC100 (*sigB*::*lacZ*), MJH005 (*fur*::*lacZ*)), PC203 (*spa*::*lacZ*), PC322 (*hla*::*lacZ*) and MJH006 (*katA*::*lacZ*) had previously been constructed (Fairhead, 1998; Chan and Foster, 1998a; Chan *et al.*, 1998; Horsburgh and Foster, unpublished data) using a suicide vector (pAZ106) resulting in a single-crossover integrational event. The fusions contain an intact copy of the gene plus a truncated copy fused to *lacZ* stably inserted into the chromosome of 8325-4. However, SPW1 (*sodA*::*lacZ*) and SPW3 (*ctaA*::*lacZ*) were created by Tn*917* insertion (Watson *et al.*, 1998b), and thus do not contain intact copies of the fused genes. The expression of these nine known components was investigated in BHI and serum, under both aerobic and microaerobic conditions. BHI is a nutrient-rich growth medium and thus differences in expression pattern between BHI and serum, represent differential regulation.

3.2.5.1 Expression of *lacZ* transcriptional fusions on solid media

Initially all *lacZ* fusion strains along with 8325-4 were grown on BHI and serum agar plates containing X-Gal (Chapter 2.11.1). X-Gal is hydrolysed by β-galactosidase, resulting in β-D-galactopyranoside and 5-bromo-4-chloro-3-indolyl. 5-bromo-4-chloro-3-indolyl has a blue colouration and is thus a visible indicator of β-galactosidase.

LacZ Fusion Strain	BHI Aero 24h	BHI Aero 48h	Sera Aero 24h	Sera Aero 48h	BHI Micro 24h	BHI Micro 48h	Sera Micro 24h	Sera Micro 48h
SH109 *hld::lacZ*	++++	++++	++	+++	++++	++++	+++	+++
PC161 *sarA::lacZ*	+++	+++	++	+++	+++	+++	+++	+++
MC100 *sigB::lacZ*	+++	+++	++	+++	+++	+++	++	+++
MJH005 *fur::lacZ*	+++	+++	++	+++	+++	+++	+++	+++
PC203 *spa::lacZ*	-	-	-	-	-	-	-	-
PC322 *hla::lacZ*	++++	++++	++	+++	++++	++++	+++	+++
MJH006 *katA::lacZ*	+++	+++	++	+++	+++	+++	++	+++
SPW1 *sodA::lacZ*	-	-	-	-	-	-	-	-
SPW3 *ctaA::lacZ*	+++	+++	+	++	+++	+++	+	++
8325-4	-	-	-	-	-	-	-	-

Table 3.6

Expression of *lacZ* transcriptional fusions of known components on solid media (BHI and serum agar plates) under aerobic and microaerobic conditions

lacZ strains were grown on BHI and serum agar plates containing X-Gal, (a visible indicator of β-galactosidase activity) under both aerobic and microaerobic conditions. Plates were scored for blue colouration after 24 and 48 h. A scale to score colouration was designed: dark blue ++++; blue +++; light blue ++; very light blue +; white -. The results are representative of at least two separate experiments.

Strains were incubated at 37 °C under both aerobic and microaerobic conditions. Plates were scored for blue colouration after 24 and 48 h, and the results obtained are summarised in Table 3.6. All *lacZ* fusion strains except *spa* and *sodA* had a blue colouration on both BHI and serum agar, under both aerobic and microaerobic conditions after 48 h. As neither *spa* nor *sodA* produced blue colonies under any conditions, X-Gal is not sufficiently sensitive to study expression of these genes. Strain 8325-4, which was used as a control, was not blue on X-Gal under any conditions. The results (Table 3.6) indicate that there is little difference in transcriptional control between growth aerobically or microaerobically, or on BHI/serum. However, from this work there are two apparent differences in the expression of the known components under the different conditions. The first is that both *agr* and *hla* have a greater blue coloration (greater expression) on BHI than serum. The second is that *agr, sarA, fur* and *hla* appear to be expressed to a greater level on serum microaerobically after 24 h compared with the level of expression aerobically.

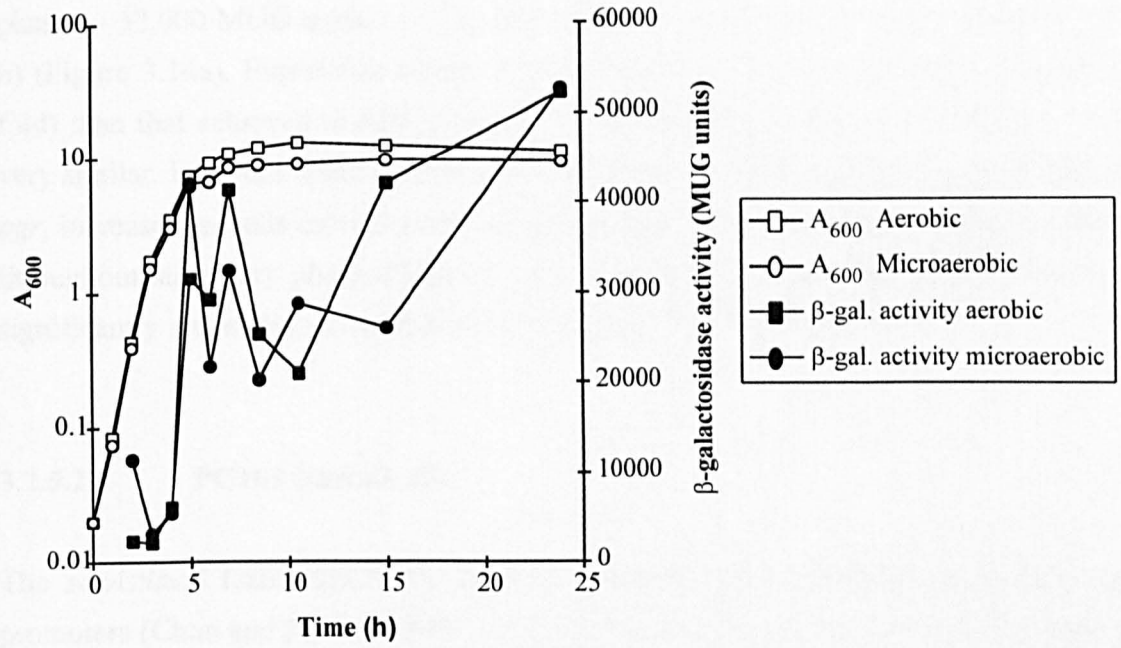
3.2.5.2 Expression of *lacZ* transcriptional fusions in liquid media

The plate assay is very insensitive and does not allow temporal expression to be measured. The use of liquid cultures with the fluorogenic substrate MUG (Chapter 2.11.2) is sensitive, quantitative and allows temporal expression to be measured. The expression of each fusion strain was measured during growth in BHI and serum, under both aerobic and microaerobic conditions. Previous studies (Chan and Foster, 1998b; Lindsay and Foster, 1999) have shown that slight alterations in growth conditions have a significant effect on the expression of virulence determinants. Hence, specific growth conditions defined in Chapter 2.4.1.1 were strictly adhered to. All fusion strains showed identical growth kinetics to each other and to the parental strain, 8325-4 (results not shown). Strain 8325-4 was used as a control in all the following experiments. The background level measured from 8325-4 samples (control samples), were deducted from the β-galactosidase activity values from the *lacZ* fusion strains.

3.2.5.2.1 SH109 (*hld::lacZ*)

The SH109 (*hld::lacZ*) fusion strain had been previously constructed (Fairhead, 1998). Production of β-galactosidase in this strain, allowed the convenient assay of *hld* and hence RNAIII expression, and therefore allows the expression of *agr* to be monitored during growth in BHI and serum under both aerobic and microaerobic conditions.

a)



b)

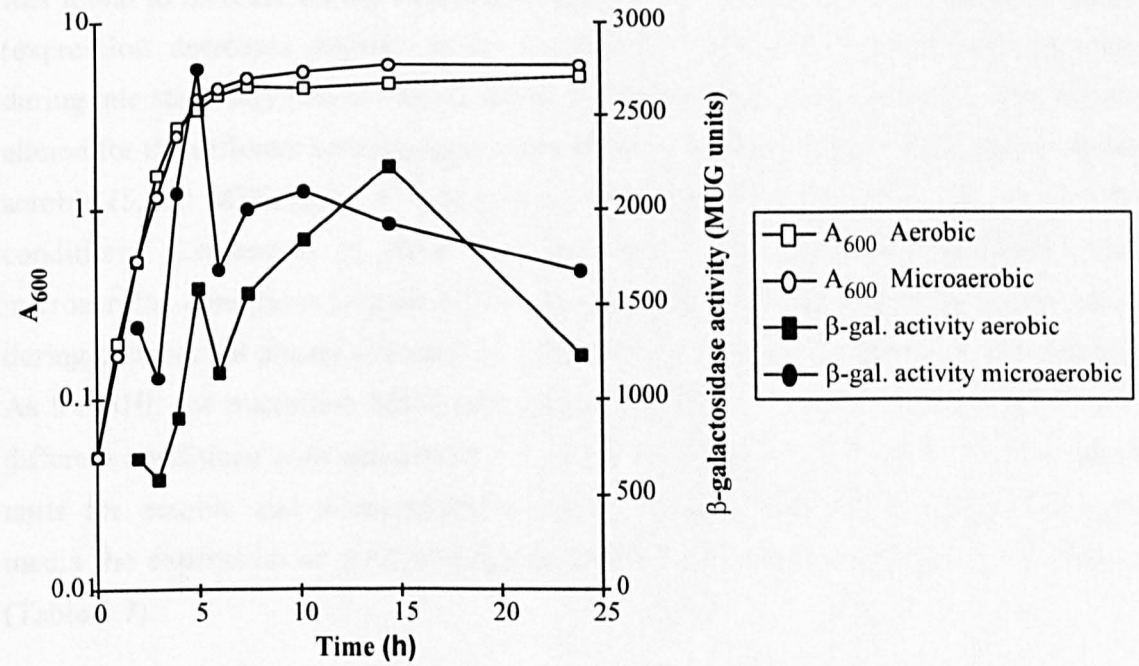


Figure 3.14

Expression of *hld*::*lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus SH109 (*hld*::*lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

Expression in BHI was shown to increase dramatically as cells entered post-exponential phase (~ 52,000 MUG units; t = 5 h) achieving a maximum in stationary phase (t = 24 h) (Figure 3.14a). Expression of *agr* in serum however was considerably lower (>15-fold) than that achieved in BHI, (Figure 3.14b) although the expression kinetics, were very similar. In serum under both aerobic and microaerobic conditions the expression of *agr*, increased as cells entered post-exponential phase (t = 5 h), where it remained high throughout stationary phase (Figure 3.14b). Expression of *agr* in both media was not significantly altered by growth under the different conditions (Table 3.7).

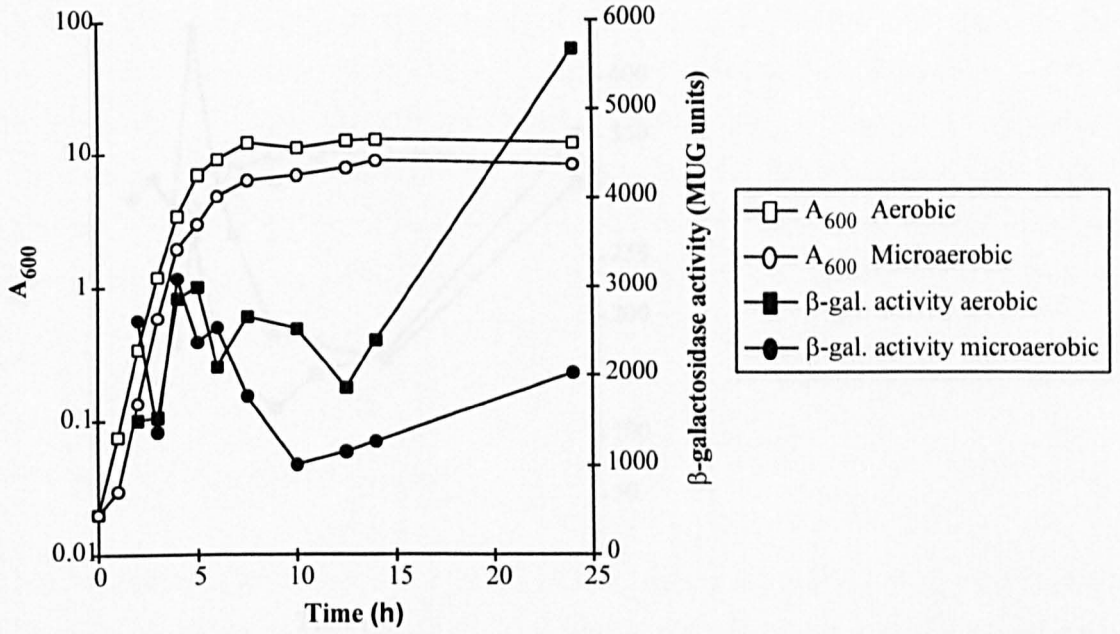
3.2.5.2.2 PC161 (*sar*::*lacZ*)

The *sarA*::*lacZ* fusion (PC161), measures transcription initiated from all three *sar* promoters (Chan and Foster, 1998a). The pattern of *sarA* expression was determined to be similar in BHI under both environmental conditions (Figure 3.15a). *sarA* expression was found to increase during exponential phase, level off during early stationary phase (expression decreases slightly under microaerobic conditions) and finally increase during late stationary phase. The maximum MUG unit expression values however were altered for the different aeration conditions with expression being ~ 2-fold greater under aerobic (5,800 MUG units; t = 24 h) than microaerobic (3,000 MUG units; t = 4 h) conditions. Expression of *sarA* in serum was also similar under aerobic and microaerobic conditions (Figure 3.15b). The general trend was constitutive expression during exponential phases followed by an increase in relative levels in stationary phase. As for BHI, the maximum MUG unit expression values in serum were altered for the different conditions with maxima of ~ 2,300 (t = 12.5 h) and 1,700 (t = 12.5 h) MUG units for aerobic and microaerobic conditions, respectively. Interestingly, for both media the expression of *sarA* was higher under aerobic than microaerobic conditions (Table 3.7).

3.2.5.2.3 MC100 (*sigB*::*lacZ*)

Growth of *S. aureus* MC100 allows the expression of *sigB* to be determined (Chan *et al.*, 1998). The expression of *sigB* when grown in BHI was found to be very similar under both aerobic and microaerobic conditions, exhibiting comparable patterns of expression and similar levels of gene expression (β-galactosidase production) (Figure 3.16a). Under both aerobic and microaerobic conditions, *sigB* showed a peak of expression as cells entered the post-exponential phase, following this, there was a

a)



b)

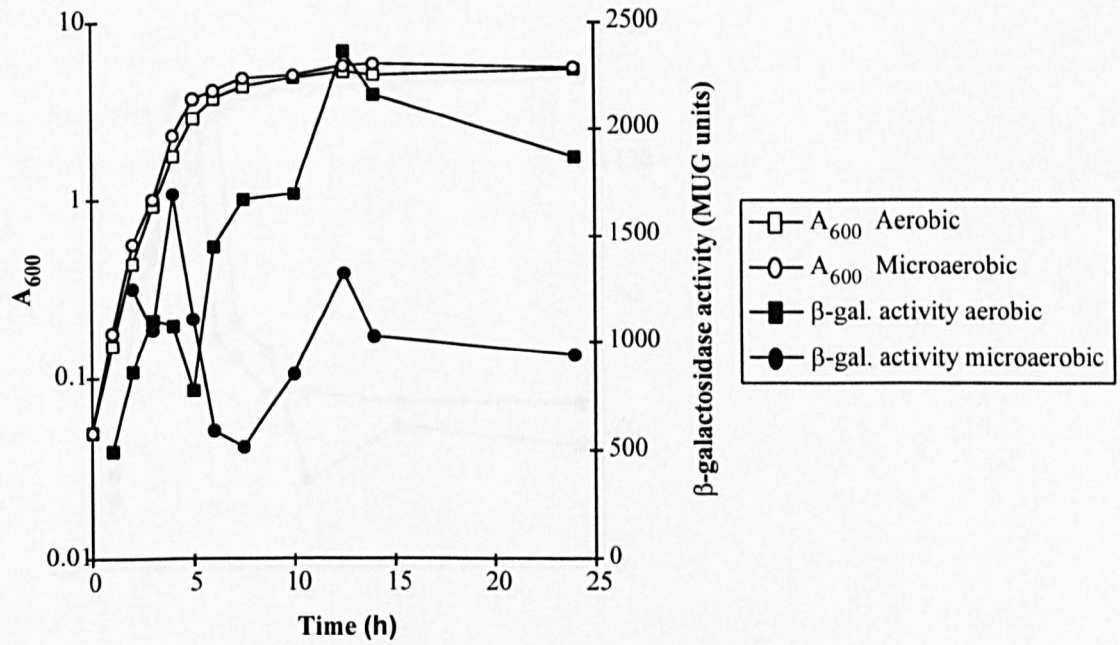
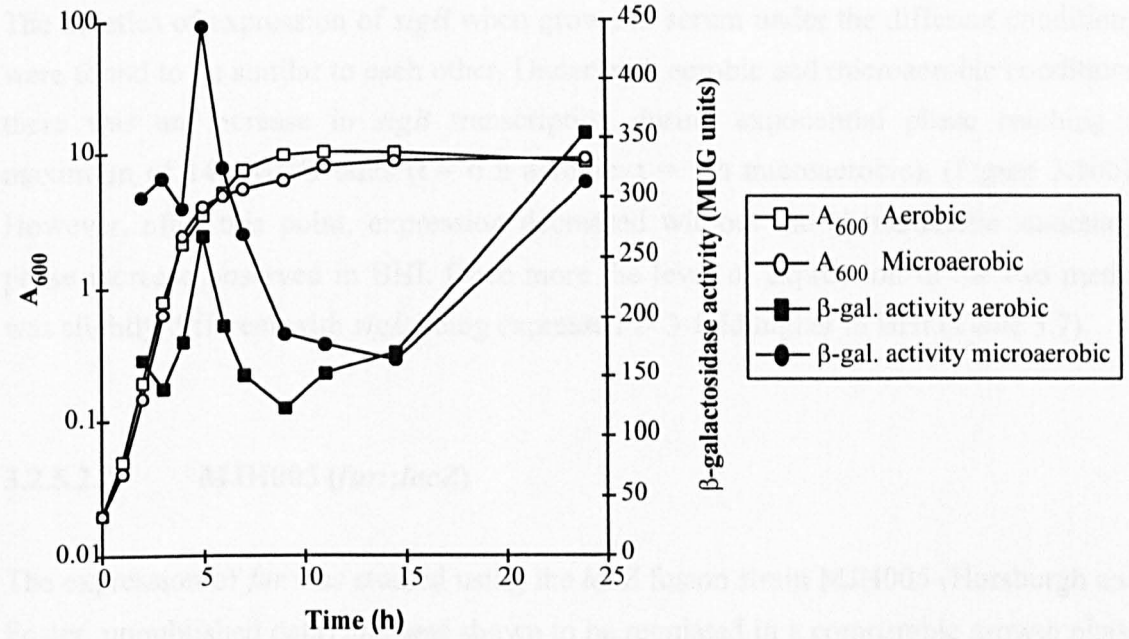


Figure 3.15

Expression of *sarA::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus PC161 (*sar::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

a)



b)

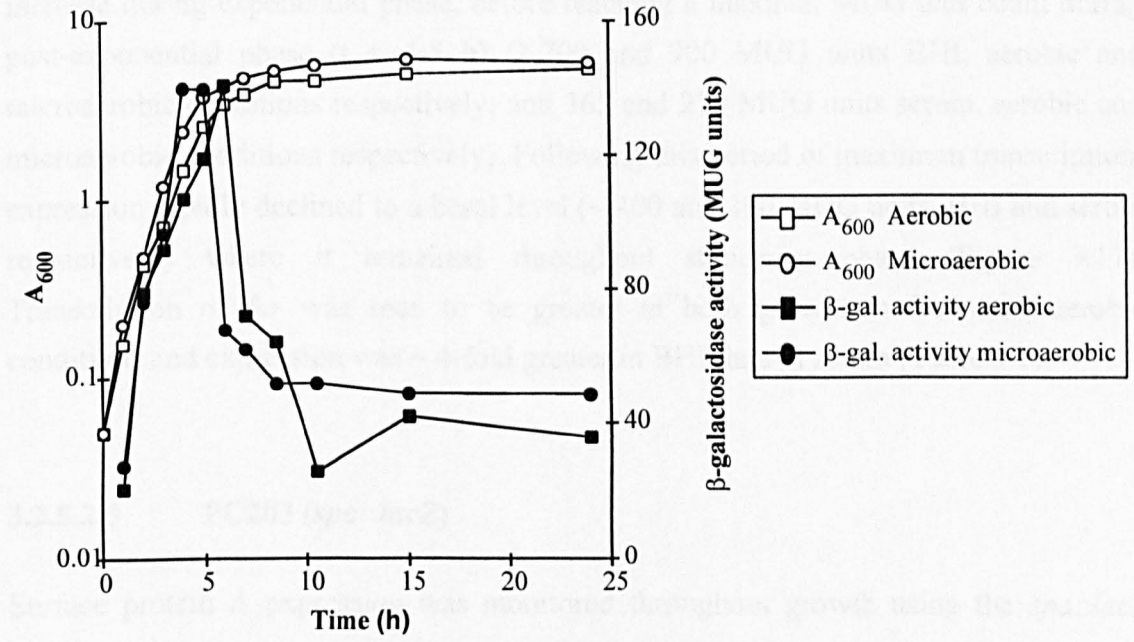


Figure 3.16

Expression of *sigB::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus MC100 (*sigB::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

reduction before a further increase in expression during stationary phase (~ t = 15 h). The kinetics of expression of *sigB* when grown in serum under the different conditions were found to be similar to each other. Under both aerobic and microaerobic conditions there was an increase in *sigB* transcription during exponential phase reaching a maximum of 140 MUG units (t = 6 h aerobic, t = 5 h microaerobic), (Figure 3.16b). However, after this point, expression decreased without the characteristic stationary phase increase observed in BHI. Once more the level of expression in the two media was slightly different with *sigB* being expressed 2- 3-fold higher in BHI (Table 3.7).

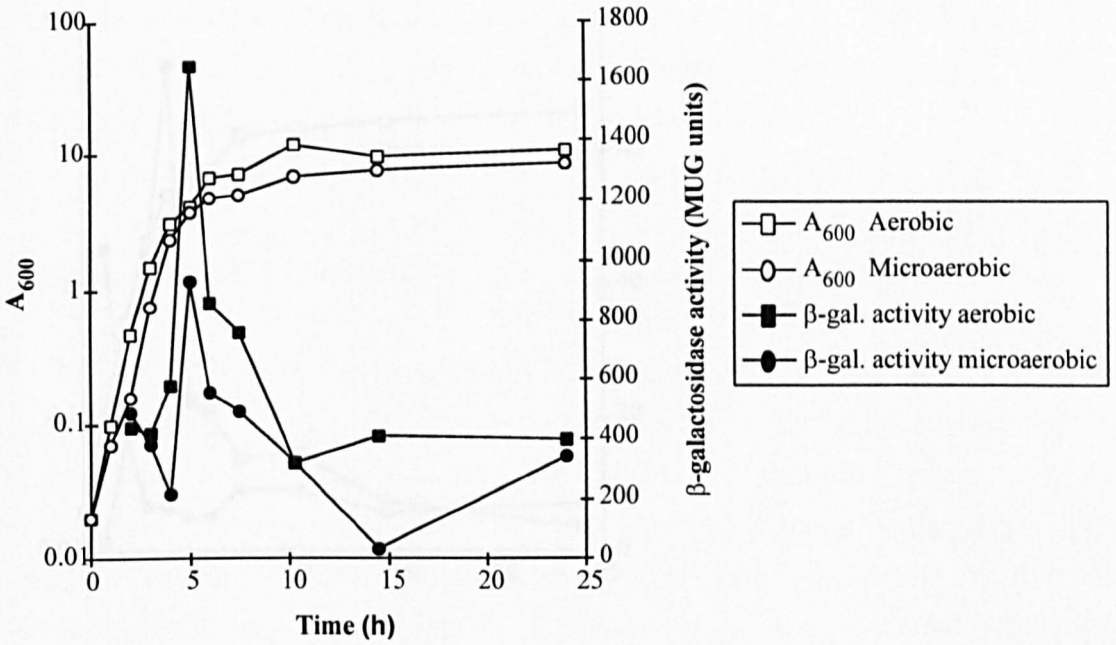
3.2.5.2.4 MJH005 (*fur*::*lacZ*)

The expression of *fur* was studied using the *lacZ* fusion strain MJH005 (Horsburgh and Foster, unpublished data) and was shown to be regulated in a comparable growth phase dependent manner, under all four experimental conditions. *fur* expression was shown to increase during exponential phase, before reaching a maximal MUG unit count during post-exponential phase (t = 4-5 h) (1,700 and 900 MUG units BHI, aerobic and microaerobic conditions respectively; and 365 and 270 MUG units serum, aerobic and microaerobic conditions respectively). Following this period of maximum transcription, expression rapidly declined to a basal level (~ 400 and 100 MUG units, BHI and serum respectively) where it remained throughout stationary phase (Figure 3.17). Transcription of *fur* was seen to be greater in both growth media under aerobic conditions and expression was ~ 4-fold greater in BHI than in serum (Table 3.7).

3.2.5.2.5 PC203 (*spa*::*lacZ*)

Surface protein A expression was monitored throughout growth using the *spa*::*lacZ* strain PC203 (Chan and Foster, 1998b). In both BHI and serum under aerobic conditions, levels increased during exponential phase, achieving maximal expression of 75 and 210 MUG units for BHI and serum respectively, during late exponential phase (t = 4 h). Following this period of maximal expression, the expression decreased dramatically as cells entered stationary phase, (< 5 and < 20 MUG units for BHI and serum respectively) (Figure 3.18). *spa* expression in serum followed exactly the same trend microaerobically as it did aerobically, with maximal expression occurring in late exponential phase (t = 4 h) followed by a decline on entry to stationary phase (Figure 3.18b). Maximum expression was however only ~ 15 % (30 MUG units) of that

a)



b)

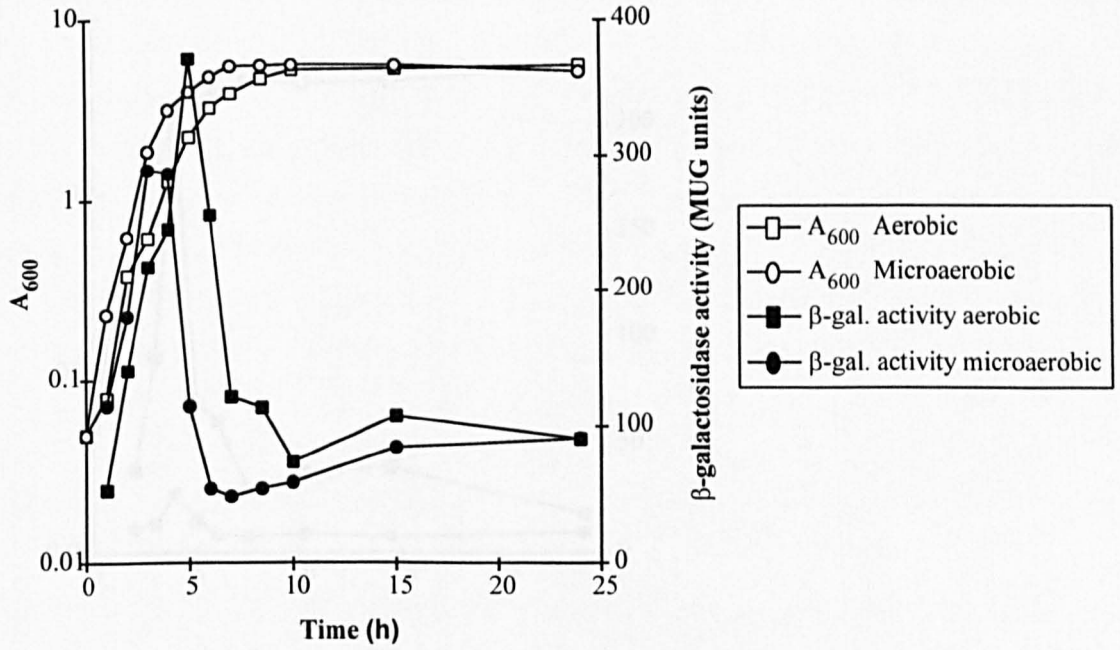
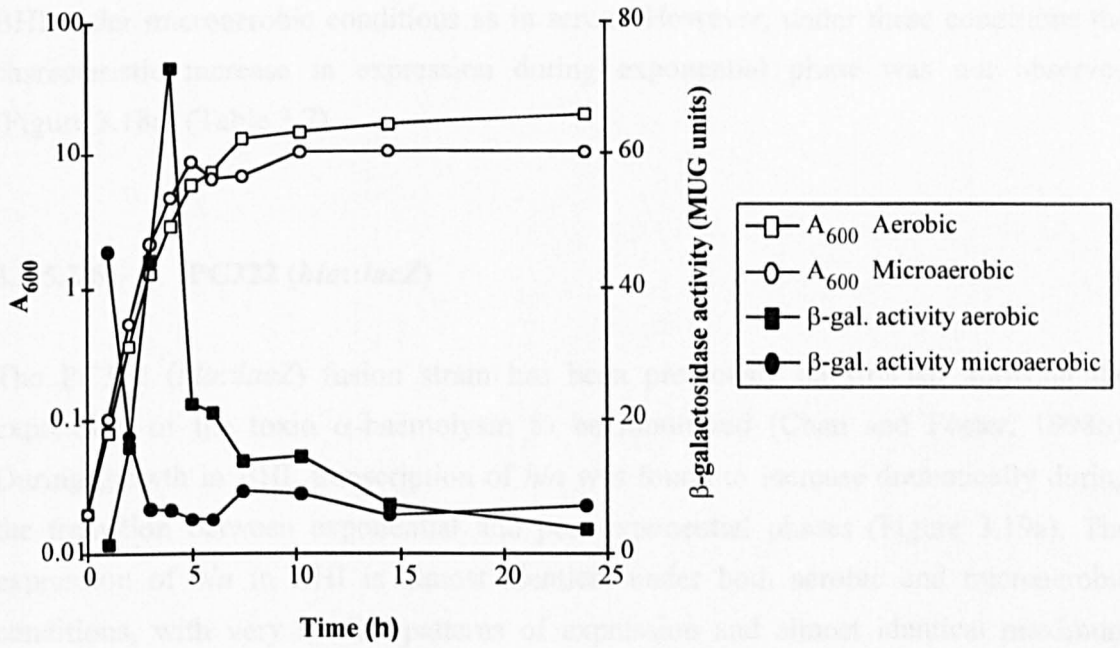


Figure 3.17

Expression of *fur::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus MJH005 (*fur::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

a)



b)

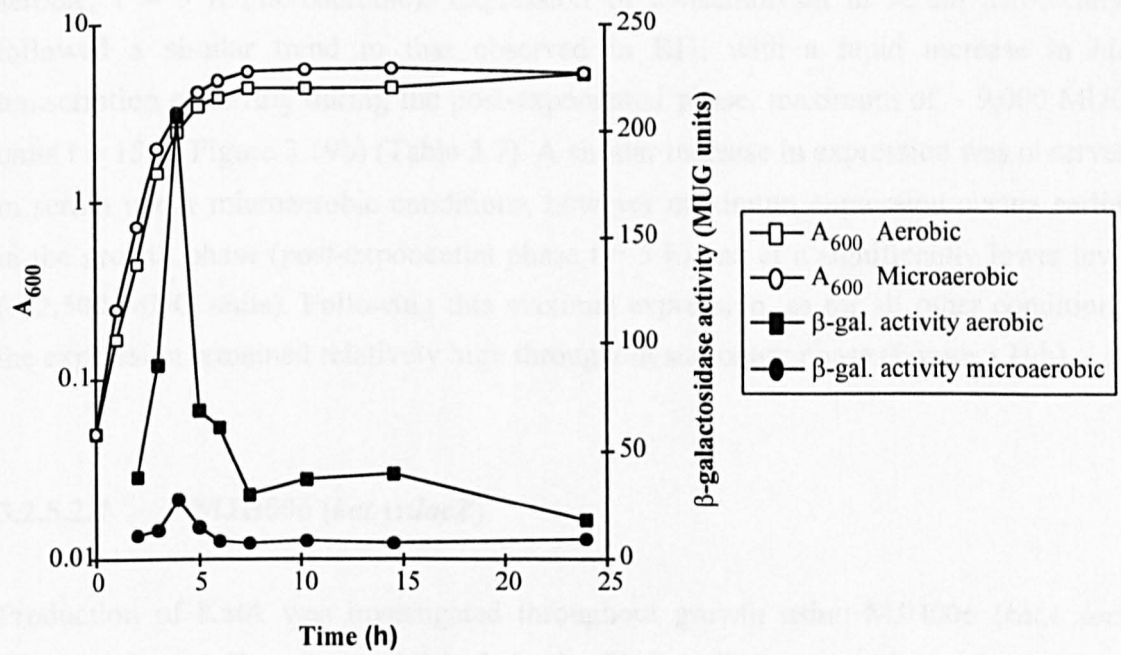


Figure 3.18

Expression of *spa::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus PC203 (*spa::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

obtained under aerobic conditions. The same reduction in expression was witnessed in BHI under microaerobic conditions as in serum. However, under these conditions the characteristic increase in expression during exponential phase was not observed (Figure 3.18a) (Table 3.7).

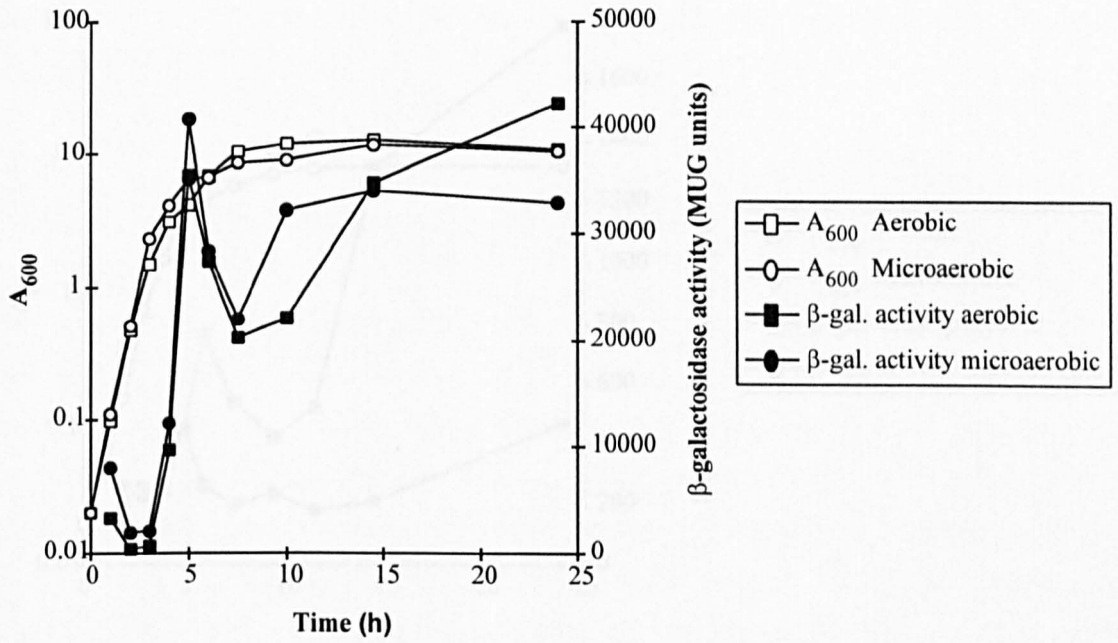
3.2.5.2.6 PC322 (*hla*::*lacZ*)

The PC322 (*hla*::*lacZ*) fusion strain has been previously constructed allowing the expression of the toxin α-haemolysin to be monitored (Chan and Foster, 1998b). During growth in BHI, transcription of *hla* was found to increase dramatically during the transition between exponential and post-exponential phases (Figure 3.19a). The expression of *hla* in BHI is almost identical under both aerobic and microaerobic conditions, with very similar patterns of expression and almost identical maximum levels of gene expression (β-galactosidase activity) of ~ 42,000 MUG units (t = 24 h aerobic, t = 5 h microaerobic). Expression of α-haemolysin in serum aerobically, followed a similar trend to that observed in BHI with a rapid increase in *hla* transcription occurring during the post-exponential phase, maximum of ~ 9,000 MUG units t = 15 h (Figure 3.19b) (Table 3.7). A similar increase in expression was observed in serum under microaerobic conditions, however maximum expression occurs earlier in the growth phase (post-exponential phase t = 5 h) and at a significantly lower level (~ 2,500 MUG units). Following this maximal expression, as for all other conditions, the expression remained relatively high throughout stationary phase (Figure 3.19b).

3.2.5.2.7 MJH006 (*katA*::*lacZ*)

Production of KatA was investigated throughout growth using MJH006 (*katA*::*lacZ* (Horsburgh and Foster, unpublished data)). Under all four experimental conditions studies revealed that maximal expression occurred during stationary phase. Expression of *katA* is significantly higher in both growth media under aerobic than microaerobic conditions (> 4-fold, BHI and > 2.5-fold, serum) and higher in BHI than serum (Figure 3.20; Table 3.7). Furthermore, the kinetics of expression of *katA*, were almost identical within the same media. Expression in BHI showed an increase as cells entered post-exponential phase and then a second increase (a large increase aerobically) in stationary phase, reaching maximal MUG unit values of 1,800 and 475 when grown aerobically and microaerobically respectively (Figure 3.20a). In serum the first increase in expression seen in BHI, as cells entered post-exponential phase was not apparent but

a)



b)

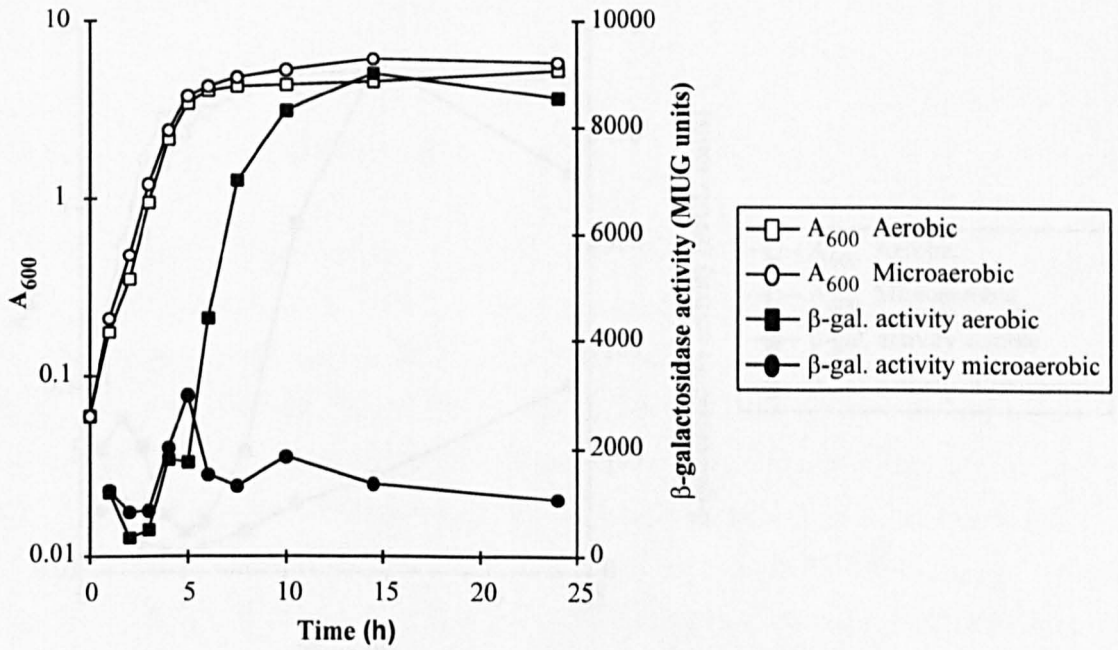
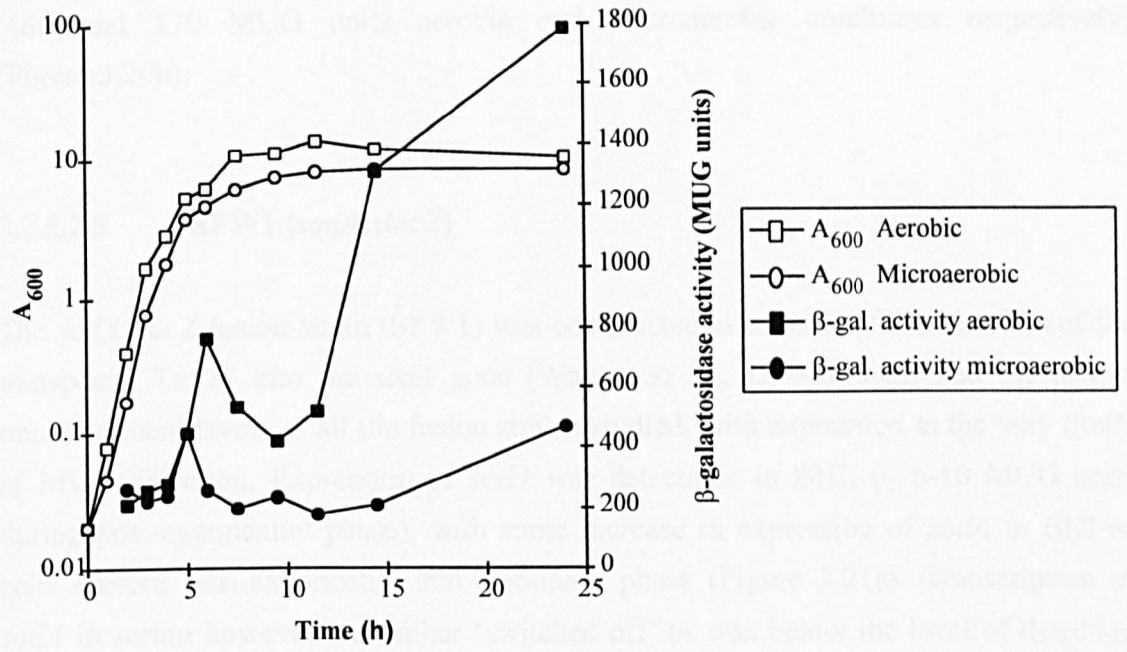


Figure 3.19

Expression of *hla::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus PC322 (*hla::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

a)



b)

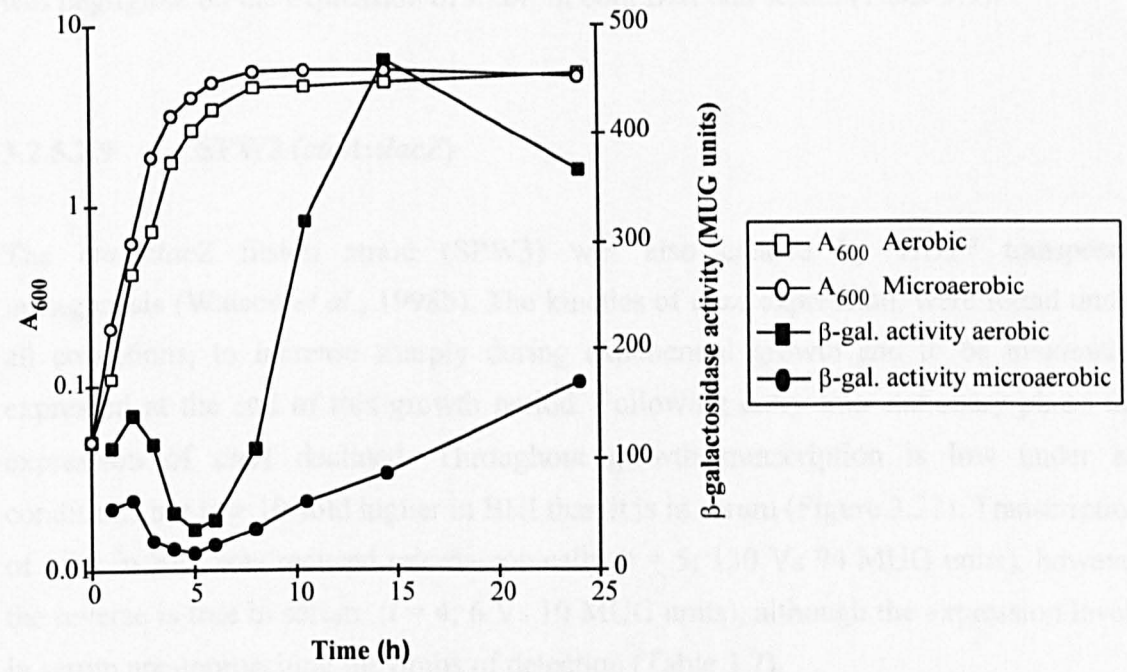


Figure 3.20

Expression of *katA::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus MJH006 (*katA::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

the large increase during stationary phase to a maximal expression value was observed, (460 and 170 MUG units aerobic and microaerobic conditions respectively) (Figure 3.20b).

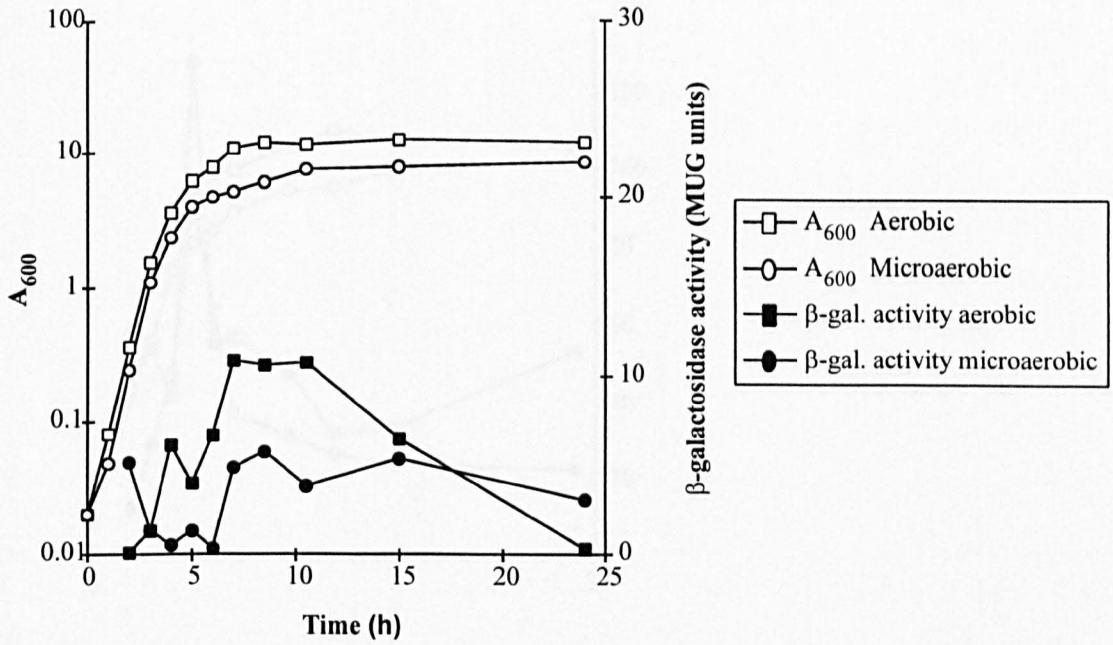
3.2.5.2.8 SPW1 (*sodA*::*lacZ*)

The *sodA*::*lacZ* fusion strain (SPW1) was constructed as a result of the insertion of the transposon Tn*917* into the *sodA* gene (Watson *et al.*, 1998b). *sodA* had the lowest transcriptional levels of all the fusion strains studied, with expression at the very limits of MUG detection. Expression of *sodA* was detectable in BHI, (~ 6-10 MUG units during post-exponential phase), with some increase in expression of *sodA* in BHI as cells entered post-exponential and stationary phase (Figure 3.21a). Transcription of *sodA* in serum however was either 'switched off' or was below the level of detection (Figure 3.21b). The effect of incubation under either aerobic or microaerobic conditions was negligible on the expression of *sodA,* in both BHI and serum (Table 3.7).

3.2.5.2.9 SPW3 (*ctaA*::*lacZ*)

The *ctaA*::*lacZ* fusion strain (SPW3) was also created by Tn*917* transposon mutagenesis (Watson *et al.*, 1998b). The kinetics of *ctaA* expression, were found under all conditions, to increase sharply during exponential growth and to be maximally expressed at the end of this growth period. Following entry into stationary phase the expression of *ctaA* declined. Throughout growth transcription is low under all conditions but is > 10-fold higher in BHI than it is in serum (Figure 3.22). Transcription of *ctaA* in BHI was reduced microaerobically (t = 5; 130 Vs 94 MUG units), however the reverse is true in serum (t = 4; 6 Vs 10 MUG units), although the expression levels in serum are approaching the limits of detection (Table 3.7).

a)



b)

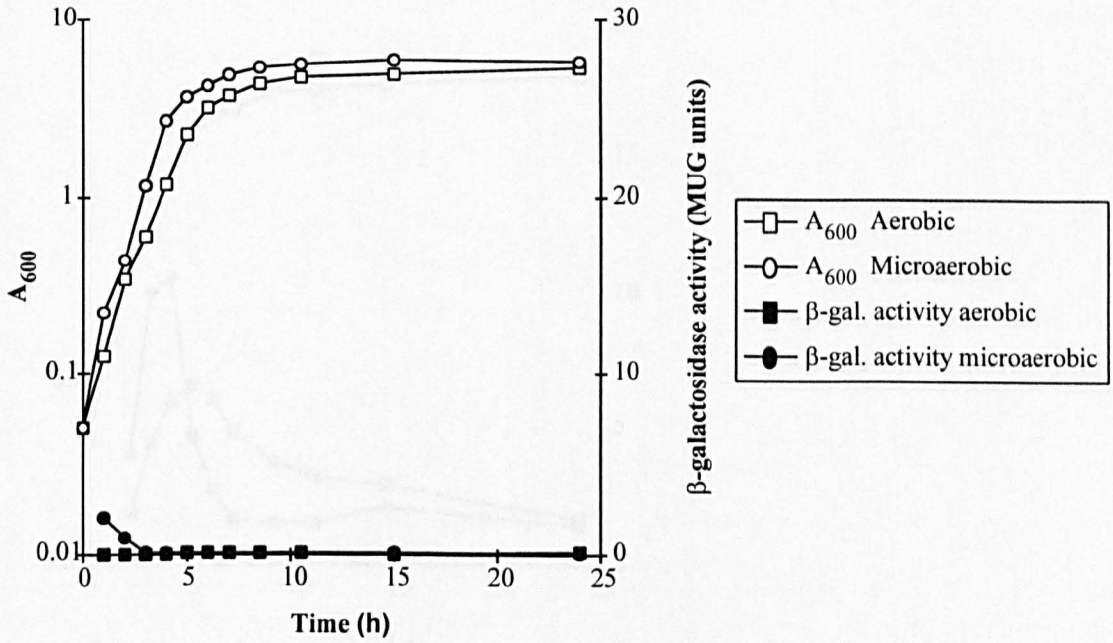
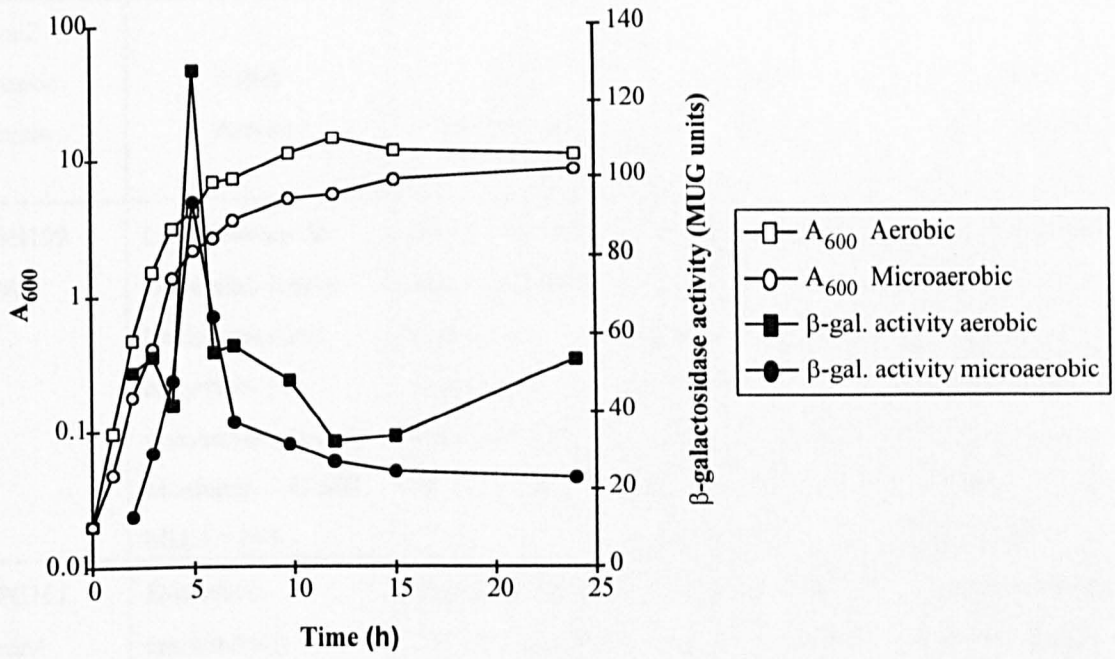


Figure 3.21

Expression of *sodA::lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus SPW1 (*sodA::lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

a)



b)

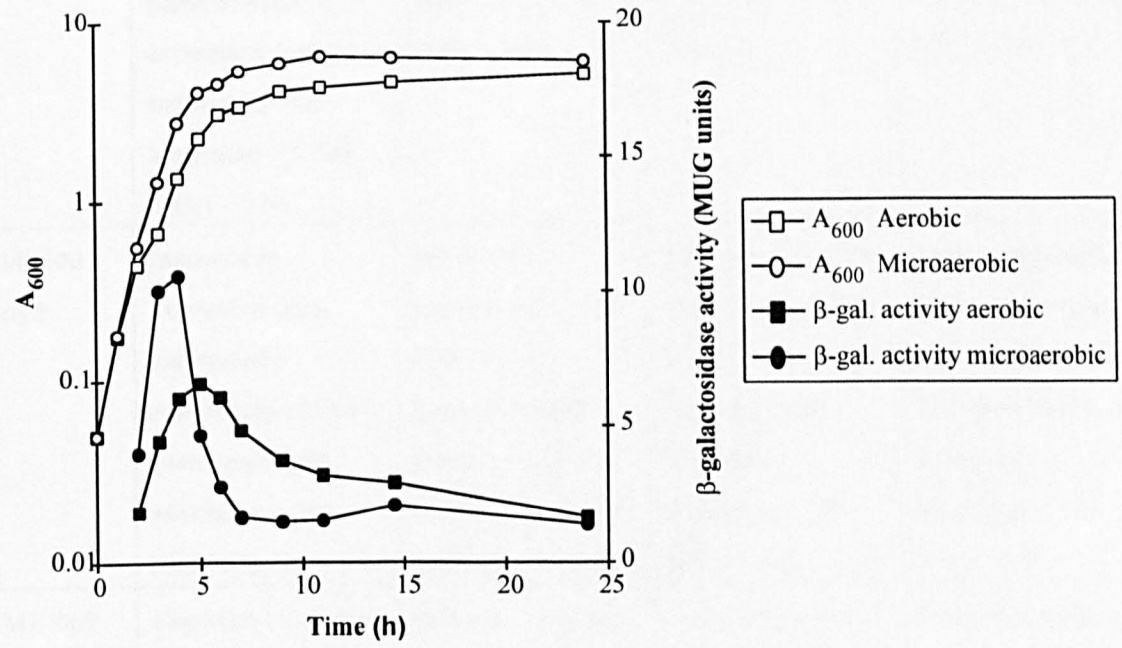


Figure 3.22

Expression of *ctaA*::*lacZ* during growth in BHI (**a**) and serum (**b**) under aerobic and microaerobic conditions

S. aureus SPW3 (*ctaA*::*lacZ*) was grown at 37 °C in both BHI and serum under aerobic and microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

lacZ fusion strain	BHI Aerobic	BHI Microaerobic	Sera Aerobic	Sera Microaerobic
SH109 *hld*	Large increase in expression during late exponential/ early post-exponential phase. Maximum ~ 53,000 MU, t = 24 h.	Almost identical expression kinetics to aerobic conditions. Maximum ~ 53,000 MU, t = 24 h.	Expression reduced > 20-fold throughout post-exponential and stationary phases. Maximum ~ 2,250 MU, t = 14 h.	Expression reduced > 20-fold throughout post-exponential and stationary phases. Maximum ~ 2,700 MU, t = 5 h.
PC161 *sarA*	Expressed constitutively throughout growth. Large increase in expression during stationary phase Maximum ~ 5,700 MU, t = 24 h.	Expression reduced 1.5- 2.5-fold during stationary phase. Maximum ~ 3,200 MU, t = 4 h.	Expression reduced 1.5- 2.5-fold during stationary phase. Maximum ~ 2,300 MU, t = 13 h.	Expression reduced 2- 5-fold during stationary phase. Maximum ~ 1,700 MU, t = 4 h.
MC100 *sigB*	Increases in expression occur during both exponential and late stationary phase. Maximum ~ 360 MU, t = 24 h.	Increased expression (1.5- 2-fold) during exponential and post-exponential phases. Maximum ~ 450 MU, t = 5 h.	Reduced expression during exponential phase (~ 2-fold) and stationary phase (~ 9-fold). Maximum ~ 140 MU, t = 6 h.	Reduced expression during exponential phase (~ 2-fold) and stationary phase (~ 9-fold). Maximum ~ 140 MU, t = 5 h.
MJH005 *fur*	Expression increased during exponential phase (to ~ 1,700 units), followed by a rapid decrease to a basal level on entry to post-exponential phase. Maximum ~ 1,650 MU, t = 5 h.	Expression reduced ~ 2-fold during post-exponential phase. Maximum ~ 900 MU, t = 5 h.	Expression reduced ~ 4-fold during post-exponential phase. Maximum ~ 470 MU, t = 6 h.	Expression reduced ~ 5-fold during post-exponential phase. Maximum ~ 280 MU, t = 4 h.

lacZ fusion strain (cont.)	BHI Aerobic	BHI Microaerobic	Sera Aerobic	Sera Microaerobic
PC203 *spa*	Expression increased during exponential growth (maximum ~ 70 MU, t = 4). Expression decreased as cells entered post-exponential phase.	Reduced expression. No characteristic peak of expression during exponential phase. Maximum ~ 10 MU, t = 7 h.	Increased expression during exponential phase (3-fold). Maximum ~ 210 MU, t = 4 h.	Reduced expression during exponential phase (> 2-fold). Maximum ~ 25 MU, t = 4 h.
PC322 *hla*	Expression increases rapidly during post-exponential phase, levels remained high during stationary phase. Maximum ~ 42,500 MU, t = 24 h.	Expression reduced slightly during late statuary phase. Maximum ~ 41,000 MU, t = 5 h.	Expression reduced ~ 5-fold throughout post-exponential and stationary phases. Maximum ~ 9,000 MU, t = 14 h.	Expression reduced ~ 15-fold throughout post-exponential and stationary phases. Maximum ~ 2,500 MU, t = 5 h.
MJH006 *katA*	Large increase in expression during stationary phase, Maximum ~ 1,780 MU, t = 24 h.	Reduced expression throughout growth. Maximum ~ 470 MU, t = 24 h.	Reduced expression throughout growth. Maximum ~ 460 MU, t = 14 h.	Reduced expression throughout growth. Maximum ~ 170 MU, t = 24 h.
SPW1 *sodA*	Very low but constitutive expression. Maximum ~ 10 MU, on entry to stationary phase (t = 7 h).	Very similar expression kinetics. Maximum ~ 5 MU, t = 8 h.	No detectable level of expression.	No detectable level of expression.
SPW3 *ctaA*	Expression increased throughout exponential phase (maximum ~ 130 MU, t = 5 h), followed by a sharp decrease to a basal level.	Slight reduction in expression throughout growth. Maximum ~ 95 MU, t = 5 h.	Reduced expression > 10-fold throughout growth. Maximum ~ 7 MU, t = 5 h.	Reduced expression > 10-fold throughout growth. Maximum ~ 10 MU, t = 4 h.

Table 3.7

Comparison of *lacZ* fusion expression data grown under different conditions, to expression in BHI under aerobic conditions

Summary of the major difference between the expression of SH109 (*hld*::*lacZ*), PC161 (*sarA*::*lacZ*), MC100 (*sigB*::*lacZ*), MJH005 (*fur*::*lacZ*), PC203 (*spa*::*lacZ*), PC322 (*hla*::*lacZ*), MJH006 (*katA*::*lacZ*) SPW1 (*sodA*::*lacZ*) and SPW3 (*ctaA*::*lacZ*) strains, when grown in BHI under microaerobic conditions and in serum under both aerobic and microaerobic conditions, compared to that when the same fusion strain is grown in BHI under aerobic conditions, (standard laboratory conditions). All fusion strains were grown in 25 ml cultures, in 100 ml flasks for 24 h at 37 °C with shaking (250 rpm). Expression of *lacZ* fusion strains was followed throughout growth by measuring the amount of β-galactosidase produced (encoded by *lacZ*) using the fluorescent substrate MUG.

MU - β-galactosidase activity (MUG units).

3.3 Discussion

Two *in vitro* models mimicking infection have been developed. The first, synthetic cerebrospinal fluid (sCSF) proved to be a starvation-survival medium and so was of no use for the growth and culturing of bacteria. For this reason, it was deemed not to be a suitable model mimicking infection.

S. aureus is capable of surviving for a long period of time (> 35 days) in both sCSF I and II. Starvation-survival kinetics similar to those previously seen in *S. aureus* (Watson *et al.*, 1998a/b), were seen in sCSF. Results showed that survival was greater in the simpler of the two sCSF's (sCSF I). This can be explained as when *S. aureus* is nutrient limited, particularly for carbon source, it enters a dormant type state and so cells are capable of long-term survival. sCSF II however contains glucose and studies have shown that under starvation conditions, when excess glucose is present, the glucose is fermented bringing about a decrease in pH. This reduction in pH has been shown to increase the rate at which *S. aureus* dies (Watson *et al.*, 1998a).

The second model mimicking infection is the serum model. *S. aureus* was shown to grow in both solid and liquid sera (pig and human), under both aerobic and microaerobic conditions. Microaerobic conditions were discovered to have a large beneficial effect on the growth of *S. aureus* on solid serum medium, with a much smaller but still noticeable effect in liquid medium. The possible reasons for this increased growth yield on serum microaerobically are discussed later. Why the increase is so much more dramatic on solid media than liquid however was not elucidated, but is potentially linked to the fact that growth on solid serum media is very poor and so any increase in growth rate or yield is likely to be much more noticeable. However, microaerobic conditions not only improve growth rates and yields in serum, but are also postulated to make the model more akin to conditions found *in vivo* (reduced O_2 and increased CO_2 tension). Thus in general microaerobic conditions were used for the growth of *S. aureus* unless the effect of environmental conditions needed to be investigated.

Examination of the growth of *S. aureus* on different media revealed that the growth yield in serum is significantly lower than that found on BHI. However, growth is comparable to that found in a chemically defined medium (CDM), as determined by both A_{600} and cfu. Thus despite serum being a complex medium containing many different components (Tables A.1.5-A.1.7) it does not support such profuse growth as BHI. The serum model therefore proved to be a suitable model for the potential

identification of components important to the growth of *S. aureus* under conditions mimicking infection. However, as the growth yield was low, especially on solid media, there is some factor, or factors which are limiting.

The effect of environmental oxygen and carbon dioxide concentrations was investigated following the discovery that microaerobic conditions (reduced oxygen and increased carbon dioxide concentration) increased the growth yield in serum. *S. aureus* grew to a increased growth yield anaerobically than aerobically, determined by colony size, although the growth in serum was found to be greatest under microaerobic conditions. This indicated that the presence of oxygen is potentially both harmful and beneficial to the growth of *S. aureus* in serum, as optimal growth is achieved in a microaerobic environment. This therefore hints towards a threshold level above which oxygen becomes toxic and below which impairs growth. Following the discovery that oxygen levels were important in determining the growth yield of *S. aureus* in serum, the effect of increasing culture aeration was investigated. This revealed that increasing the aeration of serum cultures (by decreasing the culture volume) had little or no effect on either the growth rate or final growth yield. This was not expected, as it was presumed that increasing culture aeration would increase the oxygen concentration in serum, and thus decrease the growth yield, as seen for aerobic Vs microaerobic growth. This result is also in contrast to those for *S. aureus* 8325-4 grown in BHI, where it has been shown that increasing culture aeration (by decreasing the culture volume) increases the growth yield (Lindsay and Foster, 1999). Therefore the role of oxygen concentration (and probably carbon dioxide) in the growth of *S. aureus* is not simple and further studies are required.

To further study oxidative stress in serum, catalase was added. Interestingly, this produced different results for solid and liquid media, with an advantageous effect only being witnessed on solid media. The results of the addition of catalase to solid serum showed that while catalase promotes growth aerobically it has no significant effect microaerobically. In liquid serum an improved growth yield was not even observed aerobically, and in fact high concentrations of catalase (1000 μM) proved toxic. However, from the results obtained from growth on solid media aerobically in the presence of catalase, it can be concluded that oxidative stress is a factor limiting growth on serum. More accurately *S. aureus* appears to be sensitive to H_2O_2 toxicity, as this reactive oxygen species is destroyed by catalase. However, because the growth yield is not improved to the level that was obtained microaerobically it can be hypothosized that other toxic forms of oxygen (for example O_2^- or $^{\cdot}OH$) may also be problematic and/or

the higher levels of carbon dioxide are also important for the improved growth of *S. aureus* under microaerobic conditions.

Iron was added to serum in an attempt to determine whether iron concentration was a factor limiting growth in serum, as it had been seen that the growth yield of *S. aureus* was low, especially on solid media. The addition of iron to serum however, once more revealed differences between the growth of *S. aureus* on solid and liquid media, with an advantageous effect only witnessed on solid media. The addition of iron (up to 500 μM) to solid media was shown to improve the growth of *S. aureus* under both aerobic and microaerobic conditions although the increase was much more dramatic under aerobic conditions. In liquid cultures however the addition of iron had no obvious beneficial effect on growth. The reason for this increase in growth yield on solid media could simply be due to the concentration of freely available iron being low in serum. Trivier *et al.*, (1995) showed that *S. aureus* strains had a decreased growth rate and yield under iron restricted conditions. Alternatively the addition of iron may mean that *S. aureus* no longer has to produce large quantities of siderophores (iron scavenging proteins) (Courcol *et al.*, 1997), thus conserving energy for growth. The addition of iron to serum however did not improve the growth to that achieved with BHI. Therefore it can be concluded that while the concentration of iron in serum appears to be a limiting factor, it is by no means the only one and so the iron concentration in serum is not the vital nutrient limiting growth. Why the benefit of iron addition is only seen on solid media is unclear, although it could be linked to the use of 50 % (v/v) serum in serum agar plates and 100 % (v/v) serum in serum cultures. Thus there is twice the amount of iron ml^{-1} in liquid than solid media. However, more work is needed to further prove this hypothesis.

Investigations into the role of the major global regulators *agr* and *sar* showed that these loci are not essential for growth in the serum model, as strains containing mutations in these genes were not attenuated in growth in the serum model. These results were comparable to published data for *agr* and *sarA* mutants *in vitro* (Lindsay and Foster, 1999; Chan and Foster, 1998a). The increased growth yield seen for both *agr* and *agr sarA* mutants may be due to energy constraints being lifted, as the cell is no longer capable of producing large quantities of many of the extracellular protein that are normally produced at the end of exponential phase, in an *agr* dependent manner. Thus a greater amount of energy is available for growth, increasing the culture yield. The fact that both *agr* and *sarA* mutants have been shown to have reduced virulence in some animal models (Abdelnour *et al.*, 1993; Booth *et al.*, 1995) indicates that pathogenicity is multifactorial, and no *in vitro* model can emulate its complexity.

Like the global regulators *agr* and *sar*, the alternative sigma factor σ^B does not appear to play an important role in the serum model. This conclusion can be drawn, as there was no growth attenuation of the strain defective at the *sigB* locus (PC400), compared to the parental strain 8325-4. Therefore, although it is postulated in this study that serum is a stressful environment, particularly under aerobic conditions, the alternative sigma factor σ^B, which has been shown to control the environmental stress response in *S. aureus* (Chan *et al.*, 1998), does not play an important role in the serum model.

To continue with the analysis of the role of oxidative stress, *S. aureus* strains containing chromosomal mutations in loci known to be important in oxidative stress were analysed in the serum model. These were ST1 (*katA* (catalase)), SPW1 (*sodA* (SOD)) and SPW3 (*ctaA* (haem O monoxygenase)).

KatA has previously been shown not only to be important in the ability of *S. aureus* to successfully withstand long-term starvation, but also in its capability to form lesions in a subcutaneous mouse infection model (Horsburgh and Foster, unpublished data). This study showed ST1 (*katA*) to have a reduced growth rate when grown in serum. Interestingly, this growth defect was found to be partially alleviated by incubation under microaerobic conditions. Finally as ST1 (*katA*) showed no such grow defect in BHI, this study provided direct evidence for the increased sensitivity of *S. aureus* to toxic forms of oxygen (oxidative stress) when grown in serum. The absence of the major catalase (KatA), due to insertional inactivation may also have an effect on the production of other virulence factors as the failure to produce one substance may affect cellular regulatory circuitry.

Analysis of SPW1 (*sodA*) showed that SOD did not have an important role in the serum model, as this strain showed no growth defect when compared to the isogenic parent 8325-4. SPW1 (*sodA*), like ST1 (*katA*) was postulated to be important in the serum model, as superoxide dismutase is the key enzyme in the protection against O_2^-, a toxic form of oxygen. SPW1 (*sodA*) has been shown to have a reduced SOD activity (Clements *et al.*, 1999b), however there is at least one further SOD enzyme in *S. aureus* (Fe-SOD) whose activity must be sufficient to accommodate the postulated additional oxygen toxicity of growth on serum.

CtaA has previously been shown to be important in oxidative stress, following the discovery that the addition of catalase to media aided the recovery of starved cells deficient at the *ctaA* locus (Clements *et al.*, 1999a). Furthermore, SPW3 (*ctaA*) has been shown to have, reduced amounts of all exoproteins in the culture supernatant, including

toxins such as α-haemolysin and toxic shock syndrome toxin-1. This decrease in exoprotein production is because exoproteins are preferentially expressed in post-exponential phase when growth appears to be conditional on haem A dependent cytochromes (Clements *et al.*, 1999a). SPW3 (*ctaA*) however was not defective in growth when compared to the parental strain 8325-4 under any condition examined.

Expression analysis, using the *lacZ* transcriptional fusions to the known components using X-Gal (solid media), showed this assay to be insufficiently sensitive to detect transcription from both *spa* and *sodA*. Furthermore, as this technique does not allow the temporal expression of genes to be measured, only a few conclusions can be drawn from this work. The most obvious conclusions are that *agr* (*hld*) and *hla* are expressed to a higher level on BHI than serum. The most probable explanation for this result is the fact that growth on BHI is significantly greater than that achieved on serum. As a result the concentration of AIP (Chapter 1.7.1.3), is increased and thus gene expression (*agr* and *hla*) is greater. This results in greater production of β-galactosidase and thus a greater amount of the blue end product. It can also be seen that *sarA, fur* and *hla* are expressed to a higher level on serum under microaerobic than aerobic conditions, particularly after 24 h. Rather than this being due to an up regulation of transcription under microaerobic conditions, an increased growth rate and thus a greater number of cells producing β-galactosidase, could be the explanation for this disparity in expression under aerobic and microaerobic conditions. Therefore, the expression of all nine genes appears to be identical on BHI/serum under aerobic and microaerobic conditions. These results would appear to be preliminary data indicating that these nine genes are not differentially regulated under aerobic and microaerobic conditions.

Expression analysis of the *lacZ* transcriptional fusions of known components in liquid cultures using the fluorogenic substrate MUG is a much more sensitive and accurate method for measuring gene expression than the X-Gal assay. This technique showed that all nine *lacZ* strains (*hld* (*agr*), *sarA, sigB, fur, spa, hla, katA, sodA,* and *ctaA*) when grown in BHI, under both aerobic and microaerobic conditions had detectable levels of β-galactosidase activity. The same is essentially true for the growth of the *lacZ* strains in serum with (*hld* (*agr*), *sarA, sigB, fur, spa, hla, katA,* and *ctaA* all producing detectable levels of β-galactosidase activity. However, under both aerobic and microaerobic conditions *sodA*::*lacZ* activity in serum was below the level of detection for the assay. Interestingly, in general the pattern of expression of *lacZ* fusion strains is similar under all four experimental conditions. However, there were some important regulatory differences, which are summarised in Table 3.7 and will be discussed further here.

The kinetics of *agr* (*hld*) expression showed the expected growth phase dependent pattern expected from previous studies (Fairhead, 1999). Expression in BHI rapidly increased as cells entered post-exponential phase, correlating with an increase in cell density, or more accurately octapeptide (AIP) density, reaching the threshold level necessary to switch on *agr* expression. In serum the final growth yield is significantly reduced compared to that in BHI, thus possibly explaining the decreased levels of *agr* (*hld*) expression.

sarA expression was shown to be expressed constitutively under all four conditions in agreement with previous work by Chan and Foster, (1998a). Expression of *sarA* was shown to increase during the exponential growth phase, which is consistent with RNA analysis (Bayer *et al.*, 1996).

The kinetics of *sigB* expression in *S. aureus* have previously been reported (Chan *et al.*, 1998), and were shown to rise steadily throughout the growth period, with the highest levels attained during late stationary phase. Expression kinetics in this study however, are different to those described above. In this study, under all four conditions there was a peak of activity as cells entered post-exponential phase, followed by decreased expression during stationary phase. *S. aureus* σ^{β} has been shown to play a role in stress resistance (Chan *et al.*, 1998) and proteins including alkaline shock protein 23 and a thermonuclease have been shown to be regulated in a σ^{β} dependent manner (Kullik *et al.*, 1998a). Therefore the kinetics of σ^{β} expression, are consistent with σ^{β} expression being necessary to regulate stress response proteins, to prevent damage to cellular components as the cells entered stationary phase.

The expression of the iron uptake regulator *fur*, which has been shown to repress siderophore biosynthesis and other iron uptake mechanisms in many Gram-negative bacteria, as well as in both *B. subtilis* and *S. aureus* (Bsat *et al.*, 1998; Xiong *et al.*, 2000), was seen to be regulated in a growth phase dependent manner. The regulation was very similar under all four experimental conditions with maximum expression occurring as cells entered post-exponential phase. This peak of maximal expression just before cells entered post-exponential phase may be explained by the fact that during this transitional period, proteins involved in oxidative stress are known to be produced in large amounts while the cell is still relatively metabolically active, to combat the increased oxidative stress that occurs during stationary phase. This is important as iron in the cell can lead to oxidative stress by the Fenton reaction, thus stopping the uptake of iron into the cell, will reduce oxidative stress in the cell (Keyer and Imlay, 1996). The expression pattern of *fur* in serum was similar to that seen in BHI, however the

overall levels of expression were reduced. An explanation for this reduced level of transcription in serum compared with BHI could be that available iron levels are lower in serum and therefore there is a need for higher siderophore production.

Surface protein A has been reported to be optimally expressed in late log phase and decreased during stationary phase (Vandenesch *et al.*, 1991; Chan and Foster, 1998b), in unison with the expected expression kinetics for a gene repressed by *agr* (Chapter 1.7.1). Results for *spa*::*lacZ* (PC161) expression in BHI under aerobic conditions and for serum under both aerobic and microaerobic conditions, were in agreement with this published work. Expression in BHI under microaerobic conditions however showed no transcription during late log phase. As *spa* is repressed by *agr*, it would be expected that in BHI under microaerobic conditions, when the cell density is reduced, *spa* expression would be greater as *agr* expression is reduced. This however is not the case, as it has been shown that *sarA* is a specific regulator of *spa* under microaerobic conditions (Chan and Foster, 1998a). *spa* expression was greater in serum than in BHI (under both conditions) as expected of a gene repressed by *agr*, as the final cell yield is lower in serum and as discussed above *agr* expression is significantly lower in serum. Furthermore, surface protein A is well known for its ability to bind immunoglobulins, which is important for the induction and propagation of any infection, including bacteraemia and therefore it would be expected that *spa* would be expressed preferentially *in vivo*. This result is therefore in agreement with the prediction of greater *spa* expression in the model mimicking infection (serum), than in BHI. Interestingly, expression is lower in serum under microaerobic conditions, showing a 90 % reduction in expression compared to that found under aerobic conditions and is thus in agreement with both increased *agr* expression and microaerobic repression by *sarA* (Chan and Foster, 1998a).

Expression of *hla* occurs preferentially in the transition between post-exponential and stationary phases of growth (Janzon *et al.*, 1989; Vandenesch *et al.*, 1991; Chan and Foster, 1998b). Results here are in agreement with this pattern of growth phase dependent expression. Furthermore, levels of expression were seen to be linked to culture growth yield. In other words expression of *hla* is greatest when A_{600} values are highest (BHI aerobic; highest growth yield, highest *hla* expression value). This is expected for a gene under regulatory control (positive) of *agr*, due to the density dependent signalling function (AIP) employed by *agr*, which up regulates *hla* expression. Microaerobically in BHI, levels of Hla and *hla* expression, have been shown to be diminished (Chan and Foster, 1998b). This may be due to a decrease in growth yield, and thus reduced quorum-sensing potential (Chapter 1.6.2) compared to

aerobic growth, although a specific regulatory mechanism may also account for this reduction. However, in this study only a small decrease in *hla* expression is seen as opposed to the large decrease seen by Chan and Foster (1998b). This disparity may be explained by the employment of different growth conditions which results in a greater growth yield microaerobically in this study (A_{600} ~ 8; t = 10 h) than that obtained by Chan and Foster (1998b) (A_{600} ~ 4, t = 10 h). This indicates that a fall in quorum-sensing potential may be at the heart of the difference in α-haemolysin expression levels. Interestingly, it has been reported that an increased level of CO_2 (10 %) has a stimulatory effect on the production of α-haemolysin. The reasons for this increase are not clear, but it is thought to be due to a combination of factors including CO_2 buffering capability and the role of CO_2 fixation in the formation of key amino acids essential for α-haemolysin production (reviewed by Wiseman, 1975). However, data from this study, in particular the low level of expression detected in serum under microaerobic conditions, and that by Chan and Foster (1998b) did not find this enhancement by CO_2. These inconsistencies are almost certainly due to the employment of different growth conditions. The large difference seen in *hla* expression levels in BHI and serum may however not just be due to a reduced quorum-sensing potential as there are several reports where nutritional factors influence production of α-haemolysin. For example, α-haemolysin has been found to be produced at different levels in different *in vitro* media, with an emphasis placed on the concentration of a few key amino acids (including arginine and valine) and glucose (Wiseman, 1975). Chan and Foster (1998b) also showed that salt and sucrose concentration were important as the additions of these to BHI (1 M NaCl and 20 mM sucrose), lead to a dramatic decrease in *hla* expression. Therefore, the large discrepancy in *hla* expression in BHI and serum may be, at least in part due to the difference in nutritional status of the two media (Tables A.1.1 and A.1.5-A.1.7).

katA was preferentially expressed during stationary phase and the level of expression in both media was greater under aerobic conditions. This is in accordance with the prediction that oxidative stress would be greater under aerobic conditions, and thus a gene involved in combatting oxidative stress (H_2O_2 toxicity), would be expressed to a greater level under aerobic (21 % O_2) than microaerobic (10 % O_2) conditions.

The expression of *sodA* in BHI was low during exponential phase growth (t = 3h; 2 MUG units) but increased by ~ 5-fold aerobically (t = 9h; 11 MUG units) and 3-fold microaerobically (t = 9h 6 MUG units) during stationary phase. This trend in SOD expression has previously been witnessed in *S. aureus* (Clements *et al.*, 1999b). The fact that expression was increased aerobically compared to microaerobically, was most

probably a result of increased internal superoxide levels, leading to greater oxidative stress within the cell. Increased levels of SOD during stationary phase growth have also been observed in other bacteria including *E. coli* (Benov and Fridovich, 1995) and *B. subtilis* (Inaoka *et al.*, 1998). The increase in SOD in cells during stationary phase growth may be a mechanism to prevent the accumulation of O_2^- damage to proteins, as damaged proteins will accumulate due to the low rate of protein synthesis (Nystrom and Gustavsson, 1998), and hence increased protein turnover as opposed to an increase in O_2^- generation is responsible for the increased SOD expression during this period.

The expression of *ctaA*::*lacZ* was low under all conditions throughout growth. However, like SPW1 (*sodA*::*lacZ*), SPW3 is a mutant strain (*ctaA*), with the *lacZ* fusion created by Tn*917* insertional mutagensis, and importantly this deficiency may affect the absolute level of expression. Maximal levels of *ctaA* expression occur during the transition from exponential to stationary phases of growth. This increase in expression at the onset of stationary phase has previously been documented in *B. subtilis* (Van der Oost *et al.*, 1991) and *S. aureus* (Clements *et al.*, 1999a). This pattern of expression links in well with the theory that in post-exponential phase, growth appears to be dependent on haem A dependent cytochromes (Clements *et al.*, 1999a). Therefore *ctaA* is maximally expressed when CtaA is required to enable energy production for cellular maintenance.

In this chapter it has been shown that while *S. aureus* shows classical starvation-survival kinetics in sCSF, it is capable of growth in serum. The level of oxygen and/or carbon dioxide is important for the growth of *S. aureus* in serum as an increased growth yield is observed microaerobically. The addition of both iron and catalase to solid serum media demonstrated an increase in growth, particularly under aerobic conditions, although additions of these had no apparent beneficial effect in liquid culture. None of the major global regulators of stress resistance or virulence determinant production had an important role in growth in serum. Finally the transcriptional pattern of several proven and potential virulence determinants were not altered dramatically in serum Vs BHI. The following chapters seek to determine components, which are important in the serum model.

CHAPTER 4

THE IDENTIFICATION OF *S. AUREUS* GENES SPECIFICALLY EXPRESSED IN THE SERUM MODEL

4.1 Introduction

The isolation and characterisation of mutants attenuated in pathogenicity or with altered virulence determinant production has proved an invaluable tool for dissecting and understanding the mechanisms of disease. Transposon mutagenesis has been used extensively in the study of pathogenesis and other functions in many organisms. For example, the three major (known) global regulators of *S. aureus* virulence factor production, *agr, sar* and *sae* were all isolated following transposon mutagenesis (Recsei *et al.*, 1986; Cheung *et al.*, 1992; Giraudo *et al.*, 1994). *agr* was identified as a transposon mutant displaying a pleiotropic exoprotein deficient phenotype (Recsei *et al.*, 1986) (Chapter 1.7.1). A screen for fibrinogen-binding-negative mutants, led to the identification of *sar* (Cheung *et al.*, 1992) (Chapter 1.7.2). Finally a transposon screen for mutants with decreased β-haemolysin production led to the identification of *sae* (Giraudo *et al.*, 1994) (Chapter 1.7.3).

Reporter gene fusions have also given important insights into the processes controlling gene expression. A combination of random transposon mutagenesis linked with reporter gene analysis, provides a powerful tool for the identification of genes important in and/or expressed in a given process (Camilli *et al.*, 1990). An example of a commonly used transposon with fusion generating capability is Tn*917-lac*. This transposon has been shown to be capable of generating fusions that connect the transcripts of chromosomal genes to the coding sequence of the *lacZ* gene of *E. coli* (Perkins and Youngman, 1986). Furthermore, Tn*917-lac*-mediated transcriptional fusions have been demonstrated to produce levels of β-galactosidase that reflects accurately the regulated expression of interrupted genes (Perkins and Youngman, 1986).

This transposon has been modified (Camilli *et al.*, 1990) to allow the direct cloning of DNA flanking the transposon insertions thus facilitating the identification of the insertionally inactivated gene. A novel derivative of Tn*917* (Tn*917*-LTV1) (Figure 4.1) was first used in *L. monocytogenes* (Camilli *et al.*, 1990). This transposon was again used in *L. monocytogenes,* in a screen for genes preferentially expressed in mammalian cells, compared to a rich broth (Klarsfeld *et al.*, 1994). This study led to the

identification of five *pic* (preferentially intracellular expression) genes, with up to 100-fold induction inside macrophage like cells (Chapter 1.11.1.2.1).

Although Tn*917*-LTV1 was originally constructed for the genetic analysis of *L. monocytogenes* (Camilli *et al.*, 1990), owing to its enhanced transposition frequency and the convenience for cloning DNA adjacent to sites of insertion, has meant that this system has been widely used for the insertional mutagenesis of Gram-positive bacteria. For example Tn*917* has been used in both *B. subtilis* and *B. cereus,* for the identification of germination deficient mutants and the subsequent identification of these inactivated genes (Clements and Moir, 1998, Behravan *et al.*, 2000).

Importantly, this Tn*917* system has also been used in *S. aureus*. Examples of this work include the isolation of: the staphylococcal accessory regulator (*sar*); two autolysin-defective mutants; and a whole host of starvation survival mutants (Cheung *et al.*, 1992; Mani *et al.*, 1993; Watson *et al.*, 1998b). Additionally Tn*917* has been used for the identification of genes specifically expressed upon interaction with human endothelial cells and with human serum (Vriesema, 2000). The transposon Tn*917* has been shown to randomly insert into the *S. aureus* chromosome (Youngman, 1987). This makes it ideal for generating mutant libraries in *S. aureus*. Furthermore, Tn*917*-LTV1 has been developed with a number of features making it particularly useful as a genetic tool. At one end of the transposon is a promoterless *lacZ* gene with a *B. subtilis* ribosome binding site. Thus Tn*917* generates transcriptional fusions between the promoterless *lacZ* and chromosomal genes when the transposon inserts into the chromosome in the appropriate orientation (Figure 4.2). There is an erythromycin resistance gene within the transposon region of pLTL1 for selection in *S. aureus*. Finally Tn*917*-LTV1 contains an ampicillin resistance gene for selection in *E. coli* and an *E. coli* replicon (*col*E1) along with a large number of restriction sites. These components allow the 'straightforward' rescue of chromosomal DNA flanking the Tn*917* insertion as a plasmid in *E. coli*.

This chapter describes the use of Tn*917*-LTV1 in the search for genes that are important and/or expressed in the serum model.

4.2 Results

4.2.1 Tn*917* mutagenesis of *S. aureus* 8325-4

The plasmid pLTV1 (Figure 4.1) is a 20.6 kb plasmid, comprising the transposon Tn*917-lac*, and additional flanking sequences. The flanking sequences contain a tetracycline resistance gene, and a Gram-positive temperature sensitive origin of replication. The copy number of the plasmid decreases rapidly at high temperatures with replication inhibited totally above 37 °C (Villafane *et al.*, 1987). This thermoinstability enables the selection of Tn*917* insertion mutants by growth at elevated temperatures in the presence of erythromycin. Elevated temperature treatment (> 37 °C) cures the bacteria of pLTV1, and thus selects for cells containing a chromosomal transposon insertion (Figure 4.2). Transposon mutagenesis of *S. aureus* 8325-4 was performed as described in Chapter 2.12.1.

4.2.1.1 Evaluation and enumeration of Tn*917* transposon libraries

Eight Tn*917* libraries were created (Table 4.1). These contained ~ 3 x10^8 - 5 x10^9 cfu ml^{-1}. The libraries were created using different plasmid curing temperatures with varying success. The first two libraries (1a and 1b) were cured of the plasmid at 41 °C, the third and fourth (2a and 2b) at 41.5 °C and the final four (I, II, X and Z) at approximately 41.8 °C. Temperatures of greater than 41.8 °C were not used in this study as previous work had shown that higher temperatures reduced the final growth yield dramatically (Foster, unpublished work). From Table 4.1 it can clearly be seen that raising the plasmid curing temperature increases the frequency of transposon insertion (Chapter 2.12.2). Of the eight libraries created only those with > 90 % transposon insertion frequency were deemed useful for screening, the others would have too high a proportion of colonies still containing the plasmid (pLTV1). Therefore only four (I, II, X and Z) of the eight libraries generated were used in this study.

4.2.2 Screening for genes involved in the growth of *S. aureus* in the serum model

Two screens were carried out using four transposon libraries for genes involved in the growth of *S. aureus* in serum. The first, was for genes that are essential for the growth

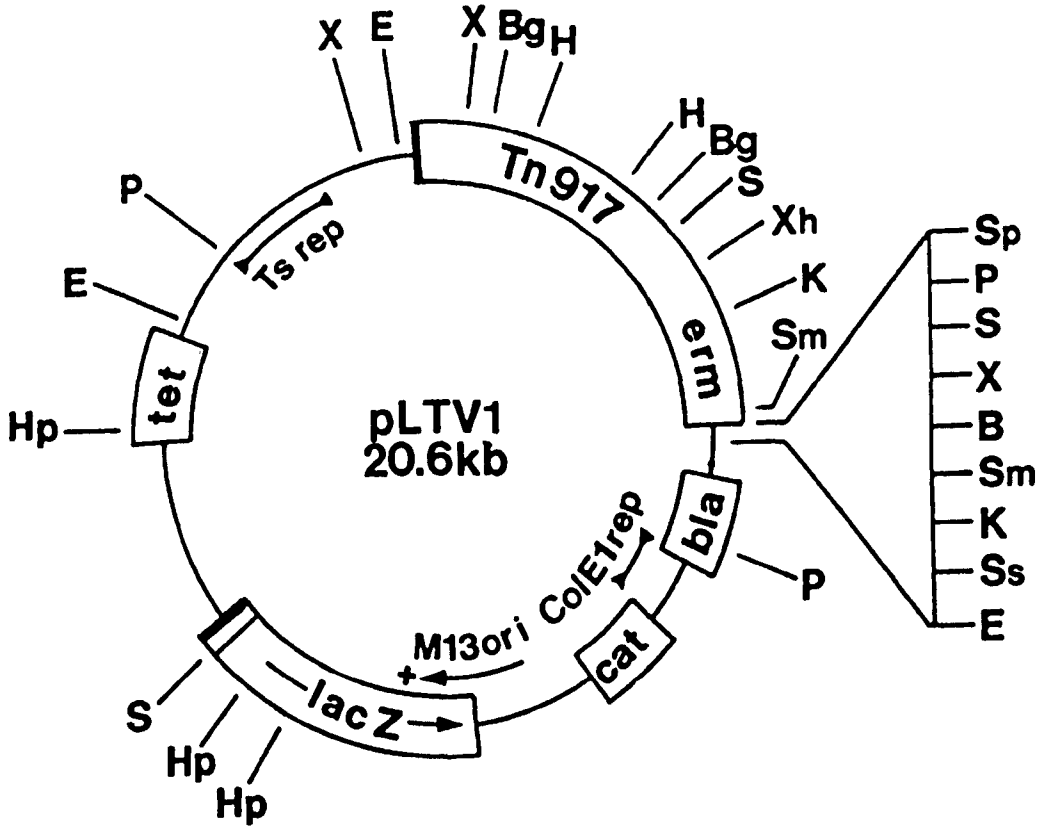


Figure 4.1

Physical map of pLTV1

Vector harbouring a transposon-proficient derivative of Tn*917*. Thicker lines represent the ends of the transposon. Features of this map: *bla*; pBR322 β-lactamase gene; ColE1 origin of replication; M13 ori, origin of replication for single stranded DNA synthesis; *erm*, erythromycin resistance gene; *cat*, a chloramphenicol acetyl transferase gene derived from pC194; *lacZ*, promotorless *lacZ* gene from *E.coli*, with translation initiation signals derived from *B. subtilis*; *tet*, tetracycline resistance gene from pAMα1Δ1; Ts rep, $pE194_{ts}$ temperature sensitive replicon. Restriction sites: B *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; Hp, *Hpa*I; K, *Kpn*I; P, *Pst*I;S, *Sal*I; Sm, *Sma*I; Sp, *Sph*I; Ss, *Sst*I; X, *Xba*I; Xh, *Xho*I.

Reproduced from Youngman (1990).

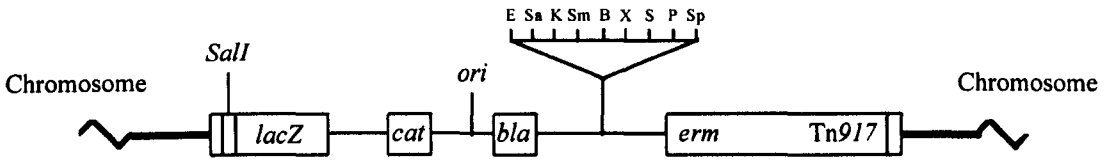


Figure 4.2

Map of Tn*917-lac* from pLTV1 inserted into the chromosome

Features shown: *lacZ*, promoterless *lacZ* gene from *E.coli*, with translation initiation signals derived from *B. subtilis*; *bla*; pBR322 β-lactamase gene; ColE1 origin of replication; *cat*, a chloramphenicol acetyl transferase gene derived from pC194; *erm*, a ribosomal RNA methyltransferase gene. Restriction sites: B *Bam*HI; E, *Eco*RI; K, *Kpn*I; P, *Pst*I; S, *Sal*I; Sm, *Sma*I; Sp, *Sph*I; X, *Xba*I.

Reproduced from Clements (1996).

Library	Plasmid curing temperature	Mean cfu ml^{-1} Ery Resistant	Mean cfu ml^{-1} Ery, tet Resistant	% Transposon insertion
1 a	41°C	3.28×10^8	1.55×10^8	52.7
1 b	41 °C	3.48×10^8	1.70×10^8	51.1
2 a	41.5 °C	4.5×10^9	5.80×10^8	87.1
2 b	41.5 °C	1.2×10^9	3.60×10^8	70.0
I	41.8 °C	5.38×10^9	2.44×10^8	**95.5**
II	41.8 °C	4.27×10^9	1.68×10^8	**96.1**
X	41.8 °C	3.65×10^9	2.35×10^8	**93.6**
Z	41.8 °C	4.12×10^8	1.76×10^7	**95.7**

Table 4.1

Evaluation and enumeration of Tn*917* transposon libraries of *S. aureus*

Table demonstrating how plasmid curing temperature affects Tn*917* transposon insertion efficiency in *S. aureus*. The mean cfu ml^{-1} values were calculated from the results of at least 2 separate experiments where colonies were plated onto BHI agar plates containing erythromycin (Ery), or erythromycin and tetracycline (Ery, tet).

Libraries (and their % transposon insertion) used for screening are shown in bold.

of *S. aureus* in serum, and the second for genes that are specifically, or preferentially, expressed in serum.

4.2.2.1 Selection of genes essential for the growth of *S. aureus* in the serum model

This first screen aimed to identify mutants based on their inability to grow on serum under microaerobic conditions (Chapter 2.13.1; Figure 4.3).

The transposon library stocks were diluted and plated onto BHI agar plates, resulting in ~ 100 colonies per plate. Colonies were then replica plated onto a BHI agar plate and a serum agar plate (Chapter 2.1.1.2 and 2.1.1.7). The BHI plate was placed under aerobic conditions, and the serum plate under microaerobic conditions. After 24 h incubation at 37 °C the serum plate was then compared to the BHI plate. Any colonies absent from the serum plate, as compared to the BHI plate were putative serum specific mutants. Putative serum specific mutants were then picked from the BHI plate and re-tested for their inability to grow on serum (Chapter 2.13.1; Figure 4.3).

Multiple screens were carried out using the above selection process, resulting in the screening of ~ 7,000 colonies from four independent Tn*917* transposon libraries. Although 39 putative serum growth defective mutants were initially identified, after re-testing, all were shown to grow on serum and thus were false positives (results not shown).

4.2.2.2 Selection of genes preferentially expressed in the serum model

The second screen was developed to identify *S. aureus* genes preferentially expressed during growth in serum Vs BHI. The screen utilises the promoterless *lacZ* reporter gene carried by Tn*917-lac* (Chapter 4.1) that on insertion into the chromosome may form a transcriptional fusion with the interrupted gene. A plate test was developed using X-Gal as the *lacZ* (β-galactosidase) substrate (Chapter 2.11.1). Colonies were examined for enhanced expression of the fused promoter (via β-galactosidase activity) on serum (Chapter 2.13.2; Figure 4.4).

The transposon library stocks were diluted and plated onto BHI agar plates resulting in ~ 100 colonies per plate. Colonies, were then replica plated onto a BHI agar plate and

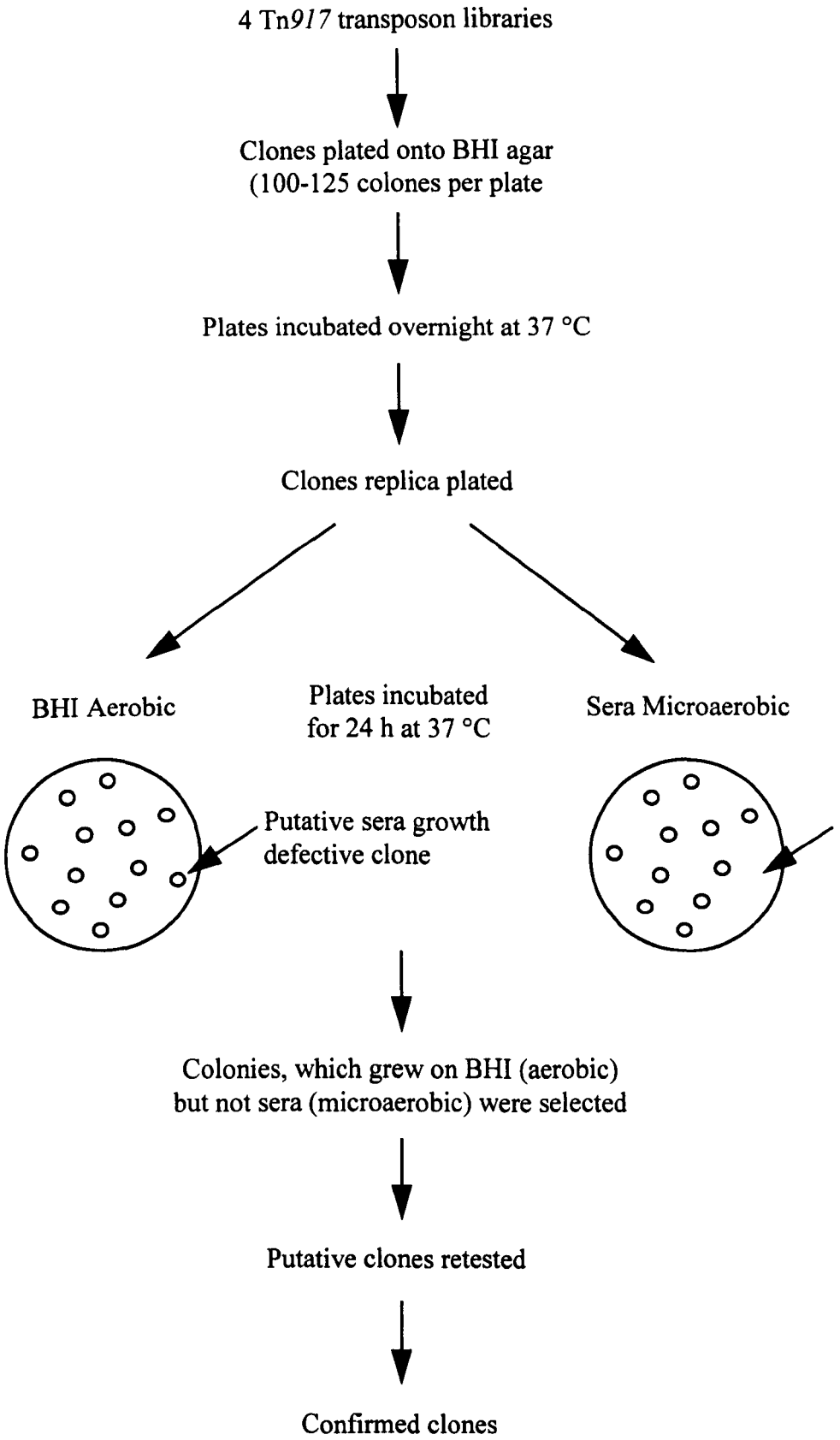


Figure 4.3

Diagrammatic representation of the screen developed for the identification of genes essential for the growth of *S. aureus* in the serum model

a serum agar plate both overlaid with X-Gal (Chapter 2.11.1). The BHI plate was incubated aerobically while the serum plate was incubated microaerobically. X-Gal is hydrolysed by β-galactosidase producing a blue coloured compound (5-bromo-4-chloro-3-indolyl), which is a visual indicator of β-galactosidase activity. After 24 h incubation at 37 °C the serum plate was compared to the BHI plate. Any colonies which appeared blue on the serum plate and white on the BHI agar plate were re-tested using the same X-Gal based screening process (Chapter 2.13.2; Figure 4.4).

In total, ~ 7,000 clones from four independent Tn*917* transposon libraries were screened. This led to the isolation of 77 clones with transposon insertions in loci that exhibited expression (increased expression) when grown on serum compared to BHI. The inactivated genes were called *seg* (Sera Expressed Gene) (Table 4.2). All four independent libraries generated clones, which should prevent the isolation of a high proportion of siblings. Importantly, none of the *seg* clones were Tet resistant, and thus do not contain an intact copy of the plasmid (pLTV1).

4.2.3 Characterisation and analysis of *seg* mutants

4.2.3.1 Expression analysis of Tn*917 lacZ* transcriptional fusions

The transposon Tn*917*, contains a promoterless *lacZ* gene, which in all *seg* mutants is inserted in the correct orientation downstream of the mutated gene's promoter, creating a transcriptional *lacZ* fusion. Thus the expression of the *seg* can be analysed by measuring β-galactosidase activity (Chapter 2.11).

All 77 *seg* mutants were grown on two BHI and serum agar plates overlaid with X-Gal (Chapter 2.3), which were incubated under aerobic or microaerobic conditions. Plates were scored for blue colouration after 24 and 48 h incubation (Table 4.3). The scoring of blue colouration on serum was difficult due to low level expression coupled with poor growth, particularly under aerobic conditions (Chapter 3.2.1.2.2). Therefore no attempt was made to distinguish between light and dark blue, instead colonies were simply scored as blue (B) or white (W).

The results show that four mutants (*seg*18, 19, 73 and 74), that were isolated following the observation that they were blue on serum but not on BHI, proved on further investigation to be white on serum, and thus were discarded (Table 4.3). Interestingly, after prolonged incubation (48 h), 16 *seg* strains that were white on BHI after 24 h had

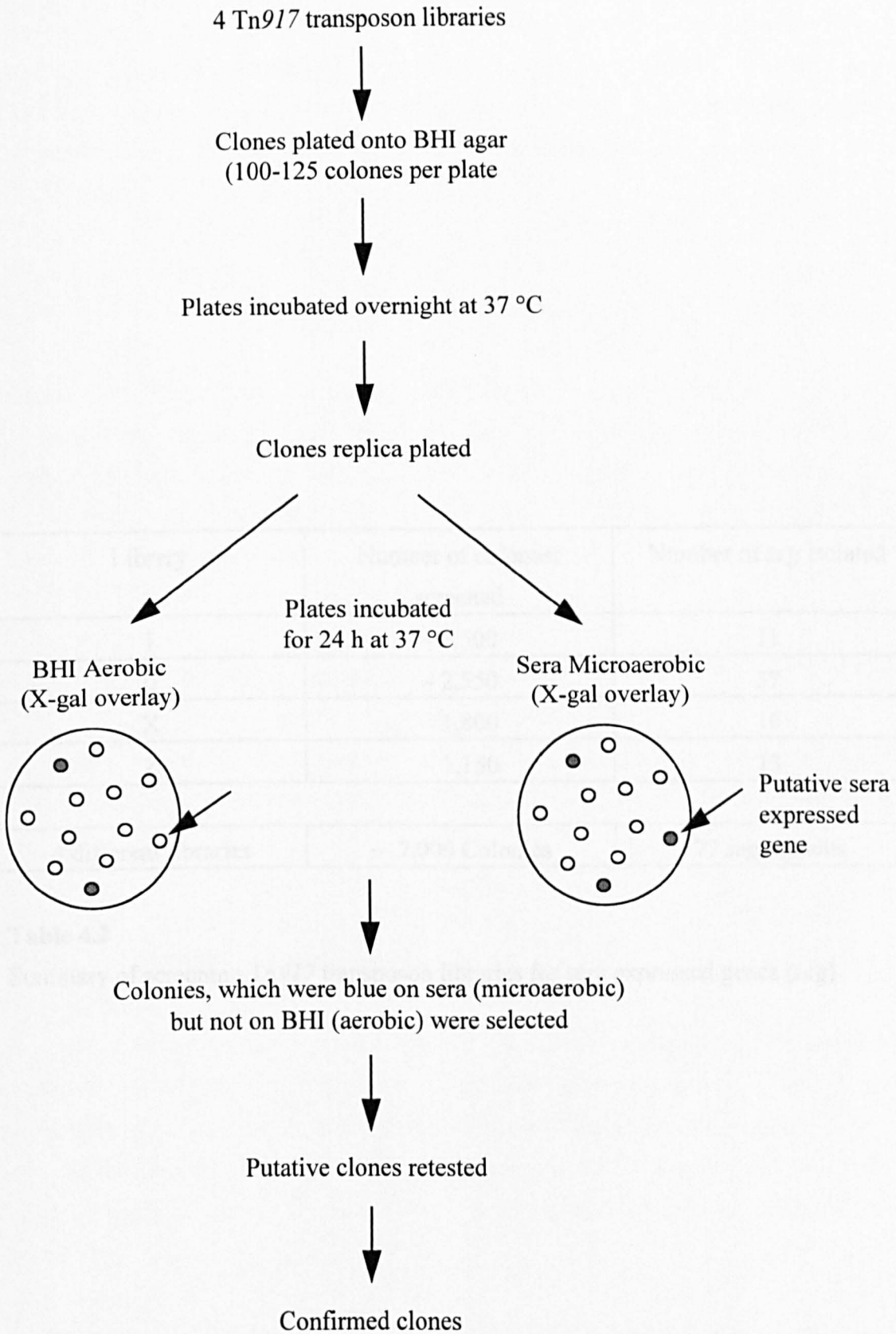


Figure 4.4

Diagrammatic representation of the screen developed for the identification of genes preferentially expressed during the growth of *S. aureus* in the serum model

Library	Number of colonies screened	Number of *seg* isolated
I	~ 1,500	11
II	~ 2,550	37
X	~ 1,800	16
Z	~ 1,150	13

4 different libraries	~ 7,000 Colonies	77 *seg* mutants

Table 4.2

Summary of screening Tn*917* transposon libraries for sera expressed genes (*seg*)

become faintly blue under both aerobic and microaerobic conditions. These mutants are still classed as *seg* strains as they are preferentially expressed (PE) on serum. Additionally, six clones were found to be blue, only on serum when grown microaerobically. Finally, all other clones were white on BHI, and blue on serum (Table 4.3).

4.2.4 Selection of *seg* mutants for further analysis

Of the 73 confirmed *seg* mutants, 25 were selected for further analysis. The selected mutants were representational of all libraries and expression classes (Table 4.3). Interestingly, the colour of the selected colonies on serum plates containing X-Gal varied from pale blue to dark blue indicating that the different mutants had various levels of β-galactosidase activity.

4.2.5 Analysis of selected *seg* mutants

A number of different techniques were used to study the 25 selected *seg* mutants at the physiological and molecular level, in order to determine the number and nature of the *seg* loci.

4.2.5.1 Phage transduction analysis

It was important to confirm that the *seg*::*lacZ* fusion created by the transposon insertion, was measuring expression of the sera expressed genes in an otherwise isogenic background to the parent (8325-4). Phage transduction was used to cross the transposon markers into the parental *S. aureus* strain 8325-4 (Chapter 2.14). 100 % cotransduction of the *seg* phenotype with the transposon indicates no other erroneous mutations.

All of the Ery/Lin resistant transductants isolated (>10 of each) from the transduction of 8325-4 with phage ϕ11 propagated on each of the original *seg* mutants had identical phenotypes (*lacZ* expression) to the corresponding donor strain. This strongly indicates that the *seg* are induced by environmental conditions and are not the result of secondary mutations, confirming linkage of the transposon antibiotic resistance to the *seg* mutation.

Class of *seg*	*seg* clone (Mutant Number)	Total number
Sera specific	**1**, 6, **8**, 9, **10**, 11, 12, **13**, **14**, **21**, **22**, 23, **24**, **25**, **26**, 27, **29**, **30**, 31, **33**, **37**, **39**, **40**, **43**, 44, **45**, 46, 47, **48**, 49, **50**, 52, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 72, 76, 77.	51
Sera-microaerobic specific	**35**, 51, 59, 70, 71, 75.	6
Preferentially expressed on sera (PE)	2, **3**, **4**, **5**, 7, 15, 16, 17, 20, 28, 32, 34, 36, 38, 41, 42.	16
Not expressed on sera (W)	18, 19, 73, 74.	4

Total = 73 *seg* mutants

Table 4.3

Summary of the classification of *seg* mutants

Those clones analysed further are shown in bold.

4.2.5.2 Growth of *seg* mutants

All the *seg* mutants grow on both BHI and serum agar plates under both aerobic and microaerobic conditions. This technique however does not detect reduced growth rates or yields. To investigate whether the transposon insertions resulted in growth defects, all *seg* clones along with the wild type strain (8325-4) were grown in BHI and serum under both aerobic and microaerobic conditions. All *seg* mutants showed indistinguishable growth rates and yields, to 8325-4 under all conditions tested (results not shown), and were comparable to previous results for the growth of *S. aureus* strains under these conditions (Chapter 3.2.5.2).

4.2.5.3 β-Haemolysin activity of *seg* mutants

The production of β-haemolysin was examined by patching strains onto blood test agar plates (Chapter 2.15). The wildtype strain (8325-4) produced zones of clearing with an average diameter of 18 mm on β-haemolysin test agar plates. All mutant strains were found to be hemolytic, producing zones of clearing of a similar size (results not shown).

4.2.5.4 Expression of Tn*917 lacZ* transcriptional fusions in liquid media

The plate method (Chapter 4.2.3.1), although very useful for studying the expression of large numbers of strains, is relatively insensitive and non-quantitative. Therefore following this analysis it was not known whether *seg* are completely switched off in BHI, or just down regulated in BHI/up regulated in serum. Furthermore, the magnitude of this difference in *seg* regulation is not known. This problem was addressed with the use of liquid cultures and the fluorogenic substrate MUG (Chapter 2.11.2). However, unlike in Chapter 3.2.5.2 where the expression of each strain was measured throughout growth, in this assay, gene expression was only measured once (during stationary phase). Cultures (5 ml) were incubated under both aerobic and microaerobic conditions for 15 hours (Chapter 2.4.1.1). β-galactosidase activity of each culture was then determined using the MUG assay (Chapter 2.11.2). This analysis showed that the difference in β-galactosidase expression observed on plates was also detected in liquid media. The results showed that all selected *seg* loci were expressed 2.5-260x higher in serum than BHI (Figure 4.5). The consequence of incubating cultures microaerobically had little or no effect on *seg* expression, and again much greater expression was afforded in serum than BHI (results not shown).

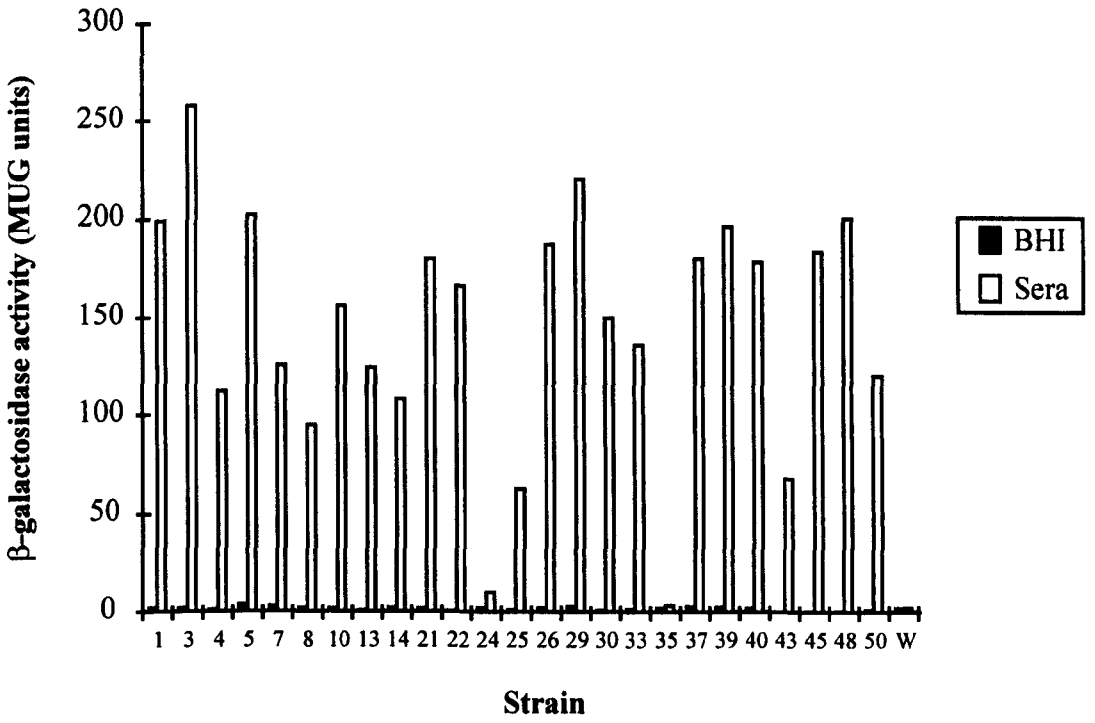


Figure 4.5

Expression of *seg*::*lacZ* transcriptional fusions in liquid media

The 25 selected *seg* strains along with *S. aureus* 8325-4 (W) were grown with shaking (250 rpm), at 37 °C, under aerobic conditions, in both BHI and serum. Specific β-galactosidase activity was determined after 15 h, and expressed in MUG units. The results are representative of two separate experiments.

There was however one exception, *seg*35 which was classified on solid media as sera-microaerobic specific (Table 4.3). This strain was also found to be induced ~5-fold on serum under microaerobic compared to aerobic conditions (results not shown).

4.2.6 Identification of sera expressed genes

4.2.6.1 Rescue of chromosomal DNA flanking Tn*917* insertion sites

In order to identify the *seg* loci, the transposon insertions were rescued along with flanking chromosomal DNA (Chapter 2.22.1.1). Chromosomal DNA from the *seg* mutants was digested with *Eco*RI or *Xba*I. This generated a fragment comprising part of the transposon (the *col*E1 replicon and the ampicillin (Amp) resistance gene (*bla*)) plus flanking chromosomal DNA out to the nearest relevant restriction site in the chromosomal DNA (Figures 4.1-4.2). By self-ligation, at a low concentration (Chapter 2.19.1.2.2), a plasmid was generated which was recovered by transformation into *E. coli* followed by selection for Amp resistant transformants (Chapter 2.16.1).

4.2.6.2 Recovery of DNA flanking Tn*917* insertions

Plasmid DNA was recovered from 2 independent transformants from each of the *seg* mutants (Chapter 2.16.1.2). Transformants were obtained from all *seg* mutants using either *Eco*RI or *Xba*I, except for *seg*3 and *seg*4, for which no transformants were obtained (Table 4.4).

The sizes of the recovered flanking chromosomal DNA fragments are shown in Table 4.4. For *Eco*RI transformants the size was determined by restriction with *Eco*RI and *Sal*I. The sizes of *Xba*I transformants were determined by restriction with *Xba*I and *Sal*I. A ~8 kb fragment was present in every plasmid digest, corresponding to the plasmid backbone. The chromosomal DNA fragment excised from the plasmids by double digestion includes, 276 bp of the transposon, due to the position of the *Sal*I restriction site, which is set back from the end of the transposon (Figure 4.2).

seg number	Plasmid number	Restriction enzyme used for rescue	Insert fragment size (Digestion with *Eco*RI and *Sal*I (bp))	Insert fragment size (Digestion with *Xba*I and *Sal*I (bp))
1	pSEG1	*Eco*RI	2700	ND
3	pSEG3	*Eco*RI or *Xba*I	No transformants	No transformants
4	pSEG4	*Eco*RI or *Xba*I	No transformants	No transformants
5	pSEG5	*Eco*RI	4300	ND
7	pSEG7	*Eco*RI or *Xba*I	2600	1800
8	pSEG8	*Eco*RI or *Xba*I	4300	1700
10	pSEG10	*Eco*RI or *Xba*I	4000	No transformants
13	pSEG13	*Eco*RI or *Xba*I	2600	5250
14	pSEG14	*Eco*RI or *Xba*I	1250	5000
21	pSEG21	*Eco*RI	2750	ND
22	pSEG22	*Eco*RI	1500	ND
24	pSEG24	*Eco*RI or *Xba*I	3000	No transformants
25	pSEG25	*Eco*RI or *Xba*I	2150	5500
26	pSEG26	*Eco*RI or *Xba*I	3250	No transformants
29	pSEG29	*Eco*RI	5000	ND
30	pSEG30	*Eco*RI or *Xba*I	1500	4600
33	pSEG33	*Eco*RI or *Xba*I	1500	4700
35	pSEG35	*Eco*RI	1250	ND
37	pSEG37	*Eco*RI	2200	ND
39	pSEG39	*Eco*RI or *Xba*I	2200	3000
40	pSEG40	*Eco*RI or *Xba*I	2200	3000
43	pSEG43	*Eco*RI	3000	ND
45	pSEG45	*Eco*RI or *Xba*I	2150	5750
48	pSEG48	*Eco*RI	2100	ND
50	pSEG50	*Eco*RI	2100	ND

Table 4.4

Fragment sizes of recovered DNA flanking Tn*917* insertions

Plasmids recovered from *seg* transformants were digested with *Eco*RI and *Sal*I or *Xba*I and *Sal*I, for transformants created with *Eco*RI and *Xba*I respectively. This produces two fragments, one is a fragment of ~ 8 kb (the plasmid backbone) and the other is the size of the cloned insert plus 276 bp (transposon).

ND - Not done.

4.2.6.3 Sequencing and analysis of the rescued *S. aureus* chromosomal DNA flanking Tn*917* insertion sites

Rescued chromosomal DNA flanking the transposon insertion sites, was sequenced from the *Eco*R1 recovered plasmids, using a 25 bp oligonucleotide as a primer (TnS1) (Table A.1.2), designed corresponding to the *lacZ*-proximal end of the transposon. The primer hybridises 181 bp upstream from the *lacZ* end of the transposon and the sequencing reaction proceeds through this end into the flanking chromosomal DNA (Watson, 1997).

Chromosomal DNA flanking Tn*917* insertion sites was sequenced (Chapter 2.21.3.1) from 23 *seg* mutants (2 *seg* strains failed to produce transformants). These sequencing reactions, in all cases produced 200-370 bp of useful chromosomal DNA sequence. This sequence was analysed to identify the Tn*917* insertion site and hence the insertionally inactivated gene (Chapter 2.23.1). As all the Tn*917* insertions are in the correct orientation to give a *lacZ* transcriptional reporter gene fusion, all the ORFs identified adjacent to the transposon in the rescued DNA are in the reverse orientation from the derived sequence. Database analysis (NCBI BLAST) was used to determine if any of the possible ORFs had homologues (Chapter 2.23.2). The results of these sequencing reactions along with their analysis are shown in Appendix A.2.1. In these figures the sequence is displayed so that the longest ORF is shown and the position of the Tn*917* insertion indicated. Where homologues were found, the percentage identity and extent of homology are shown along with the appropriate amino acid alignments between ORFs and database homologues (Figure A.2.1a-w).

Results of this analysis revealed 15 of the 23 sequenced mutants had transposons within ORFs. Upon database (NCBI) analysis, homologues were found to all except one of these ORFs (ORF5/*seg*5). The results of this analysis are shown in Table 4.5. This table shows the proposed identity of the insertionally inactivated gene, along with the % identity between possible ORFs and database homologues.

The sequence of pSEG1 revealed upon analysis to have a partial ORF (ORF1) of 76 amino acids in the same orientation as the *lacZ* (Figure A.2.1a). ORF1 is truncated by the Tn*917* insertion, and by the end of the sequenced DNA. ORF1 has a similarity to aspartokinase II (*lysC*) of *Helicobacter pylori* (41 % identity over 75 amino acids). Interestingly, chromosomal DNA sequenced and analysed from pSEG21 showed the transposon in this mutant to be inserted at exactly the same site as in *seg*1 (Figure A.2.1h). Furthermore, the sequence obtained from the rescue of DNA flanking the

seg No.	Likely identity of the insertionally inactivated gene	Identities	Organism	*seg* with identical Tn insertion site
1	Aspartokinase II	31 / 75 (41 %)	*Helicobacter pylori*	*seg*21
3	Not cloned/rescued	-	-	-
4	Not cloned/rescued	-	-	-
5	No significant homology	-	-	None
7	Hypothetical protein (surface protein)	18 / 66 (27 %)	*Staphylococcus aureus*	None
8	No significant homology	-	-	None
10	Dihydrodipicolinate synthase	38 / 73 (52 %)	*Bacillus subtilis*	None
13	Aspartokinase II	29 / 75 (38 %)	*Helicobacter pylori*	None
14	Aspartokinase II	13 / 24 (54 %)	*Mycobacterium smegmatis*	None
21	Aspartokinase II	28 / 82 (34 %)	*Bacillus subtilis*	*seg*1
22	No significant homology	-	-	*seg*30 and *seg*33
24	Aspartate semialdehyde dehydrogenase	26 / 56 (46 %)	*Campylobacter jejuni*	None
25	Aspartokinase II	48 / 96 (50 %)	*Thermotoga maritima*	*seg*50
26	Aspartate semialdehyde dehydrogenase	58 / 100 (58 %)	*Aquifex aeolicus*	None
29	Glutamate synthase large subunit	98 / 116 (84 %)	*Bacillus subtilis*	None
30	No significant homology	-	-	*seg*22 and *seg*33

seg No. cont.	Likely identity of the insertionally inactivated gene	Identities	Organism	*seg* with identical Tn insertion sites
33	No significant homology	-	-	*seg*22 and *seg*30
35	No significant homology	-	-	None
37	No significant homology	-	-	*seg*39 and *seg*40
39	No significant homology	-	-	*seg*37 and *seg*40
40	No significant homology	-	-	*seg*37 and *seg*39
43	Glutamate synthase large subunit	53 / 105 (50 %)	*Bacillus subtilis*	None
45	Aspartokinase II	38 / 76 (50 %)	*Thermotoga maritima*	*seg*48
48	Aspartokinase II	53 / 107 (49 %)	*Thermotoga maritima*	*seg*45
50	Aspartokinase II	45 / 92 (48 %)	*Thermotoga maritima*	*seg*25

Table 4.5

Summary of sequence analysis of the selected *seg* mutants

transposon in *seg*14 overlaps this sequence. However, the transposon insertion in *seg*14 is 62 bp upstream of the Tn*917* insertion in *seg*1 (Figure A.2.1a).

Chromosomal DNA sequenced and analysed from pSEG5 indicated a partial ORF (ORF 5) of 74 amino acids in the same orientation as the *lacZ* (Figure A.2.1b). The ORF is truncated by the Tn*917* insertion, as well as by the end of the sequenced DNA. ORF5 showed no similarity to any sequence in the NCBI database (BLAST search).

DNA sequence derived from pSEG7 on analysis revealed a partial ORF (ORF7) of 69 amino acids in the same orientation as the *lacZ* (Figure A.2.1c). The ORF is truncated by the Tn*917* insertion, as well as by the end of the sequenced DNA. ORF7 has 27 % identity over 66 amino acids to a hypothetical protein from *S. aureus*. Owing to the similarity of this hypothetical protein to sequences of surface proteins, ORF7 is thought to encode a surface protein.

Analysed DNA sequences derived from pSEG8 revealed no ORFs in the same orientation as the *lacZ* gene (Figure A.2.1d). Furthermore, the DNA sequence showed no striking similarity to any sequence in the NCBI database (BLAST search).

pSEG10 DNA sequence contained a partial ORF (ORF10) of 87 amino acids in the same orientation as the *lacZ* (Figure A.2.1e). ORF10 is truncated by the Tn*917* insertion and by the end of the sequenced DNA. Comparison of translated sequence with the NCBI database shows that the ORF10 partial polypeptide, has a strong similarity to dihydrodipicolinate synthase (DapA) of *B. subtilis*. ORF10 has 52 % identity over 73 amino acids to this DapA polypeptide.

Analysis of DNA derived from pSEG13 indicated a single partial ORF (ORF13) of 97 amino acids in the same orientation as the *lacZ* (Figure A.2.1f). ORF13 is truncated by the Tn*917* insertion, as well as by the end of the sequenced DNA. The ORF13 partial polypeptide is similar to aspartokinase II of *H. pylori* of (38 % identity over 75 amino acids).

The chromosomal DNA sequenced and analysed from pSEG14 showed a partial ORF (ORF14) of 34 amino acids in the same orientation as the *lacZ* gene that is truncated by the Tn*917* insertion (Figure A.2.1g). ORF14 has a similarity to the 5' end of the aspartokinase gene (*lysC*) of *Mycobacterium smegmatis* (54 % identity over 24 amino acids). The Tn*917* insertion of *seg*14 is 62 bp upstream of the transposon insertion site in *seg*1 (Figure A.2.1a).

For DNA sequences derived from pSEG22, analysis revealed no significant ORFs in the same direction as the *lacZ* gene (Figure A.2.1i). Furthermore, the DNA sequence showed no striking similarity to any sequence in the NCBI database (BLAST search). Chromosomal DNA sequenced and analysed from pSEG30 and pSEG33 showed that the transposon in these mutants is inserted at exactly the same site as for *seg*22 (Figures A.2.1i and A.2.1n-A.2.1o). These three mutants (*seg*22, *seg*30 and *seg*33) were isolated from the same library and thus are likely to be siblings.

Analysed sequences derived from pSEG24 indicated a partial ORF (ORF24) of 62 amino acids in the same orientation as the *lacZ* (Figure A.2.1j). The ORF is truncated by the Tn*917* insertion. Comparison with the NCBI database shows that ORF24 has homology to *asd*, which encodes aspartate semialdehyde dehydrogenase. ORF24 has 46 % identity over 56 amino acids to the N-terminal end of Asd from *Campylobacter jejuni*.

Chromosomal DNA derived from pSEG25 showed upon analysis to have a partial ORF (ORF25) of 100 amino acids in the same orientation as the *lacZ* (Figure A.2.1k). ORF25 is truncated by both the Tn*917* insertion, and the end of the sequenced DNA. ORF25 is strongly similar to aspartokinase II (*lysC*) from *Thermotoga maritima*. ORF25 has 50 % identity over 96 amino acids to the LysC derived from *T. maritima*. DNA sequence derived from pSEG50 showed that the transposon in this strain has inserted into an identical site in both *seg*25 and *seg*50 (Figures A.2.1k and A.2.1w). As these two mutants (*seg*25 and *seg*50) were isolated from the same library they are most likely to be siblings.

pSEG26 DNA sequence indicates the presence of a partial ORF (ORF26) of 100 amino acids in the same orientation as the *lacZ* gene (Figure A.2.1l). The ORF is truncated by the Tn*917* insertion, and by the end of the sequenced DNA. The sequence of ORF26 has strong similarity to aspartate semialdehyde dehydrogenase (Asd) of *Aquifex aeolicus* showing 58 % identity over 100 amino acids.

Analysis of DNA derived from pSEG29 indicated a partial ORF (ORF29) of 119 amino acids in the same orientation as the *lacZ* (Figure A.2.1m). ORF29 is truncated by both the Tn*917* insertion, and the end of the sequenced DNA. The ORF29 partial polypeptide shows very strong similarity to GltB of *B. subtilis* with 84 % identity over 116 amino acids. *gltB* encodes glutamate synthase (large subunit).

Analysed DNA sequence derived from pSEG35 revealed no significant ORFs (Figure A.2.1p). Furthermore, the DNA sequence showed no striking similarity to any sequence in the NCBI database (BLAST search).

Analysis of DNA sequence derived from pSEG37 failed to reveal any ORFs, in any reading frame or direction (Figure A.2.1q). Furthermore, the region of DNA sequenced showed no significant similarity to any sequences in the NCBI database (BLAST search). Interestingly, chromosomal DNA sequenced and analysed from both pSEG39 and pSEG40 showed that the transposon in these mutants was inserted at exactly the same site as in *seg*37 (Figure A.2.1q - A.1s). These three mutants (*seg*37, *seg*39 and *seg*40) were isolated from the same library, and are thus likely to be siblings.

For DNA sequence derived from pSEG43, analysis revealed a partial ORF (ORF43) of 109 amino acids in the same orientation as the *lacZ* (Figure A.2.1t). The ORF is truncated by the Tn*917* insertion, and by the end of the sequenced DNA. ORF43 has strong similarity (50 % identity over 105 amino acids) to glutamate synthase (*gltB*) large subunit of *B. subtilis*.

The chromosomal DNA sequenced and analysed from pSEG48 showed a single partial ORF (ORF45) of 86 amino acids in the same orientation as the *lacZ*. The ORF is truncated by both the Tn*917* insertion, and the end of the sequenced DNA (Figure A.2.1u). ORF45 has similarity to aspartokinase II (*lysC*) of *T. maritima* (50 % identity over 76 amino acids). Analysis of DNA sequences from pSEG48 revealed that the transposon insertion site is identical in *seg*48 and *seg*45 (Figures A.2.1u and A.2.1v). As these two mutants (*seg*45 and *seg*48) were isolated from the same library they are most likely to be siblings.

Following the analysis of the sequence recovered flanking the Tn*917* insertions, it can be seen that seven of the mutants (*seg*21, *seg*30, *seg*33, *seg*39, *seg*40, *seg*48 and *seg*50) have insertion sites identical to other mutants (Figure A.2.1a-w; Tables 4.5). For this reason no further analysis or investigation was carried out on these strains.

4.2.6.4 Further sequence analysis of *seg* mutants

Sequence analysis failed to identify homologues in the NCBI database to the region of DNA adjacent to the transposon insert for 5 of the sequenced *seg* strains, not including those with identical Tn*917* insertion sites to other strains. Thus the identity of the

promoter driving transcription of the reporter gene in these strains was unknown. Therefore to identify the region where the transposon has inserted, and to attempt to identify the sera expressed gene, further database analysis was carried out on 5 mutants: *seg*5, *seg*8, *seg*22 (*seg*30 and *seg*33), *seg*35 and *seg*37 (*seg*39 and *seg*40)). Two databases were searched, these were; the incomplete TIGR, and NCTC 8325 databases (Chapter 2.23.2). The results of this analysis are represented diagrammatically in Appendix A.2.2, and a summary of these results is shown in Table 4.6.

S. aureus database searching with the sequence produced from pSEG5 identified a contig from the TIGR database (97 % over 343 bp). Analysis of this contig confirmed the prediction that the transposon insertion of *seg*5 has inserted within an ORF (ORF5). ORF5 was found to be 363 amino acids long (Figure A.2.2a), however its amino acid sequence showed no significant homology to any know sequence in the NCBI database. Therefore the identity of ORF5 and hence of the sera expressed gene, remains unknown.

Initial analysis of sequence from pSEG8 failed to identify an ORF into which the transposon of *seg*8 had inserted. Searching the *S. aureus* databases with this sequence identified a contig with 100 % identity (over 250 bp) to the region sequenced from pSEG8. Analysis of this contig revealed that the transposon appears to have inserted 13 bp upstream of the translational start site of a gene with strong homology (49 % identity over 324 amino acids) to cystathionine γ-synthase of *B. subtilis* (Figure A.2.2b). Thus the transposon may have inserted within the control region of cystathionine γ-synthase, downstream of its promoter, from which expression of the promoterless *lacZ* must be driven.

Analysis of pSEG22 sequence with sequences in the *S. aureus* databases, revealed the transposon insertion site to be over 800 bp from a convergent ABC transporter (*ykpA*). More interestingly however was the discovery that the insertion in *seg*22 was 31 bp upstream of the translational start codon (ATG) of an aspartokinase II gene (40 % identity over 396 amino acids to *lysC* of *T. maritima*). (Figure A.2.2c). Thus it was postulated that the transposon insertion site was downstream of the aspartokinase II promoter region. Therefore in *seg*22, it is the aspartokinase II promoter that is driving production of β-galactosidase. As both *seg*30 and *seg*33 have identical insertion sites to *seg*22 the same is true of these strains.

Searching the *S. aureus* NCTC 8325 database with the sequence flanking the Tn*917* insertion in *seg*35 identified a contig with 99 % identity (over 298 bp) to this sequence.

This contig was analysed for putative ORFs, and revealed that the transposon within *seg*35 had inserted into a small putative ORF (ORF35) of 51 amino acids. ORF35 has an alternative initiation codon (TTG), which would produce a polypeptide of 47 amino acids (Figure A.2.2d). This polypeptide shows 51 % identity (over 45 amino acids) to lantibiotic gallidermin precursor (*gdmA*) from *Staphylococcus gallinarum*. Therefore it appears that the transposon insertion of *seg*35 has inserted within a lantibiotic of *S. aureus,* named auridermin owing to its significant sequence homology with both gallidermin (*S. gallinarum*), and epidermin (*S. epidermidis*). Thus it is postulated that transcription of the promoterless *lacZ* gene of *seg*35, is initiated by the promoter of the antimicrobial peptide, auridermin.

Sequence from pSEG37 showed no striking similarity to any proteins in the NCBI database. However, the transposon must have inserted downstream of an active promoter. As there are no obvious ORFs in the 250 bases sequenced from pSEG37, either the insert is just upstream of a *seg* translational start site, or transcription of the *lacZ* is being driven by read through from an ORF greater than 250 bp away. Database searching identified a contig from the TIGR database (98 % identity over 277 bp), however no ORF could be located immediately downstream of the transposon insertion site (within 250 bp). Owing to the absence of an adjacent downstream ORF, an 800 bp region upstream of this insertion site was analysed for ORFs. Although several putative ORFs were identified none of these were of any significant size (≤ 38 amino acids). Furthermore, this region has no significant homology to any sequence in the NCBI database (Figure A.2.2e). However, analysis of this contig did reveal there to be a gene in the correct orientation (convergent) to *lacZ*, with its stop codon approximately 845 bp upstream of the transposon insertion site. This gene has strong homology (76 % over 143 amino acids) to a lipoprotein (adhesin) from *Staphylococcus epidermidis* (SWISS-PROT accession number AA67571). Therefore it is possible that *lacZ* expression in *seg*37, is being driven by the promoter of the upstream adhesin gene (Figure A.2.2e). However, owing to the large distance between the surface protein (adhesin) stop codon and the transposon insertion site, the classification of the adhesin as a sera expressed gene is far from certain. Finally, as both *seg*39 and *seg*40 have identical insertion sites to *seg*37 the same is true of these strains, as for *seg*37.

4.3 Discussion

The screen for genes essential for the growth of *S. aureus* in the serum model was unsuccessful despite the screening of approximately 7,000 clones, from four different

seg No.	Chromosomal promoter driving *lacZ* expression	Sequence identity	Organism	*seg* with identical Tn insertion sites
5	Unknown	-	-	None
8	Cystathionine γ-synthase	159 / 324 (49 %)	*Bacillus subtilis*	None
22	Aspartokinase II	160 / 396 (40 %)	*Thermotoga maritima*	*seg*30 and *seg*33
35	Lantibiotic gallidermin	23 / 45 (51 %)	*Staphylococcus gallinarum*	None
37	Lipoprotein (adhesin)	109 / 143 (76 %)	*Staphylococcus epidermidis*	*seg*39 and *seg*40

Table 4.6

Summary of sequence analysis of the selected *seg* mutants following further analysis

libraries. The reason for this, is unlikely to be due to not enough clones being screened, as given the estimated 2.8 Mb genome size of *S. aureus*, an assumption of a 1 kb average gene size, suggests that complete coverage by transposon mutagenesis would require the screening of approximately 3,000 clones (Coulter *et al.*, 1998). This is particularly true for the transposon Tn*917*, which inserts randomly into the *S. aureus* chromosome (Youngman, 1987). Therefore in theory 7,000 transposons should have inserted into 'every' *S. aureus* gene at least once. Thus the most likely explanation for the failure to isolate clones essential for growth in the serum model, is that any defects in components leading to attenuated growth on serum, are so detrimental to the cell that they were also unable to grow on BHI. Therefore further work of this nature would most probably be futile.

The screen for genes preferentially expressed in the serum model was more successful, resulting in the isolation of 73 mutants that showed increased gene expression on serum compared to BHI, as determined by *lacZ* (β-galactosidase) activity. Transposon insertions in these strains are in loci that are preferentially expressed in sera, and have been termed sera expressed genes (*seg*). These 73 mutants were isolated from four different libraries with ~ 1 % of all colonies screened, being classed as *seg* strains. The 73 isolated *seg* strains were grouped according to their expression on different media (BHI/serum), under different conditions (aerobic/microaerobic), following 48 h growth, compared to 24 hours in the original screen. The discovery that sixteen *seg* strains produced colonies that were faintly blue on BHI after 48 h demonstrates that these genes are expressed at low levels on BHI, but at much enhanced levels on serum. This analysis also showed differences in the gene expression of *seg* strains, between aerobic and microaerobic conditions on solid media. This was interesting since the expression of the known components were not affected by this variable (Chapter 3.2.5).

Twenty five of the 73 *seg* mutants, were selected for further analysis, from all four transposon libraries, and representing all *lacZ* expression pattern classes. All the selected *seg* mutants showed identical growth kinetics and yields to the parental strain in liquid culture under all conditions. Thus even though the *seg* are expressed in serum they are not absolutely required for growth in that medium.

Importantly, the results of the reporter gene fusion assays in liquid culture, agreed with those for solid media. Notably, they showed a wide variety of promoter activity during stationary phase, but in all cases expression was found to be 2.5- 260-fold greater in serum than in BHI. This difference in activity during stationary phase, may be due to some *seg* promoters being more active than others, but might also infer differential

expression of these sera expressed genes. For example, a subset of these *seg* may be regulated in an *agr* dependent manner (Chapter 1.7.1) i.e. some may be expressed during exponential phase and repressed during stationary phase (thus display low activity levels during stationary phase), while others may be up regulated during post-exponential phase (therefore have significantly higher levels during stationary phase). Finally the classification of *seg*35 as a gene specifically expressed on serum under microaerobic conditions was confirmed using liquid cultures.

Recovery of flanking DNA and its subsequent sequencing, coupled with database searching, revealed that all clones (16 in total following the removal of siblings) have transposon insertions within regions homologous to others in the databases. In total, 9 putative sera expressed genes were identified. Of these, four were found to be involved in the biosynthesis of the aspartate family of amino acids (lysine, threonine, isoleucine and methionine) (Figure 5.1). These were: aspartokinase II (*lysC*) - *seg*1 (seg21), *seg*13, *seg*14, *seg*22 (*seg*30 and *seg*33), *seg*25 (*seg* 50), and *seg*45 (*seg*48); aspartate semialdehyde dehydrogenase (*asd*) - *seg*24 and *seg*26; dihydrodipicolinate synthase (*dapA*) - *seg*10; and cystathionine γ-synthase (*yjcI*) - *seg*8. The remaining insertions were within, or putatively linked to: glutamate synthase (*gltB*) - *seg*29 and *seg*43; auridermin (lantibiotc) - *seg*35; an ORF with homology to a hypothetical protein (surface protein) - *seg*7; a large ORF (ORF5) with no significant homology to any known sequence - *seg*5; and a lipoprotein (adhesin) - *seg*37 (*seg*39 and *seg*40).

Owing to the identification of four genes involved in the biosynthesis of the aspartate family of amino acids, mutations in each of these four genes (*lysC* (*seg*1), *yjcI* (*seg*8), *dapA* (*seg*10), and *asd* (*seg*24)) were selected for further analysis and will be discussed in more detail in Chapter 5. The 5 remaining putative *seg* identified in the serum screen, were not studied further, and thus will be discussed here.

Two *seg* strains, (*seg*29 and *seg*43) contained transposon insertions within genes showing significant homology to glutamate synthase (large subunit), of *B. subtilis*. Glutamate synthase is an enzyme that enables bacteria to incorporate ammonium ions into organic compounds. Only a few amino acids can be directly aminated (introduction of the amino group) by free ammonium ions. These are alanine, glutamine and glutamate catalysed by the enzymes, alanine dehydrogenase, glutamine synthetase and glutamate dehydrogenase respectively (Schlegel, 1993). As well as this the amide group of glutamine can be transferred to 2-oxogluturate by glutamate synthase. This system for the assimilation of ammonia nitrogen into organic compounds is apparently always formed and utilised by bacteria when the concentration of available ammonium ions is

low (< 1 mM) (Schlegel, 1993). The majority of the other amino acids obtain their amino groups by transamination with a primary amino acid. Therefore the incorporation of ammonium ions in which glutamate synthase plays a critical role is an important step in the production of amino acids. Due to the role of glutamate synthase in the biosynthesis of amino acids, not only directly in the production of glutamate, but also in the transamination of other amino acids, its identification as a gene specifically induced on serum is perhaps not that surprising, as amino acids may be limiting in serum (Tables A.1.6-A.1.7). This theory is further supported in light of the discovery of other *seg,* as being involved in the biosynthesis of amino acids (Chapter 5).

*seg*5 was identified as containing a transposon within an ORF (ORF5), of 363 amino acid, but which has no homology to any known protein. Therefore the serum screen has identified a novel gene, with a potentially important role in the growth of *S. aureus in vivo*. The identification of a gene with unknown function is not surprising owing to the fact that only a fraction of all bacterial genes have been studied. More importantly however, up to half the genes identified by both IVET and STM techniques of *S. aureus* have been reported to be of unknown function (Lowe *et al.*, 1998; Mei *et al.*, 1998). Further analysis of ORF5 may lead to the identification of a protein with an important role *in vivo*/pathogenicity.

The transposons of strains *seg*37, *seg*39 and *seg*40, were all found to have inserted into the same site and are believed to be siblings. Interestingly, the transposon in these strains has neither inserted within an ORF, nor inside the controlling region of a gene. Instead the insertion site is over 800 bp downstream of the translational stop codon of a gene showing significant homology to an iron-regulated lipoprotein (adhesin) of *S. epidermidis* (Personal communication, Dr M. Burnham, SmithKline Beecham). Therefore it is just possible that it is the adhesin promoter that is driving expression of the *lacZ* reporter gene, and thus the adhesin is the sera expressed gene. Interestingly, the expression kinetics of *seg*37 (*seg*39 and *seg*40) were studied, and demonstrated that in serum *lacZ* expression appeared to be switched on during exponential phase but fell rapidly during stationary phase (results not shown). Therefore these fusion strains showed classic surface protein expression kinetics (Chapter 1.6.3; Chapter 3.2.5.2.5), and thus these results provide evidence for the correct identification of the adhesin as the sera expressed gene.

The identification of a surface protein (adhesin) as a gene specifically induced upon interaction with serum is not surprising since adhesins have been shown to have an important role in infection. For example, the adhesion of *S. aureus* to human

endothelial cells has been implicated in the pathogenesis of invasive staphylococcal disease (Peacock *et al.*, 1999). This is true for all bacterial pathogens, but especially so of bacteria such as *S. aureus* which are primarily extracellular. These adhesins enable the attachment of the bacteria to components of the extracellular matrix of the host (Foster and Höök, 1998). This attachment initiates colonization, which is a vital step in the early stages of infection (Projan and Novick, 1997). Most relevant for the identification of an adhesin as being specifically induced in serum is the frequency with which *S. aureus* sepsis, seeds from the blood stream to other body sites (reviewed by Ing *et al.* 1997). This is postulated to involve interactions between circulating bacterial cells and vascular endothelial cells (Peacock *et al.*, 1999). However, although *S. aureus* has been shown to adhere to human endothelial cells *in vitro* (Ogawa *et al.*, 1985), the bacterial cell wall adhesin(s) responsible remain undefined (Peacock *et al.*, 1999). Therefore it is possible that the serum screen may have identified one of the adhesins involved in this adherence process.

*seg*7 was identified as containing a transposon within an ORF (ORF7), with homology to sequences in the databases. This clone has sequence homology to a hypothetical protein from *S. aureus*. However, these two sequences are not identical indicating that the transposon in *seg*7 has not inserted into that particular gene. Interestingly however, some of the other 'hits', although relatively low, are to surface proteins of *S. aureus,* for example 22 % homology over 63 amino acids to *sdrE,* part of a family of surface proteins (Foster and Höök, 1998). Therefore it is postulated that the transposon of *seg*7 has inserted within a gene encoding a novel surface protein, which may have an important role in attachment and in the establishment of infection *in vivo*. It is therefore possible that two of the sera expressed genes are involved in bacterial adherence to host tissues, indicating that the serum screen may be a valid screen for the identification of novel cell surface proteins that potentially play an important role in virulence.

The last of the genes identified by the serum screen, was insertionally inactivated in *seg*35. This clone was determined to have its transposon inserted within a gene showing strong homology to a lantibiotic precusor (*gdmA*) from *S. gallinarum*. Lantibiotics are a group of unique antibiotic peptides, which are produced by and primarily act on Gram-positive bacteria. The term lantibiotics was introduced as an abbreviation for lanthionine-containing antibiotic peptides (Schnell *et al.*, 1988). This term refers to the most prominent features of this group, which are the strong antibacterial activity and the intra-molecular rings formed by the thioether amino acids lanthionine and 3-methyllanthionine. In addition to lanthionine, other modified amino acids are present, which are synthesized by post-translational side-chain modifications of the ribosomally

produced precursor peptides (Sahl *et al.*, 1995). Lantibiotics are broadly divided into two groups on the basis of their structure and resulting function: type-A lantibiotics are elongated, amphiphilic peptides, which kill bacterial cells by forming pores in the targets cytoplasmic membrane; while type-B lantibiotics are compact globular proteins whose action appears to be directed at the inhibition of enzyme functions (Sahl *et al.*, 1995). Interestingly, although lantibiotics have unique structure they are functionally similar to other antimicrobial peptides and can be regarded as a structurally based subgroup of the bacteriocins (Klaenhammer, 1993). Staphylococcal lantibiotics including epidermin and gallidermin produced by *S. epidermidis* and *S. gallinarum* respectively, fall into the first group of lantibiotics (type-A) and thus form pores in the cytoplasmic membrane of susceptible cells.

Potentially the most interesting fact about the discovery of a lantibiotic of *S. aureus* (auridermin), as a gene preferentially expressed in sera, was that a lantibiotic was identified during signature tagged mutagenesis (STM) of *S. aureus*. Coulter *et al.* (1998) described the identification of a transposon insertion in a gene possessing sequence homology to the *S. epidermidis* lantibiotic gene encoding epidermin. Due to the nature of STM, genes identified by this method are attenuated for growth *in vivo,* and a mutation within the lantibiotic led to significant growth attenuation in the mouse abscess model and in an LD_{50} assay. Importantly, the lantibiotic identified in this study was also homologous to epidermin (59 % identity over 22 amino acids) and thus it is postulated that, Tn*917* has disrupted the same gene in both this study, and the STM strategy described by Coulter *et al.* (1998). If this is the case, then the serum screen has identified a gene as being preferentially expressed in serum, compared with a nutrient-rich medium that has been shown to have an important role *in vivo*. Finally it must be remembered that *seg*35 was classified as a serum-microaerobic specific strain, i.e. transcription of the gene insertionally inactivated in *seg*35 was only induced on serum under microaerobic conditions. Importantly however, this gene was not induced on nutrient-rich media under microaerobic conditions. This result therefore provides proof for the initial hypothesis, that the use of microaerobic conditions would make the model more a kin to those found *in vivo*. Finally the identification of a gene (auridermin) with an attenuated phenotype *in vivo* lends credibility to the serum screen, as a method for the identification of genes important for the growth of *S. aureus in vivo*.

The serum screen has therefore putatively identified: two surface proteins, an antimicrobial peptide (auridermin), a novel gene with unknown function, and five biosynthetic genes. The discovery that the 12 out of 16 of the *seg* mutants sequenced (not including siblings), had insertions in loci with likely biosynthetic

function was not totally unexpected as serum, is likely to be limiting for certain nutrients. Importantly, screens for genes whose products are expressed, or are important *in vivo* have also identified numerous biosynthetic loci.

IVET of *Salmonella typhimurium* led to the identification of the *carAB* operon, whose genes encode two subunits of carbamoyl-phosphate synthetase, which is involved in arginine and pyrimidine biosynthesis (Mahan *et al.*, 1993). Furthermore, in all studies for *in vivo* induced genes (Chapter 1.11.2) a large proportion of genes identified are involved in nutrient biosynthesis, in particular in the biosynthesis of amino acids. Additionally it has been reported that following STM of *S. aureus*, a significant proportion of mutants identified as being attenuated *in vivo* had transposon insertions within genes encoding enzymes involved in cellular biosynthesis pathways (Mei *et al.* 1997; Coulter *et al.*, 1998). Interestingly, the study by Mei *et al.* (1997) identified two genes involved in lysine biosynthesis, whereas the study by Coulter and colleagues (1998) identified three lysine biosynthetic genes. Importantly, both of these STM approaches isolated aspartate semialdehyde dehydrogenase (*asd*) as a mutant attenuated *in vivo*. Additionally a recent publication studying altered *S. aureus* gene expression (increased expression) upon interaction with both endothelial cells and human serum, identified four lysine biosynthetic genes including aspartokinase II and aspartate semialdehyde dehydrogenase (Vriesema *et al.*, 2000). These studies further highlight the potential importance of biosynthetic genes, especially those required for lysine synthesis, during infection. Finally these results provide solid evidence for the authenticity of the serum screen as an *in vitro* model mimicking infection.

In depth analysis of the four aspartate family amino acid biosynthetic genes; aspartokinase II (*seg*1), aspartate semialdehyde dehydrogenase (*seg*24), dihydrodipicolinate synthase (*seg*10), and cystathionine γ-synthase (*seg*8), will ensue in the following chapter.

CHAPTER 5

ANALYSIS OF THE ROLE OF THE ASPARTATE FAMILY OF AMINO ACID BIOSYNTHETIC LOCI

5.1 Introduction

5.1.1 Biosynthesis of the aspartate family amino acids

Many microorganisms are able to synthesize all 20 amino acids, which are required for protein synthesis. The amino acids lysine, threonine, methionine, and isoleucine derive the majority of their carbon atoms from L-aspartate, hence these amino acids are sometimes referred to as the aspartate family of amino acids. The biosynthesis of these amino acids, is accomplished by a complex pathway involving common intermediates from which multiple branches lead to the end products (Figure 5.1).

5.1.2 Functions of the aspartate pathway

The major function of the aspartate pathway in growing bacteria is the provision of amino acids for protein synthesis. However, the pathway also provides *meso*-diaminopimelate (DAP) and lysine, which are constituents of the cell wall peptidoglycan. *B. subtilis* utilises DAP within its cell wall, whereas the peptidoglycan of *S. aureus* contains lysine (Schleiffer and Kandler, 1972; Tipper and Strominger, 1968).

5.1.3 The common pathway

The first two catalytic steps are common to the biosynthesis of all end products of the aspartate biosynthetic pathway. These culminate in the production of aspartate semialdehyde, via aspartyl phosphate and are catalysed by the enzymes, aspartokinase (AK) and aspartate semialdehyde dehydrogenase (ASADH) (Figure 5.1).

However, differences in the regulation of the pathway give rise to multiple forms of some enzymes, and enzyme multiplicity can vary with the organism. Multiple isozymes of aspartokinase (AK), have been identified in several bacteria including: *B. subtilis, E. coli* and *Streptococcus bovis* (Zhang *et al.*, 1990; Theze *et al.*, 1974; Kalcheva *et al.*,

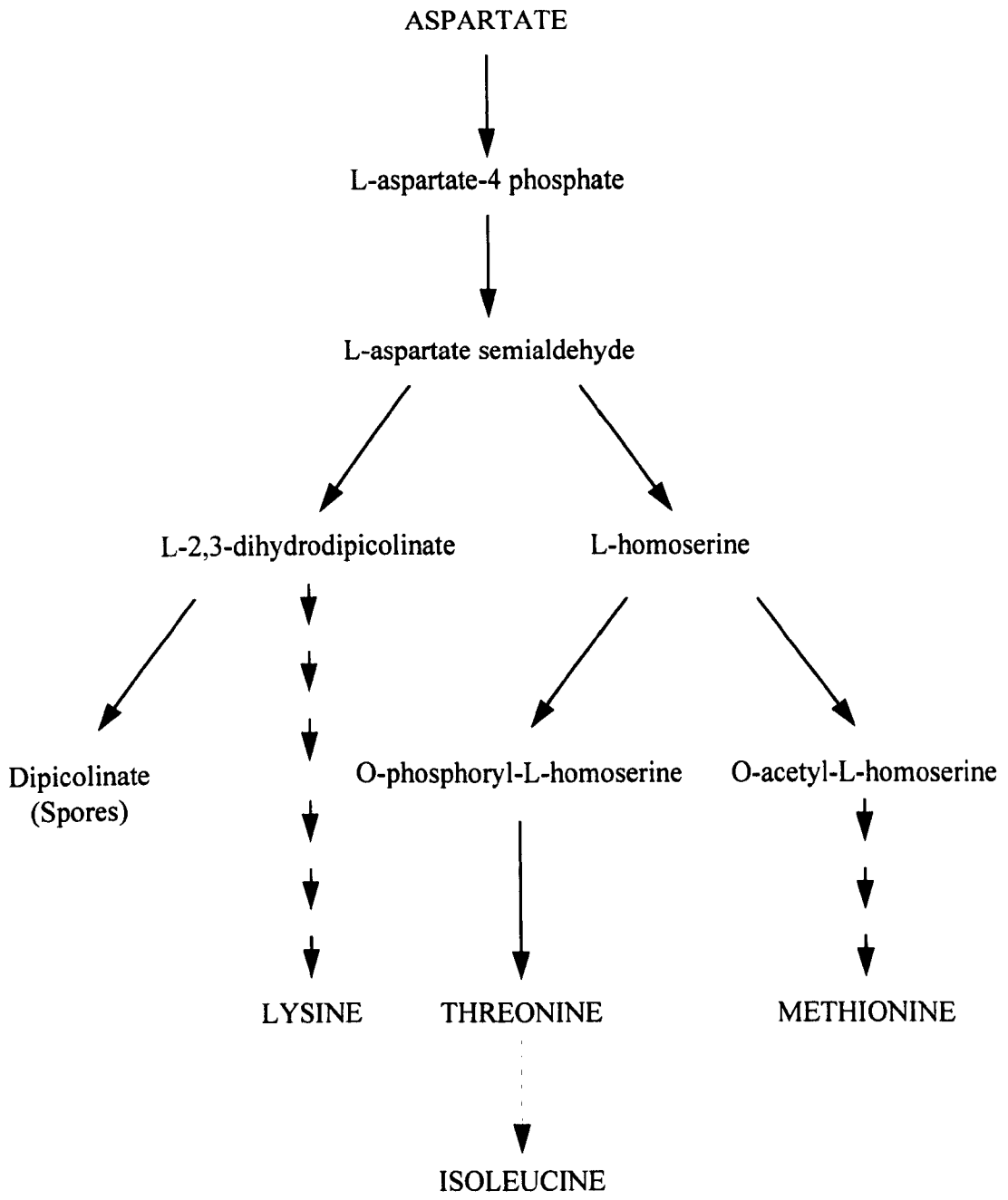


Figure 5.1
Overview of the pathway for the biosynthesis of the aspartate family of amino acids

The biosynthesis of isoleucine from threonine indicated by dotted arrows is a multi-step process. Adapted from Paulus (1993).

1993). In comparison only one AK has been discovered in pseudomonads, brevibacteria and corynebacteria (Shiio and Miyajima, 1969; Cohen *et al.*, 1969; Cremer *et al.*, 1988). *B. subtilis*, like *E. coli* has three distinct isozymes of aspartokinase (Zhang *et al.*, 1990; Theze *et al.*, 1974), and these differ in their regulation of synthesis and activity.

5.1.3.1 Aspartokinase isozymes of *B. subtilis*

The properties of the AK isozymes of *B. subtilis* (I, II, III) are summarised in Table 5.1. The putative AK identified in the present study (Chapter 4.2.2.2) showed greatest homology to AK II of *B. subtilis*.

5.1.3.1.1 Aspartokinase II (Lysine-sensitive aspartokinase)

AK II of *B. subtilis* is encoded by (*lysC*) and is regulated at both the activity and trancriptional level by lysine. Its activity has been shown to decline rapidly during glucose starvation or at the onset of stationary phase (Graves and Switzer, 1990), suggesting its primary function is the provision of precursors for the biosynthesis of lysine as a building block for proteins. AK II is a tetramer composed of two dissimilar subunits, ($\alpha_2\beta_2$) (Moir and Paulus, 1977). Interestingly, the β subunit corresponds to the carboxyl-terminal portion of the larger α subunit, thus the *lysC* subunits, α and β share a coding region. In fact the subunits are encoded by in-phase overlapping genes, which are independently translated (Bondaryk and Paulus, 1985). Problematically the function of the β subunit is unknown, for example purified β subunits yield no aspartokinase activity, whereas the α subunit alone possess activity that is virtually indistinguishable from the holo enzyme (Paulus, 1993).

5.1.3.1.2 Regulation of *lysC* in *B. subtilis*

The two in-phase overlapping coding regions are the only structural genes encoded by the *lysC* operon, being immediately followed by a putative rho-independent transcriptional terminator (Chen and Paulus, 1988). Preceding *lysC* is an unusually extensive control region. Transcript mapping (Chen *et al.*, 1987) and site-directed mutagenesis (Lu *et al.*, 1991), have shown that the expression of AK II involves a single transcriptional initiation site, about 330 bp upstream of the AK gene. Interestingly, the *lysC* leader region plays an important role in the regulation of AK II

synthesis by lysine. This leader region contains a number of potential control sites: i) an ORF encoding a potential 24 amino acid peptide (*lysC* leader peptide), with tandem lysine codons near the N-terminus, which is postulated to function as an attenuator; ii) an inverted repeat required for lysine repression; and iii) a series of four inverted repeats, the last of which resembles a rho-independent terminator, while the other three can assume alternate hairpin loops reminiscent of terminator and antiterminator structures (Chen *et al.*, 1987). The putative 24 amino acid leader peptide contains four lysine residues, and is located in the region that overlaps the protector loop, so that lysine deficiency would cause the stalling of the ribosomes in this region, thereby destabilising the protector structure and favouring the alternate base pairing pattern involving two pre-emptor loops without a functional transcriptional terminator. According to this model, lysine limitation would allow transcription to proceed beyond the attenuator site, into the structural gene for AK, thus enabling synthesis of the lysine biosynthetic enzyme (Chen *et al.*, 1987). The action of this model is supported by the finding that lysine regulates the expression of the *lysC* operon by effecting the premature termination of transcription at the rho-independent terminator site in the *lysC* leader region (Kochhar and Paulus, 1996). However, there are several problems with this model, including the fact that there is no evidence from translational fusion studies to the putative leader peptide, for translation of this ORF (Kochhar and Paulus, 1996). Additionally, the lysine induced truncation of mRNA transcripts is completely suppressed by an *aecA* mutation which is a single G-to-A substitution in the leader transcript (Lu *et al.*, 1991). However, this mutation is positioned remote from both the putative leader peptide and the terminator stem-loop structures. Therefore the mechanism of transcriptional attenuation of the *lysC* operon is unclear but seems to involve the lysine-dependent interaction between the site of the *aecA* mutation, and a remote transcription terminator element by a mechanism yet to be defined (Kochhar and Paulus, 1996).

5.1.3.2 Aspartate semialdehyde dehydrogenase

Aspartate semialdehyde dehydrogenase (ASADH), catalyses the second step in the common pathway and the *S. aureus* homologue was identified during the present study (Chapter 4.2.2.2). This enzyme catalyses the reversible substrate-dependent conversion of aspartyl phosphate, to aspartate semialdehyde. All microorganisms examined to date, have only a single ASADH encoded by *asd*.

Property	Aspartokinase I	Aspartokinase II	Aspartokinase III
Structural gene	*dapG*	*lysC*	?
Mol. wt	≥250 000	122 900	≈120 000
Subunit mol. wt	42 900	43 700 (α) 17 700 (β)	?
Feedback inhibitor(s)	*meso*-diaminopimelate	L-Lysine	L-Lysine and L-Threonine
Inhibition kinetics	Noncompetitive	Noncompetitive	Competitive
Corepressor of enzyme synthesis	None known	L-Lysine	L-Threonine
Relative enzyme level in rich and minimal media	Rich ≈ minimal	Rich << minimal	Rich << minimal
Change in enzyme level in stationary phase	None	Decreases	Decreases

Table 5.1

Aspartokinase isozymes of *B. subtilis*

Reproduced from Paulus (1993).

The regulation of ASADH by amino acids appears to vary widely in different bacteria. ASADH activity in *E. coli* is not inhibited by lysine, although the synthesis of the enzyme is subject to repression by lysine. Additionally threonine and methionine may also participate with lysine in a multivalent repression of ASADH (Boy and Patte, 1972). However, in *B. sphaericus* (Bartlett and white, 1996), *Streptomyces akiyoshiensis* (Le *et al.*, 1996) and *C. glutamicum* (Cremer *et al.*, 1988), ASADH is not influenced in its specific activity or in its synthesis by any of the aspartate amino acids (AAA).

5.1.4 The biosynthetic branch leading to lysine

The aspartate pathway splits after the synthesis of aspartate semialdehyde (Figure 5.1), with one branch leading to the biosynthesis of DAP and lysine, and the other leading to the biosynthesis of threonine (and isoleucine), and methionine (Chapter 5.1.5). The production of lysine from aspartate semialdehyde is a multi-step process for which three different pathways are known in prokaryotes (Schrumpf *et al.*, 1991). These are:

i) The acetylase variant, involving acetylated intermediates, and is the pathway found in *B. subtilis,* and the majority of other *Bacillus* species (Paulus, 1993).

ii) The succinylase variant, in which succinyl residues are used instead of acetyl, as the blocking group and is the route found in *E. coli* (Weinberger and Gilvarg, 1970). Together the acetylase and the succinylase variants are known as the epimerase pathway, and involve seven enzymatic steps.

iii) The dehydrogenase variant is present in corynebacterium, for example *Corynebacterium glutamicum* (Schrumpf *et al.*, 1991), and a few *Bacillus* species including *B. sphaericus* and *B. pasteurii* (Paulus, 1993). This pathway is a four step process, in which four of the reactions of the seven-step epimerase pathway are replaced by a single reaction catalysed by diaminopimelate dehydrogenase (Figure 5.2).

5.1.4.1 Dihydrodipicolinate synthase

The first step in DAP and lysine biosynthesis in all bacteria (independent of the pathway variant used (Chapter 5.1.4)), is catalysed by dihydrodipicolinate synthase (DHPS) (Figure 5.2). A putative DHPS was identified as part of the present study

(Chapter 4.2.2.2). DHPS is coded for by *dapA* in both *E. coli* and *B. subtilis*, and their polypeptides show a high degree of sequence similarity. This enzyme as with many of the lysine biosynthetic enzymes, is differently regulated in different bacteria. In *B. subtilis* and *C. glutamicum* DHPS is not subject to feedback inhibit by the end products lysine or DAP (Yamakura *et* al., 1974; Cremer *et al.*, 1988), whereas the corresponding enzyme in *E. coli*, is inhibited by lysine (Yugari and Gilvarg, 1965). This difference in the regulation of DHPS in *E. coli* and *B. subtilis* is important and allows DHPS of *B. subtilis* to function in the biosynthesis of DAP and dipicolinate during sporulation even at high intracellular concentrations of lysine (Paulus, 1993).

5.1.5 The biosynthetic branch leading to threonine and methionine

The first step in the committed biosynthesis of threonine and methionine is the production of L-homoserine, which is catalysed by homoserine dehydrogenase (Figure 5.1). From this point the pathway once more splits, with one route resulting in the biosynthesis of threonine (and isoleucine), and the other methionine (Figure 5.3). The mechanism of threonine biosynthesis will not be discussed further in this report.

5.1.5.1 Methionine biosynthesis and cystathionine γ-synthase

There are two known mechanisms for methionine biosynthesis in bacteria. The route leading to the biosynthesis of methionine in both *B. subtilis* and *E. coli* requires four enzymes for the conversion of L-homoserine to L-methionine, and proceeds via cystathionine (Figure 5.3). In these species the formation of cystathionine is catalysed by cystathionine γ-synthase (CTGS) and a putative CTGS was identified during the present study (Chapter 4.2.2.2). In coryneform bacteria, such as *Brevibacterium flavum*, there is an alternative route for methionine biosynthesis, in addition to the pathway with cystathionine as an intermediate (Figure 5.3). This mechanism utilises the enzyme O-acetylhomoserine sulphydrylase, to catalyse the direct conversion of O-acetylhomoserine to homocysteine (Ozaki and Shiio, 1982). Interestingly, in spite of the presence of CTGS and cystathionine β-lyase (CTBL) (Figure 5.3) in coryneform bacteria these enzymes do not play a critical role in methionine biosynthesis (Ozaki and Shiio, 1982).

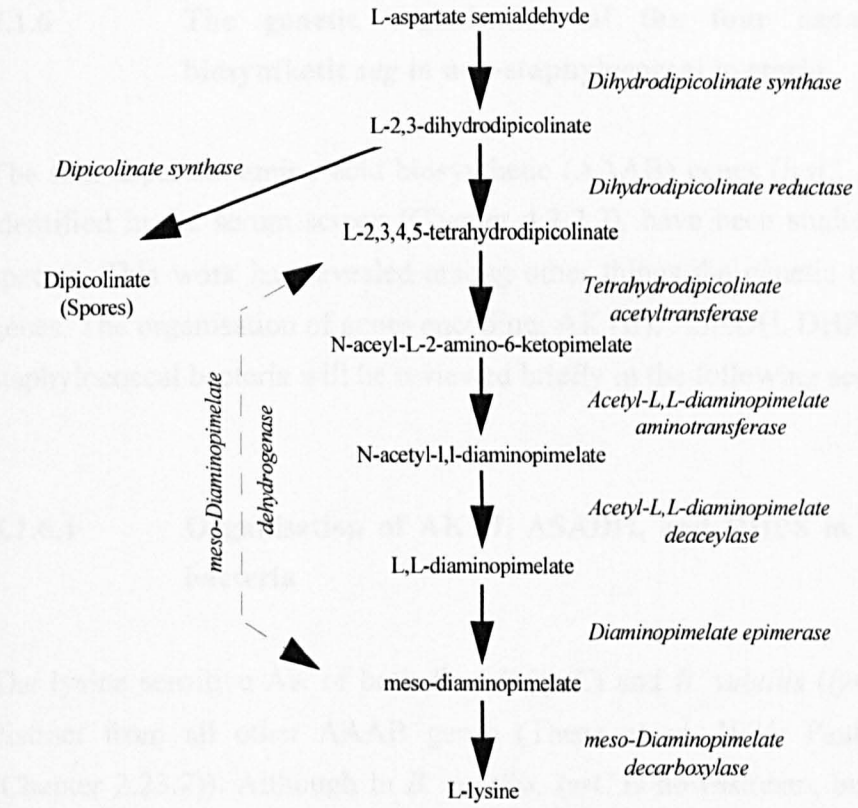


Figure 5.2

The biosynthesis of lysine in *B. subtilis*

The reaction, indicated by dotted arrows, occurs in coryneform bacteria and in a few *Bacillus* species but not in *B. subtilis*. Adapted from Paulus (1993).

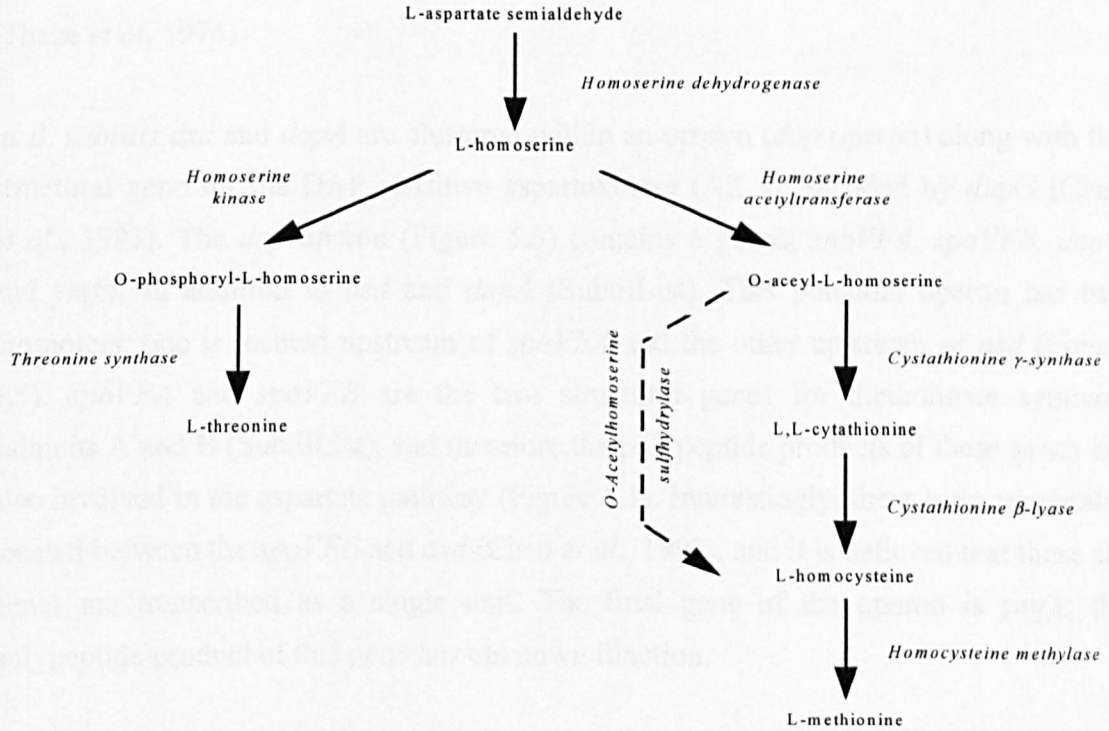


Figure 5.3

The biosynthesis of threonine and methionine in *B. subtilis*

The reaction, indicated by dotted arrows, occurs in coryneform bacteria. Adapted from Paulus (1993).

5.1.6 The genetic organisation of the four aspartate amino acid biosynthetic *seg* in non-staphylococcal bacteria

The four aspartate amino acid biosynthetic (AAAB) genes (*lysC, asd, dapA* and *yjcI*), identified in the serum screen (Chapter 4.2.2.2), have been studied in many bacterial species. This work has revealed among other things the genetic organisation of these genes. The organisation of genes encoding; AK (II), ASADH, DHPS and CTGS in non-staphylococcal bacteria will be reviewed briefly in the following sections.

5.1.6.1 Organisation of AK II, ASADH, and DHPS in non-staphylococcal bacteria

The lysine sensitive AK of both *E. coli* (*lysC*) and *B. subtilis* (*lysC*) are found at loci distinct from all other AAAB genes (Theze *et al*, 1974; Paulus, 1993; SubtiList (Chapter 2.23.2)). Although in *B. subtilis, lysC* is downstream, but the other side of a rho-independent transcriptional terminator, of a short ORF (24 amino acids), which is proposed to be the AK II attenuator, also known as the *lysC* leader peptide (Figure 5.4). In *E. coli*, ASADH (*asd*) and DHPS (*dapA*) are found at different chromosomal loci. *asd* like *lysC* is located away from all other AAAB genes, although *dapA* is found clustered with one other lysine biosynthetic gene, *dapE* encoding DAP desuccinylase (Theze *et al*, 1974).

In *B. subtilis asd* and *dapA* are clustered within an operon (*dap* operon) along with the structural gene for the DAP sensitive aspartokinase (AK I), encoded by *dapG* (Chen *et al.*, 1993). The *dap* operon (Figure 5.5) contains 6 genes; *spoVFA, spoVFB, dapG* and *ymfA,* in addition to *asd* and *dapA* (SubtilList). This potential operon has two promoters, one is located upstream of *spoVFA* and the other upstream of *asd* (Figure 5.5). *spoVFA* and *spoVFB* are the two structural genes for dipicolinate synthase subunits A and B (SubtilList), and therefore the polypeptide products of these genes are also involved in the aspartate pathway (Figure 5.1). Interestingly, there is no terminator located between the *spoVFB* and *asd* (Chen *et al.*, 1993), and it is believed that these six genes are transcribed as a single unit. The final gene of the operon is *ymfA*, the polypeptide product of this gene has unknown function.

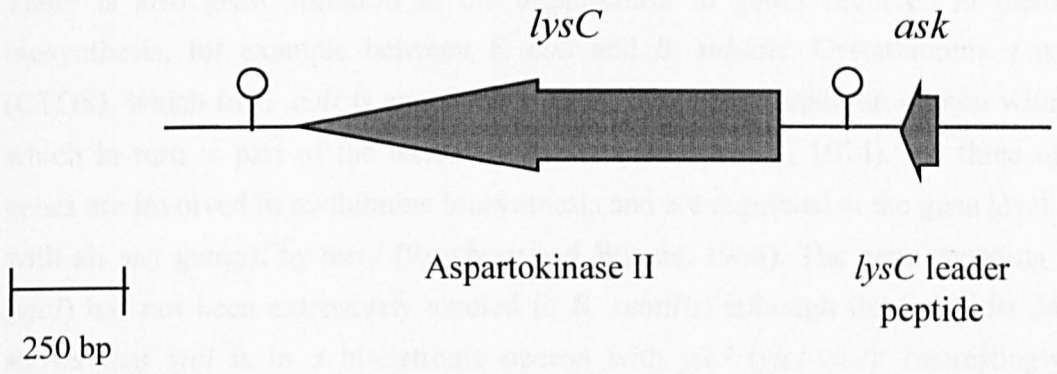


Figure 5.4
The location of the lysine sensitive AK (AK II) of *B. subtilis*

Reproduced from SubtiList.

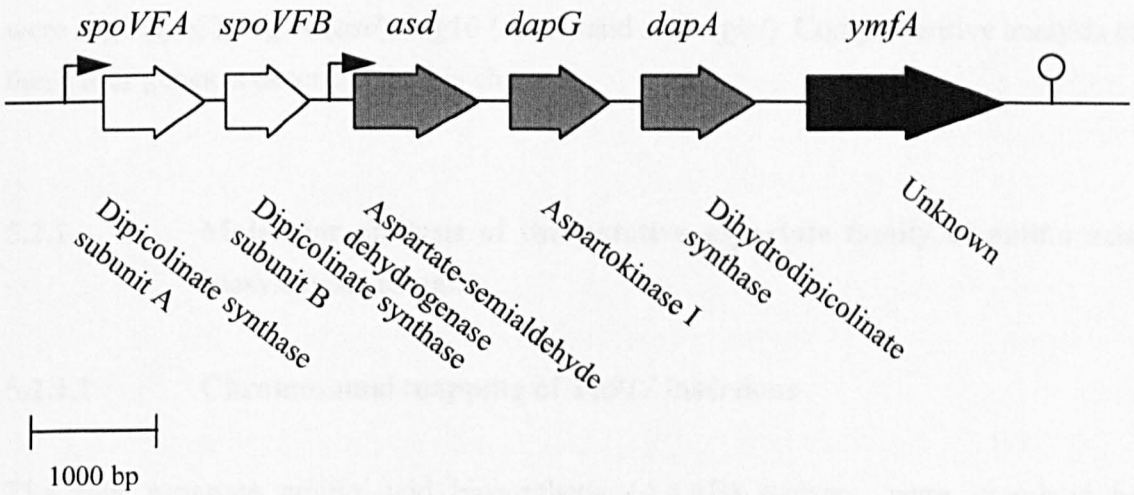


Figure 5.5
The *dap* operon of *B. subtilis*, showing the organisation of AK I (*dapG*), ASADH (*asd*) and DHPS (*dapA*)

Reproduced from SubtiList.

5.1.6.2 Organisation of cystathionine γ-synthase in non-staphylococcal bacteria

There is also great variation in the organisation of genes involved in methionine biosynthesis, for example between *E. coli* and *B. subtilis*. Cystathionine γ-synthase (CTGS), which in *E. coli* is encoded by *metB*, is located within an operon with *metL*, which in turn is part of the *metJBLF* cluster (Theze *et al.*, 1974). All three of these genes are involved in methionine biosynthesis and are regulated at the gene level (along with all *met* genes), by *metJ* (Rowbury and Woods, 1966). The gene encoding CTGS (*yjcI*) has not been extensively studied in *B. subtilis*, although the SubtiList database shows that *yjcI* is in a bi-cistronic operon with *yjcJ* (*yjcI-yjcJ*). Interestingly, *yjcJ* encodes the enzyme cystathionine β-lyase (CTBL) which catalyses the next step in methionine synthesis, following the CTGS catalysed conversion of succinyl homoserine to cystathionine (Figure 5.3).

5.2 Results

As part of the screening process for sera expressed genes, four putative aspartate amino acid biosynthetic (AAAB) genes were identified (Chapter 4.2.2.2 and 4.2.6). These were *seg*1 (*lysC*), *seg*24 (*asd*), *seg*10 (*dapA*) and *seg*8 (*yjcI*). Comprehensive analysis of these four genes is described in this chapter.

5.2.1 Molecular analysis of the putative aspartate family of amino acid biosynthesis mutants

5.2.1.1 Chromosomal mapping of Tn*917* insertions

The four aspartate amino acid biosynthetic (AAAB) mutants, were examined by Southern blot analysis in order to map the site of the transposon insertion. Chromosomal DNA purified from each strain (Chapter 2.17.1.1) was digested with *Eco*RI or *Xba*I, and then analysed by Southern blot analysis as described in Chapter 2.20. The blot was probed with a DIG labeled 8kb *Sal*I / *Eco*RI fragment from *seg*10 (pMDW10). The probe covers the region of the Tn*917* including the *lacZ* reporter gene. The probe is expected to hybridise with homologous DNA fragments in the chromosome of the mutants containing the transposon. For each of the four mutants a single fragment hybridised, indicating a single transposon insertion (results not shown).

5.2.1.2 Auxotrophy experiments

Auxotrophy experiments were carried out using chemically defined media (CDM), designed for the analysis of strains potentially auxotrophic for the aspartate family of amino acids (Chapter 2.1.1.3.1). In total 16 different CDMs were prepared, although only 10 different CDMs were used in the liquid assays (Table 5.2).

5.2.1.2.1 Solid media auxotrophy experiments

Initially auxotrophy experiments were carried out on CDM agar (Table 5.2). Results showed that *seg*1 (*lysC*) and *seg*10 (*dapA*) required lysine along with one other amino acid (threonine, methionine or isoleucine) for growth (Table 5.3). *seg*24 (*asd*) was auxotrophic for lysine, threonine and methionine for growth, while *seg*8 (*yjcI*) had a need for only methionine (Table 5.3). The auxotrophy experiments on solid media however, proved difficult to score as the fewer aspartate amino acids (AAA) present the slower the growth rate, and the lower the overall growth yield.

5.2.1.2.2 Liquid media auxotrophy experiments

Liquid CDM assays were used in an attempt to determine the exact auxotrophic requirements of the *seg* strains. This was achieved by following the growth kinetics of the AAAB strains (Figure 5.6-5.10), in selected CDM media (Table 5.2). The inclusion of isoleucine into the growth medium was found to make no difference to the growth kinetics of any of the mutant strains or of the wild type strain (results not shown), and thus was not included in any of the CDM. Furthermore, the addition of threonine alone to CDM (CDM 2) did not restore growth to any of the AAAB mutant strains (Figures 5.7-5.10). However, its addition did increase both the growth rate and yield of 8325-4, compared to CDM lacking all 4 AAA (result not shown).

The wildtype strain 8325-4 grew in the absence of lysine, threonine and methionine, and as for the AAAB mutant strains, the less amino acids included in the media, the slower the growth rate (Figure 5.6). *seg*1 (*lysC*) only grew in CDM containing lysine. When lysine alone was included in the growth media (CDM 1), growth to the same final A_{600} as in the presence of lysine, threonine and methionine (CDM 11) was observed. However, an increased lag period was witnessed in CDM 1, compared with media containing lysine, threonine and methionine (CDM 11) (Figure 5.7). *seg*24 (*asd*)

CDM No.	Amino acids
1*	+ lysine
2*	+ threonine
3*	+ methionine
4*	+ isoleucine
5	+ lysine + threonine
6	+ lysine + methionine
7	+ lysine + isoleucine
8	+ threonine + methionine
9	+ threonine + isoleucine
10	+ methionine + isoleucine
11*	+ lysine + threonine + methionine
12*	+ lysine + threonine + isoleucine
13*	+ lysine + methionine + isoleucine
14*	+ threonine + methionine + isoleucine
15*	- all four aspartate amino acids
16*	+ lysine + threonine + methionine + isoleucine

Table 5.2
Aspartate amino acid composition of CDM 1-16

For basic CDM composition see Table A.1.1.2. Note CDM 16 is identical to CDM described in Chapter 2.1.1.3.

* Denotes CDM used for both solid and liquid culture experiments. Other CDM, were only used for solid media studies.

	Strain				
CDM	8325-4	*seg*1 (*lysC*)	*seg*24 (*asd*)	*seg*10(*dapA*)	*seg*8 (*yjcI*)
1. (+ K)	+	-	-	-	-
2. (+ T)	+	-	-	-	-
3. (+ M)	+	-	-	-	+
4. (+ I)	+	-	-	-	-
5. (+ K, T)	+	+	-	+	-
6. (+ K, M)	+	+	-	+	+
7. (+ K, I)	+	+	-	+	-
8. (+ T, M)	+	-	-	-	+
9. (+ T, I)	+	-	-	-	-
10. (+ M, I)	+	-	-	-	+
11. (+ K, T, M)	+	+	+	+	+
12. (+ K, T, I)	+	+	-	+	-
13. (+ K, M, I)	+	+	-	+	+
14. (+ T, M, I)	+	-	-	-	+
15. (- All 4AA)	+	-	-	-	-
16. (+ All 4AA)	+	+	+	+	+

Table 5.3

Growth of *seg* mutants on CDM agar

Strains were inoculated onto the 16 types of CDM in triplicate, and incubated aerobically at 37 °C for 48 hours. Plates were then scored for growth: +, growth; -, no growth. The results are representative of at least two separate experiments.

K- lysine, T - threonine, M- methionine, I - isoleucine.

required lysine, threonine and methionine for growth (Figure 5.8). Interestingly however, *seg*24 (*asd*) does not appear to be auxotrophic for isoleucine, which must therefore be synthesised from threonine. *seg*10 (*dapA*) required lysine for growth, however there was an extended lag in the absence of methionine and threonine (Figure 5.9). Finally *seg*8 (*yjcI*) was unable to grow without methionine. An extended lag was seen without the inclusion of lysine or threonine, although a similar growth yield is reached (Figure 5.10).

5.2.1.3 Confirmation of the *lysC* mutant phenotype

The observation that *seg*1 (*lysC*) was auxotrophic for only lysine, and not lysine, threonine and methionine, implied the presence of at least one other functional aspartokinase (AK). More than one AK however, does not explain lysine auxotrophy especially as in *B. subtilis* a strain deficient at the *lysC* locus has a wild type phenotype (Zhang *et al.*, 1990; Chan and Paulus, 1998). Initial database searching revealed that the position of the Tn*917* insertion in *seg*1 was such that it is plausible for a near complete enzyme to be translated as the transposon has inserted at amino acid 386 leaving the mutant enzyme only 22 amino acids short of the holo enzyme (Figure A.2.1a). It was therefore necessary to confirm the phenotype of the *lysC* mutant (*seg*1). This was done using *seg*25 (*lysC*). *seg*25 (*lysC*) has an insertion which would result in at best a truncated protein of 168 amino acids (out of 408) (Figure A.2.1k). *seg*25 (*lysC*) demonstrated identical auxotrophy and growth kinetics to *seg*1 (*lysC*) in all CDM media (results not shown). Therefore the discovery that an *S. aureus* strain deficient at the *lysC* locus was a lysine auxotroph was upheld.

5.2.2 DNA sequence analysis of the lysine biosynthetic *seg* loci

The aspartate amino acid biosynthetic enzymes encoded by *lysC*, *asd* and *dapA*, were analysed in their role in the context of lysine biosynthesis. Importantly however, it should be noted that AK II (*lysC*) and ASADH (*asd*) are enzymes of the common pathway, and thus are involved in the biosynthesis of methionine, threonine and isoleucine, in addition to that of lysine.

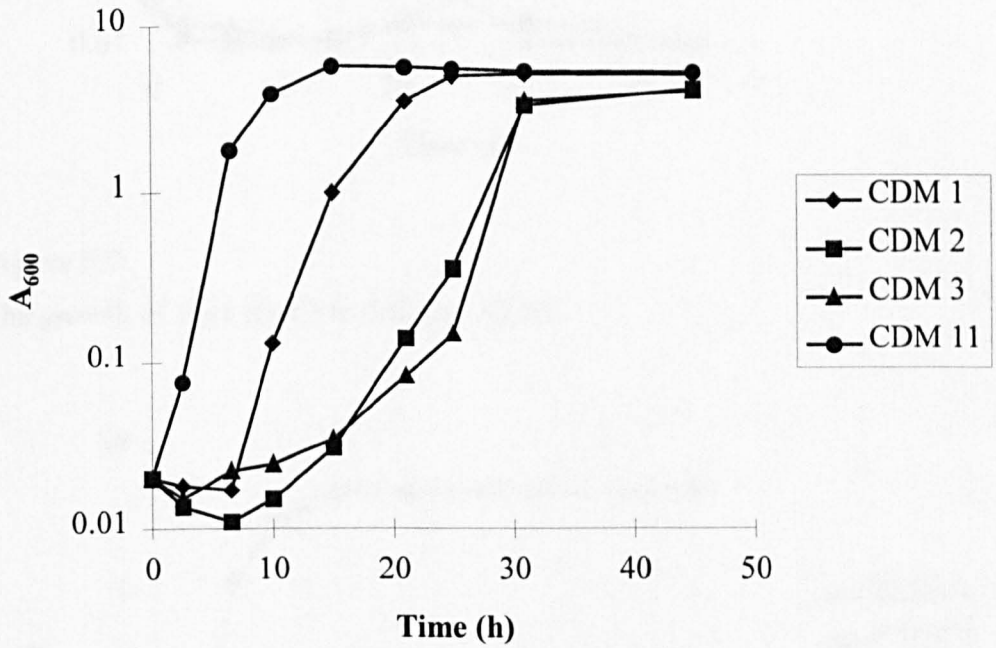


Figure 5.6

The growth of *S. aureus* 8325-4 in different CDM

S. aureus 8325-4 was grown at 37 °C in CDM 1 (+ K), CDM 2 (+ T), CDM 3 (+ M), and CDM 11 (+ K, T, M), under aerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 45 h. The results are representative of two separate experiments.

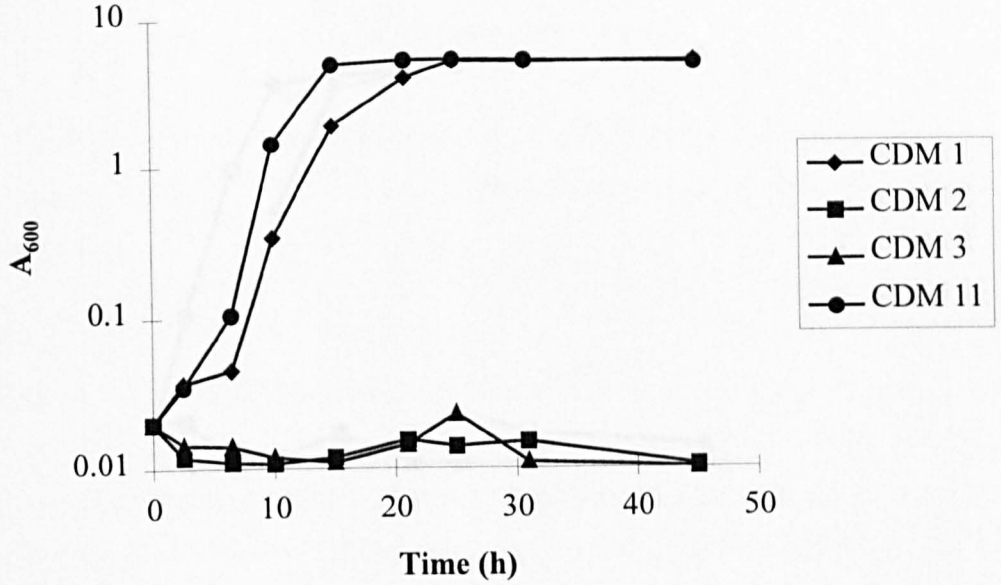


Figure 5.7
The growth of *seg*1 (*lysC*) in different CDM

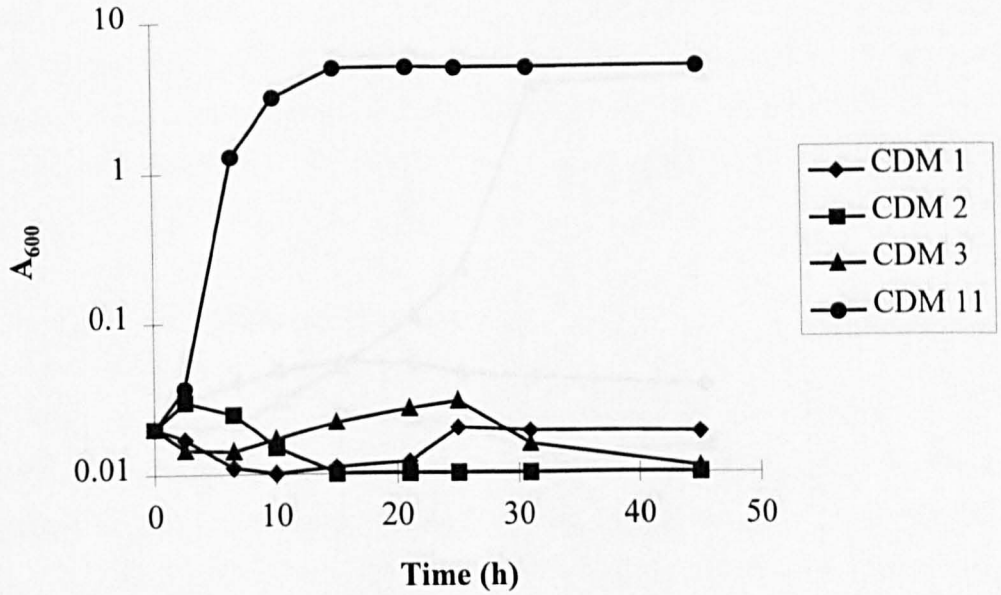


Figure 5.8
The growth of *seg*24 (*asd*) in different CDM

*seg*1 and *seg*24, were grown at 37 °C in CDM 1 (+ K), CDM 2 (+ T), CDM 3 (+ M), and CDM 11 (+ K, T, M), under aerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 45 h. The results are representative of two separate experiments.

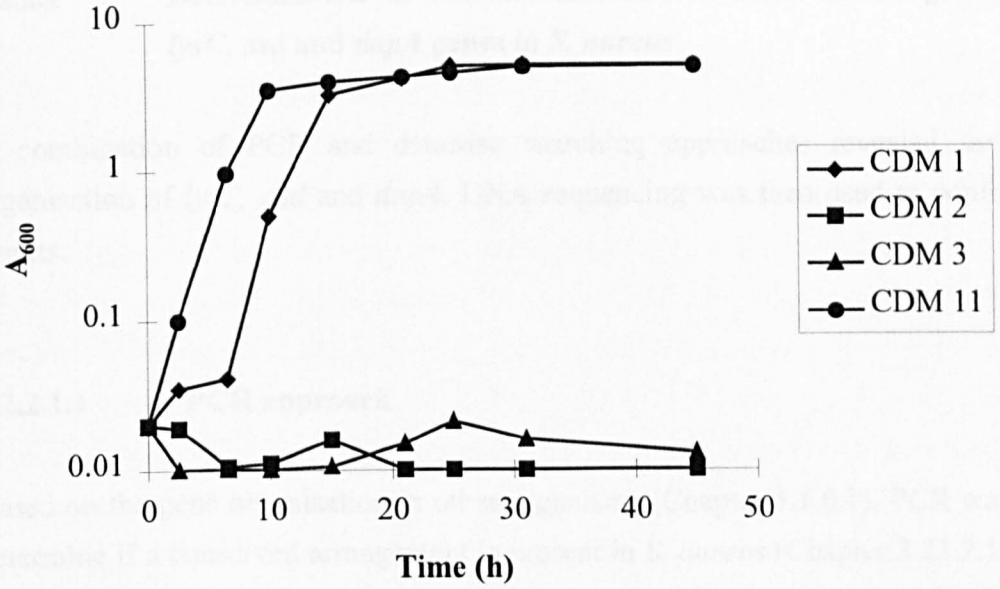


Figure 5.9

The growth of *seg*10 (*dapA*) in different CDM

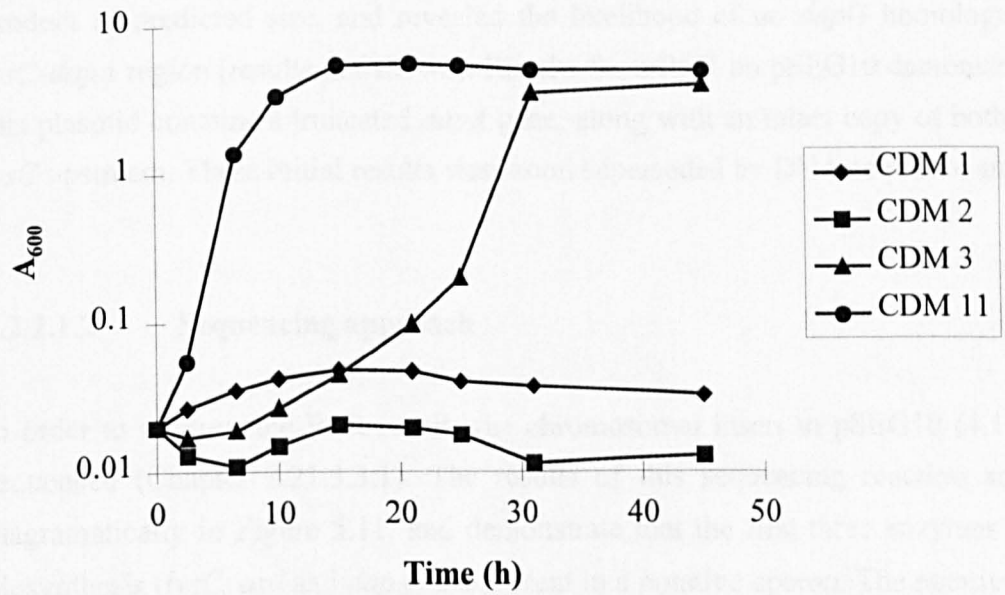


Figure 5.10

The growth of *seg*8 (*yjcI*) in different CDM

*seg*10 and *seg*8, were grown at 37 °C in CDM 1 (+ K), CDM 2 (+ T), CDM 3 (+ M), and CDM 11 (+ K, T, M), under aerobic conditions with shaking (250 rpm). Growth was measured as A_{600} for 45 h. The results are representative of two separate experiments.

5.2.2.1 Determination of the chromosomal location and organisation of *lysC, asd* and *dapA* genes in *S. aureus*

A combination of PCR and database searching approaches revealed the genetic organisation of *lysC, asd* and *dapA*. DNA sequencing was then used to confirm these results.

5.2.2.1.1 PCR approach

Based on the gene organisation in other organisms (Chapter 5.1.6.1), PCR was used to determine if a conserved arrangement is present in *S. aureus* (Chapter 2.21.2.1.1). PCR was carried out using both genomic and pSEG10 DNA (Chapter 4.2.4.2.2), as a template. All primer sequences are shown in Table A.1.2.1.

Reactions were designed to determine whether: i) *lysC* and *asd* are clustered; ii) *asd* and *dapA* are clustered; and iii) *lysC, asd* and *dapA* are clustered. Each reaction gave a product of predicted size, and revealed the likelihood of no *dapG* homologue in the *lysC-dapA* region (results not shown). Results from PCR on pSEG10 demonstrated that this plasmid contains a truncated *dapA* gene, along with an intact copy of both *asd* and *lysC* upstream. These initial results were soon superseded by DNA sequence analysis.

5.2.2.1.2 Sequencing approach

In order to confirm the PCR results the chromosomal insert in pSEG10 (4.1 kb) was sequenced (Chapter 2.21.3.3.1). The results of this sequencing reaction are shown diagramatically in Figure 5.11, and demonstrate that the first three enzymes in lysine biosynthesis (*lysC, asd* and *dapA*) are present in a putative operon. The putative operon, which may encode other downstream genes was named the *dap* operon (Figure 5.11).

5.2.2.1.3 Database construction and analysis to further study the *dap* operon of *S. aureus*

During the course of the project, the publicly available *S. aureus* genome databases (TIGR and NCTC 8325) were being released. Repeated searching over the duration of the study allowed the entire *dap* operon to be compiled (Chapter 2.23.3).

5.2.2.2 The *dap* operon of *S. aureus*

The consensus sequence of the 13 kb '*dap*' database (Figure 5.11) was analysed to identify potential open reading frames (ORFs). Using the computer program Gene Jockey II (Chapter 2.23.1), a putative 8 gene operon (*dap*) was identified. In addition to this a single ORF was discovered 3' of the *dap* operon and a further three upstream ORFs were detected. The sequences of these 12 ORFs were analysed (BLAST) to determine their relationship to sequences in the NCBI database (Chapter 2.23.2). The results of this analysis are shown in Figure 5.11 and Table 5.4. The 3' ORF shows homology to *cspC* (cold shock protein C), while the upstream ORFs show homology to *ykpA* (unknown, similar to ABC transporter), *yitL* (unknown) and *sphX* (SPHX protein precusor) (Figure 5.11; Table 5.4). Since these genes are only located close to the *dap* operon and are not apparently involved in lysine biosynthesis, no further attention will be paid to these genes.

The first three genes of the 8 gene *dap* operon as predicted (Chapter 5.2.2.1) are *lysC*, *asd* and *dapA*. Interestingly, the 4th, 5th and 8th genes also show homology to genes encoding lysine biosynthetic enzymes (LBE). These are *dapB* (dihydrodipicolinate reductase), *ykuQ* (tetrahydrodipicolinate acetyltransferase) and *lysA* (DAP decarboxylase). The final two genes of the *dap* operon (genes 6 and 7) show homology to *hipO* (hippurate hydrolase), and *dal* (alanine racemase), which do not have an apparent role in lysine biosynthesis (Figure 5.11; Table 5.4). In between *lysA* and *cspC*, is an intergenic region of approximately 650 bp (Figure 5.11). This region was found to contain several stem loop structures, one or more of which may function as rho-independent terminators (Chapter 2.23.1). However, further analysis is required to confirm function to these putative rho-independent terminator structures.

5.2.2.2.1 Sequencing of the entire *dap* operon of *S. aureus*

The *dap* database sequence has been compiled from four different strains of *S. aureus* (Figure 5.11). The organisation of the *dap* operon was confirmed in *S. aureus* strain 8325-4, by sequencing the entire *dap* locus from this strain. A PCR product was generated from genomic DNA of 5.7 kb, spanning from the 3' end of *dapA* through to *lysA* and a short region beyond this final gene (Figure 5.11), using primers MDW41 and MDW44 (Table A.1.2.1). The amplified fragment was then cloned into pCR®2.1-TOPO to produce pMTOPO9 (Chapter 2.22.2). The DNA insert was then sequenced (Chapter 2.21.3.3.2). The entire *dap* region's sequence, from 8325-4 is shown in Figure 5.12.

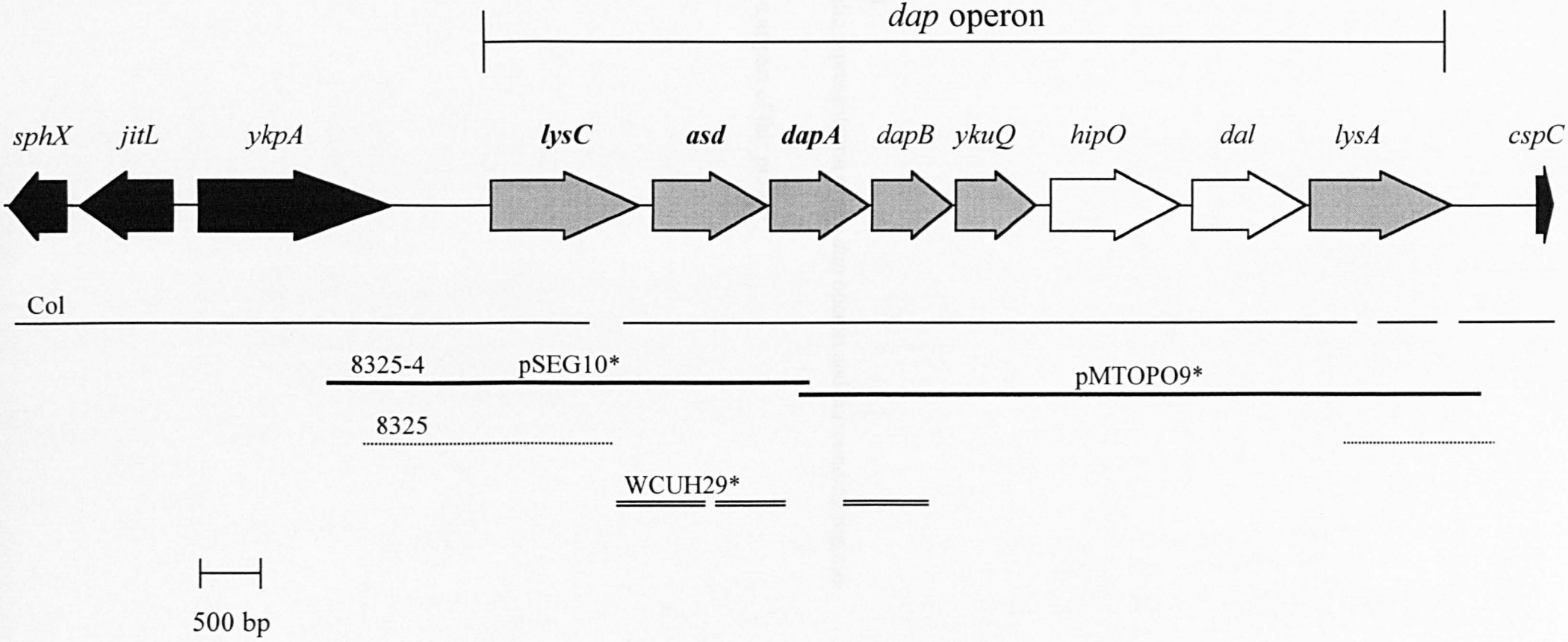

dap operon
sphX
jitL
ykpA
lysC
asd
dapA
dapB
ykuQ
hipO
dal
lysA
cspC
Col
8325-4
pSEG10*
pMTOPO9*
8325
WCUH29*
500 bp

Figure 5.11

Diagrammatic representation of the *dap* operon and surrounding regions

* Sequenced as part of this project.

Gene	Enzyme	% Identity AA / AA	Organism
lysC	Aspartokinase II	40 % 160 / 396	*Thermotoga maritima*
asd	Aspartate semialdehyde dehydrogenase	50 % 164 / 322	*Aquifex aeolicus*
dapA	Dihydrodipicolinate synthase	41 % 119 / 279	*Methanococcus jannaschii*
dapB	Dihydrodipicolinate reductase	43 % 73 / 166	*Aquifex aeolicus*
ykuQ	Unknown, similar to tetrahydrodipicolinate succinylase	60 % 142 / 233	*B. subtilis*
hipO	Hippurate hydrolase	36 % 123 / 333	*B. subtilis*
alr	Alanine racemase	50 % 91 / 308	*S. aureus*
lysA	Diaminopimelate decarboxylase	50 % 210 / 415	*B. subtilis*
cspC	Cold shock protein C	80 % 53 / 60	*S. aureus*
ykpA	Unknown, similar to ABC transporter	73 % 394 / 533	*B. subtilis*
yitL	Unknown	37 % 99 / 262	*B. subtilis*
sphX	SPHX protein precusor	36 % 62 / 171	*Synechococcus* sp. (Strain PCC 7942)

Table 5.4
Identification of genes from the *dap* database

% Identity – The number of identical amino acids, from those within the homologous region.

ykpA

P F R L Q E F F E G V N M N L V D W L R
CCCTTTCGTCTTCAAGAATTCTTTGAGGGCGTTAATATGAACCTTGTAGATTGGTTAAGA 60

Q Y A P E D E Q T E T F L R G F L G R M
CAATACGCACCAGAAGATGAGCAAACTGAAACATTTTTACGCGGCTTCTTAGGCCGTATG 120

L F S G E E V K K K A S V L S G G E K V
CTATTTAGTGGAGAAGAAGTTAAGAAAAAAGCTAGTGTACTTTCAGGTGGAGAAAAAGTA 180

R C M L S K M M L S S A N V L L L D E P
CGTTGTATGCTAAGTAAAATGATGTTATCAAGTGCAAACGTTCTTTTACTTGATGAACCA 240

T N H L D L E S I T A V N D G L K S F K
ACAAACCACTTAGACTTAGAAAGTATTACTGCTGTTAATGATGGACTTAAATCATTCAAA 300

G S I I F T S Y D F E F I N T I A N R V
GGTTCTATTATCTTTACTTCATATGACTTTGAATTTATCAATACAATCGCAAACCGTGTT 360

I D L N K Q G G V S K E I P Y E E Y L Q
ATCGATTTAAATAAACAAGGTGGCGTTTCAAAAGAAATTCCATATGAAGAATATTTACAA 420

E I G V L K *
GAAATCGGCGTTTTAAAA**TAA**ACATAATATGTAGAATTTCTTTTCGAAAATCTCTGTATT 480

GAGGCCCGACCAACTTGCACATTATTGTAAGTTGGCTTACTGTCACCTTCTATATTGGGA 540

CTTATTCCGCAACTTGTATTGTTTGTAGAAATTAGGAATCCAATTTCTCTATGTGTTGCA 600

Putative *lysC* leader peptide

M K K A S Y K V I F I Q S T T V N
CCGACAAGGA**TTG**AAAAAAGCTAGCTACAAAGTCATTTTCATTCAGTCAACTACTGTCAA 660

I T L Q G L G H *
TATAACATTGCAGGGCCTAGGGCATTGATTCATGTCCGAGACTATATTCAATTTAATAGA 720

CATTAACTGGGCGTCTCGATGCTAATATTTATAGCATCGAGGTGTCCTTTTTATTTTATT 780

TTTTAACACATCAGCATAGTGAACACCCCTTCTCTAAATTACAAATACAGTTATCAACGC 840

ATTAAAAGAGTACATTTATTTAATGTGCATATTGATTATGCTATAAGTATTTTGAATTTT 900

ATTTTTCAAAAAGTCTCATTATCACCACCTTTTAAAATATAAGCAACGCTATAACTTGAA 960

TTTTCAGGTTATTCCAAAAAATCTAAATTATCGGTTTACATTCATATTTGTCATATGTAT 1020

AATGAACACATACCAATTTAATAATATTTTGATGAGGCGCATCAATCATGAGTAAAGTTT 1080

AGATTACTGTCTGCTAACAGCTAAATTTGAAAGGGTGCGATGCCGAAGCAATTATAATAG 1140

CAGTTATAATTTGTTGGACTTTTTGGTTAAGAGCTGAGAGTTTGTCATTATTTAAAAATA 1200

ATGGAGTGCATCACTTGTATATAGATTAAGAGCAAGTTCGCATTCCGAACTTGTTCTTTT 1260

seg*22** (30,33) ***lysC **(α subunit)**

∇ M V T R S V L

TTTATTATTGTGTGCCCTTCCTAACAATT***AGGAGG***ATTAT**ATG**GTAACAAGAAGTGTGTT 1320

K F G G S S V S D F T K I K R I A E M L

GAAATTTGGCGGATCATCCGTCAGTGATTTTACAAAAATAAAAAGGATCGCTGAAATGTT 1380

***seg*14**

∇

K E R V N Q D E Q L I V V V S A M G N T

AAAGGAGCGAGTCAATCAAGATGAACAATTAATTGTCGTTGTAAGTGCTATGGGTAACAC 1440

T D Q L M T N V S T L T K A P K Q Q E L

AACAGATCAATTAATGACGAATGTATCAACCTTGACTAAAGCACCAAAACAACAAGAACT 1500

A L L L T T G E Q Q T V S Y L S M V L N

GGCATTATTATTGACAACCGGAGAGCAACAAACTGTATCTTATTTATCAATGGTATTAAA 1560

D I G M N A K A M T G Y Q A G I K T I G

TGATATCGGTATGAATGCCAAAGCAATGACTGGCTATCAAGCGGGTATTAAAACCATTGG 1620

H H L K S K I A Q I N P Q T F E Q A F Q

CCATCATTTAAAAAGTAAAATTGCTCAAATTAATCCTCAAACATTTGAACAAGCCTTTCA 1680

E N D I L V V A G F Q G I N E H Q E L T

AGAAAACGATATTTTAGTAGTTGCTGGATTTCAAGGCATCAATGAACATCAGGAATTAAC 1740

T L G R G G S D T T A V A L A V S N Q I

AACTTTAGGCAGAGGTGGTTCTGATACGACCGCTGTGGCACTTGCTGTTAGTAATCAAAT 1800

***seg*25** (50) ***seg*45** (48)

∇∇

P C E I Y T D V D G V Y A T D P R L L P

ACCTTGTGAAATTTATACCGACGTTGATGGTGTGTATGCCACTGACCCAAGACTTTTACC 1860

K A K R L D I V S Y E E M M E M S A L G

AAAGGCTAAACGACTAGACATCGTCTCATATGAAGAAATGATGGAAATGAGCGCTTTAGG 1920

A G V L E T R S V E L A K N Y N I P L Y

TGCTGGTGTACTTGAAACAAGAAGTGTTGAATTAGCTAAAAACTATAATATCCCTTTATA 1980

lysC **(β subunit)**

L G K T L S N V K G T W I M S N E E I L

TTTAGGAAAAACTTTATCGAACGTGA***AAGGA***ACATGGATT**ATG**TCAAATGAAGAAATATT 2040

E K K A V T G V A L D K H M M H V T I S

AGAGAAAAAAGCAGTTACTGGTGTGGCTTTGGATAAACATATGATGCATGTAACAATTAG 2100

Y P L P D N Q L L T Q L F T E L E E G A

TTATCCCCTACCTGACAATCAGCTACTTACCCAACTATTTACGGAACTTGAAGAAGGTGC 2160

V N V D M I S Q I V N L D G L Q L S F T

TGTAAATGTTGATATGATTTCACAAATCGTCAACTTGGATGGGCTACAACTATCCTTCAC 2220

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

GATTAAAGATAGTGATTTTCATCAAATTTCTATGATTCTTGAAACATTAAAGAATCAATA 2280

E A L A Y K I N E H Y V K I S L I G S G

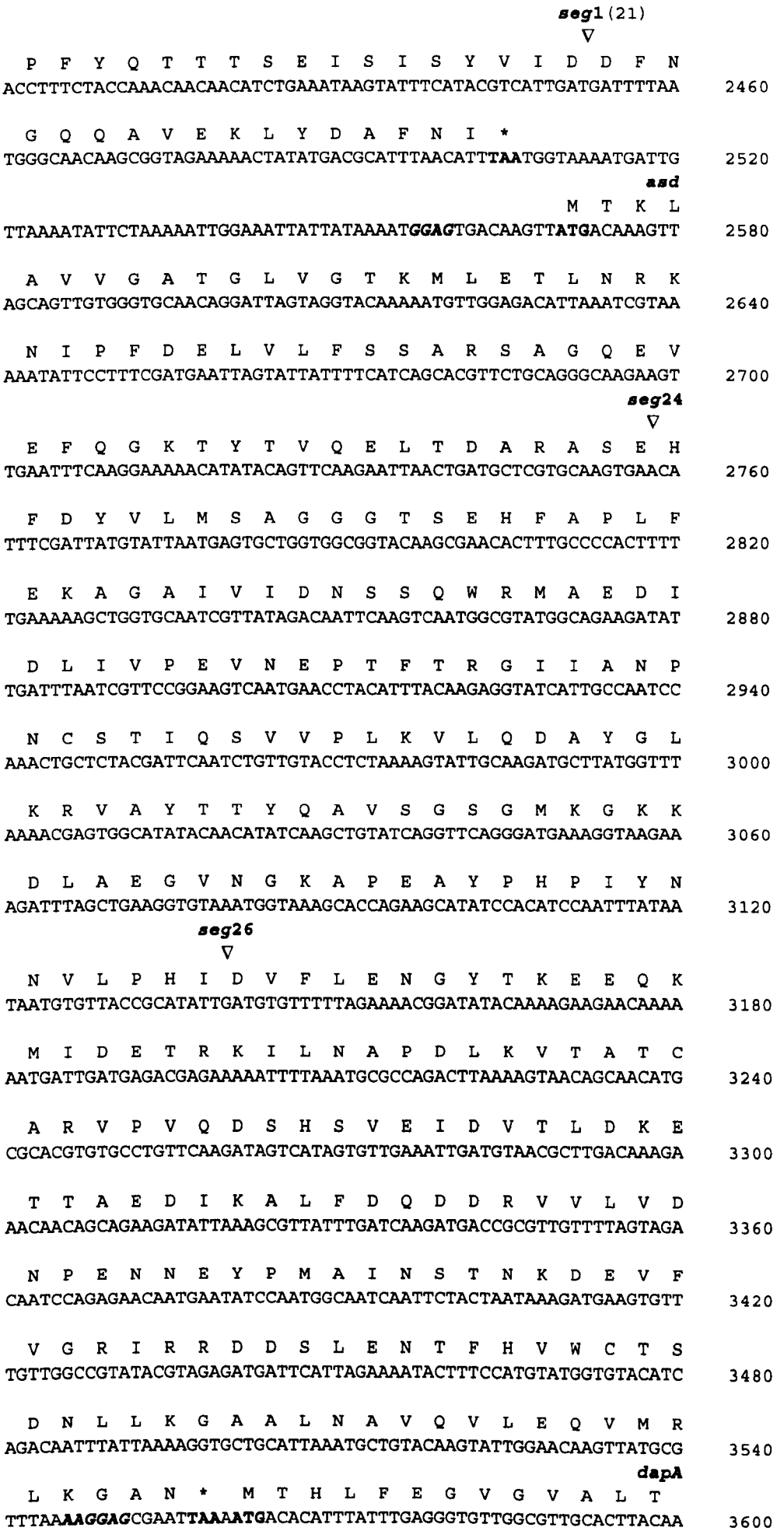
TGAAGCATTAGCTTATAAAATCAATGAGCATTATGTCAAAATTTCATTAATTGGCTCAGG 2340

***seg*13**

∇

M R D M S G V A S K A F L T L I E N N I

CATGCGTGATATGTCAGGTGTGGCATCAAAAGCATTTTTGACATTAATTGAAAATAATAT 2400


seg1(21)
P F Y Q T T T S E I S I S Y V I D D F N
ACCTTTCTACCAAACAACAACATCTGAAATAAGTATTTCATACGTCATTGATGATTTTAA 2460
G Q Q A V E K L Y D A F N I *
TGGGCAACAAGCGGTAGAAAAACTATATGACGCATTTAACATTTAATGGTAAAATGATTG 2520
asd
M T K L
TTAAAATATTCTAAAAATTGGAAATTATTATAAAATGGAGTGACAAGTTATGACAAAGTT 2580
A V V G A T G L V G T K M L E T L N R K
AGCAGTTGTGGGTGCAACAGGATTAGTAGGTACAAAAATGTTGGAGACATTAAATCGTAA 2640
N I P F D E L V L F S S A R S A G Q E V
AAATATTCCTTTCGATGAATTAGTATTATTTTCATCAGCACGTTCTGCAGGGCAAGAAGT 2700
seg24
E F Q G K T Y T V Q E L T D A R A S E H
TGAATTTCAAGGAAAAACATATACAGTTCAAGAATTAACTGATGCTCGTGCAAGTGAACA 2760
F D Y V L M S A G G G T S E H F A P L F
TTTCGATTATGTATTAATGAGTGCTGGTGGCGGTACAAGCGAACACTTTGCCCCACTTTT 2820
E K A G A I V I D N S S Q W R M A E D I
TGAAAAAGCTGGTGCAATCGTTATAGACAATTCAAGTCAATGGCGTATGGCAGAAGATAT 2880
D L I V P E V N E P T F T R G I I A N P
TGATTTAATCGTTCCGGAAGTCAATGAACCTACATTTACAAGAGGTATCATTGCCAATCC 2940
N C S T I Q S V V P L K V L Q D A Y G L
AAACTGCTCTACGATTCAATCTGTTGTACCTCTAAAAGTATTGCAAGATGCTTATGGTTT 3000
K R V A Y T T Y Q A V S G S G M K G K K
AAAACGAGTGGCATATACAACATATCAAGCTGTATCAGGTTCAGGGATGAAAGGTAAGAA 3060
D L A E G V N G K A P E A Y P H P I Y N
AGATTTAGCTGAAGGTGTAAATGGTAAAGCACCAGAAGCATATCCACATCCAATTTATAA 3120
seg26
N V L P H I D V F L E N G Y T K E E Q K
TAATGTGTTACCGCATATTGATGTGTTTTTAGAAAACGGATATACAAAAGAAGAACAAAA 3180
M I D E T R K I L N A P D L K V T A T C
AATGATTGATGAGACGAGAAAAATTTTAAATGCGCCAGACTTAAAAGTAACAGCAACATG 3240
A R V P V Q D S H S V E I D V T L D K E
CGCACGTGTGCCTGTTCAAGATAGTCATAGTGTTGAAATTGATGTAACGCTTGACAAAGA 3300
T T A E D I K A L F D Q D D R V V L V D
AACAACAGCAGAAGATATTAAAGCGTTATTTGATCAAGATGACCGCGTTGTTTTAGTAGA 3360
N P E N N E Y P M A I N S T N K D E V F
CAATCCAGAGAACAATGAATATCCAATGGCAATCAATTCTACTAATAAAGATGAAGTGTT 3420
V G R I R R D D S L E N T F H V W C T S
TGTTGGCCGTATACGTAGAGATGATTCATTAGAAAATACTTTCCATGTATGGTGTACATC 3480
D N L L K G A A L N A V Q V L E Q V M R
AGACAATTTATTAAAAGGTGCTGCATTAAATGCTGTACAAGTATTGGAACAAGTTATGCG 3540
dapA
L K G A N * M T H L F E G V G V A L T
TTTAAAAGGAGCGAATTAAAATGACACATTTATTTGAGGGTGTTGGCGTTGCACTTACAA 3600

T P F T N N K V N I E A L K T H V N F L
CCCCTTTTACAAATAACAAAGTTAATATTGAAGCTTTGAAAACACACGTTAATTTTTTAC 3660

L E N N A Q A I I V N G T T A E S P T L
TAGAAAATAATGCCCAAGCAATCATCGTTAATGGAACTACTGCTGAGAGCCCTACTTTAA 3720

T T D E K E R I L K T V I D L V D K R V
CAACAGATGAAAAAGAACGCATTCTAAAAACAGTTATTGATCTTGTAGATAAACGTGTTC 3780

P V I A G T G T N D T E K S I Q A S I Q
CTGTCATAGCAGGAACTGGCACTAATGATACTGAAAAGTCAATCCAAGCTTCAATCCAAG 3840

A K A L G A D A I M L I T P Y Y N K T N
CTAAAGCCTTAGGGGCTGATGCAATTATGTTAATTACGCCCTACTACAACAAAACGAACC 3900

Q R G L V K H F E A I A D A V K L P V V
AACGTGGTTTAGTCAAACACTTTGAAGCGATTGCAGATGCTGTGAAATTACCAGTCGTGC 3960

L Y N V P S R T N M T I E P E T V E I L
TGTACAATGTTCCTTCAAGAACGAACATGACAATTGAACCAGAAACTGTAGAAATATTAA 4020

***seg*10**
∇

S Q H P Y I V A L K D A T N D F E Y L E
GTCAACATCCTTATATAGTTGCTTTAAAAGATGCTACGAATGATTTTGAGTATTTAGAAG 4080

E V K K R I D T N S F A L Y S G N D D N
AAGTGAAAAAGCGCATTGATACAAATTCATTTGCATTATATAGTGGCAATGATGACAACG 4140

V V E Y Y Q R G G Q G V I S V I A N V I
TCGTCGAATACTATCAACGTGGCGGTCAAGGGGTTATCTCTGTTATTGCCAATGTCATTC 4200

P K E F Q A L Y D A Q Q S G L D I Q D Q
CTAAAGAATTTCAAGCGTTATACGATGCTCAACAAAGTGGATTAGATATTCAAGATCAAT 4260

F K P I G T L L S A L S V D I N P I P I
TTAAACCAATCGGCACACTGTTATCAGCTTTATCAGTTGATATTAACCCAATTCCTATTA 4320

K A L T S Y L G F G N Y E L R L P L V S
AAGCTCTAACAAGTTATTTAGGATTTGGAAATTATGAATTACGTCTACCATTGGTTAGCC 4380

L E D T D T K V L R E T Y D T F K A G E
TAGAAGATACAGATACTAAAGTGCTTCGTGAAACATATGACACATTTAAAGCGGGTGAAA 4440

dapB
N E * M N Q R V A R L A
ATGAG**TGA**AAATATTACTAATTGGCT***ATGGCG*C*AATG***AATCAACGCGTTGCTAGATTAGC 4500

E E K G H E I V G V I E N T P K A T T P
AGAAGAAAAAGGACATGAAATCGTTGGGGTCATTGAAAATACACCGAAAGCAACAACGCC 4560

Y Q Q Y Q H I A D V K G A D V A I D F S
ATATCAACAATATCAACATATTGCAGATGTTAAAGGTGCCGATGTTGCAATAGATTTTTC 4620

N P N L L F P L L D E D F H L P L V V A
AAATCCAAATCTGCTTTTCCCTTTATTAGATGAAGATTTTCATTTGCCATTAGTTGTGGC 4680

T T G E K E K L L N K L D E L S Q N M P
AACAACTGGCGAGAAAGAAAAACTACTTAATAAGTTAGATGAATTGAGTCAAAATATGCC 4740

V F F S A N M S Y G V H A L T K I L A A
TGTGTTTTTCAGCGCGAACATGAGTTATGGCGTTCATGCATTGACTAAAATTTTAGCAGC 4800

A V P L L D D F D I E L T E A H H N K K
TGCTGTTCCCCTACTTGATGATTTCGACATCGAATTGACTGAGGCACATCATAATAAAAA 4860

```
  V  D  A  P  S  G  T  L  E  K  L  Y  D  V  I  V  S  L  K  E
AGTAGATGCACCAAGTGGTACGTTAGAAAAATTGTATGATGTGATCGTATCTTTGAAAGA    4920

  N  V  T  P  V  Y  D  R  H  E  L  N  E  K  R  Q  P  Q  D  I
AAATGTAACACCTGTGTATGATAGACATGAATTAAATGAAAAACGCCAGCCACAAGATAT    4980

  G  I  H  S  I  R  G  G  T  I  V  G  E  H  E  V  L  F  A  G
TGGTATACATTCTATTCGTGGAGGTACGATTGTCGGTGAACATGAAGTTCTATTTGCTGG    5040

  T  D  E  T  I  Q  I  T  H  R  A  Q  S  K  D  I  F  A  N  G
CACTGATGAAACGATTCAAATCACGCATCGTGCACAATCAAAAGATATTTTTGCGAATGG    5100

  A  I  Q  A  A  E  R  L  V  N  G  P  N  G  F  Y  T  F  D  N
TGCAATACAAGCAGCAGAACGCTTAGTTAATGGGCCAAACGGCTTTTATACGTTTGACAA    5160
                                                        ykuQ
  L  *                                 M  V  Q  H  L  T  A  E  E
CTTATAAACATATTAAAGGAGATCGATTATTTTATGGTACAACATTTAACAGCTGAAGAA    5220

  I  I  Q  Y  I  S  D  A  K  K  S  T  P  I  K  V  Y  L  N  G
ATTATTCAATATATAAGTGATGCTAAAAAGTCTACACCAATAAAAGTATATTTAAATGGT    5280

  N  F  E  G  I  T  Y  P  E  S  F  K  V  F  G  S  E  Q  S  K
AATTTTGAAGGCATCACATATCCAGAAAGTTTTAAAGTATTTGGTTCAGAACAATCTAAA    5340

  V  I  F  C  E  A  D  D  W  K  P  F  Y  E  A  Y  G  S  Q  F
GTAATCTTTTGTGAAGCGGATGATTGGAAACCTTTTTACGAAGCATATGGTAGTCAATTC    5400

  E  D  I  E  I  E  M  D  R  R  N  S  A  I  P  L  K  D  L  T
GAAGATATAGAAATTGAAATGGATCGTCGCAATTCTGCTATTCCATTAAAAGACTTAACA    5460

  N  T  N  A  R  I  E  P  G  A  F  I  R  E  Q  A  I  I  E  D
AATACGAATGCACGAATTGAACCAGGTGCGTTTATTAGAGAACAAGCCATTATTGAAGAT    5520

  G  A  V  V  M  M  G  A  T  I  N  I  G  A  V  V  G  E  G  T
GGTGCTGTCGTTATGATGGGCGCAACAATTAATATTGGCGCAGTCGTTGGCGAAGGTACA    5580

  M  I  D  M  N  A  T  L  G  G  R  A  T  T  G  K  N  V  H  V
ATGATTGATATGAATGCTACTCTCGGTGGTCGTGCTACAACTGGTAAAAATGTACATGTA    5640

  G  A  G  A  V  L  A  G  V  I  E  P  P  S  A  S  P  V  I  I
GGGGCTGGCGCAGTATTAGCAGGTGTGATTGAACCCCCTAGTGCTTCACCGGTTATAATC    5700

  E  D  D  V  L  I  G  A  N  A  V  I  L  E  G  V  R  V  G  K
GAGGATGATGTATTAATCGGTGCAAATGCAGTTATTTTAGAAGGTGTACGTGTTGGTAAA    5760

  G  A  I  V  A  A  G  A  I  V  T  Q  D  V  P  A  G  A  V  V
GGTGCTATTGTTGCAGCTGGCGCGATTGTGACACAAGATGTACCAGCTGGTGCAGTTGTT    5820

  A  G  T  P  A  K  V  I  K  Q  A  S  E  V  Q  D  T  K  K  E
GCTGGTACACCTGCAAAAGTGATTAAGCAAGCTTCTGAAGTACAAGATACTAAAAAAGAG    5880

  I  V  A  A  L  R  K  L  N  D  *
ATTGTAGCAGCATTAAGAAAACTGAATGACTAGTGACGTCAACGTATAATAATTTCAAGG    5940


TTGAGATACGATTATGTCTCAACCCTATTGTTTAGCATTGAATAAACTTATGATCACAAA    6000
                                                        hipO
                                                      M  N
TTTCGAAATACGTTACACTATTTATATAGTAATTTAACACAAAGGATTGATAATTATGAA    6060

  E  L  E  F  V  T  K  H  R  R  H  L  H  Q  H  P  E  L  S  L
TGAATTAGAATTTGTTACGAAACATCGCCGTCATTTACATCAACATCCTGAATTAAGCTT    6120
```

```
  H  E  F  E  T  T  A  Y  I  K  A  F  L  D  S  L  N  I  K  Y
ACATGAATTTGAAACAACTGCTTATATTAAAGCGTTTTTAGATAGTTTAAATATTAAATA    6180

  D  C  P  L  E  T  G  V  I  A  Y  L  E  G  N  G  S  H  T  I
CGATTGCCCATTGGAAACTGGCGTCATTGCATACTTAGAAGGTAATGGCTCACATACGAT    6240

  A  Y  R  A  D  I  D  A  L  P  I  L  E  E  N  D  V  P  Y  R
AGCGTATAGAGCTGATATTGATGCGTTACCTATTTTAGAGGAAAATGATGTGCCTTATCG    6300

  S  Q  S  D  H  V  M  H  A  C  G  H  D  G  H  T  T  A  L  M
CAGTCAATCTGATCATGTGATGCATGCTTGTGGACATGATGGTCATACAACTGCATTAAT    6360

  L  F  V  Q  R  C  K  D  M  Q  D  A  G  Q  L  P  Q  N  V  V
GCTTTTTGTACAACGTTGCAAAGACATGCAAGATGCAGGTCAATTACCGCAAAATGTCGT    6420

  F  I  F  Q  P  A  E  E  T  G  G  G  A  N  R  L  I  K  A  G
TTTCATTTTCCAACCTGCAGAAGAAACTGGTGGCGGTGCAAATCGATTAATAAAAGCCGG    6480

  A  F  D  K  Y  P  I  E  A  V  F  G  I  H  V  N  P  F  A  D
TGCCTTTGATAAGTATCCAATTGAAGCGGTATTTGGTATTCATGTTAACCCATTTGCTGA    6540

  E  G  I  A  V  I  R  D  E  E  I  T  A  S  A  T  E  Y  R  F
TGAAGGCATTGCAGTGATAAGAGATGAAGAAATTACGGCCAGCGCAACAGAGTATCGCTT    6600

  F  L  T  G  L  S  S  H  V  A  D  K  E  Q  G  H  S  C  G  E
TTTCTTAACAGGCCTGTCAAGTCATGTTGCTGATAAAGAACAAGGTCATTCTTGTGGTGA    6660

  A  L  Q  H  V  L  T  Q  I  S  Q  I  Q  Q  F  H  L  N  G  L
AGCATTACAACATGTATTAACTCAAATATCACAAATTCAACAATTTCACCTTAACGGTTT    6720

  K  R  N  I  V  H  I  G  H  F  K  A  G  E  A  I  N  T  V  P
GAAACGAAATATTGTTCATATTGGTCATTTTAAAGCTGGTGAAGCGATTAACACTGTACC    6780

  S  N  G  Y  L  E  G  T  I  R  T  Y  D  I  D  D  L  T  I  V
AAGTAATGGCTATTTAGAAGGTACTATTCGTACATATGATATTGATGATTTAACAATCGT    6840

  K  N  Q  M  H  K  I  A  E  S  V  K  L  L  F  N  V  D  C  E
TAAAAATCAAATGCACAAGATAGCAGAAAGTGTCAAGCTTCTGTTTAATGTAGATTGTGA    6900

  V  K  F  A  E  G  Y  P  P  T  I  N  S  P  K  L  R  T  Q  I
AGTTAAATTTGCAGAAGGTTATCCCCCTACAATCAATAGTCCGAAATTACGTACTCAAAT    6960

  E  D  A  L  I  K  A  D  L  N  V  Y  D  K  P  T  P  F  L  F
AGAGGACGCCTTAATAAAAGCTGATTTAAATGTCTATGACAAACCAACGCCATTCTTATT    7020

  G  E  D  F  S  F  Y  G  Q  Q  L  A  P  A  Y  F  V  F  I  G
TGGGGAAGATTTTAGTTTTTATGGTCAACAACTAGCTCCAGCTTACTTTGTTTTTATAGG    7080

  T  R  N  E  D  K  G  F  V  T  G  L  H  T  S  H  L  N  F  D
AACACGAAATGAAGATAAAGGTTTTGTAACTGGTTTGCACACATCACATTTAAATTTTGA    7140

  E  K  V  L  I  N  V  I  N  F  Y  E  N  L  L  N  N  Y  K  E
TGAAAAAGTGTTAATAAACGTGATTAATTTTTACGAAAATTTATTAAATAATTACAAAGA    7200

  V  *
GGTGTAATACATTGACAGCAACATGGTCTGTAAATAAGAAAATATTTTTACAAAATGCAA    7260
                                                        dal
                          M  A  V  V  K  N  N  A  Y  H  Y
TCACAGTCAAAAACAATCAGCCATTAATGGCAGTTGTTAAAAATAATGCATATCACTATG    7320

  D  L  E  F  A  V  T  Q  F  I  H  A  G  I  D  T  F  S  T  T
ACCTAGAATTTGCTGTAACTCAGTTTATCCATGCAGGTATAGATACATTTAGCACAACAT    7380
```

S L R E A I Q I R Q L A P D A T I F L M
CACTACGAGAAGCAATTCAAATTAGACAACTTGCTCCAGATGCAACAATCTTTTTAATGA 7440

N A V Y E F D L V R E H Q I H M T L P S
ATGCAGTTTACGAGTTTGATTTAGTCCGTGAACATCAAATACACATGACTTTGCCGTCGT 7500

L T Y Y Y N H K N D L A G I H V H L E F
TGACATATTACTATAACCATAAAAATGATTTAGCTGGTATTCATGTTCACTTAGAATTTG 7560

E N L L H R S G F K D L N E I K E V L K
AAAATTTATTACATCGGTCTGGATTTAAAGATTTAAACGAAATTAAAGAAGTATTGAAAG 7620

D H H H N Q N A K M I I S G L W T H F G
ATCACCATCATAATCAAAATGCAAAAATGATTATTAGTGGTTTATGGACCCATTTTGGAT 7680

Y A D E F D V S D Y N V E R S Q W M E I
ATGCTGATGAATTCGATGTGTCAGATTATAATGTTGAACGTTCACAATGGATGGAAATTG 7740

V E A L L S E G Y Q F D L I H A Q N S A
TTGAAGCACTTTTATCTGAAGGTTATCAGTTCGACCTAATCCATGCTCAAAATAGTGCGA 7800

S F Y R E G Q I L L P H H T H A R V G I
GTTTTTATCGGGAAGGACAAATATTACTACCCCACCATACACATGCGCGTGTAGGTATTG 7860

A L Y G S R P Y S S L N Q H D I V Q S L
CGTTATACGGTTCAAGACCATATAGTTCACTGAATCAACATGATATAGTTCAGTCATTAA 7920

T L K A H V I Q V R E V Q A G D Y C G Y
CTTTAAAAGCACATGTTATTCAAGTGCGCGAAGTACAAGCTGGTGATTATTGCGGTTATA 7980

S F A F E V T K N N T K L A V V D I G Y
GCTTTGCCTTTGAAGTGACTAAAAACAATACAAAATTAGCTGTAGTTGATATCGGTTATG 8040

G D G I L R T R A K H E A L I N G K R Y
GCGATGGAATTTTAAGAACTCGTGCTAAACATGAAGCACTTATCAATGGTAAACGCTACC 8100

P I R A L M M S H M F V E V D G N V H A
CGATACGTGCATTAATGATGAGCCATATGTTTGTTGAAGTAGATGGCAATGTACATGCAC 8160

Q D E V I L Y N N D I R I D E Y T F K G
AAGATGAAGTTATTCTTTATAATAATGATATCCGCATCGATGAATATACCTTTAAAGGTG 8220

V G A N S E Q L S A M N H D S L K K E Y
TTGGTGCAAATTCTGAACAATTAAGTGCTATGAATCATGATTCTTTAAA***AAGGAG***TACA 8280

I S N D C * ***lysA***
M T V K Y N Q N G E L T M D G I S L
TTTCAA***ATG***ACTGT***TAA***ATATAATCAAAATGGCGAATTAACAATGGATGGTATTAGTTTA 8340

K T I A Q S F G T P T I V Y D E L Q I R
AAAACGATTGCACAAAGCTTTGGTACACCTACCATTGTTTATGATGAACTACAAATTAGA 8400

E Q M R R Y H R A F K D S G L K Y N I S
GAACAGATGCGCCGTTACCATCGCGCATTTAAAGATAGTGGATTAAAATACAATATTTCA 8460

Y A S K A F T C I Q M V K L V A E E D L
TACGCCTCAAAGGCATTTACTTGCATTCAAATGGTCAAACTTGTAGCTGAGGAAGATTTA 8520

Q L D V V S E G E L Y T A L E A G F E P
CAGTTAGATGTTGTTTCTGAAGGTGAATTATATACAGCTTTAGAAGCAGGTTTTGAACCG 8580

S R I H F H G N N K T K H E I R Y A L E
AGTCGCATCCATTTCCATGGTAACAATAAAACGAAACATGAAATTAGGTATGCTTTAGAA 8640

```
 N  N  I  G  Y  F  V  I  D  S  L  E  E  I  E  L  I  D  R  Y
AATAATATCGGTTATTTTGTTATAGATTCATTAGAAGAAATTGAATTAATAGACCGCTAT    8700

 A  N  D  T  V  Q  V  V  L  R  V  N  P  G  V  E  A  H  T  H
GCTAATGATACGGTTCAAGTTGTATTACGAGTTAATCCAGGTGTTGAAGCACATACACAC    8760

 E  F  I  Q  T  G  Q  E  D  S  K  F  G  L  S  I  Q  Y  G  L
GAATTTATTCAAACTGGGCAAGAAGATAGTAAGTTTGGATTATCAATTCAATATGGCTTA    8820

 A  K  K  A  I  D  K  V  Q  Q  S  K  H  L  K  L  K  G  V  H
GCTAAAAAAGCAATTGACAAAGTCCAACAATCTAAACACTTAAAATTAAAAGGTGTACAT    8880

 C  H  I  G  S  Q  I  E  G  T  E  A  F  I  E  T  A  K  I  V
TGTCATATTGGTTCACAGATTGAAGGTACAGAAGCATTTATTGAAACTGCTAAAATTGTT    8940

 L  R  W  L  K  E  Q  G  I  Q  V  E  L  L  N  L  G  G  G  F
TTACGTTGGCTTAAAGAGCAAGGCATTCAAGTTGAATTATTAAACCTTGGTGGTGGCTTT    9000

 G  I  K  Y  V  E  G  D  E  S  F  P  I  E  S  G  I  K  D  I
GGTATTAAATATGTTGAAGGTGACGAAAGTTTCCCTATCGAAAGTGGTATTAAAGATATT    9060

 T  D  A  I  K  S  E  I  K  V  L  G  I  D  A  P  E  I  G  I
ACAGACGCAATAAAATCCGAAATTAAAGTTCTAGGTATAGATGCACCAGAAATAGGTATT    9120

 E  P  G  R  S  I  V  G  E  A  G  V  T  L  Y  E  V  G  T  I
GAACCGGGACGATCAATTGTAGGTGAAGCTGGCGTTACTTTATATGAAGTTGGAACCATT    9180

 K  E  I  P  E  I  N  K  Y  V  S  I  D  G  G  M  S  D  H  I
AAAGAAATTCCAGAGATTAATAAATATGTTTCAATCGATGGCGGTATGAGTGATCATATC    9240

 R  T  A  L  Y  D  A  K  Y  Q  A  L  L  V  N  R  N  E  E  A
AGAACTGCACTTTATGACGCAAAGTATCAAGCATTGCTTGTTAATAGAAATGAAGAAGCA    9300

 D  D  S  V  T  I  A  G  K  L  C  E  S  G  D  I  I  I  K  D
GATGACAGTGTAACTATAGCTGGAAAATTATGTGAGTCTGGTGATATCATTATTAAAGAC    9360

 A  K  L  P  S  S  V  K  R  G  D  Y  L  A  I  L  S  T  G  A
GCTAAATTACCTTCATCAGTCAAACGTGGAGACTATCTTGCTATATTATCAACTGGTGCA    9420

 Y  H  Y  S  M  A  S  N  Y  N  Q  M  Q  K  P  S  V  F  F  L
TATCATTACTCTATGGCATCCAATTACAATCAAATGCAAAAGCCTTCTGTGTTTTTCTTA    9480

 K  D  G  K  A  R  E  V  I  K  R  Q  S  L  R  Q  L  I  I  N
AAAGATGGCAAAGCACGTGAAGTTATAAAGCGACAATCGTTAAGACAACTCATTATTAAT    9540

 D  T  K  *
GATACAAAATAAAAATAAACAAAGTAATCCCCGAAGCACAGAAAATTATACTGTGACTCG    9600


GGGAATTTTTATGCATTATAGATGACTGTAGAAAACATGCTTGGACATATACTC          9654
```

Figure 5.12

The nucleotide and deduced amino acid sequence of the *dap* operon and the surrounding regions of *S. aureus*

The putative ribosome binding sites are shown in italics, and the stop codons are shown in bold. ∇ represent Tn*917* insertion sites of *seg* genes. Numbers in brackets represent *seg* mutants with identical transposon insertion sites.

5.2.2.2.2 Transposon insertions in the *dap* operon in *S. aureus*

The *dap* operon is clearly a major lysine biosynthetic locus in *S. aureus*, since it encodes six of the nine enzymes required for the synthesis of lysine, not including isozymes (Figures 5.1 and 5.2). A total of nine different Tn*917* insertion sites in the *dap* operon were identified in the serum screen (Figure 5.13). This demonstrates that the *dap* operon is important in the serum model. Five of these transposons have inserted within *lysC* (AK II), two within *asd* (ASADH) and one within *dapA* (DHPS). Interestingly, one further transposon was found to have inserted 31 bp upstream of codon one (ATG) of *lysC* (Figure 5.12). This therefore provides evidence for the presence of upstream control regions for *lysC,* in *S. aureus* that are perhaps similar to that found in *B. subtilis*.

5.2.2.2.3 Analysis of the untranslated region upstream of *lysC* in *S. aureus*

As described in Chapter 5.1.4.2, preceding *lysC* in *B. subtilis* is an unusually extensive control region (untranslated leader region) of 330 bp, which is involved in the regulation of lysine synthesis. This region, termed the *lysC* leader region, has been analysed (Chen *et al.*, 1987), and shown to contain a number of potential control sites, including: a 24 amino acid polypeptide; and terminator, antiterminator structures (Chapter 5.1.4.2). The region upstream of *lysC* in *S. aureus* (859 bp) was examined for the potential presence of similar regulatory elements.

In *B. subtilis* a short ORF (72 bp) potentially encoding the *lysC* leader peptide is present upstream of *lysC* that is believed to be involved in an attenuation mechanism of *lysC* synthesis (Chen *et al.*, 1987). Analysis of the sequence upstream of *lysC* in *S. aureus* has revealed the presence of several putative ORFs including, a potential ORF of 75 bp (25 amino acids), that shows similarities to the leader peptide of *B. subtilis*. Firstly both putative leader peptides start with an alternative start codon, GTG (*B. subtilis*) and TTG (*S. aureus*). Secondly and most importantly both putative polypeptides have a dyad of lysine codons at positions two and three. In *B. subtilis* this pair of lysine codons is thought to have an important role in *lysC* regulation by lysine as under conditions of lysine deficiency, the ribosomes would be stalled, destabilising the secondary structure of the rho-independent terminator, thus allowing synthesis of the LBE enzyme. However, a BLAST search within the SubtiList database (Chapter 2.23.1), using the amino acid sequence of the proposed *lysC* leader sequence of *S. aureus*, failed to show any significant homology to that of *B. subtilis* (results not shown). Furthermore, no

putative Shine-Dalgarno sequence, could be identified immediately upstream of the putative start codon. This therefore creates doubt over whether this ORF is actually translated or not. The putative leader region was also analysed for the presence of stem-loops and potential rho-independent terminator (Chapter 2.23.1). This revealed an number of short repeats, which have the potential to form stem-loops. However, none of these structures appear to be strong candidates for terminator antiterminator structures (results not shown).

5.2.2.3 Searching for isozymes of LBE genes in *S. aureus* and *B. subtilis*

It has been proposed that a reason for the failure to identify mutations in the majority of LBE genes of *B. subtilis* is due to the presence of multiple isozymes, of these biosynthetic enzymes (Roten *et al.*, 1991). Interestingly, from searching the SubtiList database this does not appear to be the case (results not shown). It is known however that there are three isozymes of AK in *B. subtilis* (Chapter 5.1.3).

5.2.2.3.1 Identification of AK III (*yclM*) of *B. subtilis*

It is known that *B. subtilis*, has three isozymes of aspartokinase, although the gene encoding AK III has not been identified. On searching the *B. subtilis* database (SubtiList) with sequences of known AKs, three genes were identified. These were *dapG* (AK I), *lysC* (AK II) and *yclM*.

Analysis of *yclM* within SubtiList demonstrated that this gene encodes a protein of 454 amino acids with similarity to homoserine dehydrogenase (HDH). Two of the AK isozymes in *E. coli* are bifunctional proteins having both AK and HDH activity (AK I-HDH I, and AK II-HDH II) (Cohen and Saint-Girons, 1987). However, none of the *B. subtilis* AKs have HDH activity (Parsot and Cohen, 1988). Two pieces of evidence confirmed that *yclM* could not be a bifunctional protein formed by the fusion of genes encoding an AK and a HDH. Firstly, AK I-HDH I (*thrA*) of *E. coli* is 820 amino acids in length, and secondly yclM shows no significant sequence homology to HDH (*hom*) of *B. subtilis* (results not shown).

YclM shows greatest homology to AK I-HDH I, of *Serratia marcescens* (32 % over 171 amino acids). This is a bifunctional enzyme, but the region of homology with *yclM* only corresponds to the AK domain. Thus *yclM* is predicted to be AK III of *B. subtilis*.

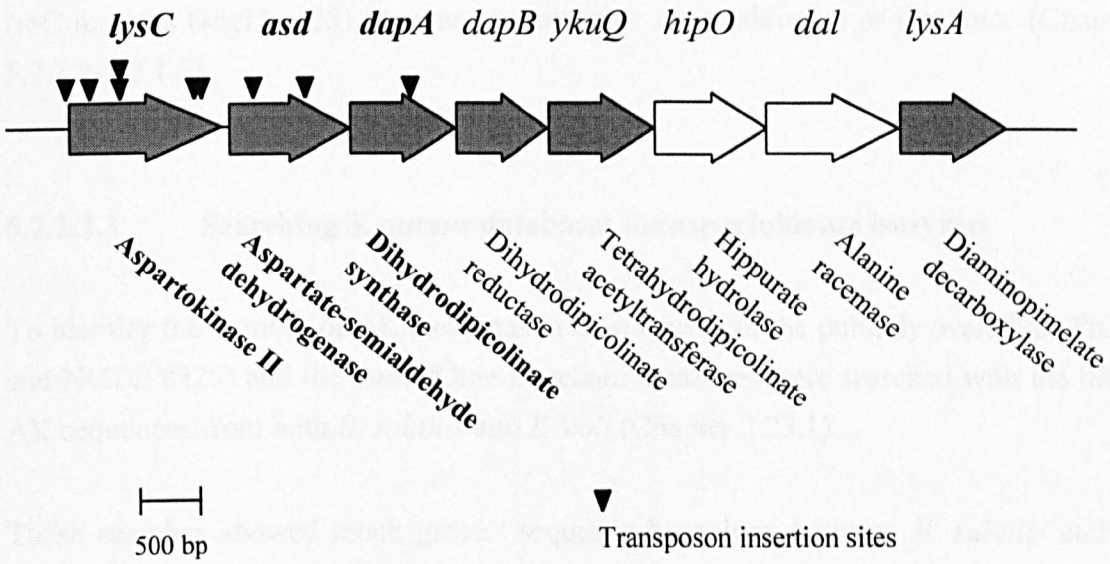


Figure 5.13
Diagram showing Tn*917* insertion sites in the *dap* operon

Tn*917* insertions from left to right: Aspartokinase II (*lysC*) control/leader region (putative): *seg*22 (*seg*30, *seg*33); Aspartokinase II (*lysC*): *seg*14, *seg*25 (*seg*50), *seg*45 (*seg*48), *seg*13, *seg*1 (*seg*21); Aspartate semialdehyde dehydrogenase (*asd*): *seg*24, *seg*26; Dihydrodipicolinate synthase (*dapA*): *seg*10.

5.2.2.3.2 Searching *S. aureus* databases for isozymes of LBE genes

During this study, it was discovered that *S. aureus* encodes two isozymes of DAP dehydrogenase. One copy is located within the *dap* operon (*lysA*), while the other (named *dapH*) is found clustered with *iucC* (aerobactin siderophore biosynthesis protein), and ORF256, which shows homology to hypothetical aldolases, in the order *iucC*-ORF256-*dapH* (results not shown). However, due to the absence of isozymes of LBE genes in *B. subtilis*, a large scale search for LBE isozymes in *S. aureus* was not undertaken. There are however likely to be at least two AK isozymes in *S. aureus*, as *lysC* mutants (*seg*1/*seg*25) were not auxotrophic for methionine or threonine (Chapter 5.2.1.2-5.2.1.3).

5.2.2.3.3 Searching *S. aureus* databases for aspartokinase isozymes

To identify the number of AK isozymes in *S. aureus* both the publicly available (TIGR and NCTC 8325) and the SmithKline Beecham databases were searched with the three AK sequences from both *B. subtilis* and *E. coli* (Chapter 2.23.1).

These searches showed much greater sequence homology between *B. subtilis* and *S. aureus,* than *E. coli* and *S. aureus* (results not shown). Searching with *B. subtilis* AK sequences identified homologues of *lysC* (AK II); 40 % over 397 amino acids, and *yclM* (AK III); 44 % over 171 amino acids, in *S. aureus*. Therefore *S. aureus* possesses at least two isozymes of AK, these being encoded *lysC* and *ask*. There is no evidence of a homologue of AK I, which in *B. subtilis* is the DAP sensitive aspartokinase (*dapG*). Searching with *E. coli* AK sequences also failed to identify a third AK isozyme (results not shown). Therefore it is unlikely that AK I exists in *S. aureus* as the SmithKline Beecham private databases are essentially complete.

5.2.2.3.4 Determination of the mechanism(s) of lysine biosynthesis in *S. aureus*

Broadly two alternative pathways exist for the biosynthesis of lysine. These are the four step epimerase pathway, and the one-step dehydrogenase pathway (Chapter 5.1.4; Figure 5.2). Previous to this study the mechanism for the biosynthesis of lysine in *S. aureus* had been poorly studied and thus it was not known which of the pathways (or both as in *C. glutamicum*), is responsible for lysine biosynthesis in *S. aureus*. The *dap*

operon of *S. aureus* encodes a protein with significant homology (60 % identity over 233 amino acids) to tetrahydrodipicolinate acetyltransferase (*ykuQ*) of *B. subtilis*, which is the first enzyme in the epimerase pathway. On searching *S. aureus* databases with the sequence of DAP dehydrogenase (*ddh*) from *C. glutamicum*, no evidence for the presence of a homologue of this enzyme was found. Thus it is likely that *S. aureus*, like *B. subtilis*, only synthesizes lysine via the epimerase pathway, and not by the one step dehydrogenase pathway.

5.2.3 DNA sequence analysis of the cystathionine γ-synthase locus

A putative cystathionine γ-synthase (*yjcI*) (*seg*8) has been identified during the screen for sera expressed genes (Chapter 4.2.2.2). The question as to whether multiple pathways of methionine biosynthesis exist in *S. aureus* (Chapter 5.1.5.1), along with the genetic organisation of cystathionine γ-synthase (CTGS) and the other methionine biosynthetic enzymes (MBE) genes are studied here.

5.2.3.1 Database searching for methionine biosynthetic enzymes of *S. aureus*

Corynebacteria, for example *C. glutamicium* and *B. flavum*, possess the enzymes to synthesize methionine via two different pathways: One pathway requires cystathionine γ-synthase (CTGS) and cystathionine β-lyase (CTBL) for the conversion of O-acetylhomoserine to homocysteine. This pathway proceeds via cystathionine and is the pathway used by both *B. subtilis* and *E. coli*. The other pathway requires only a single enzyme, O-acetylhomoserine sulphydrylase to perform the above enzymatic conversion (Chapter 5.1.5.1; Figure 5.3).

Auxotrophy analysis revealed *seg*8 (*yjcI*) to be a methionine auxotroph. This strongly suggests that *S. aureus* only possesses the enzymes (CTGS and CTBL) for the biosynthesis of methionine via cystathionine. In an attempt to confirm this, *S. aureus* databases were searched using the DNA sequences of methionine biosynthetic genes from both *B. subtilis* and *Thermotoga maritima*.

5.2.3.2 Determination of the mechanism(s) for the biosynthesis of methionine in *S. aureus*

S. aureus databases were searched with DNA sequences for CTGS (*yjcI*) and CTBL (*yjcJ*) from *B. subtilis*, and O-acetylhomoserine sulphydrylase from *T. maritima*. This analysis confirmed the presence of CTGS (*yjcI*) in *S. aureus* (49 % identity over 331 amino acids to *yjcI* of *B. subtilis*) and revealed the presence a homologue to CTBL (*yjcJ*) (48 % identity over 316 amino acids to *yjcJ* of *B. subtilis*). However, none of the databases searched contained homologs to O-acetylhomoserine sulphydrylase. Unfortunately as both the *S. aureus* databases searched (NCTC 8325 and TIGR) are currently incomplete this does not provide conclusive proof for the absence of this enzyme. This result however, coupled with the auxotrophic nature of *seg*8 (*yjcI*) suggests that like both *B. subtilis* and *E. coli, S. aureus* only encodes the enzymes for the synthesis of methionine via cystathionine, and not by the direct conversion of O-acetylhomoserine to homocysteine using O-acetylhomoserine sulphydrylase as in *C. glutamicium*.

Interestingly, while searching the *S. aureus* databases for CTGS and CTBL, a third gene, (named *metX*), was identified as having significant homology to both CTGS (*yjcI*) (44 % identity over 374 amino acids) and CTBL (*yjcJ*) (48 % identity over 374 amino acids) of *B. subtilis*. Searching the *B. subtilis* genome (SubtiList) with the DNA sequence of *metX*, revealed a third gene, in addition to *yjcI* and *yjcJ* with significant homology to this sequence: *yrhB* (unknown, similar to CTGS), 58 % identity over 380 amino acids. Thus *B. subtilis* encodes a homologue of *metX*. The fact that both *S. aureus* and *B. subtilis* potentially possess two isozymes of CTGS was a surprising discovery since if *S. aureus* encode two isozymes of CTGS, then why is *seg*8 (*yjcI*) a methionine auxotroph? Furthermore, there have been no previous reports of any bacterium encoding multiple isozymes of methionine biosynthetic enzymes, with the exception of AK / AK-HDH (Chapter 5.1.3) and methionine synthase (Figures 5.14-5.17).

Three explanations for *seg*8 (*yjcI*) being a methionine auxotroph are outlined here: The first is that *metX*, despite its high similarity to CTGS and CTBL (of both *S. aureus* and *B. subtilis*), could encode an enzymatically inactive protein. The second explanation for the methionine auxotrophy of *seg*8 (*yjcI*) in the presence of two isozymes of CTGS, is that *yrhB* and its *S. aureus* homologue (*metX*) do not encode an isozyme of CTGS, perhaps instead it encodes an isozyme of CTBL, and thus *yrhB* has been wrongly classified in the *B. subtilis* database (SubtiList). The final explanation is that the genes

encoding MBE are organised in an operon, in such an arrangement that, a mutation in *yjcI* would be polar on one, or more downstream biosynthetic genes. Under this arrangement in spite of the presence of a second CTGS isozyme (encoded by *metX*) at a distinct locus, *seg*8 (*yjcI*) would still be a auxotrophic for methionine, as the pathway to methionine would remain blocked.

5.2.3.3 Investigation into the identity of *metX* of *S. aureus*

In an attempt to determine the identity of *metX* and hence its role in methionine biosynthesis, the DNA sequence of both *yrhB* (*B. subtilis*) and *metX* (*S. aureus*) were compared against known sequences using NCBI BLAST (Chapter 2.23.1).

The results obtained searching with *yrhB,* revealed that this gene has been previously been entered in the NCBI database, as cystathionine γ-lyase (CTGL) (100 % identity over 379 amino acids). The next closest homologue is CTBL from *Lactococcus lactis* (63 % identity over 374 amino acids), followed by CTGS from *Helicobacter pylori* (60 % identity over 379 amino acids). Additionally a BLAST search within SubtiList showed that *yrhB* has slightly greater homology to *yjcJ* (CTBL) (52 % identity over 376 amino acids) than *yjcI* (CTGS) (49 % identity over 373 amino acids). Therefore the identity of *yrhB* is not clear, as it is similar to both CTGS and CTBL, and has been entered into the NCBI database as CTGL. Without enzyme assay data it is impossible to determine with 100 % certainty that *metX* or *yrhB* encode a protein with CTGS or CTBL activity. If *metX* is in fact an isozyme of CTBL, then the methionine auxotrophy of *seg*8 (*yjcI*) is simply explained due to the lack of CTGS ((insertionally inactivated by the Tn*917* insertion of *seg*8). Therefore the conversion of O-acetylhomoserine to cystathionine is blocked, preventing the biosynthesis of methionine (Figure 5.15). Interestingly however, in a NCBI BLAST search, *metX* showed greatest homology to cystathionine γ-lyase (*yrhB*) from *B. subtilis* (58% identity over 380 amino acids). Therefore since *metX* shows the greatest homology to *yrhB*, the two homologues are almost certainly isozymes of the same enzyme.

5.2.3.4 Organisation of methionine biosynthetic genes in *B. subtilis*

In *B. subtilis* CTGS (*yjcI*) and CTBL (*yjcJ*) are in a bi-cistronic operon, at a site distant from all other methionine biosynthetic enzymes (Chapter 5.1.6.2; Figure 5.14). *yrhB* (CTGL), is part of a five gene operon in which only one other gene (*yrhA*) has been

assigned a function (Figure 5.16). *yrhA* is the gene immediately upstream of *yrhB* and is similar to cysteine synthase. Therefore *yrhB* appears to be located within a locus devoid of other methionine biosynthetic enzymes.

5.2.3.5 Organisation of putative methionine biosynthetic genes in *S. aureus*

Determination of the organisation of methionine biosynthetic genes in *S. aureus* was not straightforward, owing not only to the lack of a completed genome, but also to the absence of published work on methionine biosynthesis in *S. aureus*. However, contigs from the incomplete NCTC 8325 and TIGR databases identified in Chapter 5.2.3.2, as having homology to *yjcI* and *yjcJ* of *B. subtilis,* were examined to determine the identity and organisation of genes surrounding those involved in methionine biosynthesis.

The results of this analysis are shown in Figures 5.15 and 5.17 and show the identification of a methionine biosynthetic operon (*yjc*) in *S. aureus* and the location of the putative CTGS isozyme (*metX*). From this analysis it can be seen that owing to the genetic organisation of the *yjc* operon of *S. aureus,* a mutation in *yjcI* (*seg*8) would be polar on *yjcJ* (CTBL), and thus in spite of a second isozyme of CTGS (*metX*), *seg*8 (*yjcI*) would still be auxotrophic for methionine.

5.2.4 Analysis of the expression and regulation of *lysC*, *asd* and *dapA*

In order to accurately study the regulation of the three LBE *seg*; *lysC, asd* and *dapA*, reporter gene fusions were constructed in the wildtype (8325-4) background. This was necessary as the original *seg* strains, although reporter gene fusions, are mutants and thus do not grow under certain amino acid limiting conditions (Chapter 5.2.1.2).

5.2.4.1 The construction of *lacZ* reporter gene fusions

Three *lacZ* reporter gene fusions were constructed, one in each of the three LBE genes (*lysC, asd* and *dapA*). Suitable DNA fragments (Figure 5.18) were cloned into the reporter gene fusion vector pAZ106 (Chapter 2.22.3), using primers, OL 1-4 (Table A.1.2.1). The three plasmids generated were transformed into *S. aureus* RN4220 (Chapter 2.16.2), and the stable insertions transduced into 8325-4 (Chapter 2.14), to

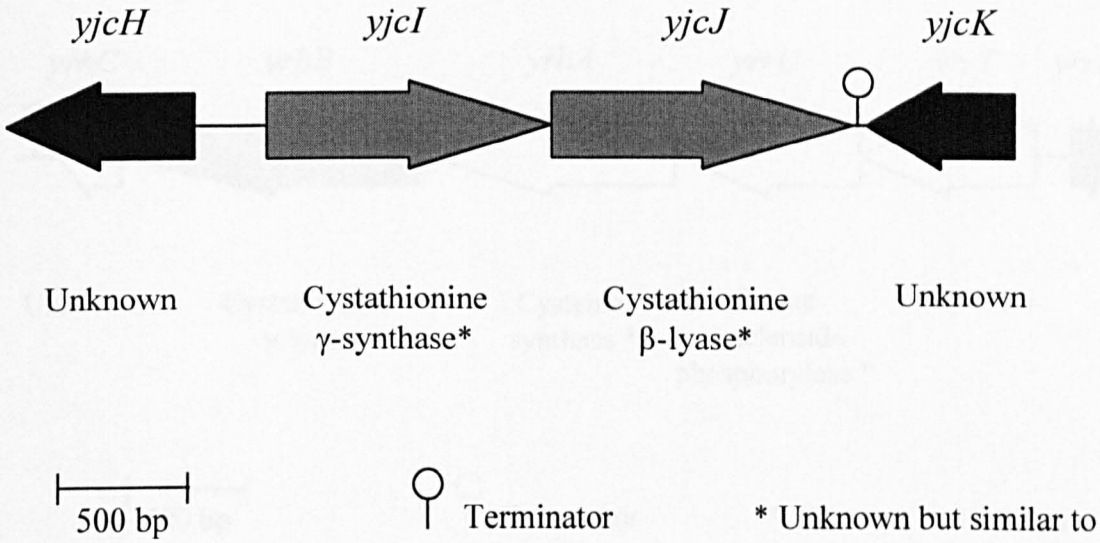


Figure 5.14

Organisation of the *yjc* locus region of the *B. subtilis* chromosome

Reproduced from SubtiList.

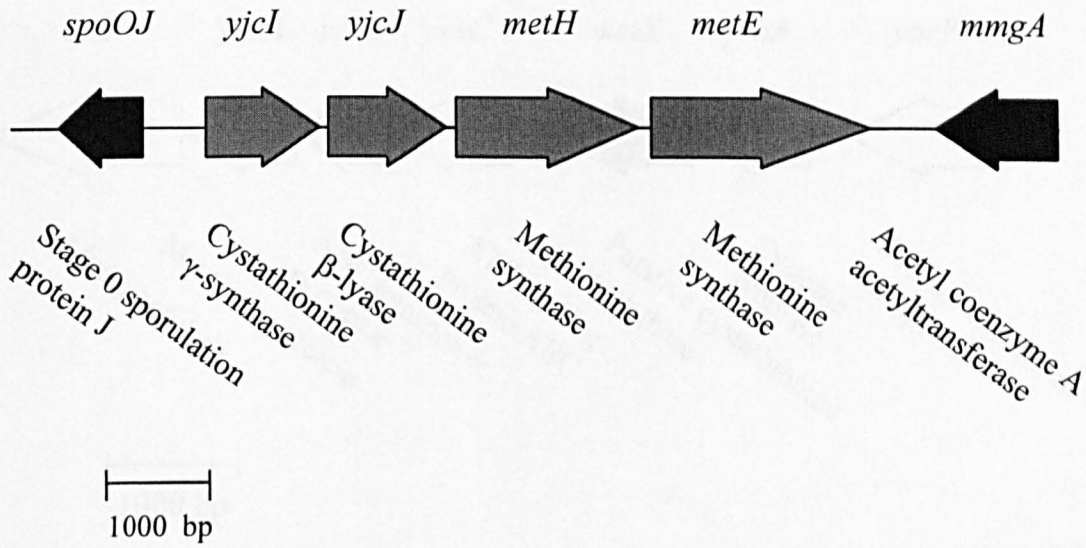


Figure 5.15

Organisation of the *yjc* locus region of the *S. aureus* chromosome

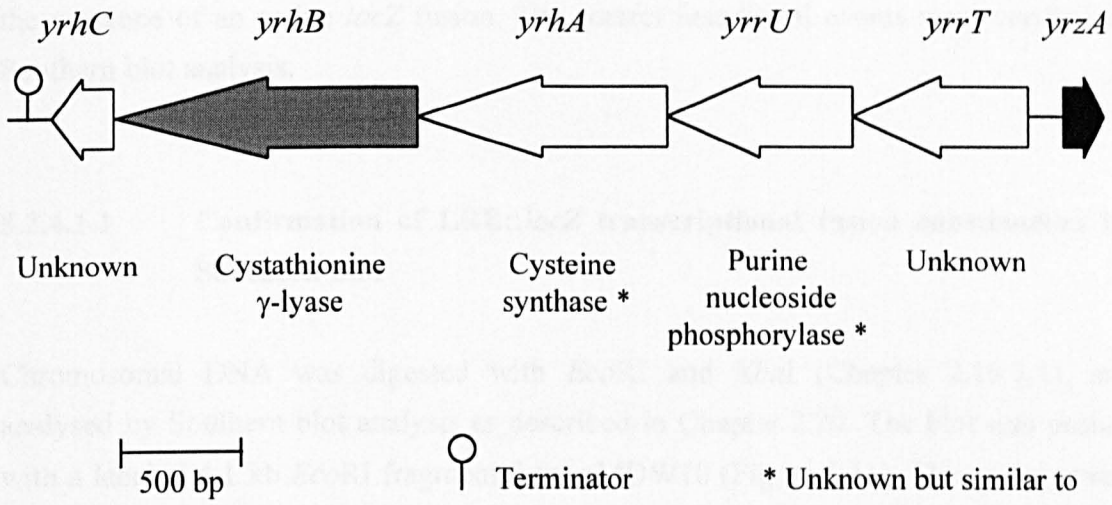


Figure 5.16
Organisation of the *yrhB* locus region of *B. subtilis*

Reproduced from SubtiList.

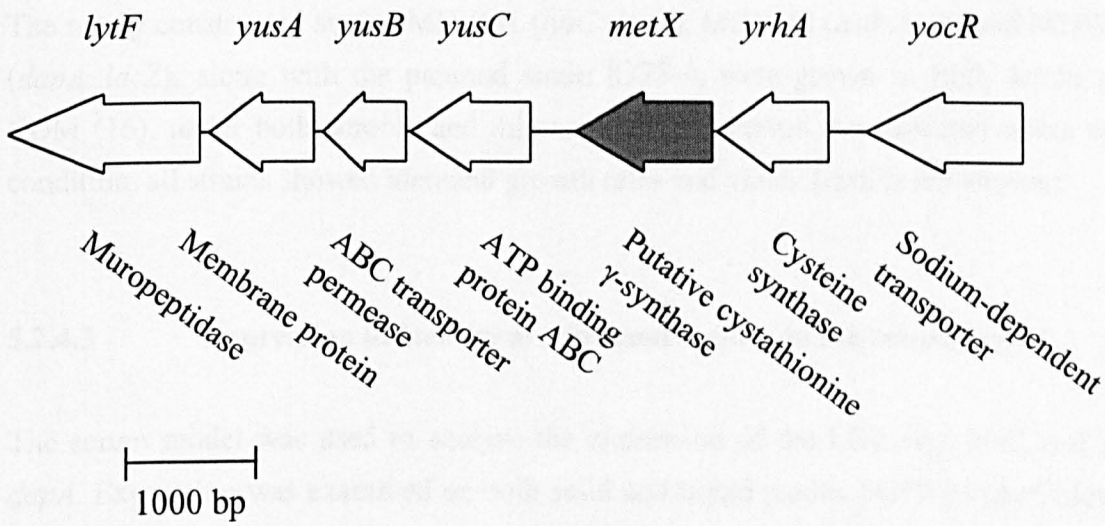


Figure 5.17
Organisation of the *metX* locus region of *S. aureus*

create strains: MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*). The construction of these *lacZ* strains is shown diagramatically in Figure 5.19. In the presence of X-Gal, all strains were blue on serum and white on BHI, which confirmed the presence of an active *lacZ* fusion. The correct insertional events were verified by Southern blot analysis.

5.2.4.1.1 Confirmation of LBE::*lacZ* transcriptional fusion construction by Southern blot

Chromosomal DNA was digested with *Eco*RI and *Xba*I (Chapter 2.19.1.1), and analysed by Southern blot analysis as described in Chapter 2.20. The blot was probed with a labeled 4.1 kb *Eco*RI fragment from pMDW10 (Figure 5.11). The probe covers the region of *lysC, asd* and the 5' end of *dapA*. For all strains, the expected fragment sizes were determined, for digestion with *Eco*RI and *Xba*I, and are shown in Figures 5.19a-d. The Southern blots for *Eco*RI and *Xba*I are shown in Figure 5.20. These blots show that all three *lacZ* strains gave the predicted patterns for correct reporter gene construction.

5.2.4.2 Growth of *lacZ* transcriptional fusion strains

The newly constructed strains MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*), along with the parental strain 8325-4, were grown in BHI, serum and CDM (16), under both aerobic and microaerobic conditions. As expected under each condition, all strains showed identical growth rates and yields (results not shown).

5.2.4.3 Expression of *lacZ* transcriptional fusions in the serum model

The serum model was used to analyse the expression of the LBE *seg*, *lysC, asd* and *dapA*. Expression was examined on both solid and liquid media. MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*), when grown on agar containing X-Gal, produced blue colonies on serum and white colonies on BHI, under both aerobic and microaerobic conditions.

MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*) along with 8325-4, were grown in BHI and serum using standard conditions (Chapter 2.4.1.1),

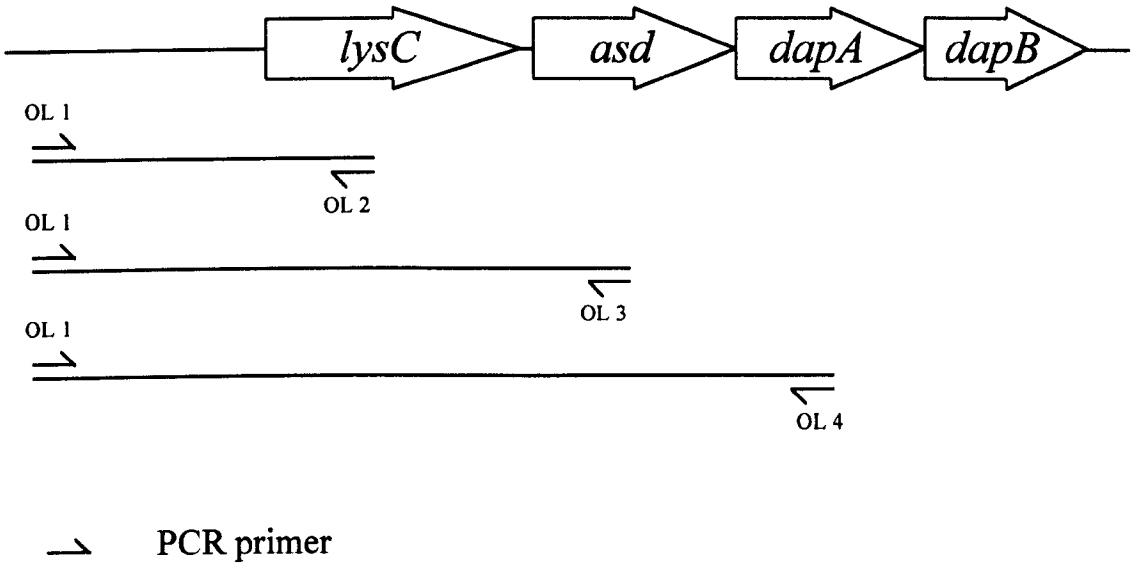


Figure 5.18

Diagrammatic representation of genomic fragments used for the construction of LBE *lacZ* fusion strains

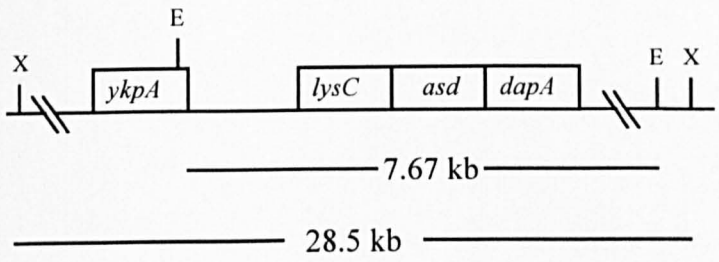


Figure 5.19a

8325-4 - Wild type

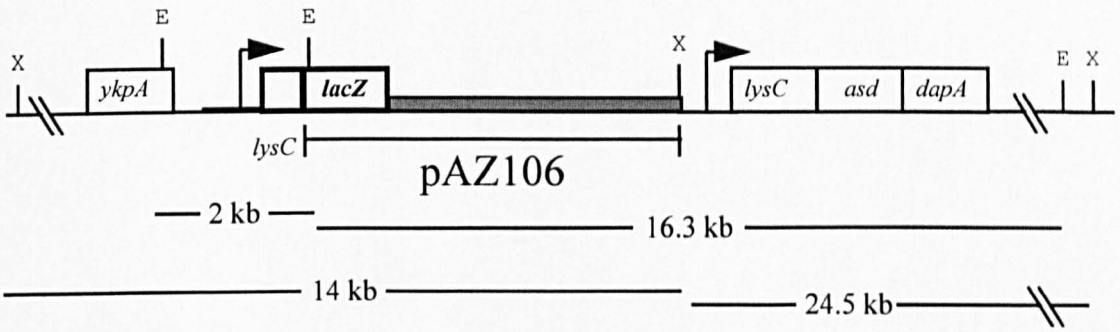


Figure 5.19b

MDW41 - *lysC*::*lacZ*, *lysC* $^{+}$

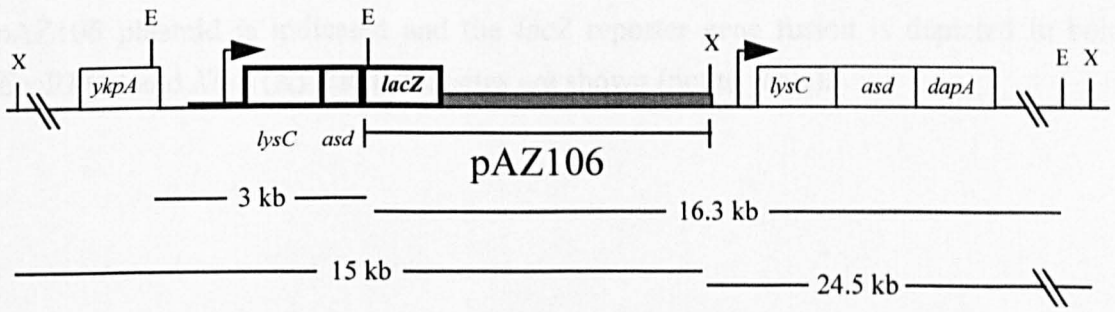


Figure 5.19c

MDW42 - *asd*::*lacZ*, *asd* $^{+}$

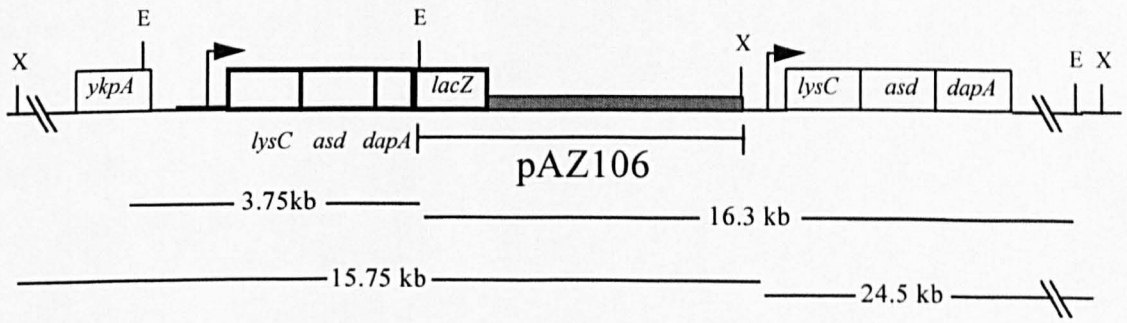


Figure 5.19d

MDW43 - *dapA*::*lacZ*, *dapA*$^{+}$

Figure 5.19a-d

lacZ construction and Southern blot fragment size predictions

pAZ106 plasmid is indicated and the *lacZ* reporter gene fusion is depicted in bold. *Eco*RI (E) and *Xba*I (X) restriction sites are shown (not to scale).

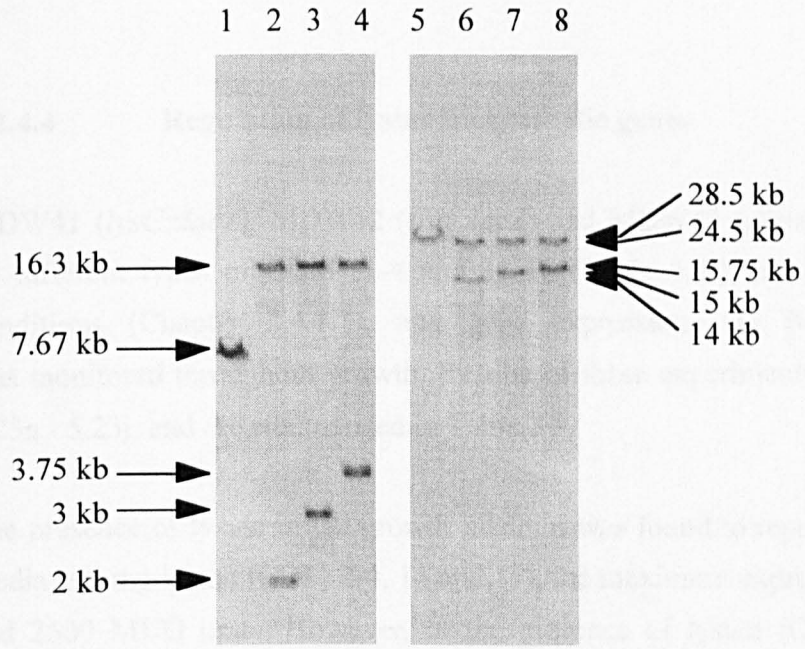


Figure 5.20

Southern blots of chromosomal DNA (*Eco*R1 and *Xba*1 digests) of 8325-4 (WT), MDW 41 (*lysC*::*lacZ*), MDW 42 (*asd*::*lacZ*) and MDW 41 (*dapA*::*lacZ*), probed for *lysC*-dapA

*Eco*R1 digests lanes: 1, 8325-4 (WT); 2, MDW 41 (*lysC*::*lacZ*); 3, MDW 42 (*asd*::*lacZ*); and 4, MDW 41 (*dapA*::*lacZ*).

*Xba*1 digests lanes: 5, 8325-4 (WT); 6, MDW 41 (*lysC*::*lacZ*); 7, MDW 42 (*asd*::*lacZ*); and 8, MDW 41 (*dapA*::*lacZ*).

under both aerobic and microaerobic conditions. Negligible expression was observed in BHI under the different conditions (Figure 5.21). In serum all three LBE genes were expressed in a growth phase dependent manner, primarily as the cultures entered stationary phase (~ 40-180 MUG units). *asd* showed the highest level of expression of these fusions under both aerobic and microaerobic conditions (180 and 170 MUG units respectively) (Figure 5.22).

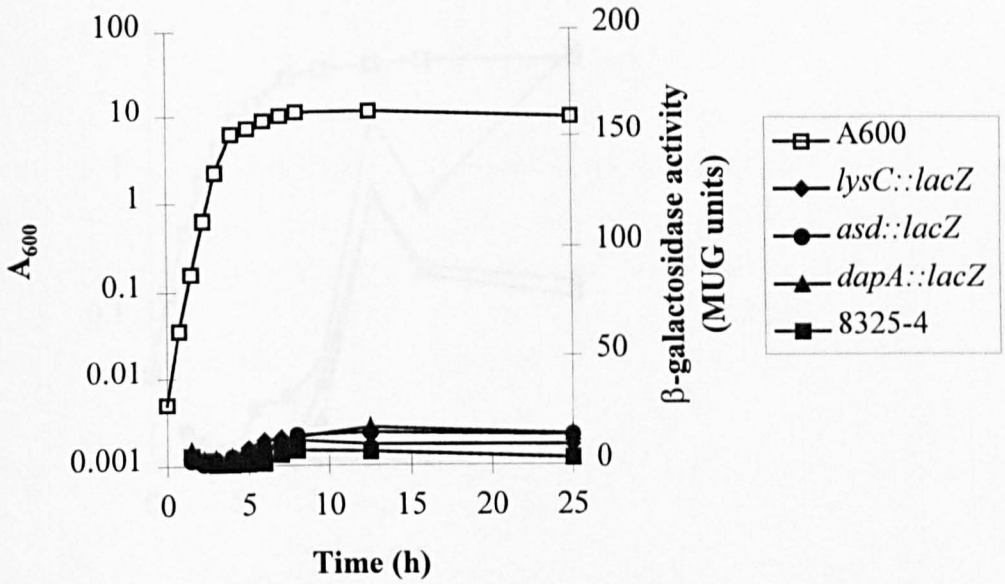
5.2.4.4 Regulation of lysine biosynthetic genes

MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*) were grown in 10 different types of CDM (1-4 and 11-16 (Table 5.2)), aerobically under standard conditions (Chapter 2.4.1.1), and gene expression, via β-galactosidase activity was monitored throughout growth. Results of these experiments are shown in Figures 5.23a - 5.23j, and are summarised in Table 5.5.

The presence of lysine in the growth medium was found to repress *lysC* expression. In media lacking lysine (CDM 2-4, 14 and 15), the maximum expression was between 650 and 2500 MUG units. However, in the presence of lysine (CDM 1, 11-13 and 16) maximum expression was between 40 and 140 MUG units. Therefore on average expression was > 10-fold higher in CDM lacking lysine. Interestingly, maximal expression occurred in all cultures as the cells entered post-exponential phase. The other AAA also appeared to play a role in the regulation of *lysC* expression. Results show that CDM cultures containing all four AAA (CDM 16), or lacking just isoleucine (CDM 11), had the lowest *lysC* expression as determined by β-galactosidase activity (Figure 5.23; Table 5.5).

Lysine was also shown to repress *asd* expression. In media where lysine was absent (CDM 2-4, 14 and 15), the maximum expression was between 475 and 2300 MUG units. In contrast when MDW42 (*asd*::*lacZ*) was grown in media containing lysine (CDM 1, 11-13 and 16), the maximum expression was significantly lower being between 40 and 150 MUG units. Therefore on average, expression was > 10-fold higher in CDM lacking lysine. Expression of *asd* occurred primarily in the post-exponential phase. Finally as seen for *lysC*, the inclusion of the other AAA in the growth medium, further reduced the expression of *asd* in the presence of lysine (Figure 5.23; Table 5.5). *dapA* expression was also found to be repressed by lysine, however in general the levels of expression were lower than those observed for both *lysC* and *asd*. In CDM media where lysine was absent (CDM 2-4, 14 and 15), the maximum expression observed was

a)



b)

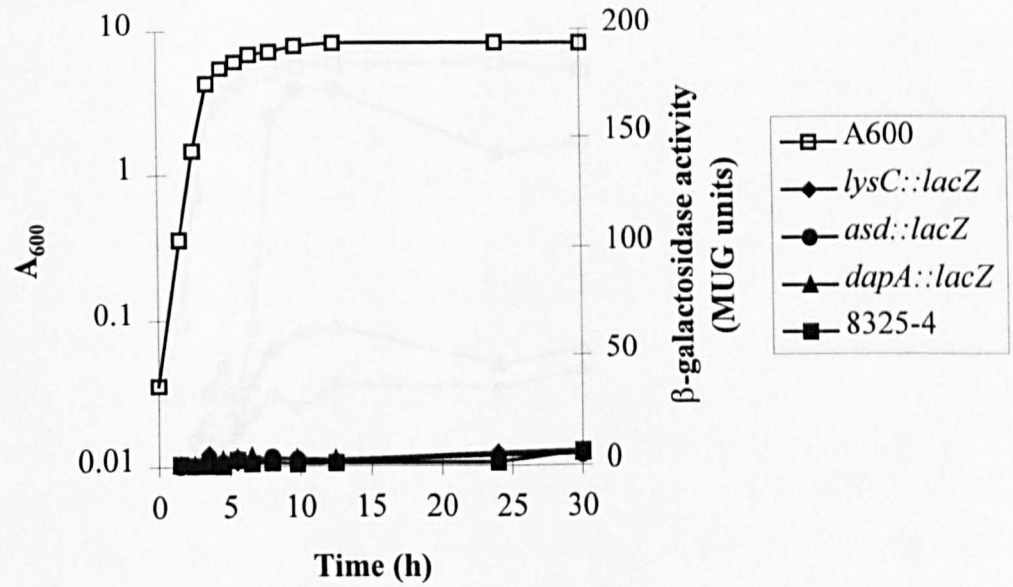
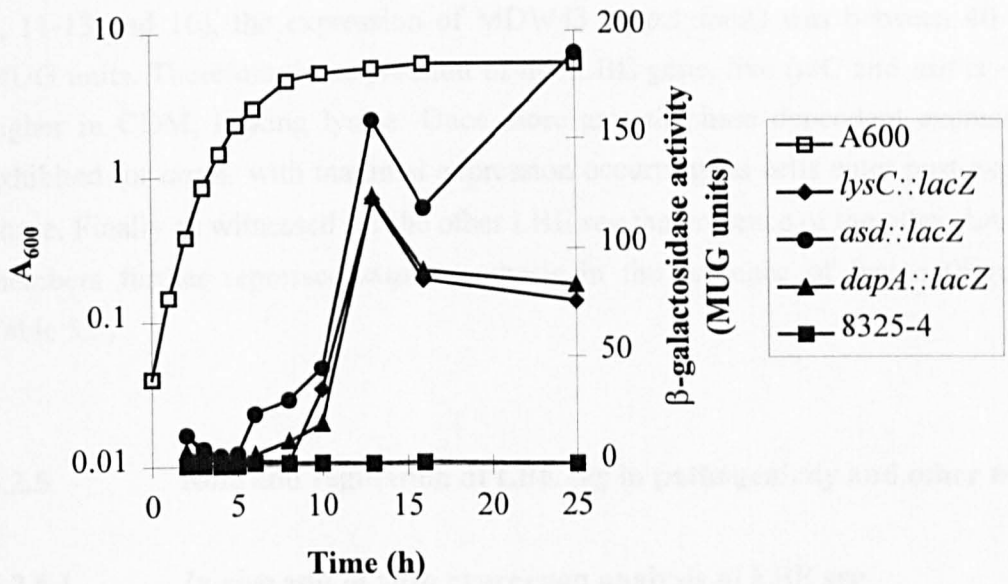


Figure 5.21

Expression of MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*) in BHI under aerobic (**a**) and microaerobic (**b**) conditions

MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*), along with the parental strain 8325-4 were grown at 37 °C in BHI under aerobic or microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of two separate experiments.

a)



b)

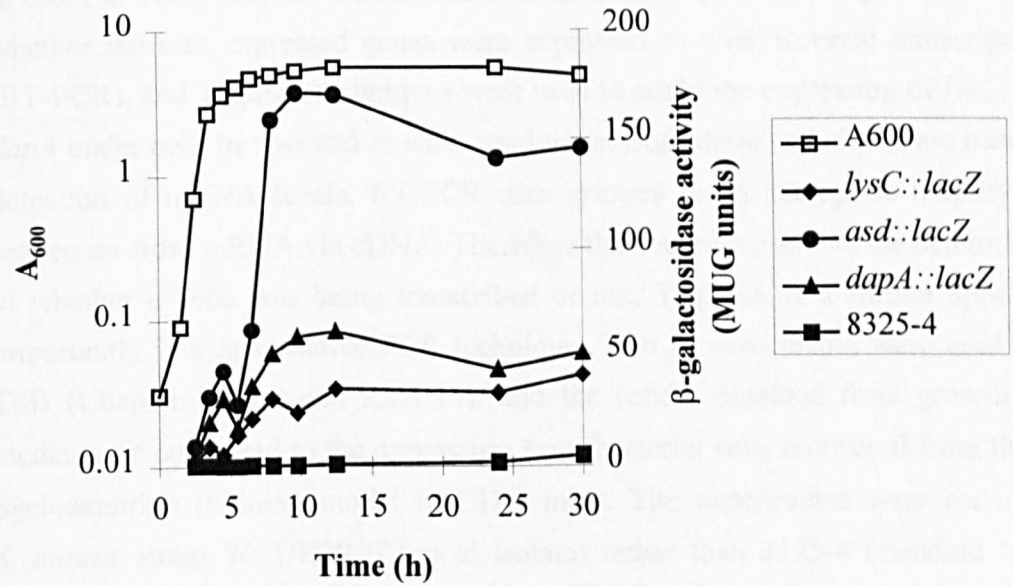


Figure 5.22

Expression of MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*) in serum under aerobic (**a**) and microaerobic (**b**) conditions

MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*) and MDW43 (*dapA*::*lacZ*), along with the parental strain 8325-4 were grown at 37 °C in serum under aerobic or microaerobic conditions with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of two separate experiments.

between 475 and 1200 MUG units. While in media in which lysine was present (CDM 1, 11-13 and 16), the expression of MDW43 (*dapA*::*lacZ*) was between 40 and 125 MUG units. Therefore the expression of this LBE gene, like *lysC* and *asd* is ~ 10-fold higher in CDM, lacking lysine. Once more growth phase dependent expression was exhibited for *dapA,* with maximal expression occurring, as cells enter post-exponential phase. Finally as witnessed for the other LBE *seg* the presence of the other AAA family members further repressed *dapA* synthesis in the presence of lysine (Figure 5.23; Table 5.5).

5.2.5 Role and regulation of LBE *seg* in pathogenicity and other functions

5.2.5.1 *In vivo* and *in vitro* expression analysis of LBE *seg*

In order to verify that the serum model mimics infection, it was important to determine whether the sera expressed genes were expressed *in vivo*. Reverse transcription-PCR (RT-PCR), and TaqMan techniques were used to study the expression of *lysC, asd*, and *dapA* under both *in vivo* and *in vitro* conditions. Both these techniques are based on the detection of mRNA levels. RT-PCR uses primers in an attempt to amplify specific sequences from mRNA via cDNA. Therefore this technique allows the determination as to whether a gene was being transcribed or not. TaqMan, is a similar approach, but importantly is a quantitative PCR technique. Two *in vitro* media were used BHI and TSB (Chapter 2.1.1.2 and 2.1.1.13), and the results obtained from growth in these media were compared to the expression from bacterial cells recovered from the murine pyelonephritis (kidney) model in CD-1 mice. The experiments were carried out in *S. aureus* strain WCUH29 (Clinical isolate) rather than 8325-4 (standard laboratory strain), as this work was conducted in collaboration with Dr. M. Burnham and Dr. L. Palmer (SmithKline Beecham (SB)).

5.2.5.1.1 Production of cDNA from *S. aureus* WCUH29

RNA was extracted from *S. aureus* WCUH29 cells, grown in both BHI and TSB (*in vitro*) along with cells recovered from infected murine kidneys (*in vivo*) (Chapter 2.17.2.1). The integrity and purity of the RNA preps were analysed by PCR and were shown to be free from DNA (results not shown). The isolated RNA was then converted to cDNA by reverse transcription (RT), in a reaction catalysed by the enzyme, reverse transcriptase (Chapter 2.19.2.2). The process of cDNA production is diagramatically

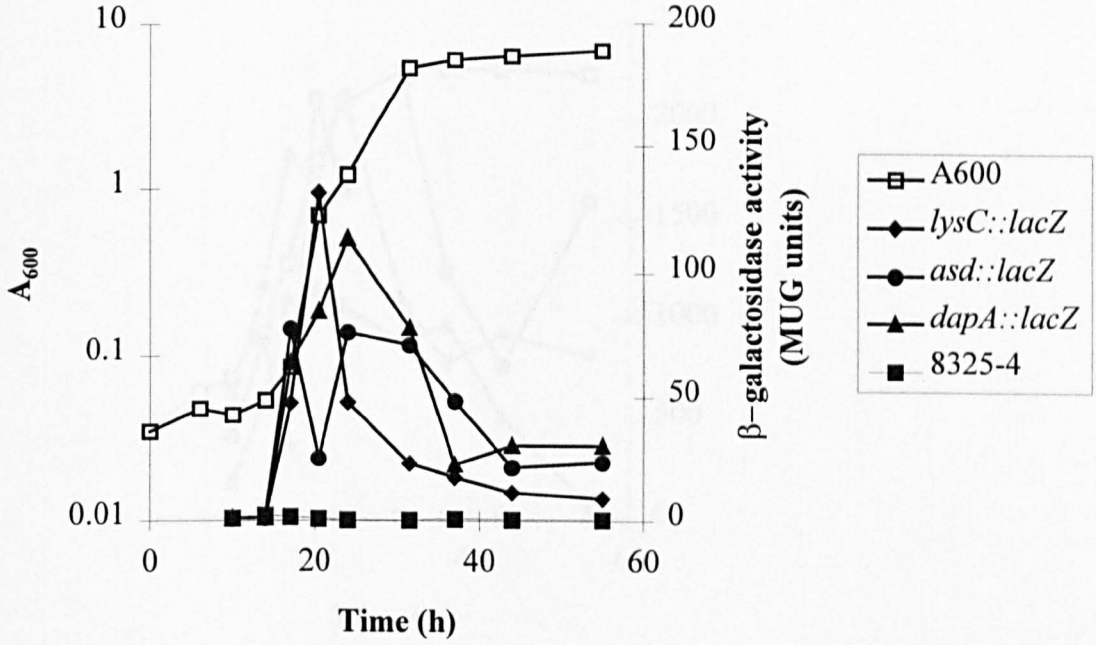


Figure 5.23a

Expression of *lysC*::*lacZ*, *asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 1 (CDM 1; + lysine)

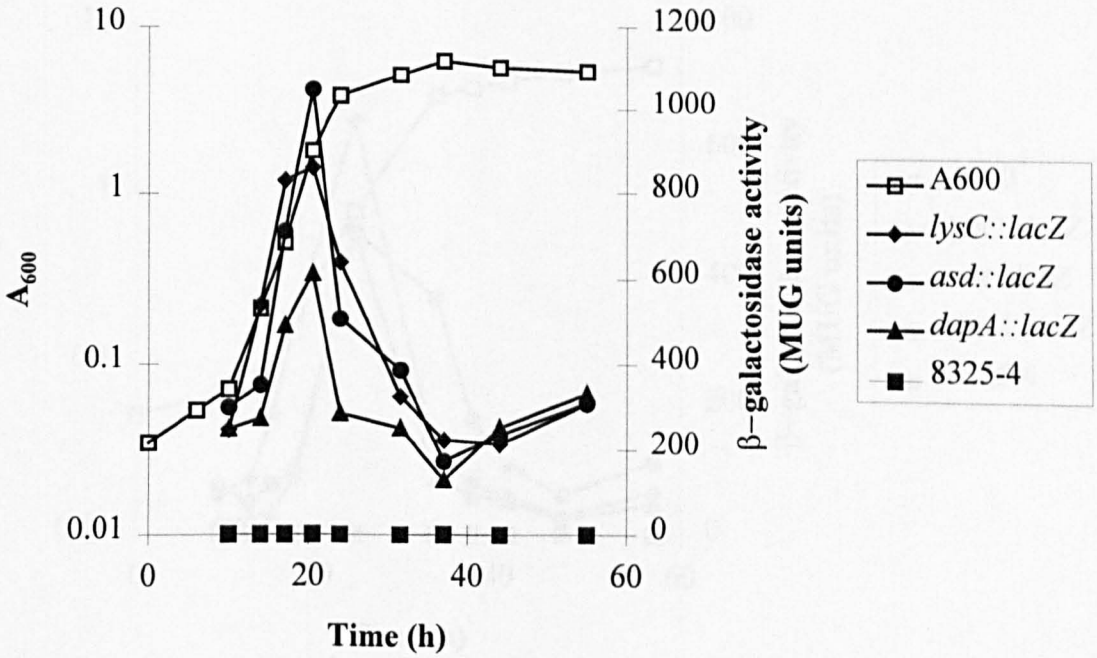


Figure 5.23b

Expression of *lysC*::*lacZ*, *asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 2 (CDM 2; + threonine)

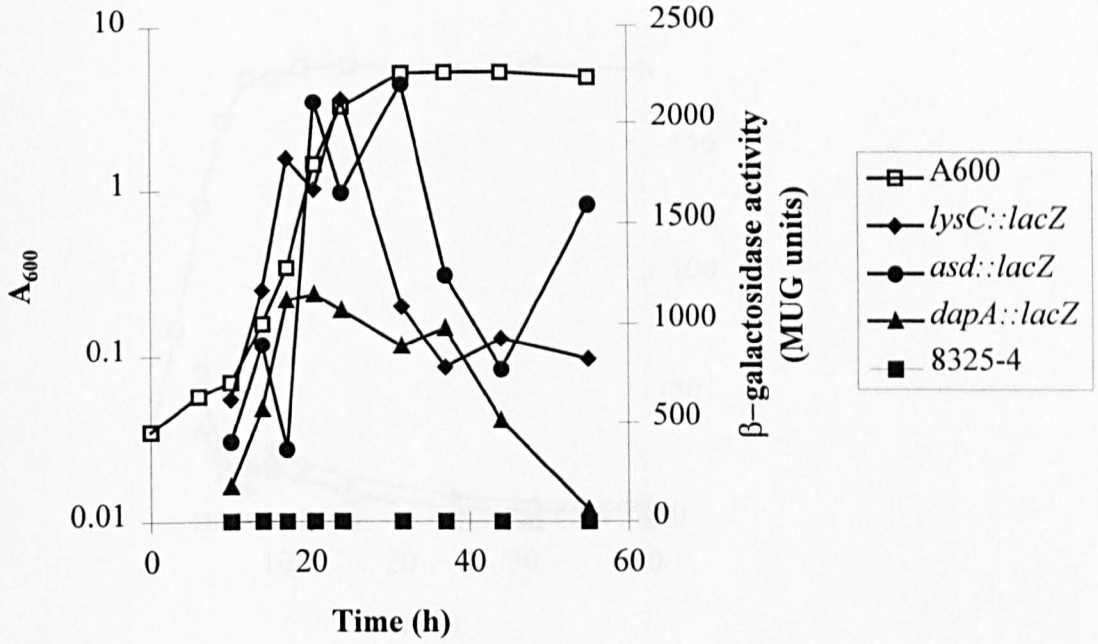


Figure 5.23c

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 3 (CDM 3; + methionine)

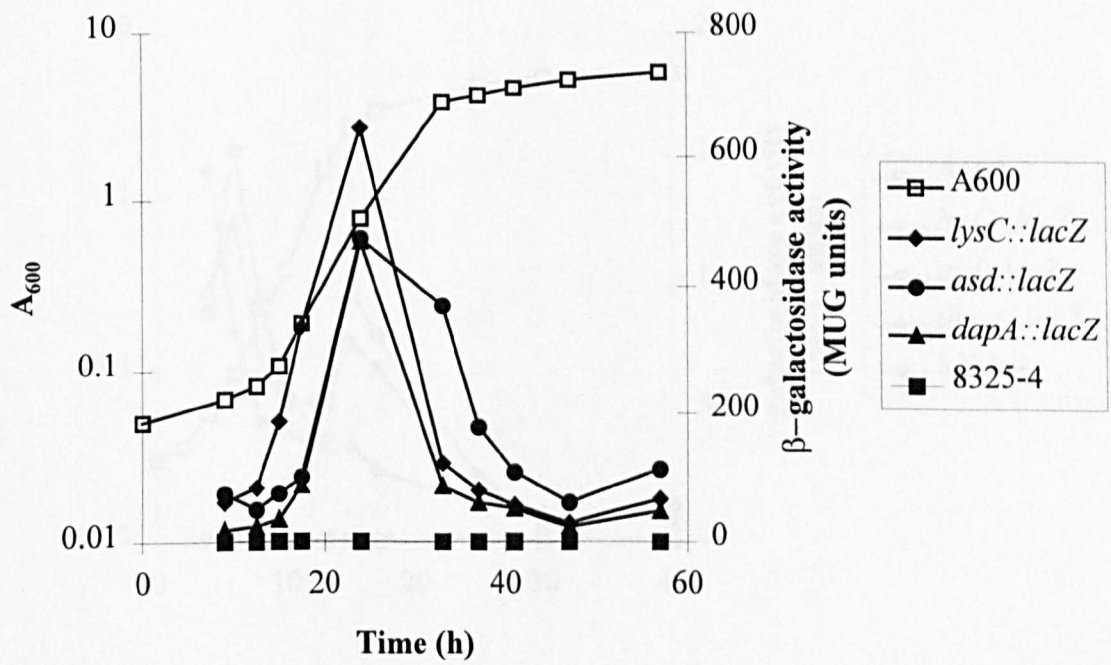


Figure 5.23d

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 4 (CDM 4; + isoleucine)

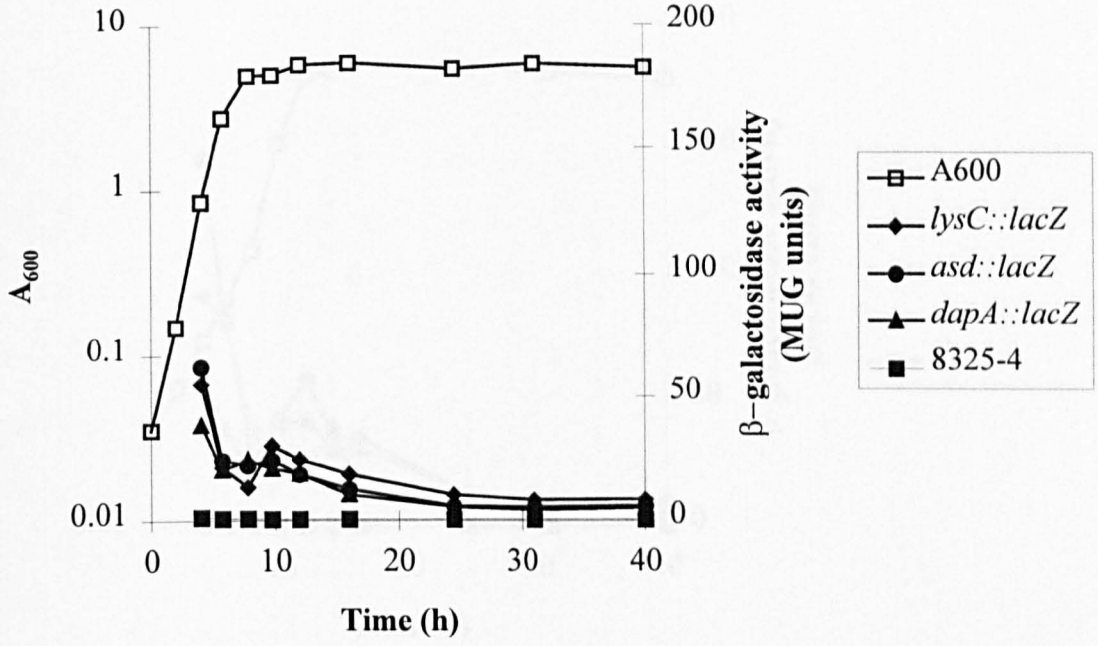


Figure 5.23e

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 11 (CDM 11; + lysine, threonine and methionine)

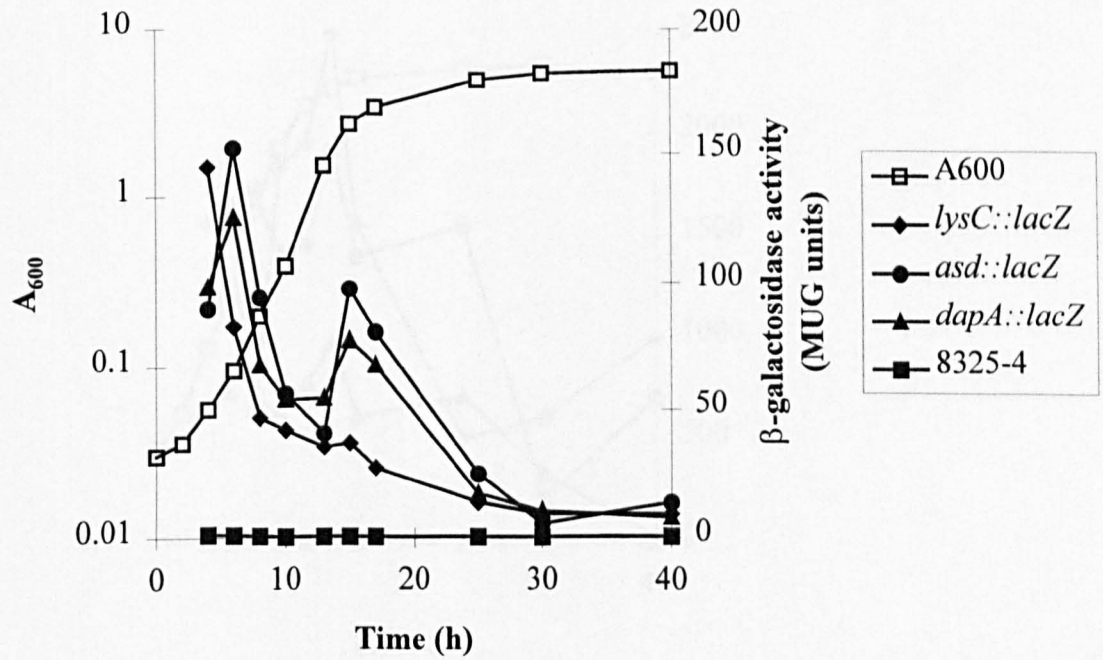


Figure 5.23f

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 12 (CDM 12; + lysine, threonine and isoleucine)

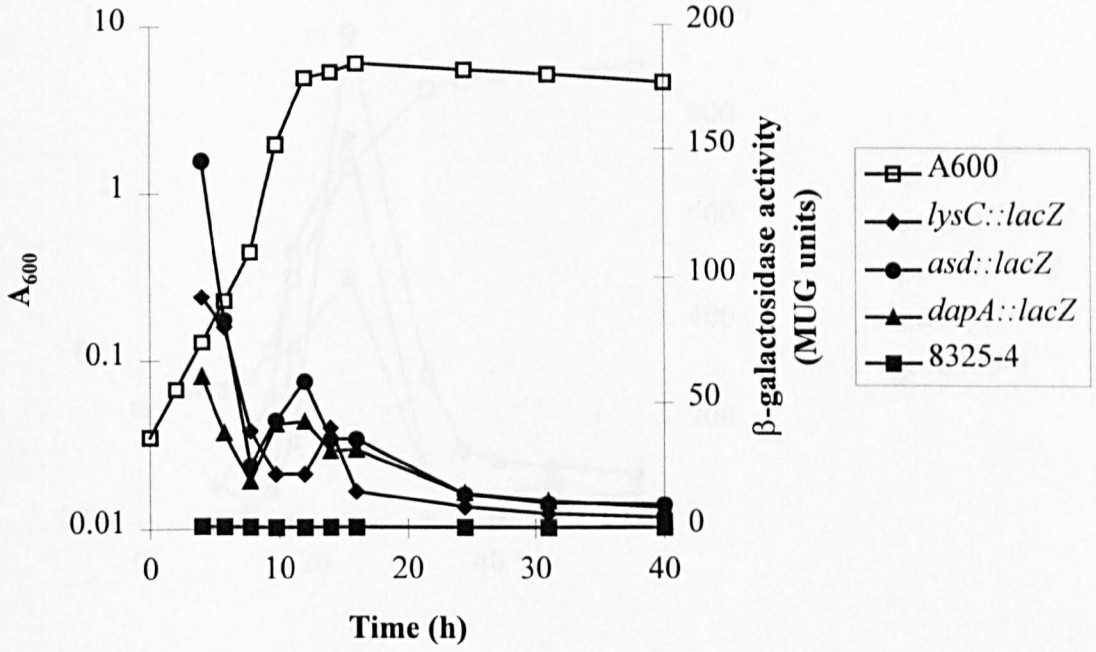


Figure 5.23g

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 13 (CDM 13; + lysine, methionine and isoleucine)

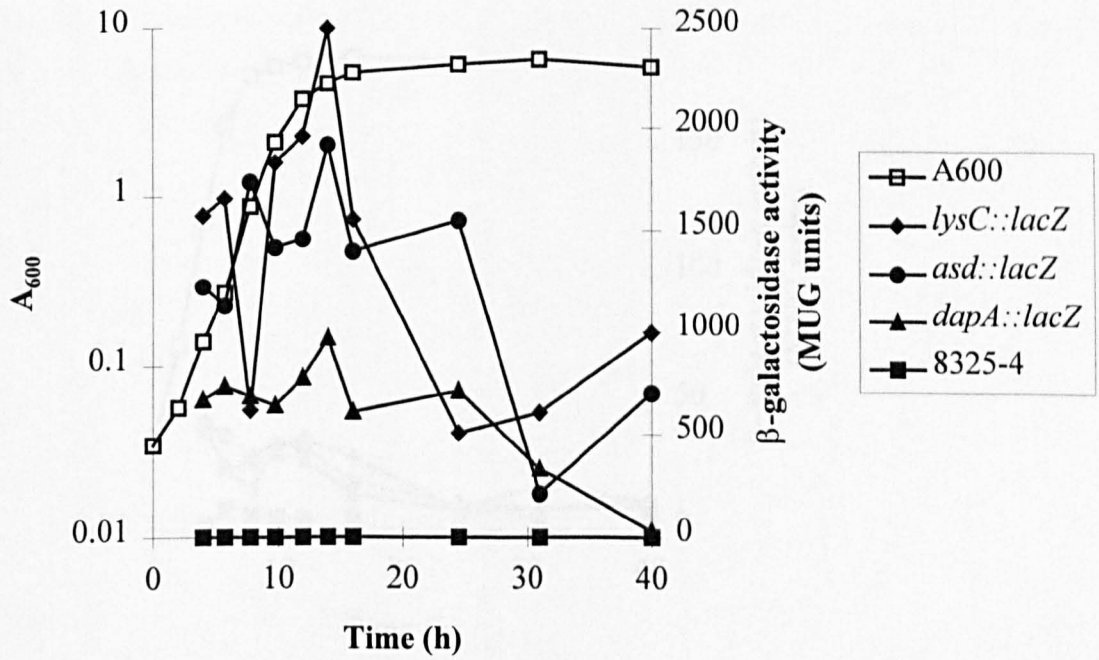


Figure 5.23h

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 14 (CDM 14; + threonine, methionine and isoleucine)

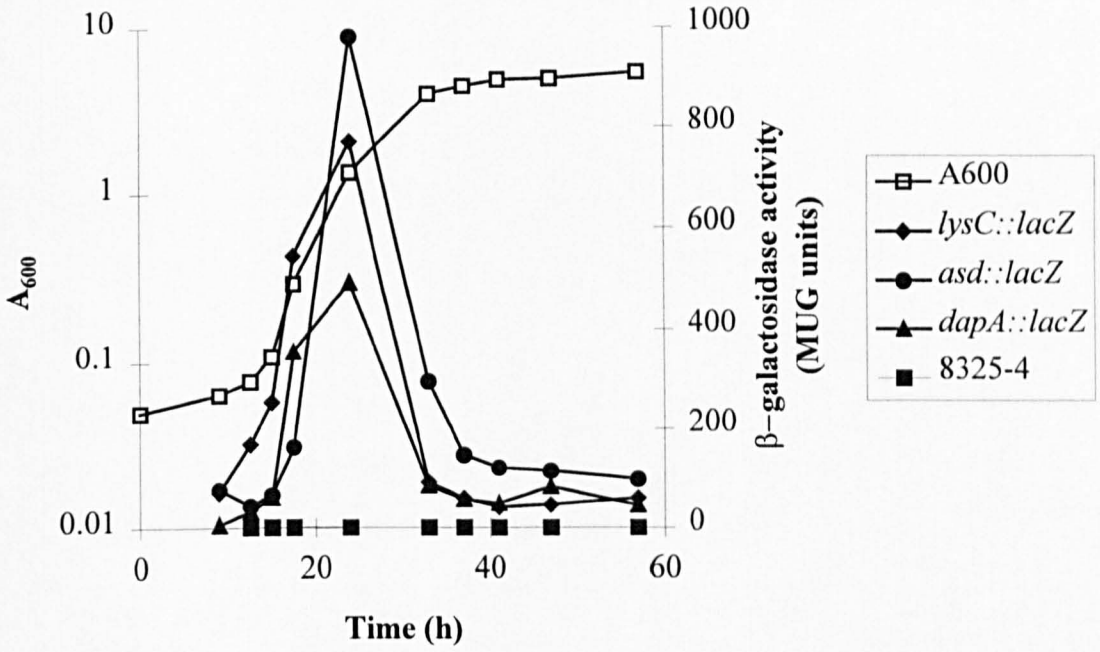


Figure 5.23i

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 15 (CDM 15; - lysine, threonine, methionine and isoleucine)

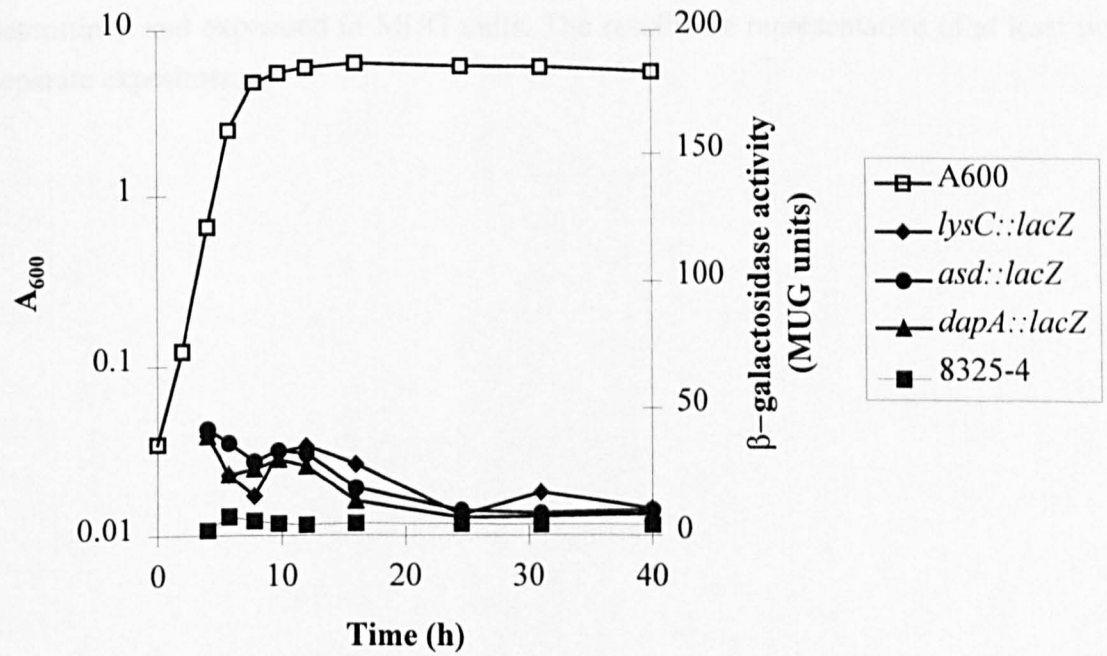


Figure 5.23j

Expression of *lysC*::*lacZ, asd*::*lacZ* and *dapA*::*lacZ* during growth in CDM 16 (CDM 16; + lysine, threonine, methionine and isoleucine)

Figure 5.23a-j

Expression of *lysC, asd* and *dapA* in different CDM

S. aureus MDW41 (*lysC*::*lacZ*), MDW42 (*asd*::*lacZ*), and MDW43 (*dapA*::*lacZ*) along with the parental strain 8325-4, were grown at 37 °C in CDM 1-4, and CDM 11-16, under aerobic conditions, with shaking (250 rpm). Specific β-galactosidase activity was determined and expressed in MUG units. The results are representative of at least two separate experiments.

CDM No.	Maximum expression (MUG units)			Gene expression On / Off	Media + / - lysine
	lysC	*asd*	*dapA*		
1	120	80	110	Off	+
2	**875**	**1050**	**620**	**On**	**-**
3	**2200**	**2300**	**1200**	**On**	**-**
4	**650**	**475**	**475**	**On**	**-**
11	55	60	40	Off	+
12	140	150	125	Off	+
13	90	140	60	Off	+
14	**2500**	**1900**	**950**	**On**	**-**
15	**775**	**1000**	**500**	**On**	**-**
16	40	40	40	Off	+

Table 5.5

Summary of the expression of *lysC, asd* and *dapA* in different CDM

represented in Figure 5.24. The cDNA samples produced were suitable for use in both RT-PCR and TaqMan procedures.

5.2.5.1.2 RT-PCR on *lysC, asd* and *dapA*

Pairs of primers designed to *lysC* (AK II), *asd* (ASADH) and *dapA* (DHPS) (sequences not provided by SB) were used to amplify fragments of each gene. PCR amplification reactions (Chapter 2.21.2.2), were prepared with one of the following templates: cDNA reverse transcribed from *S. aureus* WCUH29 total RNA (isolated from *in vitro* or *in vivo* grown cells), or *S. aureus* genomic DNA (positive control). Additionally during the RT step for each RNA sample, control reactions were set up that did not include reverse transcriptase (Figure 5.24), these samples were included as a negative control (-RT control).

The results of these RT-PCR reactions showed that *lysC, asd* and *dapA* were expressed under all three conditions (BHI, TSB and mouse kidney). Results for RT-PCR on the mouse kidney, and BHI samples, are shown in Figure 5.25. The results for cells grown in TSB were almost identical to those grown in BHI (results not shown). RT-PCR is not quantitative thus even low level expression in BHI would count as positive, owing to the exponential nature of PCR.

5.2.5.1.3 RT-PCR on *dapB, ykuQ,* and *lysA,* the other LBE genes in the *dap* operon

The other three LBE genes of the *dap* operon, *dapB* (DHPS), *ykuQ* (THPA), and *lysA* (DAP decarboxylase) were also examined. An identical set of reactions, were carried out for these three LBE genes, as for *lysC, asd* and *dapA* (Chapter 5.2.5.1.2), using pairs of primers designed to each gene (sequences not provided by SB). These RT-PCR experiments revealed once more that all genes were transcribed under both *in vivo* (mouse kidney) and *in vitro* conditions (BHI and TSB). Results for RT-PCR on the mouse kidney, and BHI samples, are shown in Figure 5.26. The results for *S. aureus* grown in TSB were almost identical to those grown in BHI (results not shown).

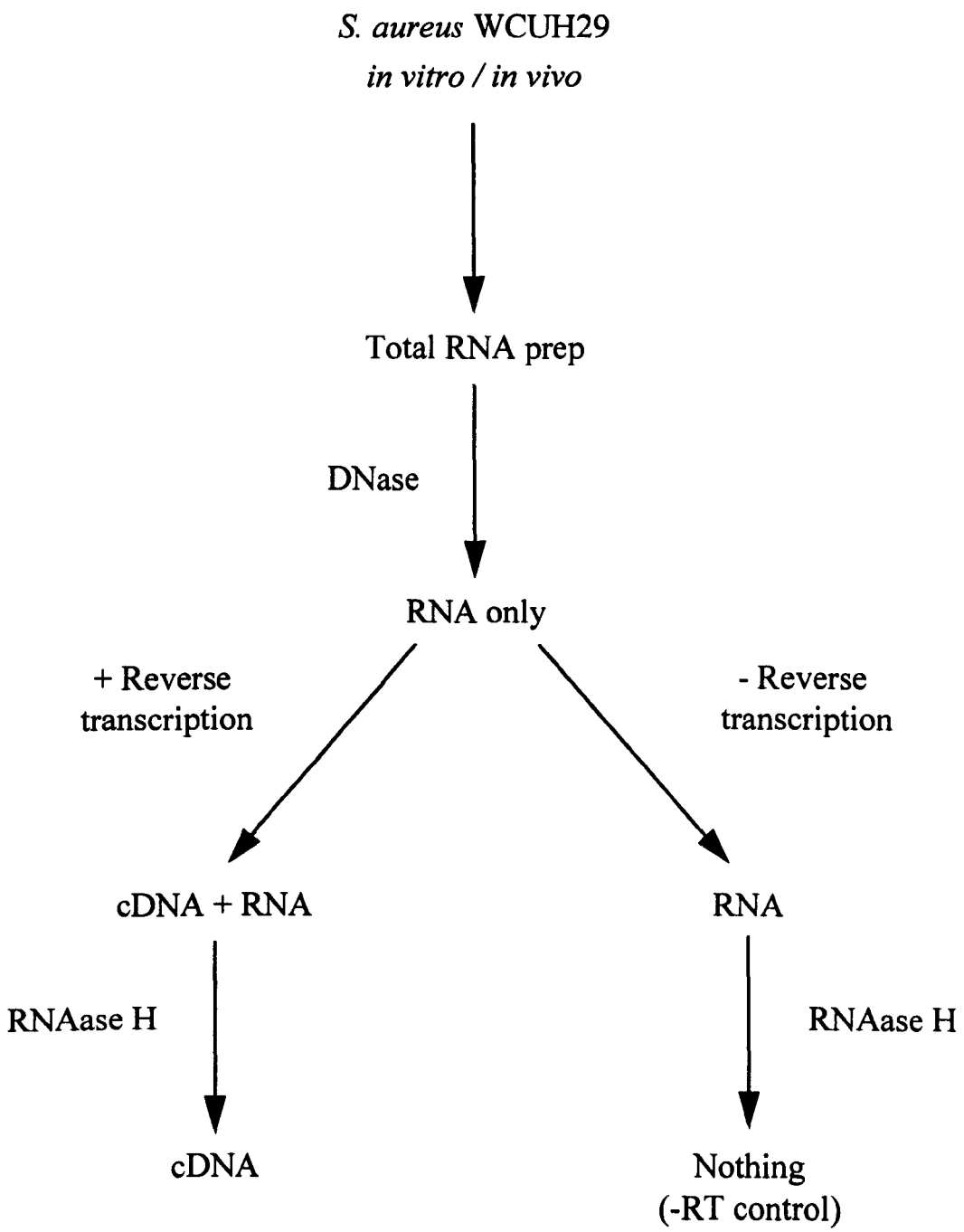
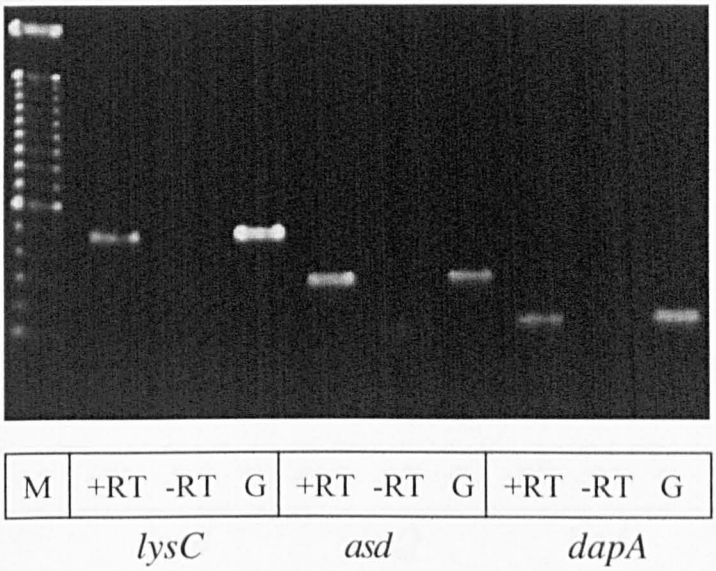


Figure 5.24

Diagrammatic representation of the production of cDNA

a)



b)

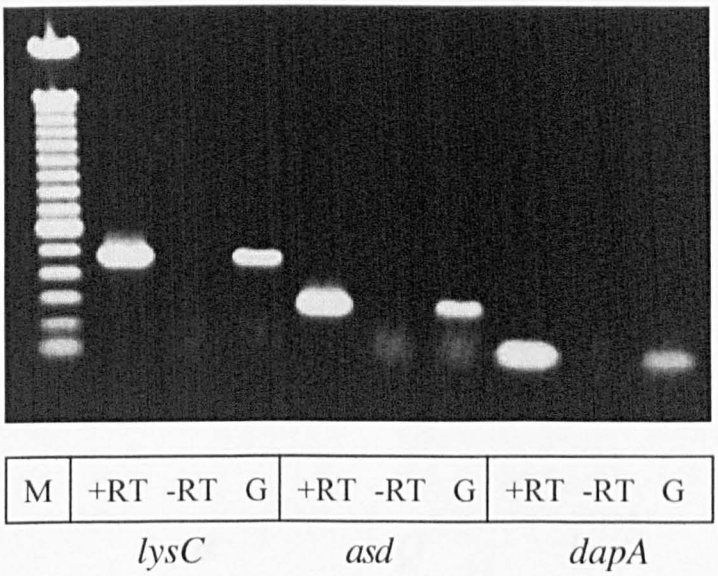


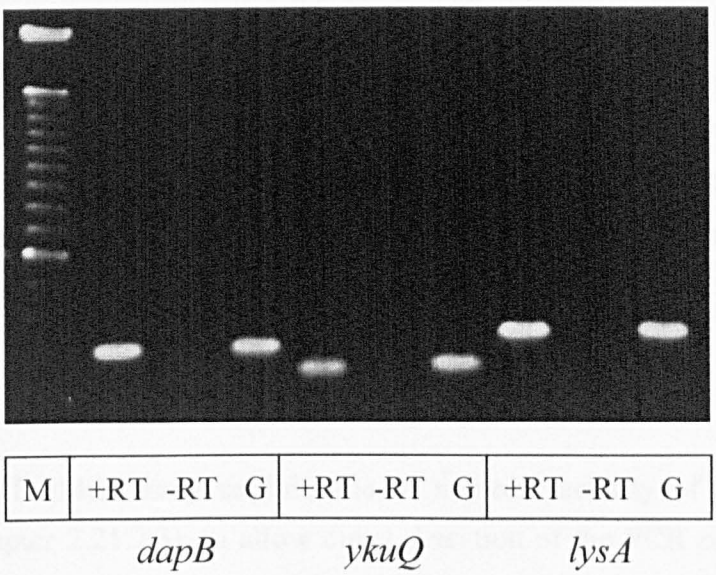
Figure 5.25

Expression of *lysC, asd* and *dapA* under *in vitro* (a) and *in vivo* (b) conditions as determined by RT-PCR

In vitro, BHI; *In vivo*, Mouse kidney (pyelonephritis); +RT, + reverse transcriptase; -RT, - reverse transcriptase; G, Genomic DNA.

These results are representative of two separate experiments.

a)



b)

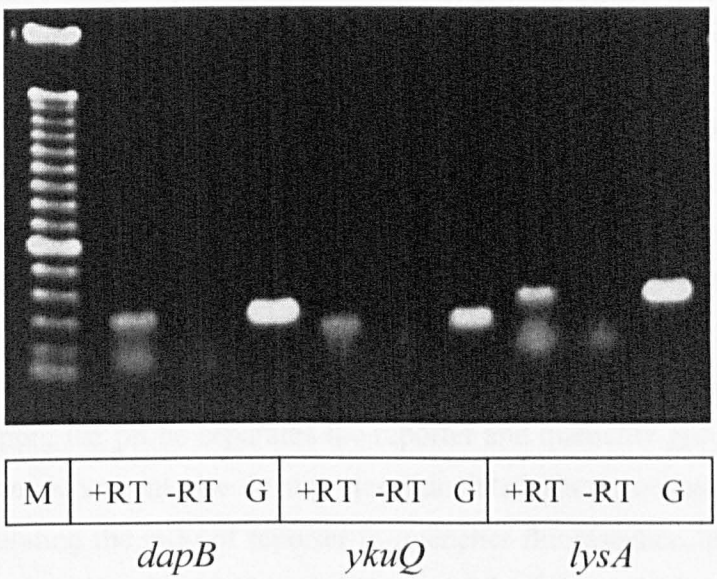


Figure 5.26

Expression of *dapB, ykuQ* and *lysA* under *in vitro* (a) and *in vivo* (b) conditions as determined by RT-PCR

In vitro, BHI; *In vivo*, Mouse kidney (pyelonephritis); +RT, + reverse transcriptase; -RT, - reverse transcriptase; G, Genomic DNA.

These results are representative of two separate experiments.

5.2.5.1.4 TaqMan

TaqMan is a fluorescence-based 5'-nuclease assay, pioneered by the Perkin-Elmer Corporation. In this assay, PCR products are detected directly without removing them from their reaction tubes. This is possible by monitoring the increase in fluorescence associated with polymerase cleavage of a quenched, dye labeled oligonucleotide DNA probe, during PCR (Figure 5.27a-c). This assay therefore provides a simple and reliable procedure for determining the presence or absence of a specific sequence (gene expression) (Perkin-Elmer, 1999).

The TaqMan assay exploits the 5' nuclease activity of TaqGold DNA polymerase (Chapter 2.21.2.3), to allow direct detection of the PCR product by the release of the fluorescent reporter during PCR amplification. The TaqMan fluorogenic oligonucleotide probes were constructed with a fluorescent reporter dye covalently linked to the 5' terminus. The 3' terminus contains both a blocking group, to prevent probe extension and a quencher, which inhibits fluorescence of the reporter group. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence (Figure 5.27b). Reactions are performed with the normal components for PCR, with the addition of the target-specific TaqMan probe (Chapter 2.21.2.3). If hybridization occurs, polymerisation of the targeted amplicon takes place. During extension of the upstream primer, Taq polymerase encounters the annealed probe and begins to displace it. This leads to 'clipping' of the probe by the 5' nuclease activity of the polymerase, which continues strand synthesis. Clipping the probe separates the reporter and quencher groups (Figure 5.27c), resulting in a concomitant rise in reporter-associated fluorescence. Results were obtained, by calculating the ratio of reporter-to-quencher fluorescence, and comparing a no-template control to an identical reaction containing the sample of interest. The difference between the two values represented the relative amount of PCR product formed (Perkin-Elmer, 1999).

TaqMan is ideally suited for gene expression analyses because quantification is based on threshold cycle (Ct), a measurement taken during the exponential phase of amplification where limiting reagents, and small differences in starting reagent amounts have yet to influence PCR efficiency. Using the fluorescent probe in combination with the appropriate instrumentation, real-time PCR experiments report a fluorescence value during every cycle that represents the amount of product amplified to that point. Ct, defined as the cycle at which fluorescence is first detectable above background, is inversely proportional to the log of the initial copy number (in this case mRNA). Thus,

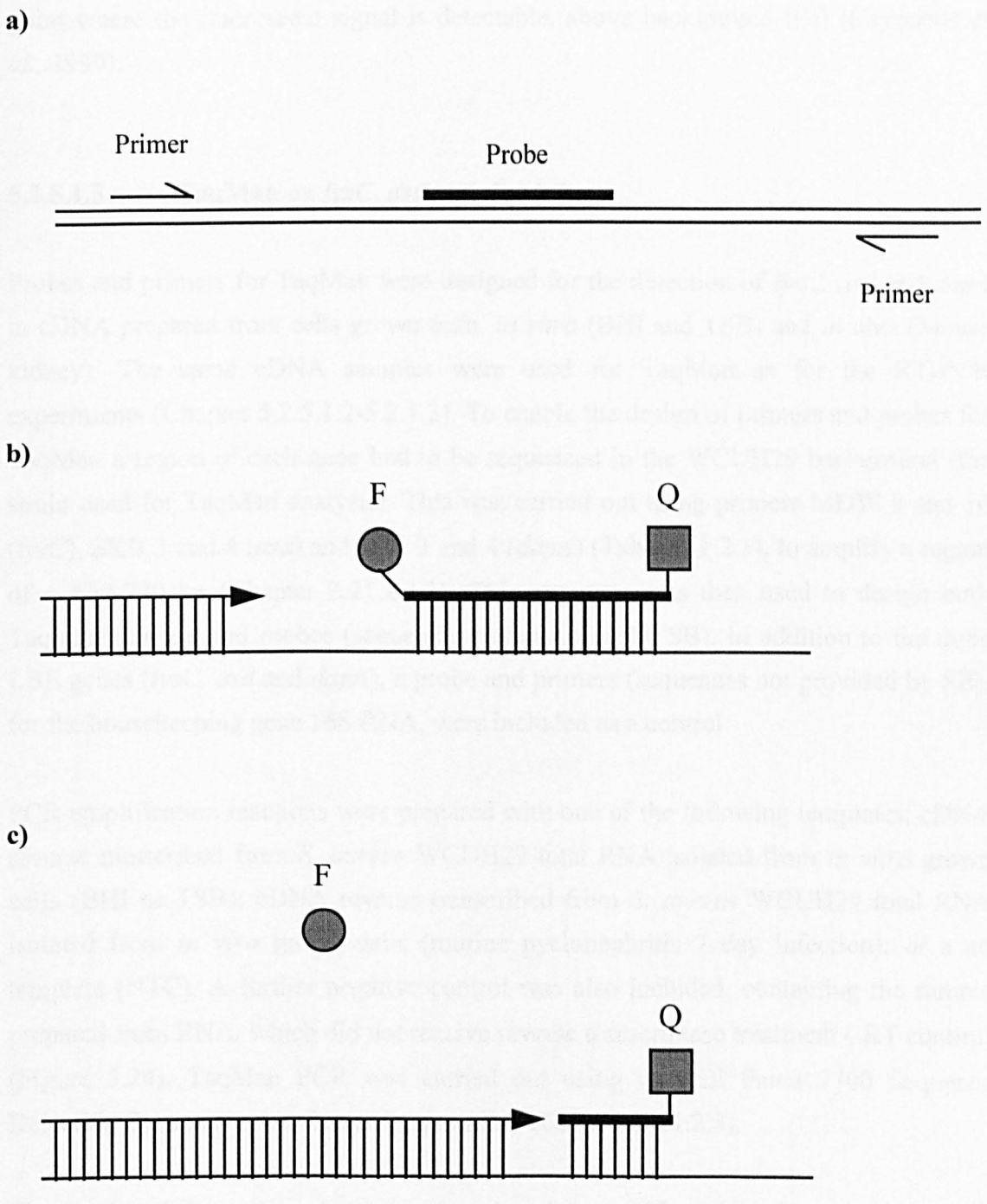


Figure 5.27
Diagrammatic representation of TaqMan PCR

F - Fluorescent reporter dye Q - Quencher reporter dye

Reproduced from Perkin-Elmer (1999).

the more template present at cycle 0, the fewer the number of cycles taken to reach a point where the fluorescent signal is detectable, above background (Ct) (Cayouette *et al.*, 1999).

5.2.5.1.5 TaqMan on *lysC, asd* and *dapA*

Probes and primers for TaqMan were designed for the detection of *lysC, asd* and *dapA* in cDNA prepared from cells grown both, *in vitro* (BHI and TSB) and *in vivo* (Mouse kidney). The same cDNA samples were used for TaqMan as for the RT-PCR experiments (Chapter 5.2.5.1.2-5.2.1.3). To enable the design of primers and probes for TaqMan a region of each gene had to be sequenced in the WCUH29 background (the strain used for TaqMan analysis). This was carried out using primers MDW 9 and 10 (*lysC*), SKB 3 and 4 (*asd*) and MW 3 and 4 (*dapA*) (Table A.1.2.1), to amplify a region of ~ 660-770 bp (Chapter 2.21.2.1.1). This sequence was then used to design both TaqMan primers and probes (sequences not provided by SB). In addition to the three LBE genes (*lysC, asd* and *dapA*), a probe and primers (sequences not provided by SB), for the housekeeping gene 16S RNA, were included as a control.

PCR amplification reactions were prepared with one of the following templates: cDNA reverse transcribed from *S. aureus* WCUH29 total RNA isolated from *in vitro* grown cells (BHI or TSB); cDNA reverse transcribed from *S. aureus* WCUH29 total RNA isolated from *in vivo* grown cells (murine pyelonephritis 7 day infection); or a no template (NTC). A further negative control was also included, containing the sample prepared from RNA, which did not receive reverse transcriptase treatment (-RT control) (Figure 5.24). TaqMan PCR was carried out using an ABI PRISM 7700 Sequence Detection System (Perkin-Elmer Corporation) (Chapter 2.21.2.3).

The results of the assay using primers and probes specific for the housekeeping gene 16S RNA, demonstrated that as predicted the gene was expressed under both *in vivo* and *in vitro* conditions (Figure 5.28a). Interestingly, the expression was significantly greater under *in vitro* than *in vivo* conditions as shown by the threshold cycle (Ct) (mean Ct: 16.5, *in vitro*; and 22.8, *in vivo*) (Table 5.6). Results from TSB and BHI were found to be very similar (results not shown).

Results of the TaqMan analysis of the three LBE *seg* showed the presence of sequences corresponding to *lysC, asd* and *dapA* in all three cDNA samples (BHI, TSB and mouse kidney). The results for the BHI and TSB samples were found to be very similar and so

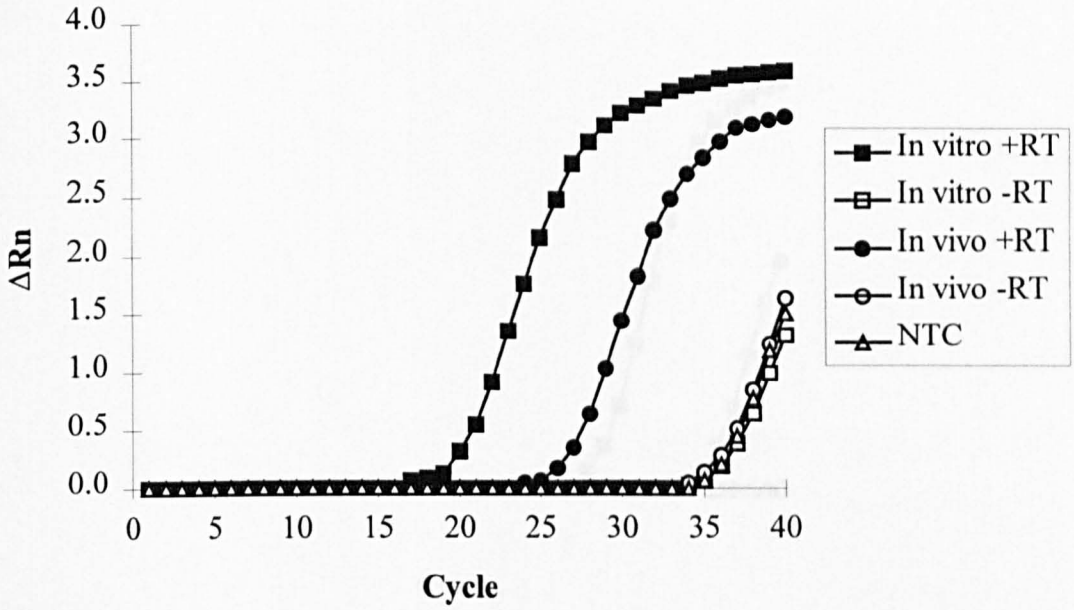
once more only results for BHI in addition to the mouse kidney samples are shown (Figures 5.28b-d). Surprisingly, these results showed that the level of LBE transcript recovered from *in vitro* grown cells (BHI) was considerably greater than the level from *in vivo* grown cell (mouse kidney). This was initially determined by the Ct value, where fluorescence is first detected (Table 5.6). Furthermore, from the Ct values, it can be determined that the level of transcript of each of the LBE genes is very similar to each other. This was determined by the fact that the Ct values are very similar, for the *in vitro* samples (25.2 - 27.3) and for the *in vivo* samples (32.3 - 34.5) (Table 5.6).

For every TaqMan reaction it is necessary to produce a calibration curve, using the same primers and probes, but using *S. aureus* genomic DNA rather than cDNA (Chapter 2.21.2.3), an example of a calibration curve is shown in Figure 5.29. By comparison of the samples with the calibration curve, the amount of mRNA of each gene is automatically calculated. These values are shown in Table 5.6. This further confirmed the very similar levels of all three LBE gene transcripts present, under the same condition (*in vivo* or *in vitro*).

5.2.5.2 Role of *lysC, asd* and *dapA* in pathogenicity

Three murine pathogenicity models were employed to study the ability of *S. aureus* to cause infection. The three models used were: subcutaneous lesion; pyelonephritis; and wound infection. These models allowed the determination as to whether *S. aureus* strains with defective AK II (*lysC*), ASADH (*asd*) or DHPS (*dapA*), were attenuated in growth/survival *in vivo,* compared to the parental wildtype strain. For the subcutaneous lesion model *lysC* (*seg*1), *asd* (*seg*24) and *dapA* (*seg*10), mutants were present in the 8325-4 background and results were compared to wildtype 8325-4. However, for both the pyelonephritis and the wound infection models, the mutations were transferred by phage transduction (Chapter 2.14) to the clinical strain WCUH29. This produced SWC1 (*lysC*), SWC24 (*asd*) and SWC10 (*dapA*), and was necessary as this work was conducted in collaboration with Dr. M. Burnham (SB). The generation of these strains was verified by Southern blot (results not shown). Thus results for these pathogenicity assays were compared to wild type WCUH29.

a)



b)

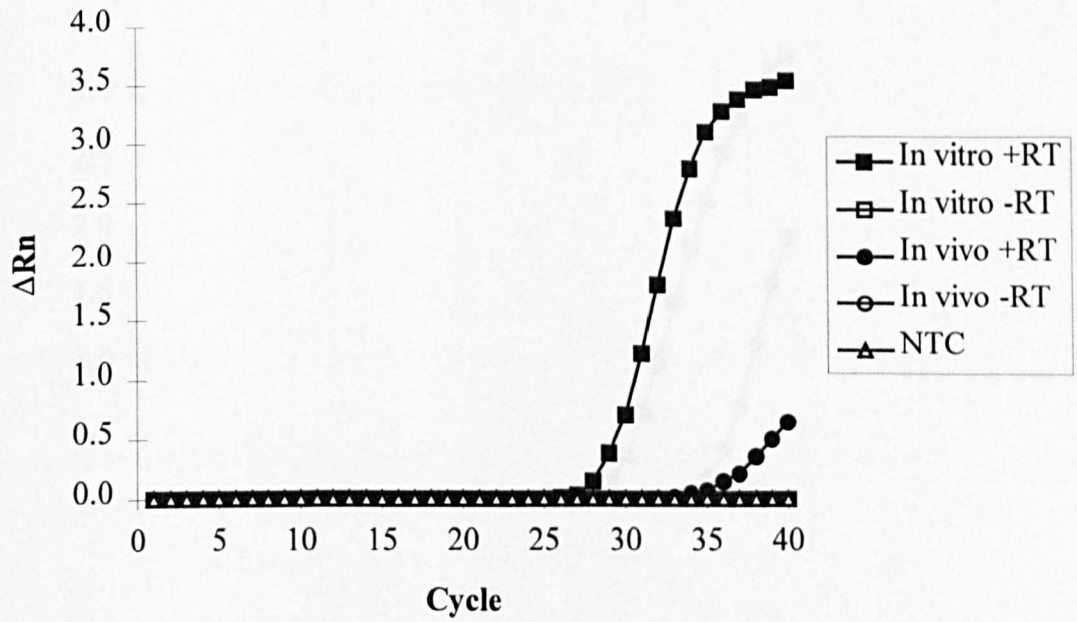


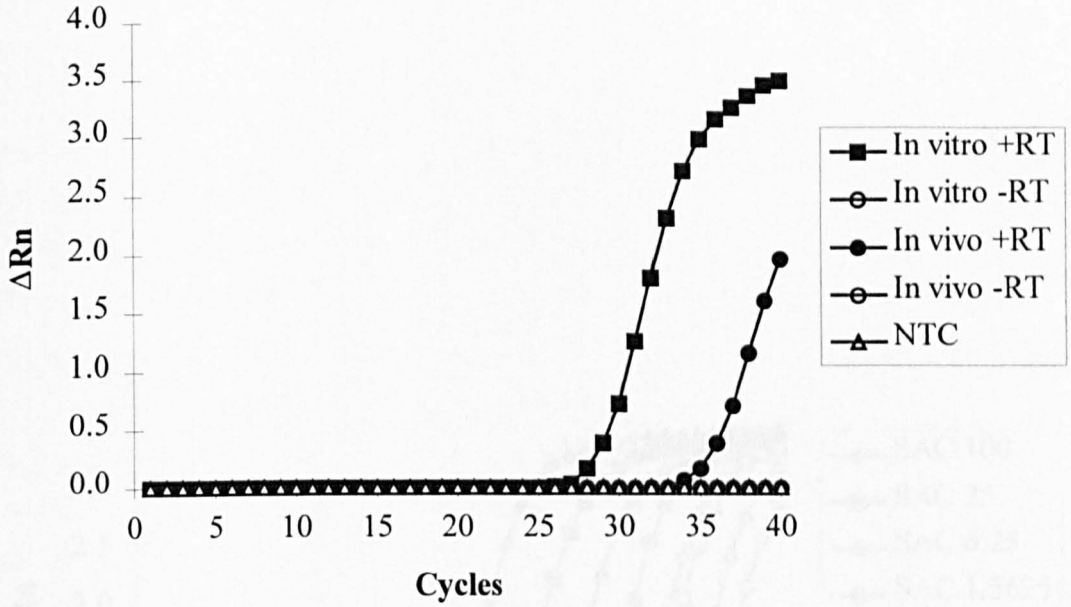
Figure 5.28

TaqMan trace detecting the presence of 16S RNA (**a**) (control), and *lysC* (**b**) sequence in the population of cDNA prepared from *S. aureus* grown *in vitro* (BHI) and *in vivo* (mouse kidney)

The results are representative of two separate experiments. ΔRn: Fluorescence of the reporter divided by the fluorescence of the passive reference dye.

Controls included: - Reverse transcriptase (-RT); No template control (NTC).

c)



d)

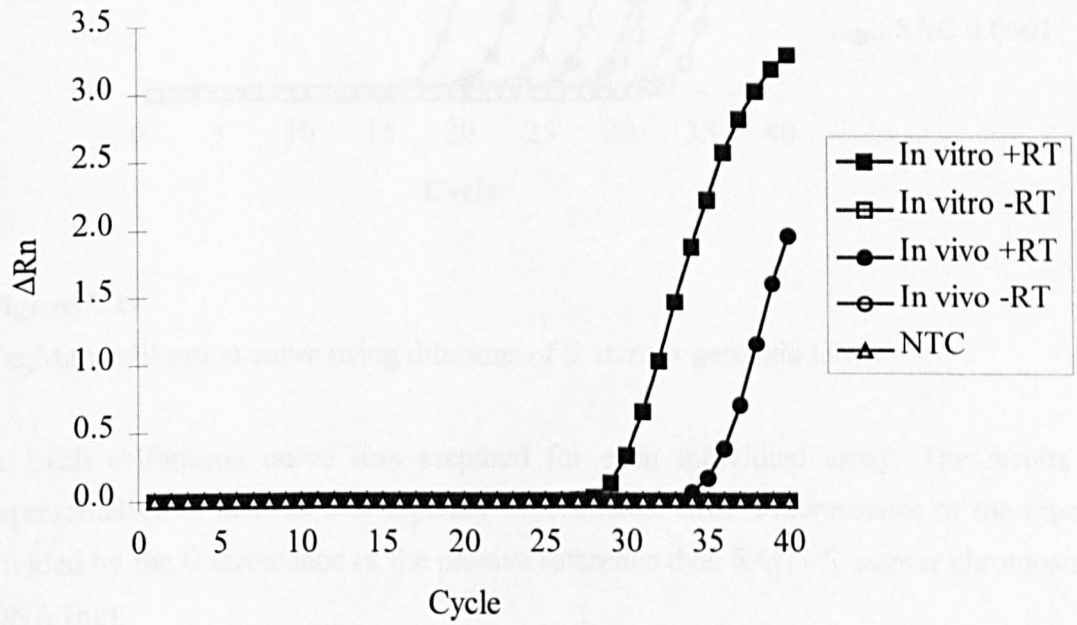


Figure 5.28

TaqMan trace detecting the presence of *asd* (**c**), and *dapA* (**d**) sequence in the population of cDNA prepared from *S. aureus* grown *in vitro* (BHI) and *in vivo* (mouse kidney)

The results are representative of two separate experiments. ΔRn: Fluorescence of the reporter divided by the fluorescence of the passive reference dye.

Controls included: - Reverse transcriptase (-RT); No template control (NTC).

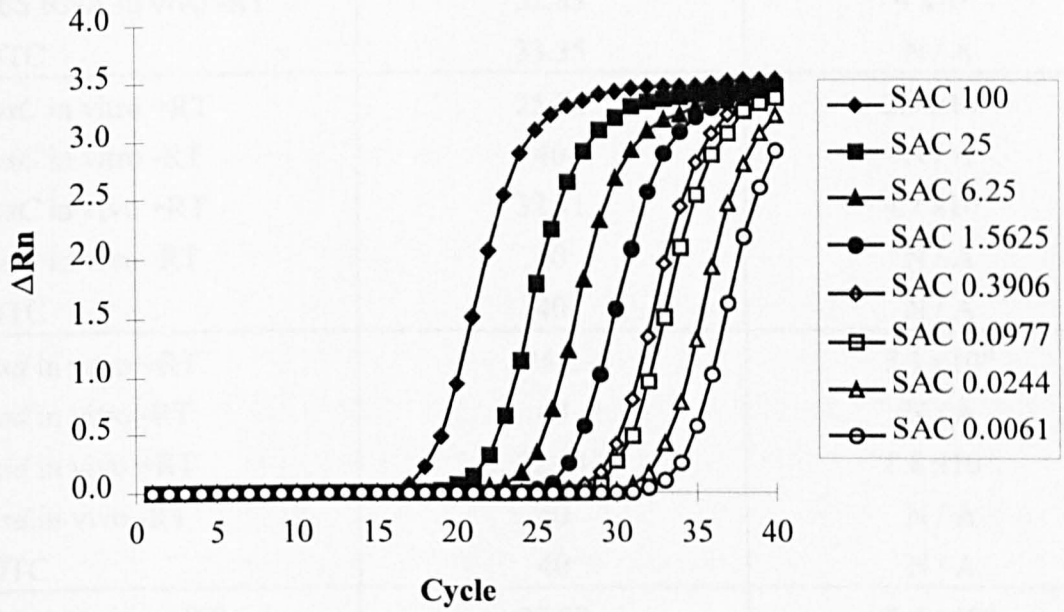


Figure 5.29

TaqMan calibration curve using dilutions of *S. aureus* genomic DNA

A fresh calibration curve was prepared for each individual assay. The results are representative of at least two separate experiments. ΔRn: Fluorescence of the reporter divided by the fluorescence of the passive reference dye. SAC - *S. aureus* chromosomal DNA [ng].

Sample Name	Average Ct	Average quantity
16S RNA in vitro +RT	16.50	$5.03 \times 10^{+1}$
16S RNA in vitro -RT	32.33	5.5×10^{-4}
16S RNA in vivo +RT	22.79	5.0×10^{-1}
16S RNA in vivo -RT	32.83	4×10^{-4}
NTC	33.35	N / A
lysC in vitro +RT	25.78	2.9×10^{-1}
lysC in vitro -RT	40	N / A
lysC in vivo +RT	32.71	4.7×10^{-3}
lysC in vivo -RT	40	N / A
NTC	40	N / A
asd in vitro +RT	25.2	5.1×10^{-1}
asd in vitro -RT	40	N / A
asd in vivo +RT	32.34	8.8×10^{-3}
asd in vivo -RT	40	N / A
NTC	40	N / A
dapA in vitro +RT	27.33	2.4×10^{-1}
dapA in vitro -RT	40	N / A
dapA in vivo +RT	34.47	4.07×10^{-3}
dapA in vivo -RT	40	N / A
NTC	40	N / A

Table 5.6

Summary of TaqMan analysis of 16S RNA, *lysC, asd* and *dapA* using cDNA prepared from both *in vitro* (BHI) and *in vivo* (mouse kidney) grown *S. aureus* cells

Average cycle numbers and quantities of transcripts are calculated from three values for +RT and two values for -RT and NTC. Results are representative of two separate experiments.

5.2.5.2.1 Role of *lysC, asd* and *dapA* in the subcutaneous lesion model

The ability of *S. aureus* to cause infection by abscess formation on the skin was investigated in collaboration with Dr. E. Ingham (University of Leeds) (Chapter 2.25.1). Following 7 days of infection 1.3 x 10^8, 1.0 x 10^8 and 7.45 x 10^7 cfu of *seg*1 (*lysC*), *seg*24 (*asd*) or *seg*10 (*dapA*) were recovered from the lesion, respectively. These values were compared to 2.51 x 10^8 cfu recovered from the 8325-4 lesion. The values for the mutant strains were shown not to be significantly different from the wild type strain (results not shown).

5.2.5.2.2 Role of *lysC, asd* and *dapA* in the pyelonephritis model

The virulence of *S. aureus* strains SWC1 (*lysC*), SWC24 (*asd*) and SWC10 (*dapA*) in comparison with the parental strain WCUH29 was determined using a murine pyelonephritis (kidney) model (Chapter 2.25.2). After 7 days of infection 5.01 x 10^5, 1.48 x 10^6 and 4.64 x 10^5 cfu of SWC1 (*lysC*), SWC24 (*asd*) or SWC10 (*dapA*) were recovered from the kidney, respectively. These four values were compared to 6.31 x 10^5 cfu recovered from the wild type (WCUH29) infected kidney. These values were found not to differ significantly from one another (results not shown).

5.2.5.2.3 Role of *lysC, asd* and *dapA* in the wound infection model

The virulence of *S. aureus* strains SWC1 (*lysC*), SWC24 (*asd*) and SWC10 (*dapA*) in comparison with the parental strain WCUH29 was determined using a murine wound infection model (Chapter 2.25.3). Following 7 days infection 5.62 x 10^6 (SWC1 (*lysC*)), 9.55 x 10^6 (SWC24 (*asd*)) or 1.12 x 10^7 (SWC10 (*dapA*)) cfu were recovered from the infected wound. These values were compared to 4.22 x 10^6 cfu recovered from the WCUH29 infected wound. Once more the LBE defective strains were not significantly attenuated in virulence, compared with the wild type (WCUH29) (results not shown).

5.2.5.3 Starvation-survival assays on LBE *seg* mutants

The role of *seg*1 (*lysC*), *seg*24 (*asd*), and *seg*10 (*dapA*), in starvation-survival under amino acid limiting conditions was investigated (Chapter 2.6.1). It has previously been

found that a *S. aureus* strain bearing a mutation in *lysA* (SPW2) has a starvation-survival defect (Watson *et al.*, 1998b). Results showed that although *seg*1 (*lysC*) and *seg*10 (*dapA*) do have a significant defect in starvation-survival compared to wild type (8325-4) after 15 days, this defect is not as pronounced as that of SPW2 (*lysA*) (Figure 5.30). Finally no starvation-survival defects were noted for any strain when incubated in phosphate buffered saline (PBS) (results not shown).

5.2.5.4 Cell wall morphology study of the three LBE mutant strains

As lysine is an essential constituent of the stem peptide in peptidoglycan of *S. aureus*, a defect in lysine metabolism may lead to alterations in the cell wall structure. Such changes could result in morphological defects to the cell. Two methods were employed to study the cell wall morphology. The first was light microscopy, which was used simply to visualize the cells in an attempt to detect any obvious morphological changes (Chapter 2.26). The second involved studying the cell wall petidoglycan using High-Performance Liquid Chromatography (HPLC) analysis.

5.2.5.4.1 Light microscopy experiments

When compared to 8325-4, none of the LBE *seg* mutants (*seg*1 (*lysC*), *seg*24 (*asd*), or *seg*10 (*dapA*)) showed any significant difference when grown in BHI or serum (results not shown).

5.2.5.4.2 Analysis on the cell wall peptidoglycan

In collaboration with Dr. A. Atrih (University of Sheffield) the peptidoglycan structure of the LBE mutants was investigated (Chapter 2.27.1). Reverse phase-HPLC analysis of muropeptides allows the determination of peptidoglycan fine structure (Atrih *et al.*, 1996). Muropeptide analysis revealed no significant difference between *seg*1 (*lysC*) and 8325-4, when grown in serum or BHI, (results not shown). Due to the lysine requirements of *seg*24 (*asd*) and *seg*10 (*dapA*), which is also essential for the growth of *seg*1 (*lysC*), muropeptide analysis was not conducted on either of the other two LBE *seg* strains.

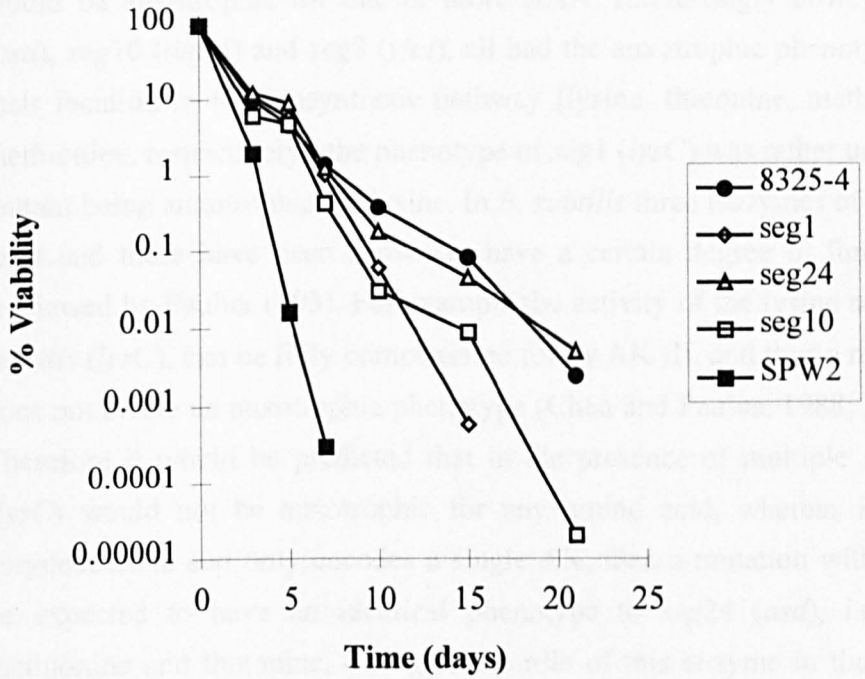


Figure 5.30

Starvation-survival kinetics of 8325-4, *seg*1 (*lysC*), *seg*24 (*asd*), *seg*10 (*dapA*), and SPW2 (*lysA*) in amino acid limiting CDM (CDM 16)

S. aureus strains, 8325-4, *seg*1 (*lysC*), *seg*24 (*asd*), *seg*10 (*dapA*), and SPW2 (*lysA*) were grown for 18 h, at 37 °C with shaking (250 rpm), in amino acid limiting CDM (CDM 16). These 6h post-exponential cultures were then incubated at 37 °C without shaking. 100 % viability corresponds to the number of viable cells counted 18 hours after inoculation (6h post-exponential), which is approximately 5 x 10^8 cfu ml^{-1}. The results are representative of two separate experiments.

5.3 Discussion

The auxotrophic nature of the AAAB *seg* strains was investigated initially on solid media, and then in liquid media owing to the difficulties in scoring growth on solid media lacking some of the aspartate amino acids (AAA). This series of experiments confirmed the predication that *seg*1 (*lysC*), *seg*24 (*asd*), *seg*10 (*dapA*) and *seg*8 (*yjcI*) would be auxotrophic for one or more AAA. Interestingly however, whereas *seg*24 (*asd*), *seg*10 (*dapA*) and *seg*8 (*yjcI*), all had the auxotrophic phenotype, predicted from their location in the biosynthetic pathway (lysine, threonine, methionine; lysine; and methionine, respectively), the phenotype of *seg*1 (*lysC*) was rather unexpected, with this mutant being auxotrophic for lysine. In *B. subtilis* three isozymes of aspartokinase (AK) exist and these have been shown to have a certain degree of functional redundancy (reviewed by Paulus 1993). For example the activity of the lysine sensitive AK II of *B. subtilis* (*lysC*), can be fully compensated for by AK III, and thus a mutation within *lysC* does not create an auxotrophic phenotype (Chen and Paulus, 1988; Zhang *et al.*, 1990). Therefore it would be predicted that in the presence of multiple AK isozymes, *seg*1 (*lysC*) would not be auxotrophic for any amino acid, whereas if *S. aureus* is like corynebacteria and only encodes a single AK, then a mutation within this gene would be expected to have an identical phenotype to *seg*24 (*asd*), i.e. requiring lysine, methionine and threonine, owing to the role of this enzyme in the biosynthesis of all AAA. Therefore the result that a *lysC* deficient strain was auxotrophic for lysine was surprising. However, the phenotype of the *lysC* mutant was confirmed using another *seg* mutant also containing an insertion within *lysC* (*seg*25). The auxotrophic nature of a *lysC* mutant can only be explained by the presence of at least one further AK isozyme in *S. aureus* that is able to rescue the methionine and threonine functions of AK II, but not that of lysine biosynthesis.

A further problem with the auxotrophy results, that only became apparent after the operonic organisation of the LBE genes was determined, was that a mutation within *lysC* (*seg*1, *seg*25) did not have a polar effect on the downstream genes including *asd*. This was determined as *seg*1/*seg*25 (*lysC*) are auxotrophic only for lysine and not for threonine and methionine. This strongly suggests the presence of a further promoter upstream of *asd*, enabling transcription of the other genes of the *dap* operon, as without this additional promoter then *seg*1/*seg*25 (*lysC*) would also have a deficiency for ASADH, and thus would be auxotrophic for lysine, threonine and methionine.

One last observation, made from these auxotrophy studies, was that the growth rate, but not the growth yield of the wild type strain 8325-4, as well as that of the mutant strains

was dramatically affected by the number of AAA included in the defined media (CDM). As a rule the fewer AAA included in the media, the greater the reduction in the growth rate. Interestingly however, lysine appears to be the most important of the AAA, as the addition of this amino acid alone increases the growth rate significantly more than the addition of any of the other AAA individually. This is probably due to the dual role of lysine in *S. aureus*, with it not only required as a building block for proteins, but also as component of the cell wall peptidoglycan.

PCR and DNA sequencing strategies, revealed the putative *dap* operon of *S. aureus* to contain a total of 8 genes, 6 of which potentially encode enzymes involved in the biosynthesis of lysine (Figure 5.11). This operon therefore represents the major locus for lysine biosynthesis in *S. aureus*. The first gene of the *dap* operon *lysC*, codes for an isozyme of AK, the first enzyme required for the conversion of aspartate to lysine (Chapter 5.1.3). Many bacterial AKs including *lysC* of *B. subtilis* (Bondaryk and Paulus, 1985) have been found to be composed of two dissimilar subunits (α and β) (Chapter 5.1.4.1). These AK subunits in the above examples, have been shown to be encoded by in-phase overlapping coding regions. Analysis of the sequence of *lysC* from *S. aureus* revealed that an identical situation may also be present for this AK. This is postulated due to the presence of a start codon (ATG) and a putative ribosome binding site (Shine-Dalgarno), at approximately the correct position as predicted from the structure of *lysC* from *B. subtilis* (Chen *et al.* 1987). This prediction is also supported by the strong sequence homology between *lysC* of *S. aureus*, and *lysC* of *B. subtilis* (40 % over 397 amino acids). The second and third genes of the *dap* operon are *asd* and *dapA* encoding the second enzyme of the common pathway (ASADH) (Chapter 5.1.3), and the first enzyme in the branch specific for lysine biosynthesis (DHPS) (Chapter 5.1.4.1). As these genes have already been reviewed extensively these genes will not be further discussed here.

The two genes immediately downstream of *dapA*, are *dapB* and *ykuQ* (Figure 5.11), which potentially encode the fourth and fifth enzymes respectively for the biosynthesis of lysine. These enzymes are, dihydrodipicolinate reductase (*dapB*) and tetrahydrodipicolinate acetyltransferase (*ykuQ*), which catalyse the conversion of dihydrodipicolinate to tetrahydrodipicolinate, and tetrahydrodipicolinate to acetyl-amino-ketopimelate (Figure 5.2).

Dihydrodipicolinate reductase (DHPR) of *B. subtilis*, encoded by *dapB* is inhibited in a noncompetitive manner by physiological concentrations of dipicolinate. This inhibition is thought to function as a regulator in sporulating cells by diverting the common

intermediate dihydrodipicolinate from the synthesis of DAP and lysine to produce more dipicolinate as this product begins to accumulate in spores (Kimura and Goto 1975). Therefore DHPR of *S. aureus* is thought to have a slightly different role in *S. aureus* with it being a non-sporulating bacterium and thus not synthesizing dipicolinate. As a result DHPR of *S. aureus* is only involved in the biosynthesis of lysine and like the DHPR enzymes of *E. coli* and *C. glutamicum,* DHPR of *S. aureus* is predicted to have no regulatory function.

Tetrahydrodipicolinate succinylase (THPS) and tetrahydrodipicolinate acetyltransferase (THPA) catalyse the first reactions in the epimerase pathway in *E. coli* and *B. subtilis* respectively. THPA of *B. subtilis* is thought from this study to be encoded by *ykuQ,* since this gene has been classed as unknown, similar to tetrahydrodipicolinate succinylase in SubtiList. However, as *B. subtilis* uses acetylated, rather than succinylated intermediates for lysine biosynthesis (Weinberger and Gilvarg, 1970), it is predicted that the polypeptide encoded by *ykuQ* is in fact THPA. Whether *S. aureus* uses acetylated intermediates, like *B. subtilis*, for the biosynthesis of lysine, rather than succinylated intermediates as in *E. coli* is not known. However, the discovery that the sequence of the fifth ORF in the *dap* operon had greatest homology to *ykuQ* of *B. subtilis* (60 % over 233 amino acids) suggests that like *B. subtilis*, *S. aureus* also uses acetylated intermediates. Therefore the gene product of *ykuQ*, of *S. aureus* like that of *B. subtilis* is predicted to be THPA.

The final lysine biosynthetic enzyme encoded by the *dap* operon (8^{th} ORF) is diaminopimelate decarboxylase (DAP decarboxylase). This enzyme is encoded by *lysA,* and catalyses the final step in the biosynthesis of lysine (Figure 5.2). This step is the pyridoxal phosphate-dependent decarboxylation of *meso*-DAP, to lysine. In *E. coli,* a positive regulatory element, *lysR,* whose product is essential for *lysA* expression, is transcribed divergently from *lysA* (Stragier *et al.*, 1983). However, none of the deduced products of the ORFs surrounding either *lysA* of *B. subtilis* (SubtiList) of *S. aureus* have any similarity to the *lysR* product. However, a gene remote from the *lysA* locus, that leads to the loss of DAP decarboxylase activity in *B. subtilis* may represent an analogous control element to *lysR* in *B. subtilis* (reviewed by Paulus, 1993). Therefore a similar system may be found in *S. aureus*. This situation however is complicated in *S. aureus* as this study revealed the presence of a DAP decarboxylase isozyme (*dapH*). This was a surprising discovery, as a homology search revealed that neither *B. subtilis* nor *E. coli* possess multiple isozymes of DAP decarboxylase (results not shown). Finally, a homologue of *lysR* was not found to be located adjacent to *dapH* in *S. aureus*. The reason for the presence of multiple isozymes of DAP decarboxylase in *S.*

aureus, but not in *B. subtilis,* may be linked to the dual role of this enzymatic step in *S. aureus*. Whereas in *B. subtilis* DAP decarboxylase is only required for the conversion of DAP to lysine, for use in protein biosynthesis (DAP is the cell wall component). In *S. aureus,* in addition to its role in protein synthesis, lysine is also an essential component of the cell wall peptidoglycan.

The remaining two genes of the *dap* operon are: *hipO* and *dal* which encode hippurate hydrolase and alanine racemase respectively (Figure 5.11). Hippurate hydrolase is the first of the two non-lysine biosynthetic genes encoded by the *dap* operon. This enzyme is an amidohydrolase also known as hippuricase, and acts to cleave benzoglycine (hippuric acid) into the constituent products, benzoic acid and glycine (Hani and Chan, 1995). Hippurate hydrolase has no apparent function in the biosynthesis of lysine and thus its presence in the *dap* operon was initially unexpected. A possible explanation behind a gene encoding hippurate hydrolase being present in the *dap* operon, was discovered following the examination of the organisation of the hippurate hydrolase genes in *B. subtilis*. Searching the SubtiList database revealed that there are possibly two isozymes of hippurate hydrolase in *B. subtilis*: *hipO* (hippurate hydrolase) and *ykuR* (unknown, similar to hippurate hydrolase). Interestingly, the second isozyme of hippurate hydrolase encoded by *ykuR* is in an operon immediately downstream of *ykuQ* (THPA). Thus the organisation of the genes encoding THPA and hippurate hydrolase is identical in *B. subtilis* and *S. aureus*. Furthermore, analysis of the *S. aureus* databases revealed that like *B. subtilis, S. aureus* also encodes at least two putative isozymes of hippurate hydrolase. This second isozyme of hippurate hydrolase (named, *amh*X), shows greater homology to *hipO* of *B. subtilis* than the isozyme encoded by the *dap* operon (results not shown), indicating that the hippurate hydrolase in the *dap* operon is the homologue of *ykuR*.

Alanine racemases are unique prokaryotic enzymes that interconvert L-alanine and D-alanine, and are the sole identified biosynthetic pathway of D-alanine for bacterial cell wall synthesis (Wasserman *et al.*, 1984). D-Ala is generally present as a dipeptide, D-alanyl-D-Ala, in the C-terminal position of UDP-N-acetylmuramyl (MurNAc)-pentapeptide precusor of the peptidoglycan. The penultimate D-Ala residue of this precusor is directly involved in the cross-linking of adjacent peptidoglycan stands in cell wall growth. Furthermore, the cell wall of *S. aureus* and other Gram-positive bacteria contains teichoic acids. These teichoic acids contain various substituents, such as D-Ala esters and glycosyl residues. Thus in *S. aureus,* D-Ala is a central molecule in the biosynthesis of two cell wall polymers, peptidoglycan and teichoic acids (Hols *et al.* 1997).

In *E. coli* two alanine racemases have been identified. The *alr* encoded alanine racemase (biosynthetic racemase), is constitutively expressed whereas the *dadX* encoded enzyme (catabolic racemase), is essential only for L-alanine catabolism (Lilley *et al.*, 1993; Wild *et al.*, 1985). The *dadX* gene product is the major source of alanine racemase (85 % of total activity) and is probably a secondary source of D-Ala for cell wall biosynthesis (Wasserman *et al.*, 1984). In *B. subtilis* a single racemase gene (*dal*) has been identified, however a Dal$^-$ mutant is not dependent on D-Ala under all conditions suggesting that a second alanine racemase is present in *B. subtilis* (Ferrari *et al.*, 1985). The second isozyme of alanine racemase in *B. subtilis* was identified during the course of this study, by searching the SubtiList database; *yncD* (unknown, similar to alanine racemase). Interestingly, both *dal* and *yncD* showed significant homology to the *dap* operon encoded alanine racemase (*dal*) of *S. aureus* (27% over 331 amino acids (*dal*); and 23% over 371 amino acids (*yncD*)).

The identification of an alanine racemase in the *dap* operon confirms the fact that *S. aureus* like *E. coli* and *B. subtilis* encodes two alanine racemases. The first alanine racemase of *S. aureus* (*alr*) was identified 5' to the *sigB* operon by Kullik and colleagues (1998b). Their work suggested that a mutation polar on this gene didn't affect the D-Ala substitution of the wall teichoic acid, and thus it was proposed that a second alanine racemase is present in *S. aureus*. Furthermore, identification of the second alanine racemase (*dal*) is in agreement with the non-essential nature of *alr* in *S. aureus*, which has been found to be essential in some bacteria, for example *Lactobacillus plantarum* (Hols *et al.*, 1997). The alanine racemase of the *dap* operon has greater homology to *dal* of *B. subtilis* than *yncD* (homologue of *alr* from *E. coli*) and thus in accordance has been named *dal*. The reasoning for the inclusion of *dal* in a lysine biosynthetic operon is not exactly clear. However, as described above, alanine racemase has an important role in cell wall biosynthesis, which may explain its presence in the *dap* operon, owing to the involvement of lysine in the cell wall peptidoglycan.

This study has identified the gene encoding the third AK isozyme of *B. subtilis* (*yclM*). Interestingly, a homologue of *yclM* was found in *S. aureus* (*ask*), however comprehensive analysis of *S. aureus* databases revealed there to be no isozyme of *dapG* (AK I) in *S. aureus*. Therefore from this work it would appear that unlike both *B. subtilis* and *E. coli*, *S. aureus* encodes only two AK isozymes. Due to the homology of these genes with other AKs in the database, it is predicted that, the *dap* operon encoded AK (*lysC*) is a lysine regulated AK, while the other (*ask*), is threonine-plus-lysine regulated. The presence of at least two AKs isozymes in *S. aureus* was initially

predicted, from the fact that mutants deficient in *lysC* (*seg*1, and *seg*25) were auxotrophic for only lysine. This theory was further supported by the fact that in bacteria with only one AK, for example *C. glutamicium* and *Streptomyces akiyoshiensis*, this enzyme (encoded by *ask*) is regulated by threonine-plus-lysine (Cremer *et al.*, 1991; Le *et al.*, 1996), whereas the AK identified by the serum screen, was predicted to be only sensitive to lysine. This dual amino acid regulation is essential for organisms with only one AK as it affords control to both sides of the biosynthetic pathway (Paulus, 1993). Therefore the discovery of the second AK of *S. aureus* was in keeping with the above predictions. The fact that *S. aureus* does not encode a DAP sensitive AK, is not that surprising considering that the only role that DAP plays in *S. aureus* is as a precusor for lysine, whereas in *B. subtilis*, DAP is an essential component of the cell wall peptidoglycan.

The presence of THPA (*ykuQ*) in the *dap* operon, suggests that *S. aureus* like *B. subtilis* and *E. coli*, synthesize lysine via the epimerase pathway. However, conclusive proof that *S. aureus*, unlike *C. glutamicum*, is unable to also synthesize lysine via the dehydrogenase pathway was not ascertained owing to the incomplete nature of the publicly available databases (TIGR and NCTC 8325). However, from published data it would appear that the presence of both the epimerase and the dehydrogenase pathways is limited to certain coryneform bacteria, such as *C. glutamicum* and *Brevibacterium lactofermentum* (Schrumpf *et al.*, 1991). Although in *Bacillus marcerans*, DAP dehydrogenase has been detected in addition to two of the epimerase variant enzymes (Bartlett and White, 1985). Therefore the possibility of two mechanisms of lysine biosynthesis existing side by side in *S. aureus* cannot at this time be ruled out.

Analysis of the mechanism of methionine biosynthesis in *S. aureus*, although not totally conclusive, very strongly suggests that *S. aureus*, like *B. subtilis* is capable of only synthesizing methionine via cystathionine (Figure 5.3). This argument is two fold: Firstly *seg*8 (*yjcI*) is a methionine auxotroph; and secondly, database searches failed to identify a homologue of O-acetylhomoserine sulphydrylase.

The identification of a third gene (*metX*) with homology to both *yjcI* (Cystathionine γ-synthase (CTGS)) and *yjcJ* (cystathionine β-lyase (CTBL)), of *S. aureus* and *B. subtilis*, complicate the issue of methionine auxotrophy for *seg*8 (*yjcI*), due to the tentative classification of *metX* as an isozyme of CTGS. This was due to the fact that *metX* showed greatest homology to *yrhB* of *B. subtilis*, which has been classified as unknown, similar to CTGS (SubtiList). However, this classification was never completely confirmed mainly due to the considerable identity between, *yjcI*, *yjcJ* and *metX*/*yrhB*,

pointing to a common ancestor for these proteins. This tentative classification of *metX* as an isozyme of CTGS did however hold throughout the course of the study, owing to the arrangement of methionine biosynthetic genes in *S. aureus*. However, the possibility still remains that *metX* and *yrhB* both encode enzymatically inactive proteins.

In *S. aureus* the organisation of methionine biosynthetic genes is different from that in *B. subtilis* (Figures 5.14-5.17). Analysis of *S. aureus* databases revealed that in *S. aureus* CTGS (*yjcI*) and CTBL (*yjcJ*) are located in an operon (*yjc*) with two further genes involved in methionine biosynthesis. Interestingly, both these genes show significant homology to methionine synthase, which is the final enzyme in the biosynthesis of methionine (Figure 5.3). Closer inspection revealed that these two genes encode different isozymes of methionine synthase. The first methionine synthase isozyme, (named *metH*), shows greatest homology (46 % identity over 597 amino acids) to *yitJ* of *B. subtilis*. However, *yitJ* is classified as unknown, similar to unknown proteins in the SubtiList database. Importantly, this synthase has strong homology (33 % over 281 amino acids) to the vitamin-B12 dependent methionine synthase isozyme of *Mycobacterium tuberculosis* (*metH*). The second methionine synthase isozyme, (named *metE*), shows greatest homology (52 % identity over 697 amino acids) to methionine synthase, vitamin-B12 independent isozyme of *B. subtilis* (*metC*). The discovery of two isozymes of methionine synthase in *S. aureus* was interesting since previously *E. coli* was thought to be an exception, in that it possesses both methylase isozymes (*metE* and *metH*), (Cohen and Saint-Girons, 1987). However, a BLAST search on the *B. subtilis* chromosome also showed the presence of both isozymes of methionine synthase; *yitJ* (*metH* homologue), and *metC* (*metE* homologue) (results not shown). Thus in regards to methionine synthase, *E. coli* is perhaps not as exceptional as once thought.

Despite the differences in the organisation of the methionine biosynthesis genes of *S. aureus* and *B. subtilis*, in both organisms, CTGS and CTBL are adjacent, potentially within an operon (Figure 5.14-5.15). As CTGS is upstream of CTBL, a mutation within this gene (*seg*8) will be polar on CTBL. Thus a second functional CTGS (*metX*) will make no difference to the growth of the mutant strain (*seg*8), on media lacking methionine, since without CTBL the biosynthetic pathway to methionine remains blocked. The polar effect is particularly significant for *S. aureus*, which has four putative genes encoding MBE, encoded by the putative *yjc* locus (Figure 5.15). The organisation of these genes is such that if they represent a single transcriptional unit, then a mutation in CTGS (*seg*8) would be polar on the three downstream MBE genes, thus producing a methionine auxotrophic phenotype despite the presence of a second

isozyme of CTGS. Therefore the original classification of *metX* as a second isozyme of CTGS was upheld.

The role of the three LBE sera expressed genes, was analysed in terms of their role in the starvation-survival response of *S. aureus*. Results of this analysis in amino acid limiting CDM (16), demonstrated that both *lysC* and *dapA* mutants showed a starvation-survival defect. This analysis also confirmed the previous observation, that a strain with a defective *dap* operon encoded *lysA* (SPW2), had a significant starvation-survival defect (Watson *et al.*, 1998b). The conformation of the starvation-survival defective phenotype was surprising in light of the discovery that *S. aureus* encodes two isozymes of DAP decarboxylase (*lysA*; *dapH*). Furthermore, as SPW2 (*lysA*) is not auxotrophic for lysine (results not shown), *dapH* must be active under some conditions.

The LBE::*lacZ* transcriptional fusion strains constructed in the wild type background (MDW41 (*lysC*::*lacZ*); (MDW42 (*asd*::*lacZ*); and (MDW43 (*dapA*::*lacZ*)), were used initially to study the expression of the three LBE *seg* in the serum model. These results confirmed the preferential expression of *lysC, asd* and *dapA* in serum compared to BHI, with expression levels between 4- and 10-fold higher in serum. Interestingly, the expression patterns were very similar for all three LBE genes, with maximal expression occurring during the transition from post-exponential to stationary phase. Interestingly, this analysis appeared to reveal that while *lysC* and *asd* were expressed at very similar levels, *dapA* was expressed at a slightly reduced level. The reason for this increased expression of *lysC* and *asd* would be argued to be due to the involvement of these two genes in the biosynthesis of all four amino acids, and not just in that of lysine as for *dapA*. However, result from the analysis of transcript levels using TaqMan showed that all three genes were expressed to an almost identical level to each other under both *in vivo* and *in vitro* conditions. As TaqMan is thought to be a more sensitive and reliable method for the determination of transcript levels, the apparent decrease level of *dapA* transcription, may be an artifact.

Using the transcriptional fusions to *lysC* (MDW41), *asd* (MDW42), and *dapA* (MDW43), the role of each of the four AAA (lysine, threonine, methionine and isoleucine) was investigated. The major findings of this series of experiments were that: i) all three LBE *seg* were regulated (repressed) at the transcriptional level by lysine; and ii) the other AAA (threonine, methionine and isoleucine) also appear to have a secondary role in the regulation of these LBE genes.

The aspartokinase (AK) identified in the serum screen, showed strong homology to the lysine sensitive AK of *B. subtilis* (*lysC*), (40 % over 397 amino acids). As a result from homology analysis alone it was predicted that *lysC* of *S. aureus* would also be regulated by lysine. This assumption proved correct, and thus this work describes the identification of a lysine sensitive (regulated) AK in *S. aureus*. Importantly however, the genetic location of the lysine sensitive AK (*lysC*), is different in *B. subtilis* and *S. aureus*. Therefore the important question exists, as to whether these two lysine sensitive AK isozymes are regulated by lysine in the same manner?

The mechanism for lysine regulation of *lysC* in *B. subtilis* has been the subject of much work, but despite this, the exact details of the regulation are not understood. As described in Chapter 5.1.3.1.2, *lysC* is regulated at the transcriptional level by lysine, and when lysine is present the *lysC* operon is repressed by > 100-fold (Kochhar and Paulus, 1996). Work has revealed that lysine has no effect on the molar amounts of total *lysC* RNA synthesized, but rather caused a quantitative shift of full-length 1.6 kb *lysC* mRNA to a truncated 0.27 kb RNA species (Kochhar and Paulus, 1996). In other words lysine regulates the expression of the *lysC* operon by effecting the premature termination of transcription, at a rho-independent terminator site within the *lysC* leader region, upstream of the translational start site of *lysC*. The results of the reporter gene fusion assays using MDW41, revealed a decreased transcription of *lysC* in the presence of lysine. However, owing to the nature of the *lysC*::*lacZ* fusion (Chapter 5.2.4.1), regulation by premature transcript termination, would also be observed as a reduction in β-galactosidase production. Thus these results are in keeping with an identical mechanism of *lysC* regulation by lysine in both *B. subtilis* and *S. aureus*.

Central to the regulation of *lysC* (premature transcript termination) in *B. subtilis* is an extensive untranslated region upstream of *lysC*. This region is 330 bp long and contains several control features, including terminator, antiterminator structures and a putative 24 amino acid ORF (the *lysC* leader peptide), which is thought to play a role in the regulation of *lysC* (Chen *et al.*, 1987).

S. aureus has the potential for the production of an extensive untranslated leader region, as the closest gene to *lysC* (*ykpA*) has its stop codon, > 850 bp upstream of the translational start of *lysC*. Unfortunately, the position of the *lysC* promoter was not mapped during this study therefore it is not known whether *S. aureus* is in fact like *B. subtilis* in possessing an extensive untranslated leader region. However, owing to the discovery that a transposon insertion, 31 bp upstream of the start codon of *lysC*, is within a transcribed region (active *lacZ* gene on serum) indicates a potential role for the

upstream region, in the regulation of *lysC* of *S. aureus*. The putative *lysC* leader region of *S. aureus* has the potential for the production of several short polypeptides, owing to the presence of multiple ORFs. One such ORF has the potential for the production of a 25 amino acid peptide, with a dyad of lysine codons in an identical position to the *lysC* leader peptide of *B. subtilis* (positions 2 and 3). Therefore this putative peptide has tentatively been classed as, the *lysC* leader peptide. However, three factors are against this ORF encoding a leader peptide. Firstly, other than the dyad of lysine codons at the start of the putative peptide, there is no significant similarity between the two ORFs at the amino acid level, indicating that they may not have evolved from the same sequence, and thus the ORF may not encode a *lysC* leader peptide. Secondly, the ORF is located approximately 700 bp upstream of the start site (ATG) of the *lysC* gene. This would require the untranslated leader region to be ≥ 700 bp long, making it over twice the length of the leader regions found in both *B. subtilis* and *E. coli* (Chen *et al.*, 1988; Cassan *et al.*, 1983). Finally, even if the putative leader peptide ORF, is located downstream of the *lysC* promoter, owing to the absence of an apparent ribosome binding site upstream of the putative leader peptide of *S. aureus*, there is some doubt as to whether this peptide is translated or not. Importantly, despite the lack of evidence for the translation of the leader peptide of *B. subtilis* (Kochhar and Paulus, 1996), a consensus ribosome binding site can be identified upstream of the leader peptide ORF. Further evidence for a different mechanism of regulation of *lysC* by lysine in *S. aureus* and *B. subtilis* came from the failure to identify structures potentially resembling terminator, antiterminator elements within the putative *lysC* leader region. However, further analysis is required to confirm this postulation.

Therefore despite *lysC* of *S. aureus* being regulated at the transcriptional level by lysine and the possible involvement of upstream sequences, owing to the apparent absence of sequences resembling those of terminator, antiterminator elements, the regulation of *lysC* is predicted from sequence analysis to be different to that which occurs in *B. subtilis*. It is however of interest that the lysine sensitive AK (*lysC*) of *E. coli* also has an extensive untranslated leader transcript (308 bp). This region however, differs from that of *B. subtilis* by the absence of elements resembling a transcription attenuator (Cassan *et al.*, 1983). Therefore it is possible that *lysC* of *S. aureus* is regulated in the same manner as found in *E. coli*, which although is not known, is believed to occur at the transcriptional level (Richaud *et al.*, 1980). Alternatively, yet another different regulatory mechanism may be employed for the control of *lysC* transcription in *S. aureus*.

The discovery that *asd* and *dapA* of *S. aureus* are regulated by lysine is not unexpected, especially in view of their inclusion in the *dap* operon of *S. aureus* immediately downstream of the lysine regulated *lysC*. Importantly however, results from auxotrophy studies have revealed that *asd* must have its own promoter, as a mutation in *lysC* does not have a polar effect on *asd*.

As mentioned above threonine, methionine and isoleucine also appear to play a secondary role in the regulation of *lysC*, *asd* and *dapA* gene expression. For example all three LBE *seg* were repressed by a further 2- 3-fold in the presence of all four AAA, compared to the inclusion of lysine alone. Therefore, it appears that the AAA have an additive (multivalent) effect on the repression of the LBE *seg*. This type of cooperative regulation of biosynthetic genes is not unusual and is in agreement with the type of control witnessed for *asd* of *E. coli* (Boy and Patte, 1972). Conversely, there is evidence for the fact that methionine, far from repressing transcription of *lysC*, *asd* and *dapA* induces the synthesis of all three of these LBE genes by > 2-fold compared to media in which methionine is absent. This would appear to be the first observation of an induction of AK, ASADH and DHPS by methionine, and is the complete reverse of the effect of methionine on *asd* from *B. megaterium* (Chatterjee and White, 1982). Importantly however, in the presence of both methionine and lysine, these LBE genes are still strongly repressed indicating that activation by methionine is a secondary regulatory function.

From analysis of the distances between genes, it could be predicted that the entire 8 gene *dap* operon, could be transcribed as a single transcriptional unit despite the potential for the presence of multiple promoters. Supporting this postulation is the fact that there are apparently no strong candidate rho-independent terminator structures, within the *dap* operon, indicating the potential for the production of a single 8.5 kb transcript. However, attempts at Northern blots probing for *lysC*, *asd* and *dapA* transcripts were unsuccessful, failing to identify any significant transcript. Thus the size and nature of mRNA transcripts from the *dap* locus remains unknown.

The importance of the *dap* operon was originally shown by its identification in the serum screen. As the serum screen was designed to mimic *in vivo* conditions it was predicted that the *dap* operon, or more accurately *lysC, asd* and *dapA* may be important *in vivo*. The results of the *in vivo* expression studies confirmed that serum is a valid model for the identification of genes induced *in vivo*. Using both RT-PCR and TaqMan, all three LBE *seg* were found to be expressed in the mouse kidney (pyelonephritis) model. In addition RT-PCR demonstrated that the three other LBE genes (*dapB, ykuQ*

and *lysA*) encoded by the *dap* operon of *S. aureus* are also expressed in the mouse kidney model. However, significant expression of all 6 of the *dap* operon encoded LBE was also found in BHI. Additionally, TaqMan showed significant expression *in vitro* as compared to the kidney model. Thus it is likely that expression levels in the mouse kidney are significantly lower than those in serum. *S. aureus* is able to grow in a wide range of niches and so it may be that in blood, expression of the *dap* operon would be found to be much greater than that witnessed in the mouse kidney. Interestingly, these results correspond well with those published recently by Vriesema *et al*, (2000). In this study at least three *S. aureus* LBE genes encoded by the *dap* operon (*lysC, asd* and *dapB*) were identified as expressed on contact with human endothelial cells. In addition a DAP decarboxylase (*lysA/dapH*) was also identified, although it is not possible from the published data to determine which isozyme of DAP decarboxylase was identified. Expression of these four LBE genes, was also found to be induced by human serum. Thus the results published by Vriesema *et al*, (2000) and the results reported in this study are in complete agreement, with both studies showing the preferential expression of LBE genes in sera.

The roles of AK II (*lysC*), ASADH (*asd*) and DHPS (*dapA*) in infection were investigated with the aid of three murine pathogenicity models. In all 3 models (subcutaneous lesion, pyelonephritis and wound infection), there was no significant difference in the number of cells recovered from the host for any of the mutant strains (*seg*1/SWC1 (*lysC*); *seg*24/ SWC24 (*asd*); or *seg*10/SWC10 (*dapA*)), compared to the number recovered for the wild type strain (8325-4/WCUH29). Thus it can be concluded that none of the LBE *seg* mutant are important for growth or survival under these *in vivo* conditions. However, one fact that should be kept in mind when examining data from pathogenicity studies is that the assay for virulence may not be sufficiently sensitive or even appropriate for the virulence factor being examined (Projan and Novick, 1997). This statement is particularly pertinent in this situation as these results appear to contradict previously published work in particular, the work of Mei *et al.* (1997), and Coulter *et al.* (1998) who carried out signature-tagged mutagenesis (STM) on *S. aureus*. Both of these studies identified *asd,* which by definition of the technique must be attenuated in virulence. The study by Mei *et al.* (1997) identified *asd* using a model of bacteremia. The bacteremia model would be predicted to be the most likely model to identify a virulence defect in a strain deficient in *asd,* as the serum model would presumably, most closely resemble this *in vivo* situation. The study conducted by Coulter *et al.* (1998) also identified the *dap* operon encoded *ykuQ* (THPA), in addition to *asd*. In this study three models were used for the identification of *in vivo* attenuated genes. These were bacteremia, abscess and wound, models of infection. Interestingly,

the mutant strain, unable to produce an active THPA polypeptide was found to be attenuated in all three models. Unfortunately however, the *asd* transposon mutant isolated in this study was unable to grow to measurable, reproducible levels *in vitro* (A_{540}< 0.02) and thus could not be related comparatively *in vivo*. This meant that the report did not indicate which, or how many models the *asd* mutant strain was attenuated in, although presumably from comparison with the other work outlined here, the *asd* mutant was only attenuated in the bacteremia model. Interestingly however, no such problems were found growing strains with a defective *asd* in BHI media, with growth yields consistently reaching densities of >10 (A_{600}). The reason for this difference in, *in vitro* growth is not known, but it does represent a direct conflict in results between the two studies. Finally both STM strategies identified a DAP decarboxylase, as having attenuated virulence. However, which isozyme (*lysA/dapH*) was identified in each study is not known.

The discovery that ASADH, THPA and DAP decarboxylase mutants are attenuated *in vivo* (bacteremia model of infection), further highlights the importance of *S. aureus* LBE genes in pathogenicity. Thus, these published results provide additional weight for the validity of the serum model, as a means for the identification of genes important for the growth of *S. aureus in vivo*.

CHAPTER 6

GENERAL DISCUSSION

6.1 Introduction

The experiments reported in this thesis have been fully discussed at the end of each chapter. It is however now necessary to combine the important points of each discussion into a comprehensive picture of the growth and gene expression of *S. aureus* both in serum and *in vivo*.

The growth of *S. aureus* along with gene expression had not previously been investigated in serum. The growth of *S. aureus* in pig and human sera under different aeration conditions, the role of regulatory loci, and of oxygen protecting enzymes, in addition to the expression of known components, were all examined (Chapter 3). To characterise specific components of the response to growth in serum, *lacZ* reporter gene fusions to genes preferentially expressed in serum compared to a nutrient-rich growth medium were isolated (Chapter 4). Subsequent characterisation of genes preferentially expressed in serum enabled the identification of genes involved in the growth of *S. aureus* in serum and potentially *in vivo*. A subset of these sera expressed genes (*seg*) identified were involved in the biosynthesis of lysine, threonine, methionine and isoleucine (Chapter 4). Genes involved in the biosynthesis of these amino acids were comprehensively studied, investigating; the biosynthetic pathways for the production of these amino acids, the genetic organisation of these genes and their regulation, along with the role of these *seg* in pathogenicity and other functions (Chapter 5).

6.2 *S. aureus* and the serum model

The environmental concentration of O_2/CO_2 appears to be the most important factor affecting both the growth rate and yield of *S. aureus* in serum. Increased CO_2 and/or decreased O_2 concentrations (microaerobic) had a beneficial effect on the growth of *S. aureus* compared to that for aerobic conditions, in both solid and liquid media.

The possible role of oxidative stress in serum was analysed by investigating the role of genes and their enzymatic gene products, with a role in the protection against toxic forms of oxygen. Catalase was found to play an important role in combating this oxidative stress as the addition of catalase to solid media enhanced the growth yield.

Additionally, the inactivation of the major catalase of *S. aureus* (KatA) decreased the growth yield. The importance of the role of catalase was further highlighted as from gene fusion studies, *katA* was shown to be expressed, to significantly lower levels under microaerobic than aerobic conditions. Interestingly, neither the addition of catalase nor the presence of a defective KatA (ST1) made a significant difference to the growth of *S. aureus* on BHI, indicating an increased sensitivity to toxic forms of O_2 in serum. However, despite the apparent importance of O_2 toxicity in serum, a further enzyme involved in the protection against these stresses, superoxide dismutase does not have an important role in the serum model. This was deduced from the fact that a strain (SPW1) encoding a defective SOD (*sodA*) is not attenuated for growth in the serum model, and because *sodA* transcription could not be detected during growth in serum.

Significantly the major global regulators of accessory factors, *agr* and *sar* were not essential for the growth of *S. aureus* in serum. Furthermore the alternative sigma factor σ^B, which has been shown to control the environmental stress response in *S. aureus* (Chan *et al.*, 1998), does not play an important role in the serum model. The results of transcriptional fusion analysis is in agreement with the absence of an important role for these regulatory loci in the serum model, showing that *agr, sarA* and *sigB* are all expressed to a relatively low level in serum.

The screen for genes important for the growth of *S. aureus* in the serum model identified 73 clones that were either specifically or preferentially expressed in serum compared to a nutrient-rich growth medium. This increased gene expression is similar to that observed for Tn*917-lac* mutants of *S. aureus* in response to both, milk containing agar plates (Lammers *et al.*, 2000) and on interaction with human endothelial cells *in vitro* (Vriesema *et al.*, 2000). Interestingly, these 73 strains have transcriptional fusions to genes with a range of expression levels on serum, as determined by the colour of the colonies on serum agar plates containing X-Gal.

25 *seg* were selected at random, for further analysis and characterisation. These mutant strains were shown to have a growth rate that was indistinguishable from each other and from the wild type, when grown in serum. This demonstrated that although these genes were preferentially expressed in serum their disruption did not cause an attenuated growth phenotype. Importantly growth in liquid media confirmed the classification of reporter gene fusions to serum specific genes.

Cloning and sequencing of the regions flanking the Tn*917* insertion, resulted in the identification of a number of transposon disrupted ORFs with similarities to known

proteins in the NCBI database. In addition, several mutants were determined to contain the transposon within an untranslated region. However, on further analysis these regions were found to be 'linked' to genes of known function. These disrupted or linked genes are transcribed by promoters with preferential activity on serum, and thus are the sera expressed genes. These genes included; putative cell surface proteins, an antimicrobial peptide (lantibiotic), and five genes involved in nutrient biosynthesis, including four genes encoding aspartate amino acid biosynthetic enzymes. Additionally, a further gene was identified as a sera expressed gene, although this *seg* had no significant homology to known sequences.

The four genes involved in the biosynthesis of the aspartate amino acids (lysine, threonine, methionine and isoleucine), were comprehensively analysed, and their role in both the serum model and *in vivo* is discussed in the remainder of this chapter.

6.3 The aspartate amino acid biosynthetic sera expressed genes

The screening of *S. aureus* Tn*917* libraries during growth on serum identified four aspartate amino acid biosynthetic genes, as being expressed in serum but not in a nutrient-rich growth media (BHI). These genes are *lysC,* encoding an aspartokinase isozyme (AK II); *asd*, encoding aspartate-semialdehyde dehydrogenase; *dapA,* coding for dihydrodipicolinate synthase; and *yjcI*, encoding cystathionine γ-synthase.

The transposon insertion of *seg*8 was identified to be 13 bp upstream of the translational start site of cystathionine γ-synthase (*yjcI*). This enzyme is involved in the biosynthesis of methionine, catalysing the conversion of acetyl-homoserine to cystathionine. Cystathionine γ-synthase and the mechanism of methionine biosynthesis had not previously been studied in *S. aureus,* and thus were studied as part of this study. Cystathionine γ-synthase was found not only to be preferentially expressed in serum, but was found to be regulated in a growth phase dependent manner with maximum expression occurring during the transition from exponential to post-exponential phase. *seg*8 was found to have a methionine auxotrophic phenotype, demonstrating that cystathionine γ-synthase is essential for methionine biosynthesis, thus providing evidence for the presence of only a single pathway for the biosynthesis of methionine in *S. aureus*. This proposal was further supported by the failure to identify a homologue of acetylhomoserine sulphydrylase, the enzyme of the alternative mechanism for the biosynthesis of methionine (Ozaki and Shiio, 1982). However, the discovery of a putative isozyme of cystathionine γ-synthase (*metX*) in *S. aureus* indicated that the true

key to the auxotrophic nature of a strain with a defective *yjcI* gene was due to the genetic organisation of methionine biosynthetic genes in *S. aureus*.

The organisation of methionine biosynthetic genes in *S. aureus* was found to be different to that found in other bacteria. *S. aureus* possess an operon termed the *yjc* operon which encodes four enzymes, for the biosynthesis of methionine. These genes are the *yjcI* encoded cystathionine γ-synthase, cystathionine β-lyase (*yjcJ*), methionine synthase vitamin-B12 dependent isozyme (*metH*), and methionine synthase vitamin-B12 independent isozyme (*metE*). This organisation results in a polar effect of the *yjcI* mutation on *yjcJ* and potentially on the methionine synthase genes. Therefore a mutation in *yjcI* bringing about premature transcriptional termination of the *yjc* transcript creating a methionine auxotrophic phenotype even in the presence of a second isozyme of cystathionine γ-synthase (*metX*).

The two enzymes of the common pathway for the biosynthesis of the aspartate amino acids, aspartokinase (*lysC*), and aspartate-semialdehyde dehydrogenase (*asd*) along with the first enzyme specific for the biosynthesis of lysine, dihydrodipicolinate synthase (*dapA*) were all identified in the serum screen. These three genes were shown to be clustered in *S. aureus* within an eight gene operon, along with genes involved in lysine biosynthesis and other functions. This organisation of lysine biosynthetic genes was found to be different to that found in all other organisms studied. The operon was named the *dap* operon, and represents the major operon in terms of lysine biosynthesis, encoding six enzymes involved in the production of lysine. These enzymes are: aspartokinase II (*lysC*), aspartate semialdehyde dehydrogenase (*asd*), dihydrodipicolinate synthase (*dapA*), dihydrodipicolinate reductase (*dapB*), tetrahydrodipicolinate acetyltransferase (*ykuQ*), and diaminopimelate carboxylase (*lysA*). The remaining two genes of the *dap* operon are hippurate hydrolase (*hipO*) and alanine racemase (*dal*) and are located in the *dap* operon between *ykuQ* and *lysA*. These two genes are not involved in lysine biosynthesis and are discussed fully in Chapter 5.

The genes of the *dap* operon were postulated to be transcriptionally linked. This was believed to be the case, due to the short intercistronic regions between genes of the *dap* operon and the apparent absence of rho-independent terminator structures within the putative operon. Furthermore evidence that the first three genes of the *dap* operon were transcribed as a single unit, was supported by the observation that the growth phase dependent expression kinetics of all three lysine biosynthetic *seg* was very similar to each other under a specific condition. Maximal expression of these genes was generally witnessed during the transition between exponential and post-exponential phase,

although in some chemically defined media this was delayed until cells entered stationary phase. Additionally, work with the *lacZ* reporter gene fusions in the wild type background showed all three genes to be regulated in an identical manner. This regulation was found to be at the transcriptional level, primarily by lysine, although multivalent effects were also observed for threonine, methionine and isoleucine. The regulation of lysine biosynthetic genes by lysine has been witnessed in other bacteria including *B. subtilis* (Paulus, 1993) and multivalent regulation has been observed for *asd* of *E. coli* (Boy and Patte, 1972). Importantly however, in general not all three of these lysine biosynthetic enzymes are regulated in any one organism. A further regulatory difference observed with *lysC, asd* and *dapA* of *S. aureus* compared to the equivalent genes of other organisms, was the apparent induction of these genes by methionine, in the absence of lysine. These similarities in regulation therefore provide evidence for the transcription of these genes as a single transcriptional unit. The final piece of evidence for the production of a single transcript containing *lysC, asd* and *dapA* and possibly all eight genes of the *dap* operon was the very similar levels of transcript of each lysine biosynthetic enzyme as determined by TaqMan.

Importantly however, results of the auxotrophic analysis revealed that transcription of the *dap* operon was driven from at least two promoters, with the second located upstream of *asd*. This was concluded because mutations in *lysC* were found not to have a polar effect on *asd*. Finally, the regulatory mechanism controlling the synthesis of the lysine biosynthetic genes remains unknown, although there is evidence supporting the possible role of a putative untranslated leader sequence upstream of *lysC*.

Studying the auxotrophic requirements of *lysC* mutants revealed the existence of a further aspartokinase isozyme in *S. aureus*. This additional isozyme was later confirmed by database searching. The isozyme was found to be homologous to threonine-plus-lysine sensitive aspartokinase of other bacteria, for example *B. subtilis*. Notably, from the analysis of the essentially complete *S. aureus* databases, it appears that unlike both *B. subtilis* and *E. coli, S. aureus* encodes two rather than three aspartokinase isozymes.

The importance of the lysine biosynthetic pathway was highlighted by the identification of 9 transposon insertions within; *lysC* (Tn = 6), *asd* (Tn = 2), and *dapA* (Tn = 1). This poses the interesting question as to why should this particular biosynthetic pathway be so important for the growth of *S. aureus* in serum? One reason for the important nature of this pathway is possibly due to the dual role that lysine biosynthesis plays in *S. aureus,* as it not only provides lysine as a building block for protein synthesis, but lysine is also an essential component of the cell wall peptidoglycan of *S. aureus*. This

theory was investigated, with the examination of the cell wall peptidoglycan of the *lysC* mutant strain (*seg*1), by HPLC. Interestingly, the cell wall of this lysine auxotrophic strain was found to be identical to that of the wild type strain (8325-4) under all conditions tested. This demonstrated that in spite of a defective lysine synthetic pathway, the mechanism for the incorporation of lysine into the cell wall peptidoglycan was unaffected by this mutation.

An alternative, and obvious explanation, for the identification of lysine and methionine biosynthetic genes during the course of this study, was perhaps due to a shortage of free lysine and methionine in serum. This theory is borne out by the concentration of free lysine and methionine present in human serum (lysine, 196 - 222 μM; methionine, 21 - 26 μM) (Table A.1.1.6). Interestingly, these values are 6.1x (lysine) and 55x (methionine) lower than that present in CDM (16). Therefore these values appear to be low enough, to allow the expression of genes repressed by lysine or methionine.

Interestingly, *lysC, asd,* and *dapA* in addition to *lysA,* were all identified in a recent screen of *S. aureus* Tn*917* mutants, as being expressed in an *in vitro* model mimicking infection. This study used milk as a model for the selection of genes important in bovine mastitis (Lammers *et al.*, 2000) and thus demonstrates the potential importance of these lysine biosynthetic enzymes for growth in milk. Additionally, a further report recently described the identification of *lysC* and *asd,* as well as *dapB* and *lysA* following Tn*917* mutagenesis, and screening for increased gene expression upon interaction with human endothelial cells (Vriesema *et al.*, 2000). Of particular importance was the observation that the expression of *lysC*, *asd, dapB* and *lysA* was also induced by human serum (Vriesema *et al.*, 2000). These two reports combined with the results described in this study indicate an important role for lysine biosynthetic genes in three different models mimicking infection. However, an interesting observation from these three studies was that an absence of lysine and the other aspartate amino acids did not appear to be the key factor in the high expression of these genes. This was deduced from several pieces of information. Firstly although the concentration of free lysine in serum is relatively low (~ 0.21 mM) compared to that found in CDM (1.3 mM) (A.1.1.6 and A.1.1.2), the addition of lysine and the other aspartate amino acids to serum agar plates, to a final concentration of 2.75 mM did not switch off gene expression (results not shown). Secondly in the work conducted in milk by Lammers *et al.* (2000), to 'prevent' selection of amino acid biosynthetic genes 0.5 % (w/v) casein enzymic hydrolysate (casamino acids) was added to the media. However, despite this addition, four aspartate pathway enzymes were identified. Finally, the identification of four aspartate pathway enzymes upon interaction with human

endothelial cells was reported not to be due to a lysine deficiency, as the media used to culture cells (M199, Life Technologies) contains sufficient lysine (Vriesema *et al.*, 2000). These results therefore show that genes of the lysine biosynthetic pathway are expressed in a variety of *in vitro* models, where the trigger for their expression is not necessarily lysine limitation. Instead amino acid biosynthetic genes maybe specifically induced by a signal, or signals found in the host as well as in these *in vitro* models. A reason for the activation of the lysine biosynthetic pathway in the presence of lysine, is to cause alteration in the cell wall, which maybe beneficial or essential for bacterial survival *in vivo*.

6.4 The role of lysine biosynthetic genes *in vivo*

This study, along with the two described above (Lammers *et al.*, 2000; Vriesema *et al.*, 2000), demonstrated the importance of lysine biosynthetic genes *in vitro*. As these three screens were all designed to mimic *in vivo* conditions it was postulated that *lysC, asd* and *dapA* maybe important *in vivo*. *In vivo* expression analysis, using both RT-PCR and TaqMan showed that these three genes were expressed *in vivo* in the murine pyelonephritis model. However, *lysC, asd* and *dapA* were also shown by TaqMan to be expressed at a significant level *in vitro* as compared to the pyelonephritis model.

The role of aspartokinase II (*lysC*), aspartate semialdehyde dehydrogenase (*asd*), and dihydrodipicolinate synthase (*dapA*) in infection, was investigated using three murine pathogenicity models. Strains defective in these enzymes were not found to be significantly attenuated in survival in these models. Therefore despite the apparent importance of these three lysine biosynthetic genes in, *in vitro* models mimicking infection, along with the fact that they are expressed *in vivo*, these genes do not have a significant function in these pathogenicity models.

Interestingly, the lack of an important role of these three LBE genes was found to contradict reports published by both Mei *et al.* (1997) and Coulter *et al.* (1998). These reports describe the use of STM for the identification of *S. aureus* mutants with attenuated phenotypes *in vivo*. These studies identified *asd* as being important in a bacteremia model of infection, as well as *lysA* (diaminopimelate decarboxylase) and *ykuQ* (tetrahydrodipicolinate acetyltransferase) as being attenuated in both, abscess and bacteremia, models of infection. These results indicate an important role for lysine biosynthetic genes in infection particularly in bacteremia. Therefore from comparison with these results, it can be concluded that enzymes involved in the biosynthesis of

lysine do play an important role *in vivo*, and perhaps the pathogenicity models selected for this study were not suitable for the components tested.

6.5 Concluding remarks

Growth on serum, which potentially mimics the pathogenesis of *S. aureus* bacteremia, was used for the isolation and identification of preferentially expressed *S. aureus* genes (*seg*). This technique identified a number of genes, some of which may contribute to the establishment (surface proteins) and the persistence (biosynthetic genes, and an antimicrobial peptide) in the blood stream. The significance of this *in vitro* technique was confirmed by the discovery that some of the genes identified in this study have also been identified in other pathogenicity mimicking *in vitro* screens, as well as by STM. Therefore this simple *in vitro* approach has the potential to be both a cheap and effective method for the identification of genes important in the pathogenesis of *S. aureus,* in particular with respect to bacteremia. Finally, this *in vitro* model mimicking infection may be useful for the design of novel approaches for anti-staphylococcal therapies for the treatment of *S. aureus* infections.

6.6 Future directions for the serum model and *in vitro* analysis of staphylococcal pathogenesis

This study has described the development and use of an *in vitro* model for the analysis of staphylococcal pathogenesis. This investigation has lead to the identification of both known and unknown genes with a potential role in, *in vivo* growth. The role of these components can now be examined in detail, which will be facilitated greatly by the imminent completion of the *S. aureus* sequencing projects. Finally, conformation of the importance of the aspartate amino acid biosynthetic enzymes in bacteremia remains an important goal.

The serum model once again showed the inherent usefulness of *in vitro* models for the identification of components potentially involved in the establishment of infection and in the causation of disease. These methods combined with the new generation of *in vivo* techniques will increase our knowledge of the pathogenicity of *S. aureus*. This is of increasing importance owing to the increasing prominence of antibiotic resistant strains. Therefore, new anti-staphylococcal therapies based on novel approaches are required, for which a greater understanding of the pathogenicity of *S. aureus* is essential.

REFERENCES

Abdelnour A, Arvidson S, Bremell T, Ryden C, and Tarkowski A. (1993). Infect. Immun. **61**, p3879-3885.

Akerley BJ, Rubin EJ, Camilli A, Lampe DJ, Robertson HM, and Mekalanos JJ. (1998). Proc. Natl. Acad. Sci. USA. **95**, p8927-8932.

Allaoui A, Sansonetti PJ, and Parsot C. (1992). J. Bacteriol. **174**, p7661-7769.

Arbuthnott JP, Coleman DC, and de Azavedo JS. (1990). J. Appl. Bacteriol. p101-107.

Arvidson S. (1983). *In* Staphylococci and staphylococcal infections. p745-808. Easmon CSF, and Adlam C (Eds.). Academic press, London.

Arvidson S, Janson L, and Lofdahl S. (1990). *In* Molecular Biology of the Staphylococci. p421-423. Novick RP (Ed.). VCH Publishers, Inc., New York.

Atrih A, Zollner P, Allmaier G, and Foster SJ. (1996). J. Bacteriol. **178**, p6173-6183.

Bachmann BJ, and Low KB. (1980). Microbiol. Rev. **44**, p1-56.

Balaban N, and Novick RP. (1995). Proc. Natl. Acad. Sci. USA. **92**, p1619-1623.

Barber M, and Rozwadowska-Dowzenko M. (1948). Lancet. **255**, p641-644.

Bartlett ATM, White PJ. 1985. J. Gen. Microbiol. **131**, p2145-2152.

Bartlett ATM, White PJ. 1986. J. Gen. Microbiol. **132**, p3169-3177.

Bayer MG, Heinrichs JH, and Cheung AL. (1996). J. Bacteriol. **178**, p4563-4570.

Behravan J, Chirakkal H, Masson A, and Moir A. (2000). J. Bacteriol. **182**, p1987-1994.

Benov L, and Fridovich I. (1995). Arch. Biochem. Biophys. **332**, p291-294.

Bergdoll MS. (1989). *In* Bacterial food bourne pathogens. p464-523. Doyle MP (Ed.). Marcel Dekker. New York.

Bjorklind A and Arvidson S. (1980). FEMS Microbiol. Letters. **7**, p203-206.

Blevins J, Gillaspy AF, Rechtin TM, Hurlburt BK, and Smeltzer MS. (1999). Mol. Microbiol. **33**, p317-326.

Bliska JB, Galan JE, and Falkow S. (1993). Cell. **73**, p903-920.

Bluhm G. (1985). Acta Medica Scandinavica. **699** (Suppl.). p1-62.

Bondaryk RP, and Paulus H. (1985). J. Biol. Chem. **260**, p585-591.

Booth MC, Cheung AL, Hatter KL, Jett BD, Callegan MC, and Gilmore MS. (1997). Infect. Immun. **65**, p1550-1556.

Bowe F, Lipps CJ, Tsolis RM, Groisman E, Heffron F, and Kusters JG. (1998). Infect. Immun. **66**, p3373-3377.

Boy E, and Patte JC. (1972). J. Bacteriol. **112**, p84-92.

Boylan SA, Redfield AR, Brodie MS and Price CW. (1993). J. Bacteriol. **175**, p7931-7937.

Brunskill EW, and Bayles KW. (1996). J. Bacteriol. **178**, p611-618.

Brunskill EW, and Bayles KW. (1996). J. Bacteriol. **178**, p611-618.

Bsat N, Herbig A, Casillas-Martinez L, Setlow P and Helmann JD. (1998). Mol. Microbiol. **29**, p189-198.

Buchmeier NA, and Heffron F. (1990). Science. **248**, p730-732.

Camilli A, Beattie D, and Mekalanos JJ. (1994). Proc. Natl. Acad. Sci. USA. **91**, p2634-2638.

Camilli A, and Mekalanos J. (1995). Mol. Microbiol. **18**, p671-683.

Camilli A, Mekalanos JJ, and Portnoy DA. (1993). Mol. Microbiol. **8**, p143-157.

Camilli A, Portnoy DA, and Youngman P. (1990). J. Bacteriol. **172**, p3738-3744.

Casadaban MJ. (1976). J. Mol. Biol. **104**, p541-555.

Cassan M, Ronceray J, and Patte JC. (1983). Nucleic Acids Res. **11**, p6157-6166.

Cayouette M, Grismer L, Sucharczuk A, and Moores J. (1999). Strategies Newsletter. **12**, p89-92.

Chan PF, and Foster SJ. (1998a). J. Bacteriol. **180**, p6232-6241.

Chan PF, and Foster SJ. (1998b). Microbiol. **144**, p2469-2479.

Chan PF, Foster SJ, Ingham E, and Clements MO. (1998). J. Bacteriol. **180**, p6082-6089.

Chatterjee M, and White PJ. (1982). J. Gen. Microbiol. **128**, p1073-1081.

Chein Y, and Cheung AL. (1998). J. BiolChem. **273**, p2645-2652.

Chen NY, Hu FM, and Paulus H. (1987). J. Biol. Chem. **262**, p8787-8798.

Chen NY, Jiang S, Klein DA, and Paulus H. (1993). J. Biol. Chem. **268**, p9448-9465.

Chen NY, and Paulus H. (1988). J. Biol. Chem. **263**, p9526-9532.

Cheung AL, Bayer MG, and Heinrichs JH. (1997). J. Bacteriol. **179**, p3963-3971.

Cheung AL, Chein YT, and Bayer AS. (1999). J. Bacteriol. **67**, p1331-1337.

Cheung AL, Eberhardt KJ, Chung E, Yeaman MR, Sullam PM, Ramos M, and Bayer AS. (1994). J. Clin. Invest. **94**, p1815-1822.

Cheung AL, Koomey JM, Butler CA, Projan SJ and Fischetti VA. (1992). Proc. Natl. Acad. Sci. USA. **89**, p6462-6466.

Cheung AL, Nast CC, and Bayer AS. (1998). Infect. Immun. **66**, p5988-5993.

Cheung AL, and Projan SJ. (1994). J. Bacteriol. **176**, p4168-4172.

Cheung AL, Wolz C, Yeaman MR, and Bayer AS. (1995). J. Bacteriol. **177**, p3220-3226.

Cheung AL, and Ying P. (1994). J. Bacteriol. **176**, p580-585.

Chiang SL, Mekalanos JJ, and Holden DW. (1999). Ann. Rev. Microbiol. **53**, p129-154.

Cirillo JD, Weisbrod TR, Pascopella L, Bloom BR, and Jacobs WR. (1994). Mol. Microbiol. **11**, p629-639.

Clements MO. (1996). Ph.D. Thesis. University of Sheffield.

Clements MO, and Foster SJ. (1998). Microbiol.**144**, p1755-1763.

Clements MO, and Foster SJ. (1999). Trends in Microbiol. **7**, p458-462.

Clements MO, and Moir A. (1998). J. Bacteriol. **180**, p6729-6735.

Clements MO, Watson SP, Poole RK, and Foster SJ. (1999a). J. Bacteriol. **181**, p501-507.

Clements MO, Watson SP, and Foster SJ. (1999b). J. Bacteriol. **181**, p3898-3903.

Cohen GN, and Saint-Girons I. (1987). *In Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology vol. 1. p429-444. Niedhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, and Umbarger HE (Eds.). ASM Washington, D.C.

Cohen GN, Stanier RY, and LeBras G. (1969). J. Bacteriol. **99**, p791-801.

Cormack BP, Valdivia RH, and Falkow S. (1996). Gene. **173**, p33-38.

Coulter SN, Schwan WR, Ng EYW, Langhorne MH, Ritchie HD, Westbrock-Wadman S, Hufnagle WO, Folger KR, Bayer AS, and Stover CK. (1998). Mol. Microbiol. **30**, p393-404.

Courcol RJ, Trivier D, Bissingger M-C, Martin GR, and Brown MRW. (1997). Infect. Immun. **65**, p1944-1948.

Cremer J, Eggeling L, and Sahm H. (1990). Mol. Gen. Genet. **224**, p317-324.

Cremer J, Eggeling L, and Sahm H. (1991). Appl. Env Microbiol. **57**, p1746-1752.

Cremer J, Treptow C, Eggeling L, and Sahm H. (1988). J. Gen Microbiol. **134**, p3221-3229.

De Saizieu A, Certa U, Warrington J, Gray C, Keck W, and Mous J. (1998). Nature Biotectnol. **16**, p45-48.

Demerec M. (1948). J. Bacteriol. **56**, p63-74.

Deora R, Tseng T, and Misra TK. (1997). J. Bacteriol. **179**, p6355-6359.

Devriese LA, Hajek V, Oeding P, Meyer SA, and Schleifer KH. (1978). Int. J. Syst. Bacteriol. **28**, p482-490.

DiRita VJ, and Mekalanos JJ. (1989). Ann. Rev. Genet. **23**, p455-482.

Dukan S, and Nystrom T. (1998). Genes Devel. **12**, p3431-3441.

Fairhead H. (1999). Ph.D. Thesis. University of Sheffield.

Ferrari E, Henner DJ, Yang MY. (1985). Biotechnology. **3**, p1003-1007.

Fitzgerald RH. (1989). Infect. Dis. Clincs of North America. **3**, p329-338.

Follettie MT, Peoples OP, Agoropoulou C, and Sinskey. (1993). J. Bacteriol. **175**, p4096-4103.

Foster TJ, and Höök M. (1998). Trends in Microbiol. **6**, p484-488.

Fridovich I. (1995). Ann. Rev. Biochem. **64**, p97-112.

Gillaspy AF, Patti, JM, and Smeltzer MS. (1997). Infect. Immun. **65**, p1536-1540.

Giraudo AT, Calzolari A, Cataldi AA, Bogni C, and Nagel R. (1999). FEMS Microbiol. Letters. **177**, p15-22.

Giraudo AT, Cheung AL, and Nagel R. (1997). Arch. Microbiol. **168**, p53-58.

Giraudo AT, Rampone H, Calzolari A, and Nagel R. (1996). Can. J. Microbiol. **42**, p120-123.

Giraudo AT, Raspanti CG, Calzolari A, and Nagel R. (1994). Can. J. of Microbiol. **40**, p677-681.

Goerke C, Campana S, Bayer MG, Döring G, Botzenhart K, and Wolz C. (2000). Infect. Immun. **68**, p1304-1311.

Graves LM, and Switzer RL. (1990). J. Biol. Chem. **265**, p14947-14955.

Haight TH, and Finland M. (1952). Proc. Exp. Biol. Med. **81**, p183-188.

Hajek V. (1976). Int. J. Syst. Bacteriol. **26**, p401-408.

Haldenwang W, and Losick R. (1980). Proc. Natl. Acad. Sci. USA. **77**, p7000-7004.

Hanahan D. (1983). J. Mol. Biol. **166**, p157-169.

Handfield M, and Levesque RC. (1999). FEMS Microbiol. Rev. **23**, p 69-91.

Hani EK, and Chan VL. (1995). J. Bacteriol. **177**, p2396-2402.

Hart ME, Smeltzer MS, and Iandolo JJ. (1993). J. Bacteriol. 175, p7875-7879.

Hartman BJ, and Tomasz A. (1984). J. Bacteriol. **158**, p513-516.

Haziza C, Stragier P, and Patte JC. (1982). EMBO J. **1**, p379-384.

Hedges RW, and Jacob AE. (1974). Mol. Gen. Genet. **116**, p125-159.

Heinrichs JH, Bayer MG, and Cheung AL. (1996). J. Bacteriol. **178**, p418-423.

Heithoff DM, Conner CP, Hanna PC, Julio SM, Hentschel U, and Mahan MJ (1997). Proc. Natl. Acad. Sci. USA. **94**, p934-939.

Heithoff DM, Sinsheimer RL, Low DA, and Mahan MJ (1999). Science. **284**, p967-970.

Hensel M, and Holden DW. (1996). Microbiol. **142**, p1049-1058.

Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, and Holden DW. (1995). Science. **269**, p400-403.

Hensel M, Shea JE, Raupach B, Monack D, Falkow S, Gleeson C, Kubo T, and Holden DW. (1997). Mol. Microbiol. **24**, p155-167.

Hiramatsu K, Aritaka N, Hanaki H, Kawasaki S, Hosoda Y, Hari S, Fukuchi Y, and Kobayashi Y. (1997). Lancet. **350**, p1670-1673.

Hols P, Defrenne C, Ferain T, Derzelle S, Delplace B, and Delcour J. (1997). J. Bacteriol. **179**, p3804-3807.

Hussain M, Hastings JGM, and White PJ. (1991). J. Med. Microbiol. **34**, p143-147.

Imlay KR, and Imlay JA. (1996). J. Bacteriol. **178**, p2564-2571.

Inaoka T, Matsumura Y, and Tsuchido T. (1998). J. Bacteriol. **180**, p3697-3703.

Ing MB, Baddour LM, and Bayer AS. (1997). *In* The Staphylococci in human disease. p331-353. Crossley KB, and Archer GL (Eds.). New York. Churchill Livingstone.

Janzon L, Lofdahl S, and Arvidson S. (1989). Mol. Gen. Genet. **219**, p480-485.

Jevons, MP. (1961). Br. Med. J. **1**, p124-125.

Ji G, Beavis RC, and Novick RP. (1995). Proc. Natl. Acad. Sci. USA. **92**, p12055-12059.

Ji G, Beavis RC, and Novick RP. (1997). Science. **276**, p2027-2030.

Kalcheva EO, Faiziev MM, Shanskaya VO, and Maluta SS. (1993). Can. J. Microbiol. **40**, p224-227.

Kavanaugh WM, and Williams LT. (1996). *In* Signal Transduction. p3-18. Heildin C-H, and Purton M (Eds.). Chapman and Hall, London.

Kemp EH, Sammons RL, Moir A, Sun D, and Setlow P. (1991). J. Bacteriol. **173**, p4646-4652.

Keyer K, and Imlay JA. (1996). Proc. Natl. Acad. Sci. USA. **93**, p13635-13640.

Kimura K, and Goto T. (1975). J. BioChem. **77**, p415-420.

Klaenhammer TR. (1993). FEMS Microbiol. Rev. **12**, 39-86.

Klarsfeld AD, Goossens PL, and Cossart P. (1994). Mol. Microbiol. **13**, p 585-597.

Kleerebezem M, Quadri LEN, Kulpers O, and de Vos WM. (1997). Mol. Microbiol. **24**, p895-904.

Kloos WE, and Bannerman TL. (1994). Clin. Microbiol. Rev. **7** p117-140.

Kloos WE, and Lambe DW. Jr. (1991). Manual of Clin. Microbiol. 5th Ed. p222-237. Balows *et al.*, (Eds.). ASM, Washington D.C.

Kochhar S, and Paulus H. (1996). Microbiol. **142**, p1635-1639.

Kornblum J, Kreiswirth BN, Projan SJ, Ross H, and Novick RP. (1990). *In* Molecular Biology of the Staphylococci. p421-423. Novick RP (Ed.). VCH Publishers, Inc., New York.

Kreiswirth B, Lofdahl S, Bentley M, O'Reilly M, Schlievert P, Bergdoll M, and Novick RP. (1983). Nature. **305**, p709-712.

Kremery V, Trupl J, Drgona L, Lacka L, Kukuckova E, and Oravcova E. (1996). Eur. J. Clin. Microbiol. Infect. Dis. 15 p259-261.

Krieg NR, and Hoffman PS. (1986). Ann. Rev. Microbiol. **40**, p107-130.

Kullik I, Giachino P, and Fuchs T. (1998a). J. Bacteriol. **180**, p4818-4820.

Kullik I, Jenni R, and BergerBachi B. (1998b). Gene. **219**, p9-17.

Lammer A, Kruijt E, van de Kuijt C, Nuijten PJM, and Smith HE. (2000). Microbiol. **146**, p981-987.

Le Y, He J and Vining LC. (1996). Microbiol. **142**, p791-798.

Lenter C. (1984). Geigy Scientific Tables. Vol 3. Ciba Geigy Ltd., Basel.

Lilley PE, Stamford NP, Vasudevan SG, and Dixon NE. (1993). Gene. **129**, p9-16.

Lindsay JA, and Foster SJ. (1999). Mol. Gen. Genet. **262**, p323-331.

Lindsay JA, Riley TV, and Mee BJ. (1995). Microbiol. **141**, p197-203.

Lowe AM, Beattie DT, and Deresiewicz RL. (1998). Mol. Microbiol. **27**, p967-976.

Lu Y, Chen NY, and Paulus H. (1991). J. Gen. Microbiol. **137**, p1135-1143.

Mahan MJ, Slauch JM, and Mekalanos JJ. (1993). Science. **259**, p686-688.

Mahan MJ, Tobias JW, Slauch JM, Hanna PC, Collier RJ, and Mekalanos JJ. (1995). Proc. Natl. Acad. Sci. USA. **92**, p669-673.

Mallonee DH, Glatz BA, and Pattee PA. (1982). Appl. Env. Microbiol. **43**, p397-402.

Mani N, Tobin P, Jayaswal RK. (1993). J. Bacteriol. **175**, p1493-1499.

Manna AC, Bayer MG, and Cheung AL. (1998). J. Bacteriol. **180**, p3828-3836.

McNamara PJ, and Iandolo JJ. (1998). J. Bacteriol. **180**, p2609-2615.

Mei J, Nourbakhsh F, Ford CW, and Holden DW. (1997). Mol. Microbiol. **26**, p399-407.

Mekalanos JJ. (1992). J. Bacteriol. **174**, p1-7.

Mengaud J, dramsi S, Gouin E, Vazquez-Boland JA Milon G, and Cossart P. (1991). Mol. Microbiol. 5, p2273-2283.

Miller I, Maskell D, Hormaeche C, Johnson K, Pickard D, and Dougan G. (1989). Infect. Immun. **57**, p2758-2763.

Miller JH. (1972). Cold Spring Harbour Laboratory, Cold Spring Harbour. New York. Moir D, and Paulus H. (1977). J. Biol. Chem. **252**, p4648-4654.

Moran CP. Jr. (1993). *In Bacillus subtilis* and other Gram-positive bacteria: Biochemistry, Physiology and Molecular Genetics. p237-267. Sonenshein AL, Hoch JA, Losick R (Eds.). ASM, Washington, D.C.

Morfeldt E, Janzon L, Arvidson S, and Lofdahl S. (1988). Mol. Gen. Genet. **211**, p435-440.

Morfeldt E, Panova-Sapundjievaa I, Gustafsson B and Arvidson S. (1996a). FEMS Microbiol. Letters. **143**, p195-201.

Morfeldt E, Tayler D, von Gabain A, and Arvidson S. (1995). EMBO J. **14**, p4569-4577.

Morfeldt E, Tegmark K, and Arvidson S. (1996b). Mol. Microbiol. **21**, p1227-1237.

Neilands JB. (1972). Struct. Bond. **11**, p145-170.

Nicholas RO, Li T, McDevitt D, Marra A, Sucoloski S, Demarsh PL, and Gentry DR. (1999). Infect. Immun. **67**, p3667-3669.

Novick RP (1990). *In* Molecular Biology of the Staphylococci. p421-423. Novick RP (Ed.). VCH Publishers, Inc., New York.

Novick RP, and Muir TW. (1999). Current opinion in Microbiol. **2**, p40-45.

Novick RP, Projan SJ, Kornblum J, Ross HF, Ji G, Kreiswirth B, Vandenesch F, and Moghazeh S. (1995). Mol. Gen. Genet. **248**, p446-458.

Novick RP, Ross HF, Projan SJ, Kornblum J, Kreiswirth B, and Moghazeh S. (1993). EMBO J. **12**, p3967-3975.

Novick RP. (1963). Virology. **33**, p155-156.

Novick RP. (1991). Methods Enzymol. **204**, p587-636.

Nystrom T, and Gustavsson N. (1998). Biochem. Biophys. Acta. **1365**, p225-231.

Ogawa SK, Yurberg ER, Hatcher VB, Levitt MA, and Lowy FD. (1985). Infect. Immun. **50**, p218-224.

Ohlsen K, Koller KP, and Hacker J. (1997). Infect. Immun. **65**, p3606-3614.

Osbourne AE, Barber CE, and Danils MJ. (1987). EMBO J. **6**, p23-28.

Ozaki H, and Shiio. (1982). J. BioChem. **91**, p1163-1171.

Parkinson JS, and Kofoid EC. (1992). Annu. Rev. Genet. **26**, 71-112.

Parsot C, and Cohen GN. (1988). J. Biol. Chem. **263**, p14654-14660.

Pattee PA, Lee H-C, Bannantine JP. (1990). *In* Molecular Biology of the Staphylococci. p41. Novick RP (Ed.). VCH Publishers, Inc., New York.

Paulus H, (1993). *In Bacillus subtilis* and other Gram-positive bacteria: Biochemistry, Physiology and Molecular Genetics. p237-267. Sonenshein AL, Hoch JA, Losick R. (Eds.). ASM, Washington, D.C.

Pavelka JR. MS, and Jacobs JR. WR. (1996). J. Bacteriol. **178**, p6496-6507.

Peacock SJ, Foster TJ, Cameron BJ, and Berendt AR. (1999). Microbiol. **145**, p3477-3486.

Peng HL, Novick RP, Kreiswirth B, Kornblum J, and Schlievert P. (1988). J. Bacteriol. **170**, p4365-4372.

Perkin-Elmer Corperation (1999). Web site. www.pebio.com.

Perkins JB, and Youngman PJ. (1986). Proc. Natl. Acad. Sci. USA. **83**, p140-144.

Plum G, and Clarck-Curtiss JE. (1994). Infect. Immun. **62**, p474-483.

Projan SJ, and Novick RP. (1997). *In* The Staphylococci in human disease. p55-83. Crossley KB, and Archer GL (Eds.). New York. Churchill Livingstone.

Rechtin TM, Gillaspy AF, Schumacher MA, Brennan RG, Smeltzer MS, and Hurlburt BK. (1999). Mol. Microbiol. **33**, P307-316.

Recsei P, Kreiswirth B, O'Reilly M, Schlievert P, Gruss A and Novick RP, (1986). Mol. Gen. Genet. **202**, p58-61.

Regassa LB, and Bentley MJ. (1993). Infect. Immun. **61**, p1581-1585.

Regassa LB, NovicK PR, and Bentley MJ. (1992). Infect. Immun **60**, p3381-3388.

Renzoni A, Cossart P, and Dramsi S. (1999). Mol. Microbiol. **34**, p552-561.

Richaud F, Phuc NH, Cassan M, and Patte JC. (1980). J. Bacteriol. **143**, p513-515.

Roos KL and Scheld WM. (1997). *In* The Staphylococci in human disease. p413-440. Crossley KB, and Archer GL (Eds.). New York. Churchill Livingstone.

Roten CAH, Brandt C, and Karamata D. (1991). J. Gen. Microbiol. **137**, p951-962.

Rowbury RJ, and Woods DD, (1966). J. Gen. Microbiol. **42**, p155-163.

Sahl H-G, Jack RW, and Bierbaum G. (1995). Eur. J. Biochem. **230**, p827-853.

Saint-Girons I, Duchange N, Cohen GN, and Zakin MM. (1984). J. Biol. Chem. **259**, p14282-14286.

Sambrook J, Fritsch EF, and Maniatis T. (1989). Molecular Cloning (a laboratory manual). Cold Spring Harbour Laboratory, Cold Spring Harbour. New York.

Saravia-Otten P, Muller H-P, and Arvidson S. (1997). J. Bacteriol. **179**, p5259-5263.

Schlegel HG. (1992). Gen. Microbiol. Cambridge University Press.

Schleiffer KH, and Kandler O. (1972), Bacteriol. Rev. **36**, p407-477.

Schnell N, Entian K-D, Schneider U, Gotz F, Zahner H, Kellner R, and Jung G. (1988). Nature. **333**, p276-278.

Schrumpf B, Schwarzer A, Kalinowski J, Puhler A, Eggeling L, and Sahm H. (1991). J. Bacteriol. **173**, p4510-4516.

Scully RE, McNeely BU, and Mark EJ. (1986). New Eng J. Med. **314** p39-49.
Shea JE, Hensel M, Gleeson C, and Holden DW. (1996). Proc. Natl. Acad. Sci. USA. **93**, p2593-2597.

Shiio I, and Miyajima R. (1969). J. BioChem. **65**, p849-859.

Silhavy TJ, and Beckwith JR. (1985). Microbiol. Rev. **49**, p398-418.

Smeltzer MS, Gill SR, and Iandolo JJ. (1992). J. Bacteriol. **174**, p4000-4006.

Smith QR, and Johanson CE. (1983). Am. J. Physiol. **249**, p470-477.

Spector MP, Park YK, Tirgari S, Gonzalez T, and Foster JW. (1988). J. Bacteriol. **170**, p345-351.

Staden R. (1990). Methods in Enzymology. **183**, p163-180.

Stock JB, Ninfa AJ and Stock AM. (1989). Microbiol. Rev. **53**, p450-490.

Stragier P, Richaud F, Borne F, and Patte JC. (1983). J. Mol. Biol. **168**, p307-320.

Svensson B, Lubben M, and Hederstedt L. (1993). Mol. Microbiol. **10**, p193-201.

Taylor RK, Miller VL, Furlong DB, and Mekalanos JJ. (1987). Proc. Natl. Acad. Sci. USA. **84**, p2833-2837.

Theze J, Margarita D, Cohen GN, Borne F, and Patte JC. (1974). J Bacteriol. **117**, p133-143.

Tipper DJ, and Strominger JL. (1968). J. Biol. Chem. **243**, p3169-3179.

Touati D. (1988). J. Bacteriol. **170**, p2511-2520.

Tremaine M, Brockman DK, and Bentley MJ. (1993). Infect. Immun. **61**, p356-359

Trivier D, and Courcol RJ. (1995). FEMS Microbiol. Letters. **141**, p117-127.

Tuomanen E. (1996). FEMS Microbiol. Rev. **18**, p289-299

Valdivia RH, and Falkow S. (1996). Mol. Microbiol. **22**, p367-378.

Valdivia RH, and Falkow S. (1997). Science. **277**, p2007-2011.

Van DijK G, Bottone AE, Strubbe JH, and Steffens AB. (1994). Brain Res. **660**, p96-103

Van de Oost J, Von Wachenfield C, Hederstedt L, and Sarste M. (1991). Mol. Microbiol. **5**, p2063-2072.

Vandenesch F, Kornblum J, and Novick RP. (1991). J. Bacteriol. **173**, p6312-6320.

Villafane R, Bechhofer DH, Narayann CS, and Dubnau D. (1987). J. Bacteriol. **169**, p4822-4829.

Vriesema AJM, Beekhuizen H, Hamdi M, Soufan A, Lammers A, Willekens B, Bakker O, welten AGA, Veltrop MHAM, van de Gevel JS, Dankert J, and Zaat SAJ. (2000). Infect. Immun. **68**, p1765-1772.

Wang J, Mushegian A, Lory S, and Jin S. (1996a). Proc. Natl. Acad. Sci. USA. **93**, p10434-10439.

Wang J, Lory S, Ramphal R, and Jin S. (1996b). Microbiol. **22**, p1005-1012.

Wanner BL. (1993). J. Cell. BioChem. **51**, p47-54.

Wasserman SA, Daub E, Grsafi P, Botstein D, and Walsh CT. (1984). BioChem. **23**, p5182-5187.

Watson SP. (1997). Ph.D. Thesis. University of Sheffield.

Watson SP, Clements MO, and Foster SJ. (1998a). J. Bacteriol. **180**, p1750-1758.

Watson SP, Antonio A, and Foster SJ. (1998b). Microbiol. **144**, p3159-3169.

Weinberger S, and Gilvarg C. (1970). J. Bacteriol. **101**, p323-324.

Wild J, Hennig J, Lobocka M, Walacsak W, and Klopotowski T. (1985). Mol. Gen. Genet. **198**, p315-322

Wilkinson BJ. (1997). *In* The Staphylococci in human disease. p1-38. Crossley KB, and Archer GL (Eds.). New York. Churchill Livingstone.

Wiseman G.M. (1975). Bacteriol. Rev. **39**, p317-344.

Woese CR. (1987). Microbiol. Rev. **51** p221-271.

Wu S, de Lencastre H, Tomasz A. (1996). J. Bacteriol. **178**, p6036-6042.

Xiong A,Singh VK, Cabrera G, and Jayaswal RK. (2000). Microbiol. **146**, p659-668.

Yamakuru F, Ikeda Y, Kimura K, and Sasakawa T. (1974). J. BioChem. **76**, p611-621.

Young GM, and Miller VL (1997). Mol. Microbiol. **25**, p319-328.

Youngman PJ. (1987). *In* Plasmids: a practical approach. Hardy K (Ed.). p79-103. IRL Press, Oxford.

Youngman PJ. (1990). *In* Molecular Biology Methods for Bacillus. Harwood CR, and Cutting SM (Eds p221-226. John Wiley and Sons, New York.

Yost FJ, and Fridovich I. (1973). J. Biol. Chem. **248**, p4905-4908.

Yugari Y, and Gilvarg C. (1965). J. Biol. Chem. **240**, p4710-4716.

Zakin MM, Greene RC, Dautry-Varsat A, Duchange N, Ferrara P, Py MC, Margarita D, and Cohen GN. (1982). Mol. Gen. Genet. **187**, p101-106.

Zhang J-J, Hu F-M, Chen N-Y, and Paulus H. (1990). J. Bacteriol. **172**, p701-708.

APPENDIX

A.1 General

A.1.1 Growth media

Component	Weight ($g\ l^{-1}$)	Component	Weight ($g\ l^{-1}$)
Calf brain infusion solids	12.5	Glucose	2.0
Beef heart infusion solids	5.0	Sodium chloride	5.0
Proteose peptone	10.0	Di-Sodium phosphate	2.5

Table A.1.1
List of ingredients of BHI (Oxoid)

37g of BHI powder is made up to one litre with distilled water and autoclaved. 1 % agar is added prior to autoclaving to produce BHI agar.

Chemical	Weight	Chemical	Weight
Solution 1 - Dissolved in 700 ml of distilled water adjusted to pH 7.2			
$Na_2HPO_4.2H_2O$	7 g	L-Lysine	100 mg
KH_2PO_4	3 g	L-Leucine	150 mg
L-Aspartic Acid	150 mg	L-Methionine	100 mg
L-Alanine	100 mg	L-Phenylalanine	100 mg
L-Arginine	100 mg	L-Proline	150 mg
L-Cystine	50 mg	L-Serine	100 mg
Glycine	100 mg	L-Threonine	150 mg
L-Glutamic Acid	150 mg	L-Tryptophan	100 mg
L-Histidine	100 mg	L-Tyrosine	100 mg
L-Isoleucine	150 mg	L-Valine	150 mg
Solution 2 - Dissolved in 140 ml of distilled water.			
Biotin	0.02 mg	Pyridoxal	0.8 mg
Nicotinic acid	0.4 mg	Pyridoxamine di HCl	0.8 mg
D-Pantothenic acid	0.4 mg	Riboflavin	0.4 mg
Thiamine HCl	0.4 mg		
Solution 3 - Dissolved in 50 ml of 0.1 M HCl			
Adenine Sulphate	20 mg	Guanine HCl	20 mg
Solution 4 - Dissolved in 10 ml of 0.1 M HCl			
$CaCl_2.6H_20$	10 mg	$(NH_2)SO_4.FeSO_4.6H_2O$	6 mg
Solution 5 - Dissolved in 100 ml of distilled water.			
Glucose	10 g	$MgSO_4.7H_2O$	500 mg

Table A.1.2
List of ingredients of CDM as devised by Hussain *et al.* (1991)

Solutions 1, 3 and 4 are mixed and autoclaved together whereas solution 2 is filter sterilised (0.2 μm pore size) and added after cooling. Solution 5 is autoclaved separately and added after cooling.

Constituent	Amount
Chloride	12.-130 mM
Albumin	0.295 g l^{-1}
IgG	0.043 g l^{-1}
Glucose	2.8-4.2 mM
Protein:	
Lumbar	0.15-0.45 g l^{-1}
Cisternal	0.15-0.25 g l^{-1}
Ventricular	0.05-0.15 g l^{-1}

Table A.1.3
Values are for adults and refer to cerebrospinal fluid.

Scully *et al.* (1986).

	Constituent	Value
Blood volumes (Mean adults)	Blood volume	4.05 l
	Plasma volume	2.3 l
	Erythrocyte volume	1.75. l
Physiochemical data	Osmolarity	280-298 mmol kg^{-1}
	pH (arterial)	7.35-7.45
Blood gases	pCO_2 (arterial)	35-45 mmHg
	pO_2 (arterial)	75-100 mmHg

Table A.1.4
Blood volume, physiochemical and blood gas data.

Scully *et al.* (1986).

	Constituent	Amount
Inorganic constituents		
	Water	51.6-53.0 mol l^{-1}
	Sodium	134-147 mM
	Potassium	3.5-5.0 mM
	Chloride	96-106 mM
	Bicarbonate	21-30 mM
	Phosphate	0.7-1.4 mM
	Sulphate	0.25-0.38 mM
	Calcium	2.1-2.7 mM
	Magnesium	0.7-1.3 mM
	Copper	12-30 µM
	Zinc	12-24 µM
	Lead	< 2.4 µM
	Iron	10-30 µM

cont.	Constituents	Amount
Nitrogenous constituents		
	Ammonia (NH_4^+)	12-60 μM
	Urea	2.5-7.0 mM
	Creatine	53-124 μM
	Uric acid	0.18-0.42 mM
	Amino acids (Free total)*	2.56-3.95 mmol l^{-1}
Organic constituents		
	Glucose (fasting)	3.5-5.5 mM
	Lactose	0.6-1.8 mM
	Pyruvate	0.03-0.11 mM
	Citrate	0.08-0.16 mM
	Bilirbin (total)	3-15 μM
Lipids (Fasting)		
	Triglyceride	0.6-3.2 mM
	Cholesterol	3.7-6.8 mM
	Phospholipids	2.6-3.2 mM
	Fatty acids (total)	0.2-0.8 mM
Vitamins		
	Ascorbic acid (C)	20-100 μM
	Biotin	2.5-5.5 μM
	Cholecalciferol (D_3)	0.07 μM
	Cobalamin (B_{12})	0.15-0.67 nM
	Folic acid	15-80 nM
	Niacin (nicotinic acid)	50 μM
	Pantothenic acid	1.3 μM
	Retinol (A)	0.3-2.1 μM
	Riboflavin (B_2)	0.07-0.01 μM
	Thiamine (B_1)	0.2-0.3 μM
	α-Tocopherol (E)	20-45 μM
Proteins		
	Total Protein	60-84 g l^{-1}
	Albumin	35-50 g l^{-1}
	Globulins (total)	23-35 g l^{-1}
	IgG	7-15 g l^{-1}
	Lipoprotein (HDL)	3-8 g l^{-1}
	Ceruloplasmin	0.2-0.4 g l^{-1}
Clotting factors		
	Fibrinogem	2-4 g l^{-1}
	Plaminogen	0.3 g l^{-1}
	Prothrombin	0.1 g l^{-1}

Table A.1.5
Blood constituents
Values are for adults and refer to serum/plasma.

Adapted from Scully *et al.* (1986).

* For more details on the amounts of amino acids in serum see Table A.1.6 and A.1.7

Amino acids	Adults µmol l^{-1}			
	38 Men		28 women	
	Mean	*s*	Mean	*s*
Alanine	404	94	377	108
Arginine	92	22	96	25
Aspartic acid	21	5	20	5
Citrulline	46	10	40	11
Cystine / 2	109	18	107	23
Glutamic acid	226	85	214	71
Glycine	242	44	258	64
Histidine	77	12	70	16
Isoleucine	81	18	56	12
Leucine	157	30	115	23
Lysine	222	50	196	53
Methionine	26	6	21	7
Ornithine	85	24	67	19
Phenylalanine	57	13	48	11
Proline	274	62	218	63
Serine	137	35	142	35
Taurine	67	25	66	17
Threonine	207	58	193	58
Tyrosine	72	12	63	15
Valine	274	53	214	41

Table A.1.6
Free amino acids in serum of adults

Adapted from Lentner (1984).

Amino acids	µmol ml^{-1}	
	Pig	Human
Alanine	1125.1	1348.6
Arginine	560.6	545.7
Aspartic acid	368.6	260.6
Cystine	146.9	101.1
Glutamic acid	762.3	600.0
Glycine	1719.4	1107.4
Histidine	908.0	932.6
Isoleucine	1042.3	518.3
Leucine	2464.6	2270.3
Lysine	87.4	254.9
Methionine	242.3	315.4
Phenylalanine	1825.1	1982.9
Proline	6984.0	6148.0
Serine	1332.6	940.0
Threonine	2657.7	2300
Tyrosine	1565.1	1465.7
Valine	2222.9	2044.0

Table A.1.7
Total amino acids content in pig and human serum (Sigma) as determined by HPLC analysis (Chapter 2.27.2)

A.1.2 PCR primers

Primer	Primer sequence	Source/reference
TnS1	CTC ACA ATA GAG AGA TGT CAC CGT C	Watson, 1997
MDW9	TGT ATG CCA CTG ACC CAA GAC	This study
MDW10	AGA ACG TGC TGA TGA AAA TAA TAC	This study
SKB3	AAG TTA GCA GTT GTG GGT GC	This study
SKB4	CTA TCT TGA ACA GGC ACA CG	This study
MW3	AAA AAC AGT TAT TGA TTC TTG	This study
MW4	GTC CTT TTT CTT CTG CTA A	This study
MW41	GAG TAT ATG TCC AAG CAT G	This study
MW44	CCA GTC GTG CTG TAC AAT GTT CC	This study
OL 1	CGG GAT CCG GGC ATT GAT TCA TGT CCG	This study
OL 2	GGA ATT CGC CAC ACC AGT AAC TGC	This study
OL 3	GGA ATT CCC TTT CAT CCC TGA ACC	This study
OL 4	GGA ATT CGC CAG TTC CTG CTA TGA CAG G	This study
OLP3	CAA GAT CTC GAG AAG CTT CTC CTG CAT TAG GAA GCA GCC CAG	Horsburgh, unpublished
RT-PCR	Sequences not supplied	SmithKline Beecham
TaqMan	Sequences not supplied	SmithKline Beecham

Table A.1.2.1

Details of oligo nucleotide PCR primers used in this study

A.2 Sequence analysis

A.2.1 Sequence, translation and BLAST analysis of the DNA region flanking the Tn*917* insertion site

Figure A.2.1a
***seg*1**

```
                ORF1
  L  K  D  S  D  F  H  Q  I  S  M  I  L  E  T  L
ATTAAAGGATAGTGATTTTCATCAAATTTCTATGATTCTTGAAACATTAA   50

K  N  Q  Y  E  A  L  A  Y  K  I  N  E  H  Y  V  K
AGAATCAATATGAAGCATTAGCTTATAAAATCAATGAGCATTATGTCAAA  100

 I  S  L  I  G  S  G  M  R  D  M  S  G  V  A  S  K
ATTTCATTAATTGGCTCAGGCATGCGTGATATGTCAGGTGTGGCATCAAA  150
                  seg14
  A  F  L  T  L  I  E  N  N  I  P  F  Y  Q  T  T
AGCATTTTTGACATTAATTGAAAATAATATACCTTTCTACCAAACAACAA  200
                                  Tn917
T  S  E  I  S  I  S  Y  V  I
CATCTGAAATAAGTATTTCATACGTCATTGA  231
```

5 best BLAST scores of ORF1

Gene Heading	Accession No.	% Identity
Blast: lysC_helpy	4155744	40% over 75 amino acids
Blast: lysC_helpy	025827	40% over 75 amino acids
Blast: lysC_bacsu	CAA99580.1	34% over 76 amino acids
Blast: lysC_bacsu	P08495	34% over 76 amino acids
Blast: lysC_bassp	Q59229	32% over 82 amino acids

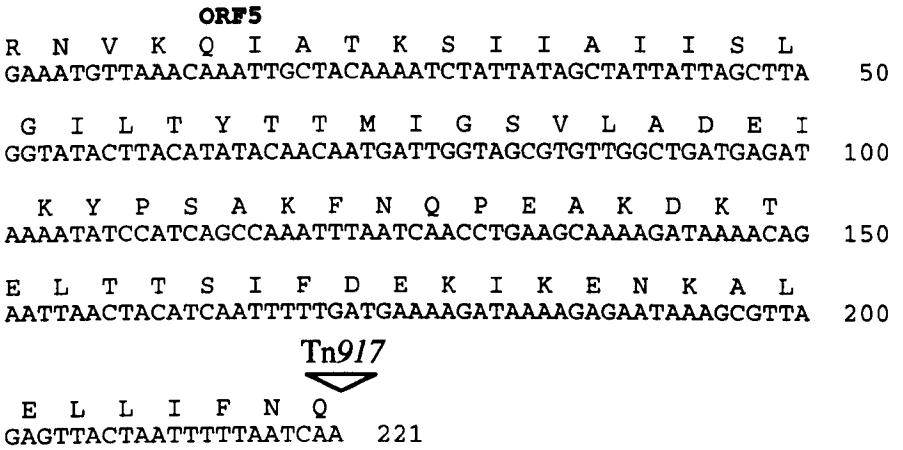
Amino acid alignment

ORF1 - 40% identity over 75 amino acids to LysC from *H. pylori* (lysC_helpy, 4155744).

```
              10                  20         30         40
ORF1    DSDFHQISMILETLKN-------QYEALAYKINEHYVKISLIGSGMRDMS
        D DF  +    +E  K         Q +++ Y  +E+  K+S++G GM+  S
lysC    DLDFTIVKTQIEETKQALKPFLAQMDSIDY--DENIAKVSIVGVGMKSHS
           310        320        330         340        350


           50         60         70
ORF1    GVASKAFLTLIENNIPFYQTTTSEISISYVID
        GVAS AF  L ++NI      +TSEI IS +ID
lysC    GVASIAFKALAKDNININMMISTSEIKISVLID
             360        370        380
```

Figure A.2.1b

seg5

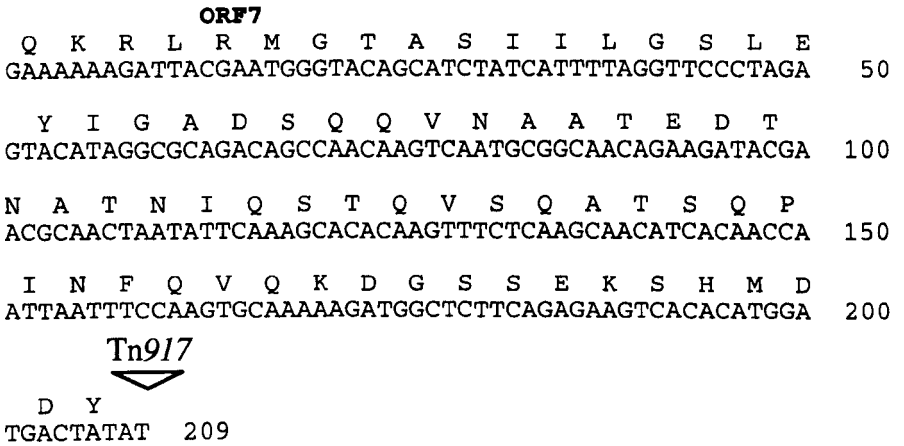


5 best scores of ORF5

No significant homology

Figure A.2.1c

seg7



5 best BLAST scores of ORF7

Gene Heading	Accession No.	% Identity
Blast: hypoth_staau	BAA86643.1	27% over 66 amino acids
Blast: sdrE_staau	CAA06652	22% over 63 amino acids
Blast: pls_staau	AAD09131.1	27% over 66 amino acids
Blast: CD44_bosta	Q294233	6% over 55 amino acids
Blast: mst33a_drome	CAB55437.1	34% over 35 amino acids

Amino acid alignment

ORF7 - 27% identity over 66 amino acids to a hypothetical protein from *S. aureus* (hypoth_staau, BAA86643.1).

```
                10        20        30        40        50
ORF7   KRLRMGTASIILGSLEYIGADSQQVNAATEDTNATNIQSTQVSQATSQPI
       +R  +GTASI++G+    G  + Q  AA  +T      S+  S+
sahy   RRFTVGTASILVGATLIFGVANDQAEAAENNTTQKQDDSSDASKVKG---
                30        40        50        60

                60
ORF7   NFQVQKDGSSEKSHMD
       N Q  +  S+  +  D
sahy   NVQTIEQSSANSNESD
        70        80
```

Figure A.2.1d

seg8

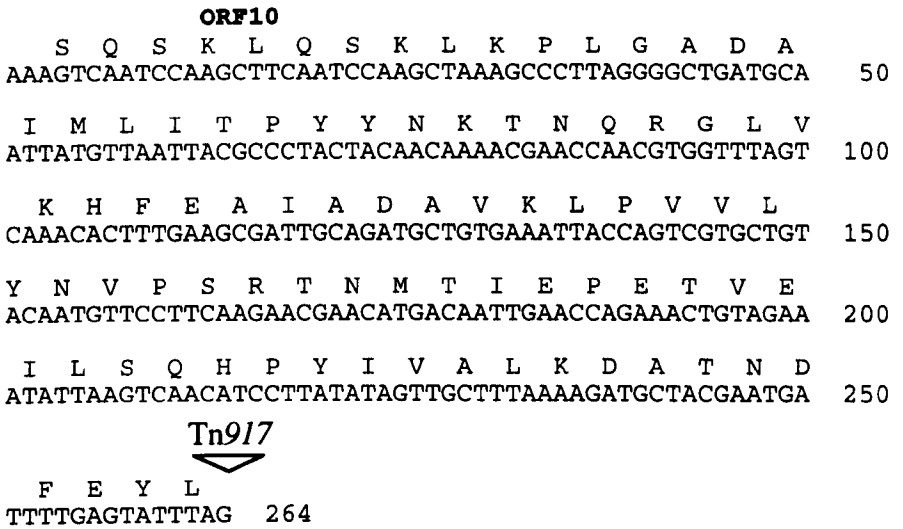
```
GCACATTACTATATCCTTACTTACTGACTTTAATTGTGATAATTGTTCAG    50

TAAGCATATTTACTTTTAATGCGTACTGAATAAGGTTATTTCAGCGATGG   100

AATAACAAATAAAGGTGGTACCGCGAAACATAAGCTTTCGTCCTTTTTAT   150
                                                Tn917
CCGATTCATTCGGGTACGAAGGACGGAAGCTTTTTTTATTTTTTCTTATA   200
```

5 best BLAST scores of ORF8

No significant homology

Figure A.2.1e

*seg*10



5 best BLAST scores of ORF10

Gene Heading	Accession No.	% Identity
Blast: dapA_bacsu	Q04796	52% over 73 amino acids
Blast: dapA_metja	Q57695	55% over 69 amino acids
Blast: dapA_strco	CAA20295	54% over 72 amino acids
Blast: dapA_myctu	CAA15549	50% over 71 amino acids
Blast: dapA_corgl	P19808	47% over 71 amino acids

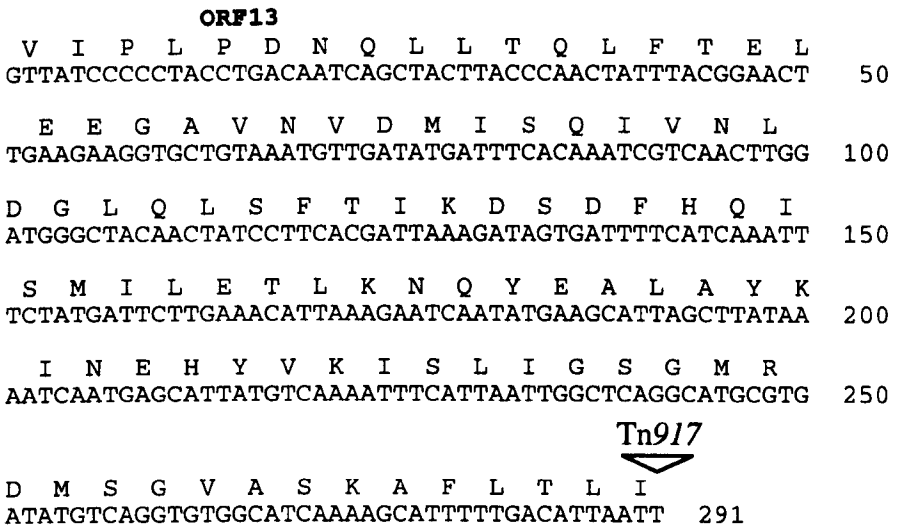
Amino acid alignment

ORF10 - 52% identity over 73 amino acids to DapA from *B. subtilis* (dapA_bacsu, Q04796).

```
              20         30         40         50         60
ORF10 GADAIMLITPYYNKTNQRGLVKHFEAIADAVKLPVVLYNVPSRTNMTIEP
      G DA+ML+TPYYNK +Q G+ +HF+AIA    LPV+LYNVP RT  ++ P
dapA  GVDAVMLVTPYYNKPSQEGMYQHFKAIAAETSLPVMLYNVPGRTVASLAP
          100        110        120        130        140

                70         80
ORF10 E-TVEILSQHPYIVALKDATNDFE
      E T+ + +  P +VA+K+A+ D E
dapA  EITTRLAADIPNVVAIKEASGDLE
          150        160
```

Figure A.2.1f

*seg*13



5 best BLAST scores of ORF13

Gene Heading	Accession No.	% Identity
Blast: lysC_helpy	4155744	38% over 75 amino acids
Blast: AK_helpy	O25827	38% over 75 amino acids
Blast: lysC_bacsu	142524	34% over 75 amino acids
Blast: lysC_bacsu	CAA99580.1	34% over 75 amino acids
Blast: AK2_bassp	P08495	32% over 75 amino acids

Amino acid alignment

ORF13 - 38% identity over 75 amino acids to LysC from *H. pylori* (lysC_helpy, 4155744).

```
              30        40        50        60        70
ORF13 VNVDMISQIVNLDG-LQLSFTIKDSDFHQISMILETLKNQYEALAYKINE
      +NVDMI Q +  DG   L FTI  +   +    L+    Q +++ Y  +E
lysC  INVDMIVQTIGRDGKTDLDFTIVKTQIEETKQALKPFLAQMDSIDY--DE
               300       310       320       330

              80        90
ORF13 HYVKISLIGSGMRDMSGVASKAFLTL
      +  K+S++G GM+  SGVAS AF  L
lysC  NIAKVSIVGVGMKSHSGVASIAFKAL
      340       350       360
```

Figure A.2.1g

*seg*14

```
TAGCAGTTATAATTTGTTGGACTTTTTGGTTAAGAGCTGAGAGTTTGTCA    50

TTATTTAAAAATAATGGAGTGCATCACTTGTATATAGATTAAGAGCAAGT   100

TCGCATTCCGAACTTGTTCTTTTTTTATTATTGTGTGCCCTTCCTNACNA   150
                          ORF14
                 M  V  T  R  S  V  L  K  F  G  G  S
TTAGGAGGATTATATGGTAACAAGAAGTGTGTTGAAATTTGGCGGATCAT   200

S  V  S  D  F  T  K  I  K  R  I  A  E  M  L  K  E
CCGTCAGTGATTTTACAAAAATAAAAAGGATCGCTGAAATGTTAAAGGAG   250
               Tn917
 R  V  N  Q  D
CGAGTCAATCAAGAT  265
```

5 best BLAST scores of ORF14

Gene Heading	Accession No.	% Identity
Blast: AK_mycsm	P41403	54% over 24 amino acids
Blast: lysC_thema	AAD36585.1	42% over 28 amino acids
Blast: AK_myctu	P97048	54% over 22 amino acids
Blast: AK_metja	Q57991	36% over 30 amino acids
Blast: AK1H_ecoli	P00561	12% over 28 amino acids

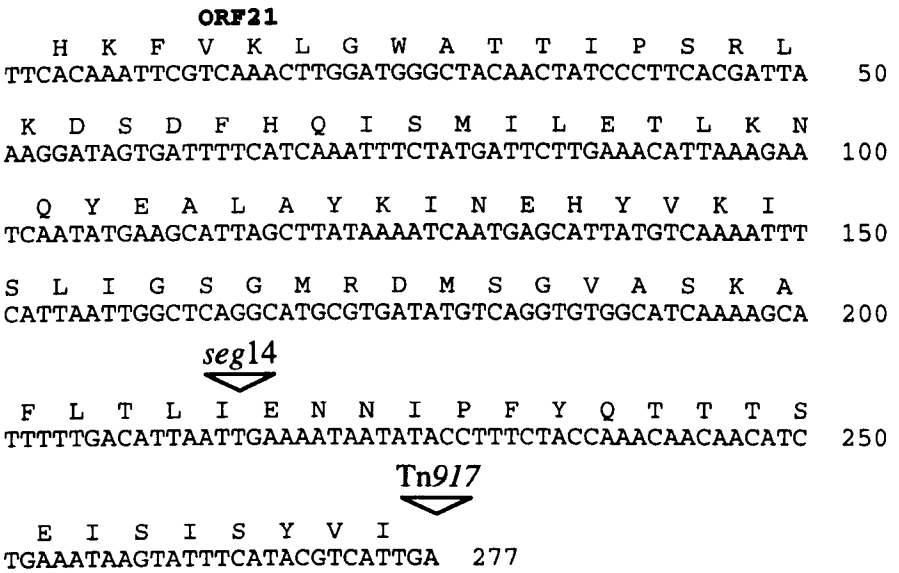
Amino acid alignment

ORF14 - 54% identity over 24 amino acids to LysC from *M. smegmatis* (AK_mycsm, P41403).

```
            10        20
ORF14  VLKFGGSSVSDFTKIKRIAEMLKE
       V+KFGG+++SD  K++++AE + +
lysC   VQKYGGSSVADAERIRRVAERIVE
            10        20
```

Figure A.2.1h

*seg*21



5 best BLAST scores of ORF21

Gene Heading	Accession No.	% Identity
Blast: lysC_bacsu	CAA99580.1	34% over 82 amino acids
Blast: lysC_ bacsu	142524	34% over 82 amino acids
Blast: AK2_ bacsu	P08495	34% over 82 amino acids
Blast: AK2_ bacsp	Q59229	32% over 82 amino acids
Blast: lysC_synsp	BAA18676	36% over 74 amino acids

Amino acid alignment

ORF21 - X% identity over Y amino acids to LysC from *B. subtilis* (lysC_bacsu, CAA99580.1).

```
      10         20         30         40         50
ORF21 TTIPSRLKDSDFHQISMILETLKNQYEALAYKINEHYVKISLIGSGMRDM
      T I    +K  D  Q    +LE  K+  E     +       K+S++GSGM
lysC  TGISFSVKTEDADQTVAVLEEYKDALEFEKIETESKLAKVSIVGSGMVSN
               310        320        330        340        350

      60         70         80         90
ORF21 SGVASKAFLTLIENNIPFYQTTTSEISISYVI
       GVA++ F  L + NI      +TSEI +S V+
lysC  PGVAAEMFAVLAQKNILIKMVSTSEIKVSTVV
               360        370        380
```

Figure A.2.1i

***seg*22**

```
GGCGCATCAATCATGAGTAAAGTTTAGATTACTGTCTGCTAACAGCTAAA   50

TTTGAAAGGGTGCGATGCCGAAGCAATTATAATAGCAGTTATAATTTGTT  100

GGACTTTNTNGAAAAGAGCTGAGAGTTTGTCATTATTTAAAAATAATGGA  150

GTGCATCACTTGTATATAGATTAAGAGCAAGTTCGCATTCCGAACTTGTT  200
```

Tn*917*

```
CTTTTTTTATTATT  214
```

5 best BLAST scores of ORF22

No significant similarity found

Figure A.2.1j

seg24

```
AACGTCATTGATGATTTTAATGGGCAACAAGCGGTAGAAAAACTATATGA    50

ACGCATTTAACATTTAATGGTAAAATGATTGTTAAAATATTCTAAAAATT   100
                                                ORF24
                                        M  T  K  L  A  V  V
GGAAATTATTATAAAATGGAGTGACAAGTTATGACAAAGTTAGCAGTTGT   150

  G  A  T  G  L  V  G  T  K  M  L  E  T  L  N  R
GGGTGCAACAGGATTAGTAGGTACAAAAATGTTGGAGACATTAAATCGTA   200

K  N  I  P  F  D  E  L  V  L  F  S  S  A  R  S  A
AAAATATTCCTTTCGATGAATTAGTATTATTTTCATCAGCACGTTCTGCA   250

 G  Q  E  V  E  F  Q  G  K  T  Y  T  V  Q  E  L  T
GGGCAAGAAGTTGAATTTCAAGGAAAAACATATACAGTTCAAGAATTAAC   300
                Tn917
  D  A  R  A  S
TGATGCTCGTGCAAGTG  317
```

5 best BLAST scores of ORF24

Gene Heading	Accession No.	% Identity
Blast: dhas_camje	Q59291	46% over 56 amino acids
Blast: dhas_aquae	O67716	51% over 54 amino acids
Blast: asd_thema	AAD36590.1	50% over 56 amino acids
Blast: dhas_bascu	Q04797	57% over 52 amino acids
Blast: asd_synsp	S49978	27% over 59 amino acids

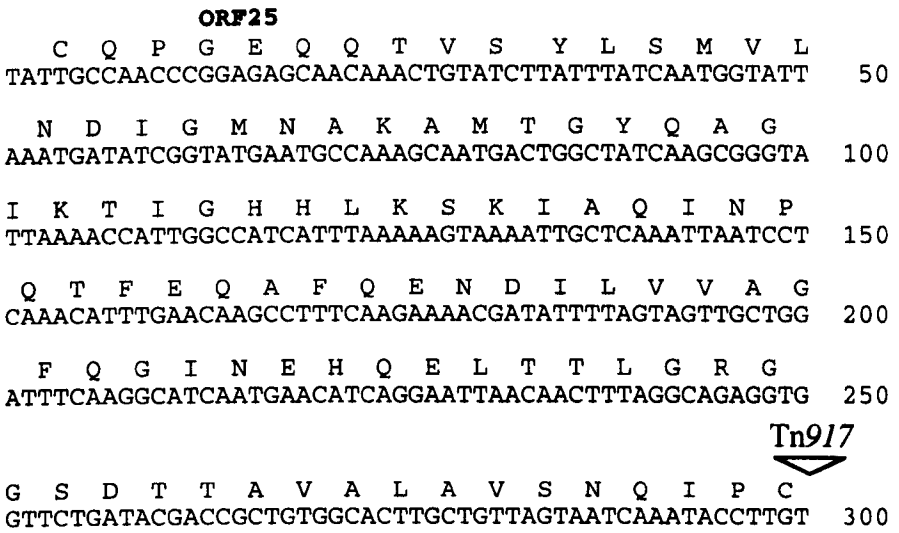
Amino acid alignment

ORF24 - 46% identity over 56 amino acids to Asd from *C. jejuni* (dhas_camje, Q59291).

```
             10        20        30        40        50
ORF24  KLAVVGATGLVGTKMLETLNRKNIPFDELVLFSSARSAGQEVEFQGKTYT
       K+A+VGATG VG ++L  L+  + P + ++  +SA+SAG ++EF+GK+YT
asd    KIAIVGATGAVGEELLNVLDELDFPVESILPLASAKSAGNQIEFRGKSYT
           10        20        30        40        50

            58
ORF24  VQELTD
       ++ELT+
asd    IKELTE
           60
```

Figure A.2.1k

***seg*25**



5 best BLAST scores of ORF25

Gene Heading	Accession No.	% Identity
Blast: lysC_thema	AAD36585.1	50% over 96 amino acids
Blast: AK2_ bacsp	Q59229	53% over 91 amino acids
Blast: AK2_ bacsu	P08495	48% over 94 amino acids
Blast: lysC_ bacsp	CAA99580.1	50% over 91 amino acids
Blast: lysC_synsp	BAA18676	50% over 91 amino acids

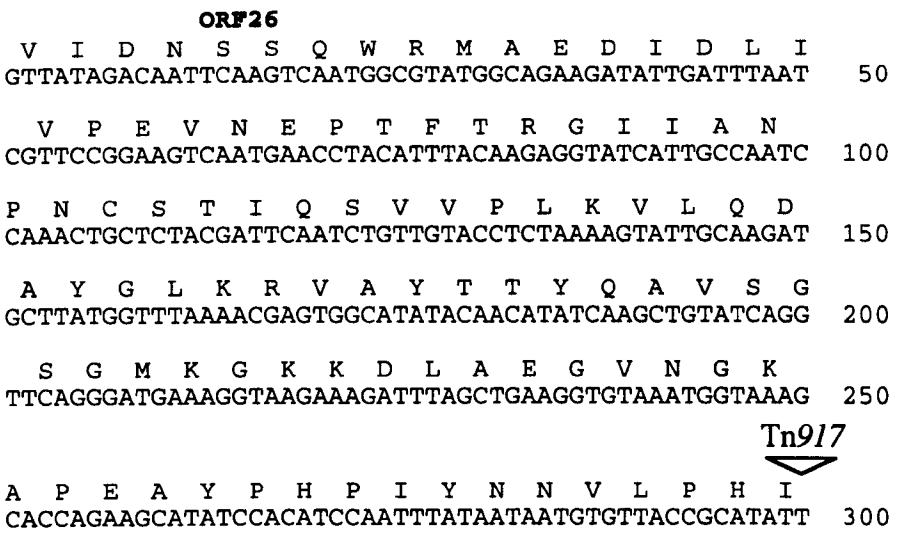
Amino acid alignment

ORF25 - 50% identity over 96 amino acids to LysC from *T. maritima* (lysC_thema, AAD36585.1).

```
           10         20         30         40         50
ORF25 GEQQTVSYLSMVLNDIGMNAKAMTGYQAGIKTIGHHLKSKIAQINPQTFE
      GEQ + + LSM L D+G+ AK++  +Q  IKT  HH  ++I  I+
lysC  GEQVSAALLSMALKDLGVKAKSLNAFQVKIKTTPHHTSARIVDIDDSVIR
             80         90         100        110        120


           60         70         80         90         100
ORF25 QAFQENDILVVAGFQGINEHQELTTLGRGGSDTTAVALAVSNQIPC
      +  ++ D+LVV GFQG+NEH +LTTLGRGGSDT+AVALA   ++PC
lysC  ENLKDYDVLVVTGFQGVNEHGDLTTLGRGGSDTSAVALAAKLRVPC
             130        140        150        160
```

Figure A.2.1l

***seg*26**



5 best BLAST scores of ORF26

Gene Heading	Accession No.	% Identity
Blast: dhas_aquae	O67716	58% over 100 amino acids
Blast: asd_thema	AAD36590.1	48% over 100 amino acids
Blast: dhas_camje	Q59291	50% over 100 amino acids
Blast: asd_shevi	BAA08490.1	52% over 100 amino acids
Blast: dhas_legpn	O31219	52% over 100 amino acids

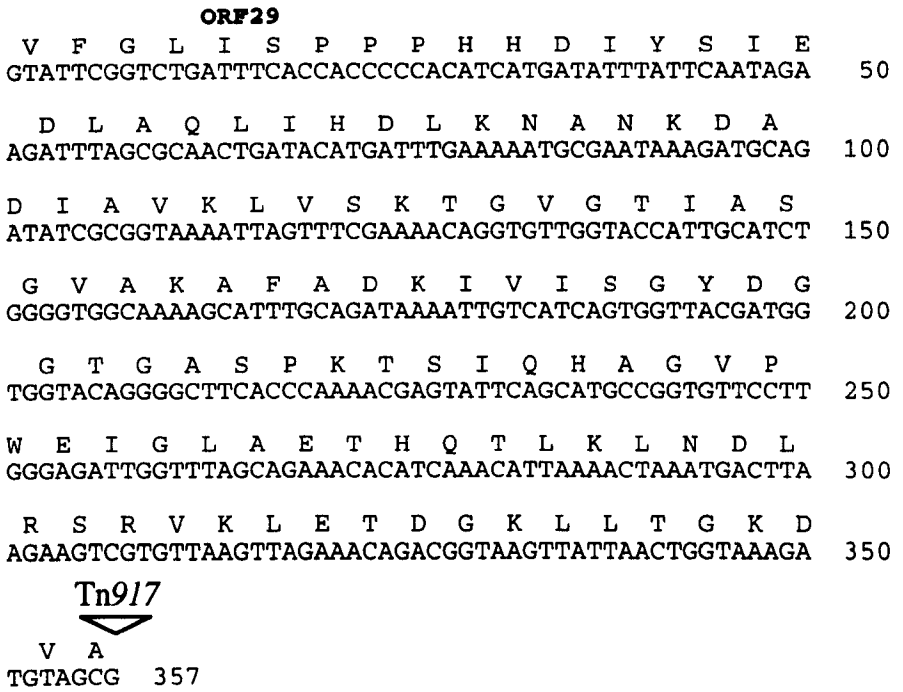
Amino acid alignment

ORF26 - 58% identity over 100 amino acids to Asd from *A. aeolicus* (dhas_aquae, O67716).

```
                10        20            30        40
ORF26 VIDNSSQWRMAEDIDLIVPEVNEPTFT----RGIIANPNCSTIQSVVPLK
      VIDNSS WRM  D+ L+VPEVN          +GIIANPNCSTIQ VV LK
asd   VIDNSSAWRMDPDVPLVVPEVNPEDVKDFKKKGIIANPNCSTIQMVVALK
              100       110       120       130       140

          50        60        70                80
ORF26 VLQDAYGLKRVAYTTYQAVSGSGMKGKKDL-------AEGVNGKAPEAYP
       + D  G+KR V +TYQAVSG+G K  +DL         EG       + +P
asd   PIYDKAGIKRVVVSTYQAVSGAGAKAIEDLKNQTKAWCEGKEMPKAQKFP
              150       160       170       180       190

      90        100
ORF26 HPIYNNVLPHI
      H I  N LPHI
asd   HQIAFNALPHI
              200
```

Figure A.2.1m

seg29



5 best BLAST scores of ORF29

Gene Heading	Accession No.	% Identity
Blast: gltb_bacsu	P39812	84% over 116 amino acids
Blast: gltb_syny3	P55037	77% over 117 amino acids
Blast: gltb_plebo	BAA12741	77% over 116 amino acids
Blast: gltb_cloac	AAD41675.1	75% over 117 amino acids
Blast: gltb_deira	AAF09770.1	76% over 116 amino acids

Amino acid alignment

ORF29 - 84% identity over 116 amino acids to GtaB from *B. subtilis* (gltb_bacsu, P39812).

```
              10        20        30        40        50
ORF29 GLISPPPHHDIYSIEDLAQLIHDLKNANKDADIAVKLVSKTGVGTIASGV
      GLISPPPHHDIYSIEDLAQLIHDLKNAN+DA I+VKLVSK GVGTIA+GV
gltb  GLISPPPHHDIYSIEDLAQLIHDLKNANRDARISVKLVSKAGVGTIAAGV
           980       990       1000      1010      1020

              60        70        80        90        100
ORF29 AKAFADKIVISGYDGGTGASPKTSIQHAGVPWEIGLAETHQTLKLNDLRS
      AKA AD IVISGYDGGTGASPKTSI+H G+PWE+GLAE HQTL LN LR
gltb  AKATADVIVISGYDGGTGASPKTSIKHTGLPWELGLAEAHQTLMLNGLRD
           1030      1040      1050      1060      1070

              110
ORF29 RVKLETDGKLLTGKDV
      RV LETDGKL+TG+DV
gltb  RVVLETDGKLMTGRDV
           1080      1090
```

Figure A.2.1n

***seg*30**

```
GATGAGGCGCATCAATCATGAGTAAAGTTTAGATTACTGTCTGCTAACAG    50

CTAAATTTGAAAGGGTGCGATGCCGAAGCAATTATAATAGCAGTTATAAT   100

TTGTTGGACTTTTTGGTTAAGAGCTGAGAGTTTGTCATTATTTAAAAATA   150

ATGGAGTGCATCACTTGTATATAGATTAAGAGCAAGTTCGCATTCCGAAC   200
```

Tn*917*

```
TTGTTCTTTTTTTATTATT  219
```

5 best BLAST scores of ORF30

No significant similarity found

Figure A.2.1o

*seg*33

```
CAATCATGAGTAAAGTTTAGATTACTGTCTGCTAACAGCTAAATTTGAAA   50

GGGTGCGATGCCGAAGCAATTATAATAGCAGTTATAATTTGTTGGACTTT  100

TTGGTTAAGAGCTGAGAGTTTGTCATTATTTAAAAATAATGGAGTGCATC  150

ACTTGTATATAGATTAAGAGCAAGTTCGCATTCCGAACTTGTTCTTTTTT  200
```

Tn*917*

```
TATTATT  207
```

5 best BLAST scores of ORF33

No significant similarity found

Figure A.2.1p

***seg*35**

```
TCTAACTTATGGAAGTTGACTGGCAAAACCCATACTGTTAAATTAATTCG    50

GGTACGGGATTCTAAAGGAAACTAATCCTGGAGTATAACTGTTAGAAATA   100

GAGGAAGTTAAACCCGAAATTTAAGAGTTAAACAAAAAATATTCACAACA   150

TATTAAAAAAAGAAGTTACATTATTTGTTTAAACTAACAATAGTACCATA   200

AAATTAATGACGATTTACAAATTAATAACAAATAATTATTATTAAATTCA   250

TTGATTTATTGTGTGGTAATGATTAATATGACAATTGAAGTGGGTCTTAA   300

CTTCAACTTTTAAATAGGGGGTTTTATCTTTGGAAAAAGTTCTTGATTTA   350
```

Tn*917*

```
GACGTGCAAGTT  362
```

5 best BLAST scores of ORF35

No significant similarity found

Figure A.2.1q

seg37

```
GATGCGTTAGGAAGTTCTTCTAATACAATTTATACAACGCCATTCCCTAC   50

ACACTCTTATAAAAGAGATTCACGCGCGTCAATAAATTGTATTACATACT  100

AACTAAAAAAGCTTTTCTTAATCGTACTAACGAAGTTAGAGGTTCTTATGT  150

GATTGGACAACTGTTCTCCCTACTAGACTTTTAGCATAGCTGGTTAAAAG  200

TACTATTACAGATTGTTCAATTGTATGCCCCTTTCCTTTGAACGGAATAT  250
                                                Tn917

TAAAAGCTATGTGCAGTAACTTTTAATAGGCATACCACTTTTGCATCAA  249
```

5 best BLAST scores of ORF37

No significant similarity found

Figure A.2.1r

***seg*39**

```
AATACAATTAATACAACGCCATTCCCTACACACTCTTATAGTAGAGATTC   50

ACGCGCGTCAATAAATTGTATTACATACTAACTGAAAAAGCTTTTGTTAA  100

TCGTACTAACGAAGTTAGAGGTTCTTATGTGATTGGACAACTGTTCTCCC  150

TACTAGACTTTTAGCATAGCTGGTTAAAAGTACTATTACAGATTGTTCAA  200

TTGTATGCCCCTTTCCTTTGAACGGAATATTAAAAGCTATGTGCAGTAAC  250
```


```
TTTTAATAGGCATACCACTTTTGCATCAA  279
```

5 best BLAST scores of ORF39

No significant similarity found

Figure A.2.1s

*seg*40

```
AAAGAGATTCCACGCGCGTCAATAAATTGTATTACATACTAACTAAGAAG   50

CTTTTGTTAATCGTACTAACGAAGTTAGAGGTTCTTATGTGATTGGACAA  100

CTGTTCTCCCTACTAGACTTTTAGCATAGCTGGTTAAAAGTACTATTACA  150

GATTGTTCAATTGTATGCCCCTTTCCTTTGAACGGAATATTAAAAGCTAT  200
```

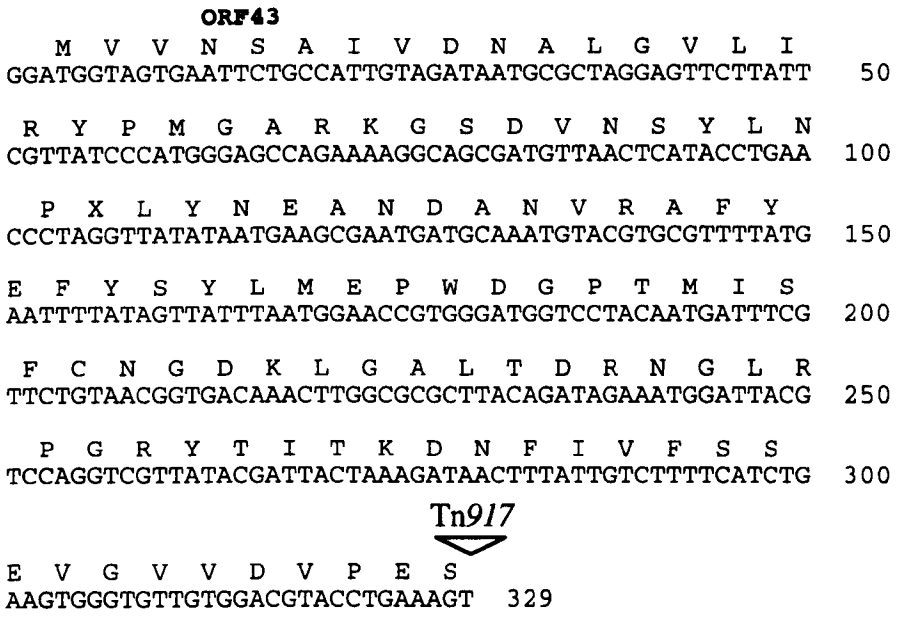

```
GTGCAGTAACTTTTAATAGGCATACCACTTTTGCATCAA  239
```

5 best BLAST scores of ORF40

No significant similarity found

Figure A.2.1t

*seg*43



5 best BLAST scores of ORF43

Gene Heading	Accession No.	% Identity
Blast: gltb_bacsu	P39812	50% over 105 amino acids
Blast: gltb_medsa	1066499	62% over 67 amino acids
Blast: glsn_medsa	Q03460	62% over 67 amino acids
Blast: gltb_strco	CAB52861.1	58% over 65 amino acids
Blast: gltb_orysa	BAA35120.1	50% over 104 amino acids

Amino acid alignment

ORF43 - 50% identity over 105 amino acids to GtlB from *B. subtilis* (gltb_bacsu, P39812).

```
            10        20        30        40          50
ORF43 NSAIVDNALGVLIRYPMGARKGSDVNSYLNPXLYNEANDANV--RAFYEF
      +S+I+DNA    +   M  RK +     L P  + E    +   RAFYE+
gltB  DSSILDNAFEFFV---MAGRKPAHTAMMLIPEPWTENTHMSKEKRAFYEY
             300          310       320       330

             60        70        80        90        100
ORF43 YSYLMEPWDGPTMISFCNGDKLGALTDRNGLRPGRYTITKDNFIVFSSEV
      +S LMEPWDGPT ISF +G ++GA+ DRNGLRP RY +TKD++I+FSSEV
gltB  HSSLMEPWDGPTAISFTDGKQIGAILDRNGLRPARYYVTKDDYIIFSSEV
       340       350       360       370       380

            108
ORF43 GVVDVPE
      GV++V +
gltB  GVIEVEQ
       390
```

Figure A.2.1u

seg45

```
                ORF45
   G  I  K  M  I  S  V  C  M  P  K  A  M  T  G  Y
ATGGTATTAAAATGATATCGGTATGCATGCCAAAAGCAATGACTGGCTAT    50

 Q  A  G  I  K  T  I  G  H  H  L  K  S  K  I  A  Q
CAAGCGGGTATTAAAACCATTGGCCATCATTTAAAAAGTAAAATTGCTCA  100

  I  N  P  Q  T  F  E  Q  A  F  Q  E  N  D  I  L
AATTAATCCTCAAACATTTGAACAAGCCTTTCAAGAAAACGATATTTTAG  150

V  V  A  G  F  Q  G  I  N  E  H  Q  E  L  T  T  L
TAGTTGCTGGATTTCAAGGCATCAATGAACATCAGGAATTAACAACTTTA  200

 G  R  G  G  S  D  T  T  A  V  A  L  A  V  S  N  Q
GGCAGAGGTGGTTCTGATACGACCGCTGTGGCACTTGCTGTTAGTAATCA  250
          Tn917
  I  P  C
AATACCTTGTG  261
```

5 best BLAST scores of ORF45

Gene Heading	Accession No.	% Identity
Blast: lysC_thema	AAD36585.1	50% over 76 amino acids
Blast: lysC_synsp	BAA18676	54% over 70 amino acids
Blast: AK2_bacsp	Q59229	55% over 68 amino acids
Blast: AK2_bacst	P53553	51% over 70 amino acids
Blast: AK2_bacsu	P08495	49% over 71 amino acids

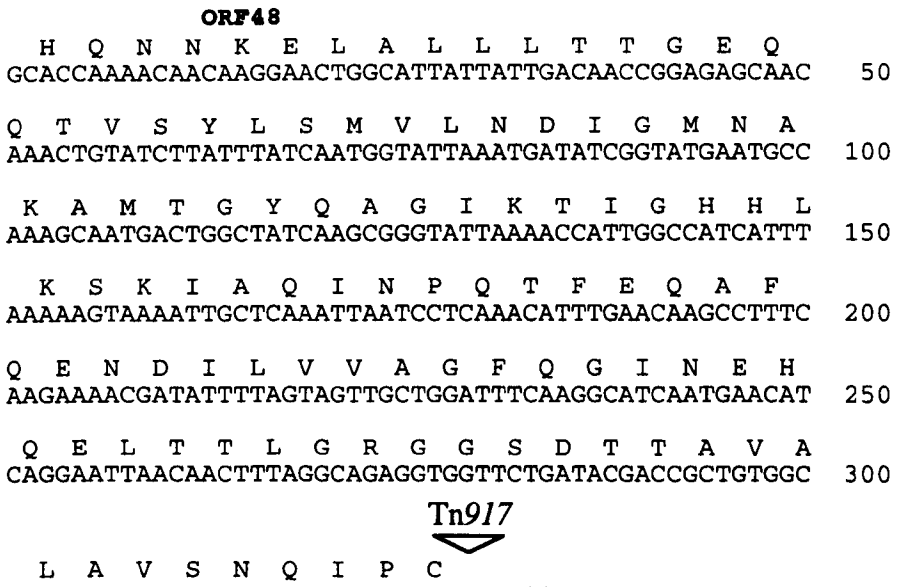
Amino acid alignment

ORF45 - 50% identity over 76 amino acids to LysC from *T. maritima* (lysC_thema, AAD36585.1).

```
                20        30        40        50        60
ORF45 KAMTGYQAGIKTIGHHLKSKIAQINPQTFEQAFQENDILVVAGFQGINEH
      K++  +Q  IKT  HH  ++I  I+     +  ++ D+LVV GFQG+NEH
lysC  KSLNAFQVKIKTTPHHTSARIVDIDDSVIRENLKDYDVLVVTGFQGVNEH
              100       110       120       130       140

                70        80
ORF45 QELTTLGRGGSDTTAVALAVSNQIPC
       +LTTLGRGGSDT+AVALA   ++PC
lysC  GDLTTLGRGGSDTSAVALAAKLRVPC
              150       160
```

Figure A.2.1v

*seg*48



5 best BLAST scores of ORF48

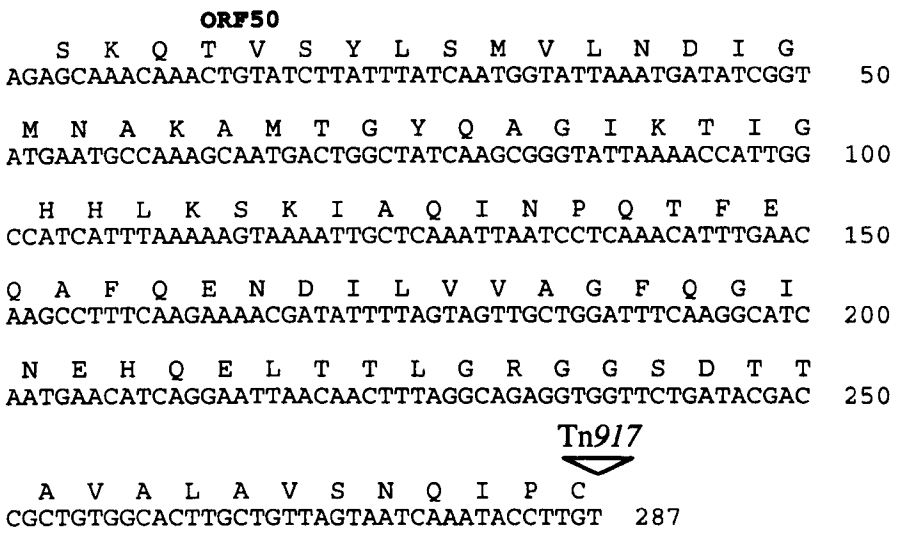
Gene Heading	Accession No.	% Identity
Blast: lysC_thema	AAD36585.1	49% over 107 amino acids
Blast: AK2_bacsp	Q59229	52% over 104 amino acids
Blast: AK2_bacsu	P08495	47% over 105 amino acids
Blast: lysc_bacsu	CAA99580.1	49% over 102 amino acids
Blast: AK2_bacst	P53553	47% over 104 amino acids

Amino acid alignment

ORF48 - 49% identity over 107 amino acids to LysC from *T. maritima* (lysC_thema, AAD36585.1).

```
              10        20        30        40        50
ORF48 NNKELALLLTTGEQQTVSYLSMVLNDIGMNAKAMTGYQAGIKTIGHHLKS
      +++EL +LL TGEQ + + LSM L D+G+ AK++  +Q  IKT  HH  +
lysC  DSRELDMLLATGEQVSAALLSMALKDLGVKAKSLNAFQVKIKTTPHHTSA
               70        80        90        100       110

              60        70        80        90        100
ORF48 KIAQINPQTFEQAFQENDILVVAGFQGINEHQELTTLGRGGSDTTAVALA
      +I  I+      +  ++ D+LVV GFQG+NEH +LTTLGRGGSDT+AVALA
lysC  RIVDIDDSVIRENLKDYDVLVVTGFQGVNEHGDLTTLGRGGSDTSAVALA
              120       130       140       150       160

             109
ORF48 VSNQIPC
        ++PC
lysC  AKLRVPC
             168
```

Figure A.2.1w
*seg*50



5 best BLAST scores of ORF50

Gene Heading	Accession No.	% Identity
Blast: lysC_thema	AAD36585.1	48% over 92 amino acids
Blast: AK2_bacsp	Q59229	52% over 87 amino acids
Blast: AK2_bacsu	P08495	47% over 90 amino acids
Blast: lysc_myctu	1899205	44% over 83 amino acids
Blast: AK_myctu	P97048	44% over 83 amino acids

Amino acid alignment

ORF50 - 48% identity over 92 amino acids to LysC from *T. maritima* (lysC_thema, AAD36585.1).

```
            10        20        30        40        50
ORF50 TVSYLSMVLNDIGMNAKAMTGYQAGIKTIGHHLKSKIAQINPQTFEQAFQ
      + + LSM L D+G+ AK++  +Q  IKT  HH  ++I  I+     +  +
lysC  SAALLSMALKDLGVKAKSLNAFQVKIKTTPHHTSARIVDIDDSVIRENLK
         80        90        100       110       120

            60        70        80        90
ORF50 ENDILVVAGFQGINEHQELTTLGRGGSDTTAVALAVSNQIPC
      + D+LVV GFQG+NEH +LTTLGRGGSDT+AVALA   ++PC
lysC  DYDVLVVTGFQGVNEHGDLTTLGRGGSDTSAVALAAKLRVPC
         130       140       150       160
```

Figure A.2.1a-w

Sequence, translation and BLAST analysis of the DNA region flanking the Tn*917* insertion site

The best 5 scores and amino acid alignments, of predicted ORFs were the product of a BLAST database search.

A.2.2 Further sequence analysis of *seg*

Figure A.2.2a

seg5

```
AGTTGCATCCTTAGCATTATTAATCGCAATCACATTGTTTGTTAGAAAAAAATCTAAAGG      60
                  ORF5
               M  R  N  V  K  Q  I  A  T  K  S  I  I  A  I  I
CAATGTGGAATAATGAGAAATGTTAAACAAATTGCTACAAAATCTATTATAGCTATTATT     120

 S  L  G  I  L  T  Y  T  T  M  I  G  S  V  L  A  D  E  I  K
AGCTTAGGTATACTTACATATACAACAATGATTGGTAGCGTGTTGGCTGATGAGATAAAA     180

 Y  P  S  A  K  F  N  Q  P  E  A  K  D  K  T  E  L  T  T  S
TATCCATCAGCCAAATTTAATCAACCTGAAGCAAAAGATAAAACAGAATTAACTACATCA     240
                                                      seg5
                                                       ∇
 I  F  D  E  K  I  K  E  N  K  A  L  E  L  L  I  F  N  Q  E
ATTTTTGATGAAAAGATAAAAGAGAATAAAGCGTTAGAGTTACTAATTTTTAATCAAGAA     300

 N  K  N  V  T  E  E  Q  Q  L  V  D  E  K  A  Q  L  I  S  D
AATAAAAATGTAACTGAGGAACAACAACTAGTTGACGAAAAGGCGCAATTGATTTCAGAT     360

 M  T  G  K  I  Y  L  Q  V  K  L  K  G  Q  I  D  K  D  Q  R
ATGACTGGTAAAATTTACTTGCAAGTAAAGCTAAAAGGTCAAATAGATAAAGACCAACGT     420

 V  F  Q  N  D  K  N  E  E  F  P  L  V  I  K  D  E  K  D  D
GTTTTTCAAAATGACAAAAATGAAGAATTTCCTTTGGTAATAAAAGATGAAAAGGATGAC     480

 T  I  V  R  I  L  I  E  Q  H  M  D  K  I  N  M  H  V  K  T
ACAATAGTAAGAATTTTAATTGAACAGCATATGGATAAAATCAATATGCATGTTAAAACG     540

 L  A  E  K  K  N  L  D  N  K  E  M  V  Y  S  I  H  F  K  E
TTGGCTGAAAAGAAAAATCTAGATAACAAAGAAATGGTGTATTCTATTCATTTTAAAGAG     600

 K  K  V  Q  H  D  D  A  K  E  V  P  S  K  H  Q  N  Q  E  N
AAAAAAGTACAACATGATGATGCAAAAGAAGTGCCTTCAAAACATCAAAATCAAGAAAAT     660

 N  Q  D  Q  L  K  K  D  I  D  D  K  K  D  S  Q  K  S  D  T
AATCAAGATCAGCTTAAAAAAGATATTGATGACAAAAAAGATAGTCAAAAATCAGATACT     720

 K  E  R  R  T  S  L  F  T  E  K  G  L  N  D  I  P  V  Q  K
AAGGAAAGACGTACTAGCCTTTTTACTGAAAAAGGATTAAATGATATTCCTGTACAAAAA     780

 D  K  V  Q  Q  D  S  N  K  K  I  E  N  E  R  P  K  A  S  G
GATAAAGTGCAACAAGACAGTAATAAAAAGATTGAAAATGAGCGACCTAAAGCATCAGGT     820

 T  L  K  V  E  N  S  P  P  T  I  K  K  V  E  N  N  H  K  E
ACATTAAAAGTTGAAAATAGCCCTCCAACAATAAAAAAGGTTGAAAATAATCACAAAGAG     900

 Q  P  K  H  K  D  E  K  S  K  K  E  K  K  K  V  V  E  K  E
CAACCGAAACATAAAGATGAAAAATCAAAAAAGGAAAAGAAAAAAGTAGTTGAAAAAGAA     960

 K  A  L  P  A  F  N  R  D  D  D  S  K  N  S  S  Q  L  S  S
AAAGCGTTACCAGCTTTTAATAGAGATGATGATAGCAAGAATAGTAGTCAATTATCTAGT    1020

 D  I  K  E  L  D  E  P  N  H  K  K  Q  Y  M  L  F  A  A  G
GATATTAAAGAACTTGATGAACCAAATCATAAAAAGCAATATATGTTATTTGCAGCTGGC    1080

 I  V  L  A  T  I  L  L  I  S  A  H  L  Y  S  R  K  R  G  N
ATTGTGTTAGCAACTATTTTACTTATTTCGGCACATTTATACAGCAGAAAGAGAGGTAAC    1140

 Q  V  *
CAAGTTTGAGAATCATAAAGTATTTAACCATTTTAGTGATAAGCGTCGTTATCTTAACCA    1200
```

5 best Blast scores of ORF5

No significant homology found

Figure A.2.2b
***seg*8**

```
GTACCGCGAAACATAAGCTTTCGTCCTTTTTATCCGATTCATTCGGGTACGAAGGACGGA      60
                          seg8                          ORF8
                           ∇                M  K  D  T  Q  L  A  Q
AGCTTTTTTTATTTTTTCTTATAGGAGGGTCTTAATATGAAGGATACACAGTTAGCCCAA     120

 I  T  L  T  D  D  S  T  G  A  I  A  N  P  I  H  L  S  T  A
ATCACATTAACCGATGATTCAACCGGTGCTATAGCGAATCCAATCCATTTATCTACTGCC     180

 Y  K  H  P  K  L  G  Q  S  T  G  F  D  Y  T  R  T  K  N  P
TACAAGCATCCAAAACTAGGACAATCGACAGGTTTTGATTATACACGTACTAAAAATCCT     240

 T  R  S  T  F  E  T  C  F  A  K  L  E  H  G  I  A  S  F  A
ACACGCTCAACATTTGAAACCTGTTTTGCCAAACTTGAGCATGGTATTGCATCATTCGCT     300

 T  S  S  G  M  S  A  I  Q  L  I  C  N  L  F  K  P  H  D  E
ACATCAAGTGGAATGTCAGCCATTCAATTAATATGTAATCTATTTAAACCTCATGATGAA     360

 I  L  V  S  F  D  L  Y  G  G  T  F  R  L  F  E  F  Y  E  Q
ATTTTAGTTTCATTCGATTTATATGGTGGCACATTTAGATTATTTGAATTTTACGAGCAA     420

 Q  Y  D  I  K  F  K  Y  V  D  F  T  D  Y  E  Q  V  E  K  E
CAATACGATATCAAATTTAAGTACGTTGATTTTACAGATTATGAACAAGTTGAAAAAGAA     480

 I  T  D  K  T  V  A  L  F  I  E  P  I  S  N  P  Q  M  I  A
ATCACTGATAAAACAGTTGCATTATTCATTGAACCAATATCTAACCCACAAATGATTGCT     540

 I  D  V  K  P  Y  Y  Q  L  C  K  A  K  G  L  L  S  I  I  D
ATTGATGTAAAGCCATACTATCAACTTTGTAAAGCTAAAGGCTTATTGTCAATTATCGAC     600

 N  T  F  L  T  P  Y  L  S  T  P  L  A  E  G  A  D  I  V  L
AATACTTTTTTAACACCTTATCTTTCAACACCACTAGCAGAAGGTGCTGATATAGTCTTA     660

 H  S  A  T  K  Y  I  G  G  H  N  D  V  L  A  G  V  V  T  V
CATTCAGCCACGAAATATATTGGCGGACATAACGATGTACTAGCAGGTGTCGTAACCGTC     720

 K  D  E  S  L  A  Q  Q  L  F  D  F  H  N  M  T  G  A  T  L
AAAGATGAATCACTCGCGCAACAGTTGTTTGATTTTCACAACATGACTGGCGCAACACTT     780

 S  P  I  D  S  Y  L  L  L  R  G  L  K  T  L  H  L  R  I  E
TCACCAATAGATAGTTATTTGTTGTTACGTGGACTTAAAACTTTGCATTTACGCATTGAG     840

 R  A  Q  S  N  A  R  K  L  A  K  K  C  Q  S  L  Q  A  I  D
CGTGCGCAATCAAACGCTAGAAAACTTGCTAAAAAATGTCAGTCACTTCAAGCAATTGAC     900

 E  V  L  Y  S  G  Q  T  G  M  L  S  L  R  L  N  K  A  Y  S
GAAGTACTATATAGCGGGCAAACTGGCATGCTTAGTTTAAGACTTAACAAGGCCTATAGC     960

 V  A  K  L  L  E  N  L  D  I  C  I  F  A  E  S  L  G  G  T
GTCGCTAAATTATTAGAAAATTTAGACATTTGCATTTTTGCAGAAAGTTTAGGAGGTACT    1020
```

5 best BLAST scores of ORF8

Gene Heading	**Accession No.**	**% Identity**
Blast: yjci_ bacsu	A69847	49% over 324 amino acids
Blast: metb_haein	P44502	46% over 328 amino acids
Blast: metb_ haein	G64047	46% over 328 amino acids
Blast: metc_lacla	AFF14693.1	37% over 329 amino acids
Blast: metc_lacla	AFF14695.1	38% over 329 amino acids

Amino acid alignment

ORF8 - 49% identity over 324 amino acids to YjcI from *B. subtilis* (yjci_bacsu, A69847).

```
ORF8: 104 DTQLAQI-TLTDDSTGAIANPIHLSTAYKHPKLGQSTGFDYTRTKNPTRSTFETCFAKLE 283
          +T+LAQI   +D+ TG ++ PI+LSTAY+H  +G+STGFDY RTKNPTR   E   A LE
yjcI:  6  ETKLAQIGNRSDEVTGTVSAPIYLSTAYRHRGIGESTGFDYVRTKNPTRQLVEDAIANLE  65

ORF8: 284 HGIASFATSSGMSAIQLICNLFKPHDEILVSFDLYGGTFRLFEFYEQQYDIKFKYVDFTD 463
          +G    A SSGM+AIQ I  LFK  DE++VS DLYGGT+RLFE   ++Y + F Y DF+D
yjcI:  66 NGARGLAFSSGMAAIQTIMALFKSGDELIVSSDLYGGTYRLFENEWKKYGLTFHYDDFSD 125

ORF8: 464 YEQVEKEITDKTVALFIEPISNPQMIAIDVKPYYQLCKAKGLLSIIDNTFLTPYLSTPLA 643
           + +  +IT  T A+F+E  +NP M   D++   ++ K  GLL I+DNTF TP L  PL
yjcI: 126 EDCLRSKITPNTKAVFVETPTNPLMQEADIEHIARITKEHGLLLIVDNTFYTPVLQRPLE 185

ORF8: 644 EGADIVLHSATKYIGGHNDVLAGVVTVKDESLAQQLFDFHNMTGATLSPIDSYLLLRGLK 823
           GADIV+HSATKY+GGHND+LAG+V VKDE L +++F   N  GA L P DS+LL+RG+K
yjcI: 186 LGADIVIHSATKYLGGHNDLLAGLVVVKDERLGEEMFQHQNAIGAVLPPFDSWLLMRGMK 245

ORF8: 824 TLHLRIERAQSNARKLAKKCQSLQAIDEVLYSGQTGMLSLRLNKAYSVAKLLENLDICIF 1003
          TL LR+ + Q+NA++LA   +  + I +VLY G+ GMLS RL K   V   L+ L    F
yjcI: 246 TLSLRMRQHQANAQELAAFLEEQEEISDVLYPGKGGMLSFRLQKEEWVNPFLKALKTICF 305

ORF8:1004 AESLGGTETLVTFPYTQTHVDMPD 1175
          AESLGG E+ +T+P TQTH+D+P+
yjcI: 306 AESLGGVESFITYPATQTHMDIPE 329
```

Figure A.2.2c

***seg*22 (*seg*30, *seg*33)**

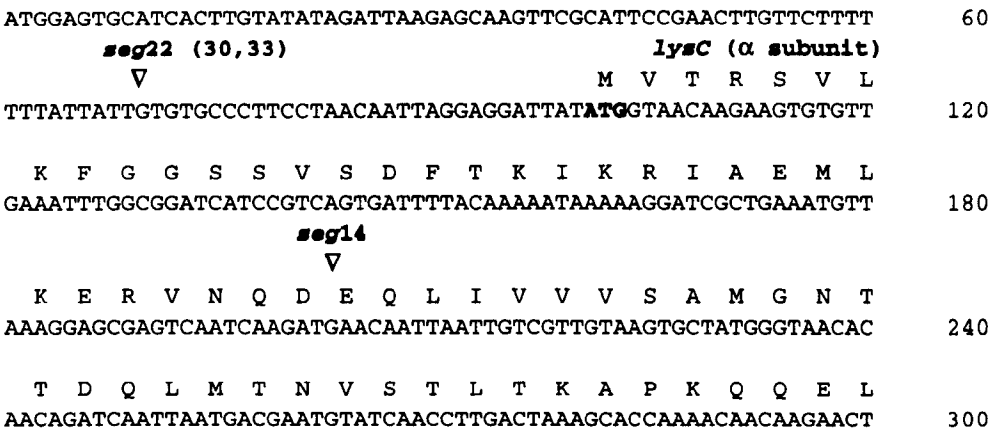
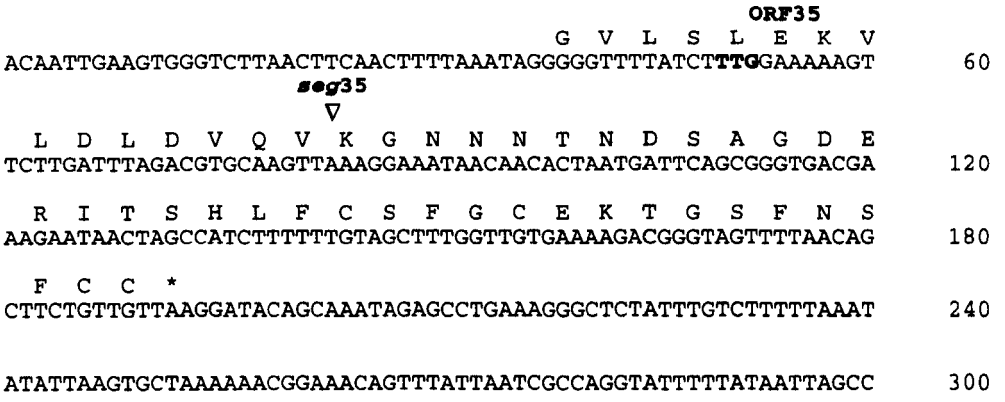


Diagram showing the location of the Tn*917* insertions in *seg*22 (*seg*30, *seg*33)

The transposon insertion sites are shown in relation to the start codon (**ATG**) of *lysC* (α subunit) and the Tn*917* insertion in *seg*14.

Figure A.2.2d

*seg*35



5 best BLAST scores of ORF35

Gene Heading	Accession No.	% Identity
Blast: lang_staga	P21838	51% over 45 amino acids
Blast: lane_staep	P08136	50% over 44 amino acids
Blast: mutA_strmu	AAC18827.1	56% over 25 amino acids
Blast: lanm_strmu	P80666	76% over 17 amino acids
Blast: epid_ staep	CAA30690.1	59% over 22 amino acids

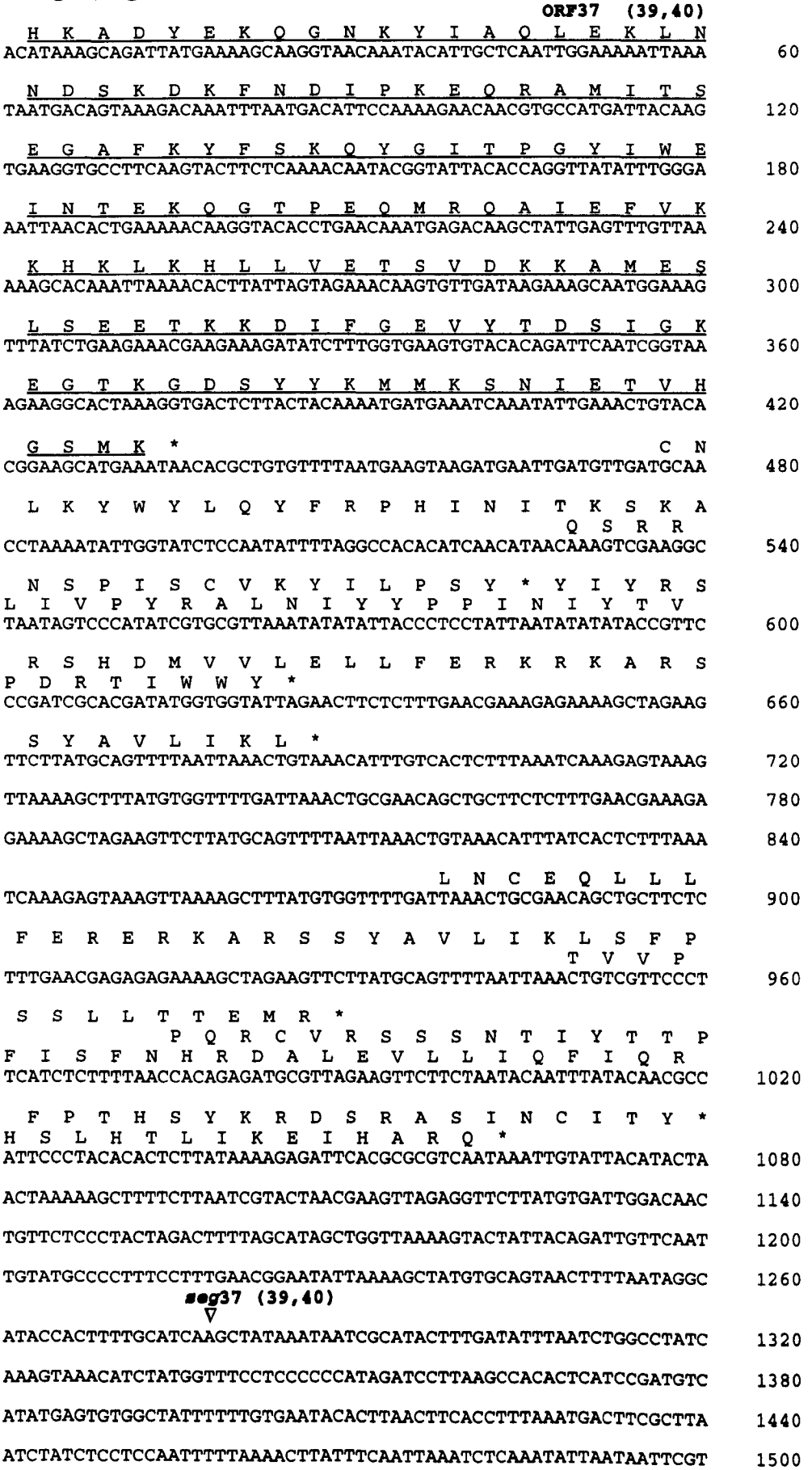
Amino acid alignment

ORF35 - 51% identity over 45 amino acids to lantibiotic gallidermin precursor (GdmA) from *S. gallinarum* (lang_staga, P21838).

```
ORF35: 28   KVLDLDVQVKGNNNTNDSAGDERITSHLFCSFGCEKTGSFNSFCC 190
           ++ DLDV+V    + NDS  + RI S   C+ GC KTGSFNS+CC
gdhM : 9    ELFDLDVKVNAKES-NDSGAEPRIASKFLCTPGCAKTGSFNSYCC 52
```

Figure A.2.2e

***seg*37 (*seg*39, *seg*40)**



5 best BLAST scores of ORF37 (39/40)

Gene Heading	Accession No.	% Identity
Blast: lipo_staep	CAA67571.1	76% over 143 amino acids
Blast: psaA_strmi	AFF055088	48% over 139 amino acids
Blast: psaA_strpn	AAD09975.1	48% over 139 amino acids
Blast: psaA_ strpn	ABB09440.1	48% over 139 amino acids
Blast: ensp_entfa	AAA70056.1	47% over 143 amino acids

Amino acid alignment

ORF37 (39/40) - 76% identity over 143 amino acids to a lipoprotein from *S. epidermidis* (lipo_staep, CAA67571.1).

```
ORF37: 2   KADYEKQGNKYIAQLEKLNNDSKDKFNDIPKEQRAMITSEGAFKYFSKQYGITPGYIWEI 61
           K+ YEKQGN YI++LE+LN DSK+KF+DIPK QRAM+TSEGAFKYF++Q+ + PGYIWEI
lipo : 167 KSTYEKQGNAYISKLEELNKDSKNKFDDIPKNQRAMMTSEGAFKYFAQQFDVKPGYIWEI 226

ORF37: 62  NTEKQGTPEQMRQAIEFVKKHKLKHLLVETSVDKKAMESLSEETKKDIFGEVYTDSIGKE 121
           NTEKQGTP QM+QAI+FVK + LKHLLVETSVDKKAM+SLSEETKKDI+GEV+TDSIGKE
lipo : 227 NTEKQGTPGQMKQAIKFVKDNHLKHLLVETSVDKKAMQSLSEETKKDIYGEVFTDSIGKE 286

ORF37: 122 GTKGDSYYKMMKSNIETVHGSMK 144
           GTKGDSYYKMMKSNI+T+HGSMK
lipo : 287 GTKGDSYYKMMKSNIDTIHGSMK 309
```

Figure A.2.2a-e

Further sequence analysis of *seg*

*seg*5 (**a**), *seg*8 (**b**), *seg*22 (*seg*30, *seg*33) (**c**), seg35 (**d**), and *seg*37 (*seg*39, *seg*40) (**e**), were all analysed further by searching *S. aureus* databases for contigs with homology to the *seg* sequences. These contigs were then analysed to determine the identity of the sera expressed gene.