Paterson, Gavin Kirkwood (2003) *Identification and characterisation of novel pneumococcal virulence factors.*

PhD thesis

http://theses.gla.ac.uk/3576/

Identification and characterisation of novel pneumococcal virulence factors

Gavin Kirkwood Paterson

A thesis submitted to the University of Glasgow for the degree of Ph.D.

Institute of Biomedical and Life Sciences

Division of Infection and Immunity

Joseph Black Building

University of Glasgow

Glasgow

G12 8QQ

Submitted December 2003

UNIVERSITY
of
GLASGOW

Contents

3

List of Tables

List of Figures

Declaration of authorship

This thesis is the original work of the author unless otherwise stated

Gavin K. Paterson

Acknowledgements

Thank you to everyone who helped in various ways but especially to Clare for her support, encouragement and advice.

Abbreviations

APS	Ammonium persulphate
ATP	Adenine tri-phosphate
BAB	Blood agar base
BHI	Brain heart infusion media
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine serum albumin
Cbp	Choline binding protein
cDNA	Complementary DNA
CFU	Colony Forming Units
CSP	Competence stimulating peptide
^{o}C	Degrees Celsius
d	Day(s)
DFI	Differential fluorescence induction
dH$_2$O	Distilled water
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphates
DTT	Dithiothreitol
EDTA	Diaminoethanetetra-acetic acid disodium salt
gDNA	Genomic DNA
g	Gram(s) or g-force (when referring to centrifugation speeds)
hr	Hour(s)
Kb	Kilobase
K.Da	Kilodaltons
L	Litre(s)
LB	Luria broth
µg	Microgram(s)
µl	Microlitre(s)
mg	Milligram(s)
ml	Millilitre(s)
mM	Millimolar
min	Minute(s)
M	Molar
MLST	Multilocus sequence typing
mRNA	Messenger RNA
n	Population size (for statistical analysis)
OD	Optical density
O/N	Overnight
%	Percent
PAFr	Platelet activating factor receptor
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
rpm	Revolutions per minute

RNA	Ribonucleic acid
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
sec	Second(s)
SDS	Sodium-dodecyl sulphate
SE	Standard error of the mean
STM	Signature-tagged mutagenesis
TIGR	The Institute of Genomic Research
TEMED	N',N',N',N'-tetramethylethylenediamine
WT	wild-type
2D	2-dimensional

Abstract

Streptococcus pneumoniae (the pneumococcus) is an important cause of human death and disease worldwide. Emphasis for acquiring a greater understanding of the host-microbe interactions in pneumococcal pathogenesis is provided by the spread of antibiotic resistance and inadequacies with current vaccines. To contribute to that understanding the aim of this thesis was to use the completed pneumococcal genome sequences to identify putative virulence factors and to characterise these further.

Sortase enzymes, responsible for the covalent anchoring of specific proteins to the cell wall of Gram-positive bacteria have been shown to be important virulence factors for several pathogens. The identification of four sortase homologues *srtA-D* in the pneumococcal genome prompted their investigation as candidate pneumococcal virulence factors. Not all of these sortase genes were present in both sequenced strains and so sortase gene distribution was investigated among a collection of clinical isolates. In contrast to *srtB*, *C* and *D*, *srtA* was found in all strains examined and so was selected for further study. It was subsequently found to contribute to pneumococcal virulence in mouse models of pneumonia, bacteraemia and colonisation. This is the first demonstration of a contribution of *srtA* to pneumococcal virulence. To complement this examination of *srtA*, some of the surface proteins known or likely to be anchored by SrtA were also investigated for a role in virulence in the animal model of pneumonia.

In addition, two genes, annotated in the pneumococcal genomes as a macrophage infectivity potentiator protein and an exfoliative toxin A were also investigated and found to be novel pneumococcal virulence factors. However, it appears they have been incorrectly annotated and these genes do not represent a macrophage infectivity

potentiator protein and exfoliative toxin A. Instead, one of them seems to be involved in the response to oxidative stress while no function for the other can yet be ascertained.

Janus mutagenesis is a novel technique for manipulation of the pneumococcal genome allowing the creation of mutations that lack a selectable marker. This provides an accessible and potentially powerful method to easily alter the genome to make informative mutations. This thesis describes the first use, to our knowledge, of Janus mutagenesis to investigate pneumococcal virulence.

Chapter One Introduction

Streptococcus pneumoniae (the pneumococcus) a major human pathogen is a Gram-positive, facultative anaerobic bacterium. First described in 1880 by the independent work of Pasteur and Sternberg (Davis BD 1990) its study has lead to several important discoveries. These include the identification of nucleic acid as the hereditary matter (Avery *et al.* 1944) with this work also highlighting the importance of bacterial polysaccharide capsules in virulence. Work on the pneumococcus lead also to recognition of the first polysaccharide antigen (Goebel and Adams 1943), the recombination mechanism of transformation (Griffith 1928), the first bacterial quorum-sensing polypeptide (Tomasz 1965) and the first bacterial autolysin (Dubos 1937). As noted by Tomasz (Tomasz 1999) the pneumococcus has a split image; on one hand a dangerous pathogen causing considerable disease and yet a useful contributor to our understanding of molecular biology.

Pneumococcal colonisation and disease

The primary niche of the pneumococcus is the human nasopharynx where it exists asymptomatically as a commensal being spread by aerosol. Colonisation can occur within hours of birth and by twelve days post-birth the carriage rates are similar to that of the babies' mothers (Gray and Dillon 1989). Colonisation is typically transient lasting several weeks to months with up to four different capsular serotypes present at one time (Gillespie 1989). Virtually all individuals are likely to be colonised at least at one stage in their life. Carriage rates are highest in young children and in crowded environments, reaching up to 70%. In addition to humans, the pneumococcus has been isolated from horses. This phenomenon appears to be restricted to a clonal group of serotype 3 strains

genetically distinct from their human counterparts (Whatmore *et al.* 1999). Farm workers are colonised with strains distinct from those found in horses and so cross-infection appears unlikely (Burrell *et al.* 1986). Interestingly, equine isolates lack pneumolysin and autolysin, two virulence factors reportedly found in all human isolates tested to date (Whatmore *et al.* 1999).

The recent acquisition of a colonising strain is regarded as the prerequisite to pneumococcal disease. The most common diseases caused by *S. pneumoniae* are pneumonia, otitis media, bacteraemia and meningitis. In the USA, the pneumococcus is estimated to cause seven million cases of otitis media, over 500 000 cases of pneumonia, 50 000 cases of bacteraemia and 3000 cases of meningitis per year, representing a substantial health care burden (Obaro 2000). Additionally, in developing countries, pneumococcal pneumonia is estimated to cause one million deaths in children under the age of five (Denny and Loda 1986). However, the true disease burden caused by the pneumococcus is uncertain because these aliments can be caused by a variety of different organisms and are often treated without bacteriological confirmation of the cause. This is particularly true in the developing world where the disease burden may be highest but diagnostic capacity limited. Improved diagnostics or a highly efficacious vaccine is suggested to provide a truer reflection of the pneumococcal disease burden and to show current values to be underestimates (Obaro 2000).

In addition to these common diseases, the pneumococcus appears able to cause infection in most tissues it is able to reach, with a long list of additional aliments including; cellulitis, pericarditis, endocarditis, septic arthritis, osteomyelitis, haemolytic uremic syndrome and sinusitis. Most of these invasive diseases are rare and appear to arise as a

complication of systemic infection. For example, two studies looking at pneumococcal septic arthritis found that 69% of patients where positive for blood cultures of *S. pneumoniae* (James and Thomas 2000; Ispahani *et al.* 1999). Additionally, the pneumococcus has been isolated from the sputum of up to 28% of patients with exacerbations of chronic obstructive pulmonary disease, although its aetiological role, if any, is unclear (Sethi and Murphy 2000).

Identification and Epidemiology

When grown on blood agar pneumococcal colonies display characteristic draughtman-like colony morphology and produce a zone of α-haemolysis. Sensitivity to optochin (hydrocupreine hydrochloride) is used to give a presumptive identification of *S. pneumoniae* while bile solubility offers further confirmation. Optochin sensitivity is thought to be over 98% sensitive with a specificity of 100% (Kellogg *et al.* 2001).

Over 90 different serotypes of *S. pneumoniae* have been identified based on the antigenically distinct polysaccharide capsule surrounding the bacterium (Kalin 1998). Serotypes can be distinguished using the Quellung reaction. This simple test involves mixing serotype-specific anti-capsule antisera with bacterial cultures. The clumping of bacterial cells indicates a positive result, showing presence of that specific capsule on the bacteria. In addition to capsule serotyping, multilocus sequence typing (MLST) has more recently been used to characterise pneumococcal isolates (Enright and Spratt 1998). This entails sequencing a ~450 base pair fragment from each of seven housekeeping genes and provides a highly discriminatory scheme to characterise isolates. In addition to its ability to sensitively discriminate between a large numbers of isolates, MLST has the particular

advantage over other typing methods of the unambiguity and electronic portability of DNA sequence data allowing direct comparison between laboratories.

Although most, if not all, capsular serotypes have been associated with infection, some are more commonly recovered from cases of serious disease than others, with 90% of invasive disease worldwide caused by 16 serotypes (Feldman and Klugman 1997). Furthermore, the prevalent serotypes vary with age group, geography and can change over time. In the 20th century serotype 2 was an important contributor to disease in the USA and Europe but is now rarely recovered from disease (Kalin 1998). In adults a greater number of different serotypes are isolated than from children with the prevalence of serotypes varying between even apparently similar countries such as the USA and UK (Kalin 1998). The reasons for such changes and differences are not well-studied. Examining the invasive disease and carriage rates of pneumococci among children in Oxford it was found that serotypes 1, 4 and 14 had a high propensity to cause invasive disease. Additionally, it appeared that capsule was a more important marker for invasive disease than sequence type (as assessed by MLST) (Brueggemann *et al.* 2003). In support of this was the finding that genetically divergent clones with the same serotype had similar abilities to cause invasive disease while strains with the same sequence type but different serotypes had invasiveness matching that of their serotype (Brueggemann *et al.* 2003). Why this should be is unclear but may relate to differences between serotypes in terms of activation of and/or resistance to complement and/or the ability to induce antibodies and clearance by lectin-like structures on phagocytes (AlonsoDeVelasco *et al.* 1995).

Prevention and treatment

Antibiotics

Following the introduction of penicillin in the 1940s, the treatment of pneumococcal disease was so successful that vaccine development was neglected for years. Indeed, early vaccines were withdrawn from the market in the 1950s as a result of the efficacy of penicillin. However, even with antibiotics and improvements in intensive care mortality remains high (Hook *et al.* 1983; Kramer *et al.* 1987). The first report of penicillin resistance came from New Guinea in 1969 and prompted an editorial comment in *The New England Journal of Medicine* the following year (Finland 1969). However, while this editorial recognised the serious implications of antibiotic resistance it suggested that the likelihood of resistance spreading was remote. This belief turned out to be misplaced as the frequency and level of resistance to penicillin and other antibiotics increased worldwide particularly during the 1980s in South Africa, Spain, Hungry and France (Tomasz 1999). Antibiotic resistance is thought to be due to selective pressure from increased prescription of antibiotics and its spread is facilitated by the natural transformability of the pneumococcus allowing horizontal gene transfer from other strains and organisms. Resistant strains are more common in children, which may reflect the widespread use of antibiotics in this age group. The main mechanisms of resistance to penicillin are alterations to penicillin binding proteins resulting in their reduced affinity for penicillin and other β-lactams. The full impact of antibiotic resistance upon clinical outcome remains to be determined (Tan 2003).

New antimicrobials against the pneumococcus in the form of bacteriophage cell wall lytic enzymes show potential in animal models and may have a useful clinical role in future

(Loeffler *et al.* 2001; Loeffler *et al.* 2003). The available genome sequence of the pneumococcus may facilitate the identification of targets for antimicrobials (Paton and Giammarinaro 2001; Di Guilmi and Dessen 2002)

Vaccines

With antibiotic resistance rising and the greater economic costs of treatment compared to prevention, vaccination is an important issue in pneumococcal disease. Indeed even with antibiotic treatment mortality can be high. A pneumococcal vaccine comprising purified polysaccharide preparations from 14 serotypes was first licensed in the United States in 1977. This was replaced, in 1983, with the 23-valent polysaccharide vaccine that is now widely in use (Pneumovax II, Pnu-Immune 23). This vaccine contains purified polysaccharide antigens from 23 serotypes implicated in the majority of invasive disease in the developed world (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F and 33F). It is recommended for those at risk of pneumococcal infection and has been shown to be highly cost effective when used for vaccinating elderly persons (Ament *et al.* 2000). The protection conferred is serotype specific and so it does not protect against serotypes not included in the vaccine. Although effective in adults, this vaccine fails to induce a good response in children younger than 2 years, the age group with the highest incidence of infection (Wuorimaa and Kayhty 2002). This is because the polysaccharide antigen activates an antibody response independent of T cell help via direct stimulation of antigen specific B cells. However, these T cell independent responses are weak in the immature immune system of young infants thereby undermining the value of vaccination. Furthermore, 20 % of the elderly

18

population also respond poorly to polysaccharide antigens due to a weak IgG response (Wuorimaa and Kayhty 2002). The polysaccharide vaccine does not provide any protection against carriage of the organism and so does not provide herd immunity.

Due to the ineffectiveness of polysaccharide vaccines in infants, conjugate vaccines were developed. These are polysaccharide antigens covalently linked to a carrier protein rendering the polysaccharide a T-cell dependent antigen and thereby able to invoke a response in infants. The utility of such an approach had previously been seen with a highly successful glycoprotein conjugate vaccine against *Haemophilus influenza* type B (Hib).

Several pneumococcal conjugate vaccines have been developed and are in various stages of testing. One of these preparations, a 7-valent conjugate (Prevnar™ / Prevenar®, Wyeth Lederle Vaccines), was licensed in the United States in February, 2000. The polysaccharides included in this vaccine are from serotypes 4, 6B, 9V, 14, 18C, 19F and 23F. These are conjugated to the non-toxic CRM197 diphtheria toxoid carrier protein. Other conjugate vaccines in trials include 9- and 11-valent preparations. Studies with Prevnar™ have shown it to be protective against otitis media, pneumonia and invasive disease caused by vaccine serotypes (Black *et al.* 2000 Eskola *et al.* 2001). Importantly, the vaccine has been shown to be safe, immunogenic and protective in infants less than 2 years of age. In the first efficacy trial of infants in America, the vaccine gave 73% protection against pneumonia, 8.9% against otitis media and was 97% effective against invasive disease caused by vaccine types. Although the protection against otitis media appears small this would still translate to a huge reduction in the number of cases and their associated health care burden given the high prevalence of otitis media (Tan 2000;

Obaro and Adegbola 2002). In addition, the conjugate vaccine appears to reduce colonisation by serotypes included in the vaccine. However, although this might be a valuable feature, resulting in reduced transmission and therefore herd immunity it seems to be due in part, to replacement by serotypes not included in the vaccine (Tan 2000; Obaro and Adegbola 2002). The conjugate vaccine has not yet been licensed for use in the UK.

Despite the success and potential of conjugate vaccines, they have several drawbacks. In particular the production costs are considerably increased over the current polysaccharide vaccine and may prohibit use in developing countries where the burden of disease is especially high. For example, although the Hib vaccine is effective and has been available for several years it is not routinely available in developing countries despite the positive impact it would have there. As discussed aboved there is the concern that a decrease in disease by vaccine serotypes will be countered by an increase in disease by non-vaccine types. This has already been demonstrated in a Finnish study of otitis media (Eskola *et al.* 2001), and may also extend to increased colonisation and disease by other organisms. For example, *S. pneumoniae* has been shown to antagonise ecological competitors through the production of hydrogen peroxide (Pericone *et al.* 2000) and can desialylate the lipopolysaccharide of *H. influenzae* and *Neisseria meningitidis* (Shakhnovich *et al.* 2002) potentially rendering them more susceptible to clearance by the immune system. The absence of the pneumococcus may therefore allow enhanced colonisation and disease by these and other organisms. It will also be necessary to provide different formulations of serotypes in the vaccine to meet geographical differences in serotype prevalence. For example, although the nine-valent conjugate

vaccines cover the serotypes causing 80-90% of pneumococcal disease in many areas it would cover only 66% of those in Asia (Obaro 2001). There may also a problem with horizontal gene transfer allowing vaccine types to acquire a non-vaccine serotype capsule and thereby persist in the population. Evidence for this exchange of capsule type has already been described in a natural pneumococcal populations (Tomasz 1999). It will therefore be important to monitor the serotype and genotype make-up of the pneumococcal population.

Due to the inefficiency of the 23-valent polysaccharide vaccine in infants and the problems with cost and serotype coverage discussed with the conjugate vaccines there is a need for improved pneumococcal vaccines with increased serotype coverage and at a lower cost. While intense epidemiological investigations may increase our understanding of disease causing serotypes and their geographical prevalence it would not be practical to perform this worldwide and to tailor-make conjugate vaccines to suit individual populations particularly in the developing countries. Much research is currently focussed upon identifying suitable candidate proteins for a new vaccine. A pneumococcal protein or mixture of proteins could have the advantage of providing cross-protection against many serotypes provided it is conserved between many/all strains. Several pneumococcal proteins, including pneumolysin and pneumococcal surface protein A (PspA), are currently being investigated as potential vaccine candidates (reviewed by (Swiatlo and Ware 2003).

Many new techniques have been applied to the identification of virulence factors and putative vaccine candidates of *S. pneumoniae*, including signature tagged mutagenesis (STM), differential fluorescence induction (DFI) and motif searches for surface

expressed molecules (Pearce *et al.* 1993; Polissi *et al.* 1998; Wizemann *et al.* 2001; Hava and Camilli 2003; Marra *et al.* 2002). These studies have identified a number of genes that have potential as vaccine candidates, although detailed analysis of these must be performed to assess their suitability for inclusion in a future vaccine. A potentially important development in pneumococcal vaccine research is the use of human experimental colonisation to identify pneumococcal immunogens and protective responses (McCool *et al.* 2002; McCool *et al.* 2003).

Host-microbe interactions

A variety of animal models exist for pneumococcal diseases and these have proved useful in studying the pneumococcus. Most prominently, mouse models are used for pneumonia, bacteraemia, meningitis and colonisation. In addition, chinchillas are used for colonisation and otitis media studies with infant rats and rabbits are also frequently used as models. From these models a range of pneumococcal virulence factors have been identified, too numerous to discuss in detail here. The polysaccharide capsule is regarded as a key virulence factor protecting the bacterium from phagocytosis in the absence of anti-capsular antibodies. The specific disruption of the capsule renders the organisms effectively avirulent (Watson and Musher 1990). Signature-tagged mutagenesis has been the most proliferate method for identifying pneumococcal virulence factors. Three independent screens have been performed on the pneumococcus in mouse pneumonia infections (Polissi *et al.* 1998; Lau *et al.* 2001; Hava and Camilli 2002). Together these have identified over 230 genes involved in virulence. Interestingly, there was little overlap in the genes identified in the different screens. This may be due to the use of

different bacterial and mouse strains, sampling time point, mutagenesis strategy and stringency of cut-offs. Additionally, the limited overlap suggests that not all pneumococcal virulence factors were identified. This is supported by the inability of the three STM screens to identify the capsule as a virulence factor despite its demonstrated role in virulence. In addition to the large number of genes involved in virulence, complexity is added by differences in the contribution to virulence of certain factors between different strains. For example, while the response regulator *rr09* contributes to virulence in a serotype 3 strain it was absolutely essential for virulence in a serotype 2 strain (Blue and Mitchell 2003). Additionally, strain differences in gene content, for example the *rlrA* pathogenicity island which is present is some but not all strains (Tettelin *et al.* 2001) may affect virulence as might sequence variation in genes between strains. The contribution to virulence of a particular factor may also vary with the type of infection. For example, the transcriptional regulator *rlrA* contributes to virulence in pneumonia and colonisation but not to bacteraemia (Hava and Camilli 2002). Furthermore, some virulence factors may have many different activities making it difficult to interpret their full role in virulence, for example pneumolysin (Rubins and Janoff 1998; Cockeran *et al.* 2002).

Adhesion and invasion

In the nasopharynx pneumococci adhere to host cells through the recognition of *N*-acetyl-D-galactosamine β1-3 galactose (GalNAcβ1-3Gal) that forms part of a glycolipid receptor on the epithelial cell surface (Andersson and Svanborg Eden 1988). The presence of this sugar in human milk is suggested to contribute to the ability of breast-feeding to reduce

pneumococcal colonisation and otitis media (Cunningham 1979), although other factors in human milk such as antibodies may also contribute to this effect (Finn *et al.* 2002).

In addition, the pneumococcal surface protein, choline binding protein A (CbpA) mediates binding of to the human polymeric immunoglobulin receptor, an interaction that may contribute to trancyctosis of the bacterium and tissue invasion (Zhang *et al.* 2000). The significance of this finding has been questioned however, given that it appears to be bacterial and cell line specific event (Brock *et al.* 2002). In particular, of the bacterial strains tested the phenomenon appears restricted to acapsular, non-pathogenic laboratory strains and was not seen in any clinical isolates examined (Brock *et al.* 2002). This highlights the potential danger of over-reliance on laboratory strains to study pathogenesis.

The factors permitting spread from the nasopharynx to the lungs are poorly understood. A combination of pneumococcal strain and host factors likely contributes. In particular preceding viral infection especially with influenza A is associated with increased incidence of pneumococcal disease. An important factor in this affect is the viral neuraminidase removing the terminal sialic acids from epithelium to expose potential pneumococcal receptors (McCullers and Bartmess 2003). Within the lungs, pneumococci adhere to the alveolar wall, probably to type II pneumocyte cells. Bacterial spread to the blood suggests the ability to bind and cross through vascular endothelial cells also although spread through the lymphatic system may be another route of systemic spread. Attachment to these cells is via binding to host glycoconjugate receptors containing *N*-acetyl-$_D$-galactosamine β1-4 galactose (GalNAcβ1-4Gal) or N-acetyl galactosamine β1-3 galactose (GalNAc β1-3 Gal). The pneumococcal ligands responsible remain to be fully

investigated. A number of putative adhesins have been identified (Berry and Paton 1996;

Spellerberg et al. 1996; Cundell et al. 1995; Wizemann et al. 1996) however, many of

these are likely to have indirect roles in adhesion. For example, a mutant in the

pneumococcal pyruvate oxidase, SpxB was attenuated in the ability to bind a variety of

cells (Spellerberg et al. 1996). However, this defect was complemented by the addition of

acetate showing that SpxB was not acting directly as an adhesion but that acetyl-CoA the

product of SpxB activity was regulating the adhesion properties of the bacterium in some

way.

Cytokine activation of endothelial and lung cells leads to the up-regulation of the receptor

for platelet activating factor (PAFr) an additional receptor for the pneumococcus. The

pneumococcal ligand for PAFr appears to phosphorylcholine in the teichoic acid which

may structurally mimic PAF (Cundell et al. 1995). While invasion from blood to the

brain was thought to be the route leading to pneumococcal meningitis it now appears that

the pneumococcus can also reach the central nervous system directly from the

nasopharynx via retrograde axonal transport along olfactory neurons (Van Ginkel et al.

2003). This has important implications for vaccination because vaccines that block

bacteraemia but not carriage will fail to prevent this type of infection. The binding of

plasminogen and fibronection by the pneumococcal proteins alpha-enolase and PavA

respectfully may also play a role in tissue invasion (Bergmann et al. 2001; Holmes et al.

2001).

An important factor in pneumococcal colonisation and disease is that of phase variation

in colony opacity (reviewed by Weiser 2000). When viewed with oblique, transmitted

light and magnification on transparent medium it is possible to observe both opaque and

transparent colonies in the colonies derived from the same strain. This phenomenon appears common to all strains but may be more readily seen in certain serotypes possibly due to differences in capsule, which may obscure the opacity variation between colonies. The basis of this reversible phase variation is unclear but its consequences are becoming apparent. Transparent colonies appear to have a selective advantage in colonisation and opaque colonies an advantage in systemic infection. In line with these findings transparent colonies bind host cells in greater numbers and produce increased amounts of hydrogen peroxide, which may be important in out-competing other organisms during colonisation. In addition, the two opacity phenotypes differ in the amount of phosphorylcholine in the cell wall teichoic acids. The decreased amounts of phosphorylcholine in opaque colonies may explain their advantage in systemic infection by promoting avoidance of natural antibodies to phosphorylcholine and C-reactive protein that appears to recognise this structure and thereby promote bacterial clearance. In support of this is the finding that opaque colonies are more resistant to phagocytic killing *in vitro* (Kim *et al.* 1999). It has been suggested the two variants may also differ in the amount of polysaccharide capsule they produce but this has not been conclusively studied (Weiser 2000). In addition, the expression profile of choline binding proteins, surface proteins tethered to phosphorylcholine in teichoic acids varies between colony phase variants and may also contribute their different phenotypes (Weiser 2000).

Phase variation in many pneumococcal genes is suggested by the presence of sequence repeats similar to sites of phase variation in other organisms. Although understudied to date, phase variation may therefore be an important feature of pneumococcal biology beyond that seen with colony opacity. In support of this is the finding of a strain from

otitis media carrying a mutation in PspA in such a repetitive sequence that caused the secretion of a truncated version of this surface protein (McCool *et al.* 2002; McCool *et al.* 2003). Despite this mutation which rendered the protein unable to bind lactoferrin, its only known ligand, the strain was still able colonise human volunteers. The release of this highly immunogenic protein may have the advantage of allowing evasion of the immune response.

Environmental sensing and gene regulation

The ability to sense and respond rapidly (Lange *et al.* 1999; Throup *et al.* 2000) to their environment is important for bacterial survival. Little is known about the regulation of pneumococcal gene expression in response to environmental stimuli especially in the host setting. Differential fluorescence induction identifies gene promoters induced under specific conditions. It has been used on the pneumococcus to discover genes up-regulated under various conditions (Bartilson *et al.* 2001; Marra *et al.* 2002). Likewise microarray technology has also been used to investigate gene expression and will likely be heavily employed in future (de Saizieu *et al.* 2000; Robertson *et al.* 2002; Sebert *et al.* 2002; Ng *et al.* 2003). Two component systems have a central role in bacterial sensing and subsequent adaptation. Thirteen such systems have been identified in the pneumococcus with various important and complex roles in bacterial survival and pathogenesis. These systems are viewed as a potentially good target for novel antimicrobials due to their absence in the host and importance for bacterial biology.

Numerous other factors not discussed here are important in pneumococcal biology and pathogenesis such as the ability to acquire nutrients in the host, host defence and susceptibility, evasion of the immune response, resistance to stress and competence.

Genome sequencing and pneumococcal research

The genome sequence of two pneumococcal strains has been completed with a third almost finished (Hoskins et al. 2001; Tettelin et al. 200). The fully annotated genomes from two of these strains; R6 an acapsular, avirulent laboratory strain and TIGR4, an invasive serotype 4 strain are available publicly (see www.tigr.org). In addition genome sequencing of a serotype 23F (23F-1) and serotype 6 strain (670-6B) are underway. Together the availability of this data along with the various animal models of different disease and carriage states and the natural competence of the pneumococcus allowing its genetic manipulation make it a favourable organism to work with. It is therefore worthy of study not only in its own right as an important human pathogen but also as an accessible model for understanding host-pathogen interactions that may be applicable to other infections.

The availability of these genome sequences allows the *in silico* identification of putative virulence factors. Examination of the two completed genomes identified four pneumococcal homologues of sortase enzymes. At that time sortase had been shown to be a virulence factor in *Staphylococcus aureus*. It was therefore deemed likely these pneumococcal sortase homologues were novel virulence factors and were investigated further.

Sortase enzymes

Sortase (Srt) are bacterial enzymes responsible for the covalent attachment of specific proteins to the cell wall of Gram-positive bacteria (reviewed by Paterson and Mitchell 2004). Focussing on the processing and cell wall anchoring of protein A from *Staphylococcus aureus*, many of the major steps in this pathway have been elucidated (Mazmanian *et al.* 2001). Proteins targeted for sortase-mediated cell wall anchoring contain a number of features essential for their localisation. An amino terminal signal peptide directs these proteins to the secretory pathway with three critical features for cell wall anchoring found at their carboxyl terminal. These features, referred to collectively as the cell wall sorting signal are, moving towards the carboxyl terminus; a LPXTG motif (leucine, proline, X, threonine and glycine where X is any amino acid), a hydrophobic region and a tail of charged residues. During secretion the hydrophobic domain and charged residues impede membrane translocation allowing recognition of the LPXTG motif by the membrane associated sortase enzyme. In a two-step transpeptidation reaction sortase then cleaves the LPXTG motif between the threonine and glycine and covalently attaches the threonine to the amino group of pentaglycine cell wall crossbridge resulting in cell wall attached protein. Both cleavage and transpeptidation can be performed *in vitro* providing an assay in which sortase biology can be investigated (Ton-That *et al.* 1999; Ton-That *et al.* 2000). An outline of sortase-mediated cell wall anchoring is shown in Figure 1.
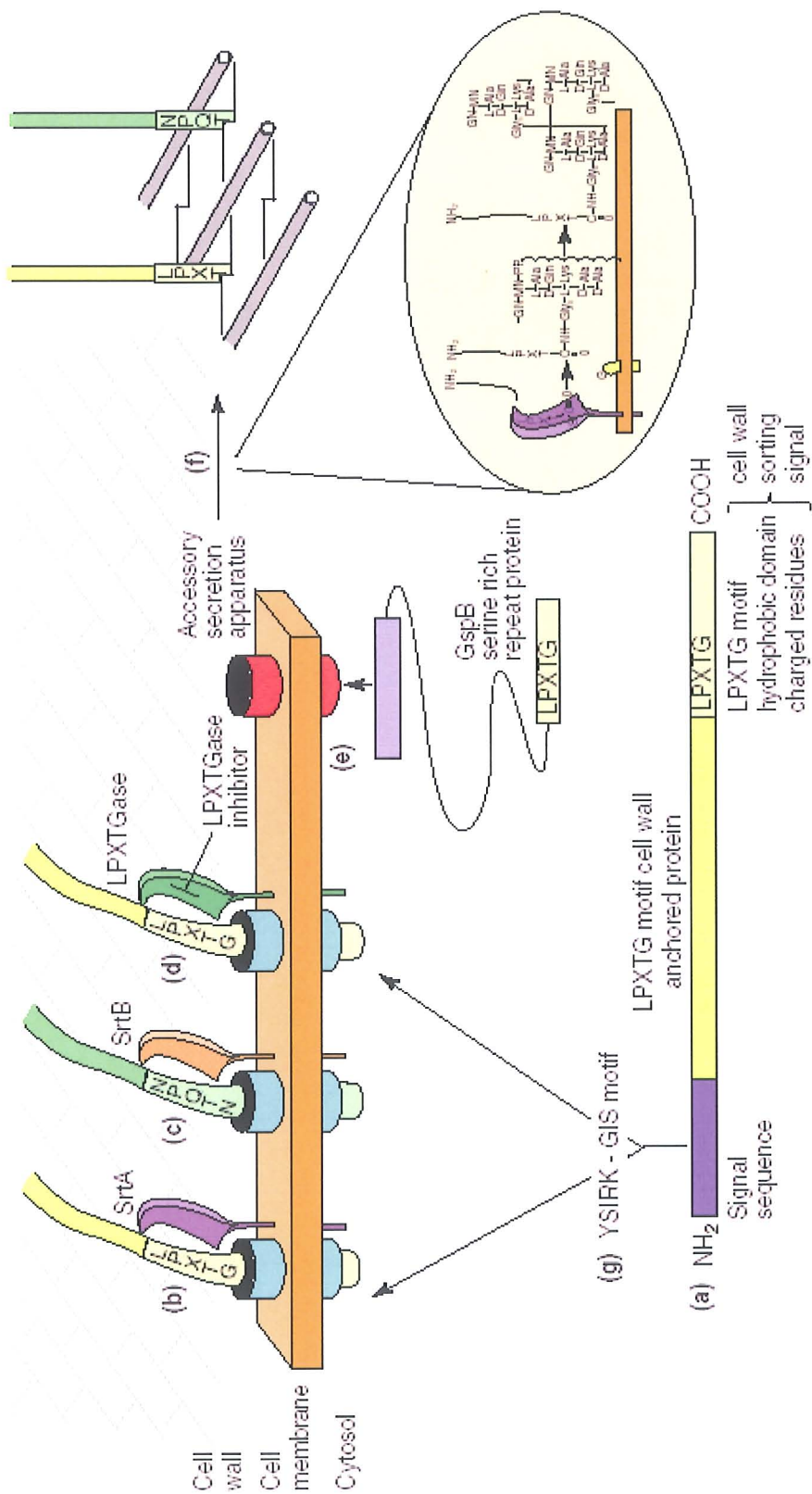
Figure 1: Summary of sortase-mediated cell wall anchoring pathway

Proteins for cell wall anchoring are secreted by the Sec pathway, b) during secretion the LPXTG motif is recognised by membrane associated SrtA. c) Additional sortase enzymes with different specificities are now recognised, in this case SrtB of *Staphylococcus aureus* identifies the NPQTN motif. d) The membrane associated enzyme LPXTGase also cleaves the LPXTG motif but with different specificity to sortase. This activity may increase cell wall anchoring. An inhibitor of LPXTGase is also present in the membrane. How these interact with each other and sortase to regulate cell wall anchoring is as yet unclear. e) Accessory secretion apparatus selectively transport specific substrates such as GspB of *Streptococcus gordonii* for cell wall anchoring. f) Following cleavage of the LPXTG motif, sortase covalently attaches the threonine of the LPXTG motif to the cell wall precursor lipidII. This is subsequently incorporated into the mature cell wall along with the cell wall anchored protein as shown in the insert. g) The YSIRK-G/S motif found in the signal sequences of some cell wall anchor proteins is needed for efficient secretion of these proteins. The factor(s) acting on this motif are unknown. Taken from Paterson and Mitchell (2004).

30

Biochemistry and structure

While searching for chemical inhibitors of sortase activity in *S. aureus*, Ton-That and Schneewind (Ton-That and Schneewind 1999) noted the importance of a sulphydryl group for sortase activity. The sulphur containing single cysteine residue of sortase at position 184 was identified as a potential active site residue acting as a nucleophile attacking the threonine-glycine bond of the LPXTG motif. The importance of this cysteine was confirmed by its substitution to alanine, which rendered the purified enzyme unable to cleave the LPXTG motif *in vitro* (Ton-That and Schneewind 1999) and unable to complement defective cell wall sorting in a sortase mutant *S. aureus* strain (Ton-That *et al.* 2002). This active site cysteine is found within a conserved TLXTC motif found in all confirmed and putative sortase enzymes (Ton-That *et al.* 1999).

Earlier work showing that LPXTG proteins are anchored to the cell wall crossbridge (Schneewind *et al.* 1995) suggested it was likely that proteins are attached to cell wall precursors rather than the mature cell wall, given the lack of free crossbridges in the latter. This was further supported by the work of Ton-That *et al.*, (Ton-That and Schneewind 1999) who suggested that lipid II was the probable substrate based on the affect on cell wall anchoring of various antibiotics. This has now been confirmed by the identification of the protein-lipid II intermediate in *S. aureus* (Perry *et al.* 2002) and the results of Ruzin *et al.* showing sortase-mediated addition of a LPXTG peptide to lipid II *in vitro* (Ruzin *et al.* 2002). However, the possibility that additional acceptor sites might exist, such as branching glycine residues in non-crosslinked peptidoglycan, has not yet been excluded (Ruzin *et al.* 2002).

To further understand sortase activity the NMR structure of a truncated recombinant version of SrtA from *S. aureus* was resolved by Ilangovan *et al* (Ilangovan *et al.* 2001a; Ilangovan *et al.* 2001b). This truncated version, lacking the amino terminal 59 amino acids (the membrane anchor), retained LPXTG peptide cleavage and transpeptidation activity *in vitro*. The results revealed a previously unseen β–barrel structure with eight β–strands aligned in an antiparallel and parallel fashion (Figure 2a). The active site was found within an elongated hydrophobic groove able to accommodate unfolded proteins, suggesting a scanning mechanism whereby sortase scans newly translocated proteins for the LPXTG motif (Figure 2b). Within the active site, a histidine residue (position 120) is seen in close proximity to the crucial cysteine suggesting the cysteine could be held in an active state through a thiolate-imidazolium ion pair interaction with the histidine, in a similar manner to the papain/cathepsin protein family (Figure 2c). A key role for this histidine was later confirmed, whereby its substitution to alanine abolished sortase activity *in vivo* and *in vitro* (Ton-That *et al.* 2002). Cysteine 184 and histidine 120 are thus both essential for SrtA activity. The conservation of these residues in all annotated sortase suggests a universal importance and mechanism (Ton-That *et al.* 1999). Interestingly, in the sortase NMR structure that was solved in the absence of substrate, the cysteine and histidine side chains do not interact. Thus it would appear that substrate binding induces conformational changes promoting interaction of the cysteine and histidine side chains leading to sortase activation. The structure indicates that only subtle rotations may be required for this interaction. This substrate-induced activation is suggested to provide a control mechanism to prevent spurious proteolysis (Ilangovan *et al.* 2001b). The existence of a thiolate-imidazolium ion pair however, appears very

doubtful in view of more recent evidence, with a general base catalysis the likely

catalysis mechanism (Connolly *et al.* 2003).
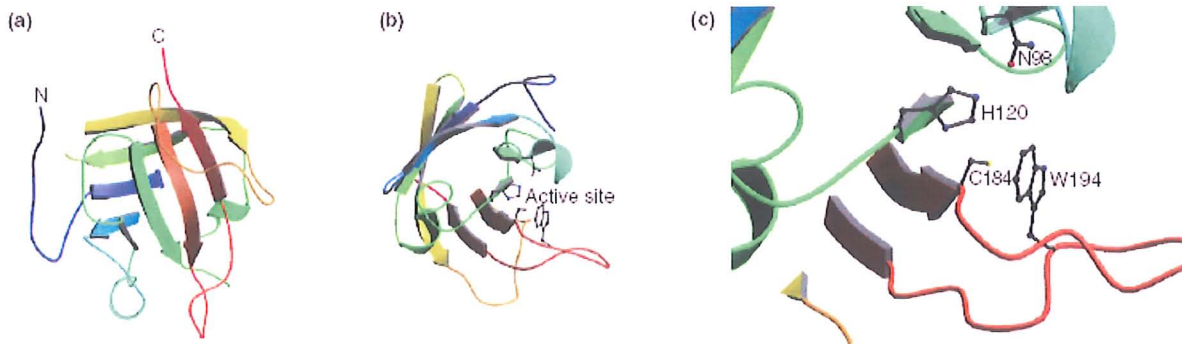
Figure 2: Structure of *Staphylococcus aureus* sortase A
Resolved by NMR SrtA forms a unique eight-stranded β-barrel fold. b) Rotated view
showing active site. c) Expanded view of active site showing the residues known or
suggested to be involved in activity. Cysteine 184 and histidine 120 are essential for sortase
activity and are conserved in all sortase proteins. Tryptophan 194 also contributes to activity
possibly through stabilising the thiolate-imidazolium ion pair originally believed to form
between the cysteine and histidine. Asparagine 98 may also do likewise but this has not been
tested experimentally. Neither of these latter two residues is absolutely conserved among
sortase proteins. However, it now appears that a thiolate-imidazolium ion pair does not form
between the cysteine and histidine but rather catalysis is through a general base mechanism.
Note that in the absence of substrate as shown here, the active site cysteine and histidine do
not appear to interact. Substrate binding is proposed to promote their interaction and sortase
activation. This may be a mechanism to prevent spurious proteolytic activity.
Numbered based on full length SrtA. Generated using Molscript, Bobscript and Raster3D
based on data generated by Ilangovan *et al.* (2000) from Pdb code 1IJA.

The structure revealed other residues potentially important to the sortase reaction. One

such residue investigated was a tryptophan at position 194 (Ton-That *et al.* 2002). A

mutant *srt*A with this residue substituted to alanine restored the anchoring defect of a

*srt*A null mutant *S. aureus* strain but processing was slower than with wild type sortase.

Additionally, *in vitro,* this purified protein cleaved a LPXTG peptide less efficiently than

wild type SrtA and performed transpeptidation with lower fidelity, failing to attach all

cleaved protein to the peptidoglycan substitute used (Ton-That *et al.* 2002). Interestingly this residue is not absolutely conserved in other sortase proteins (Ton-That *et al.* 2002). This might suggest subtle mechanistic differences between different sortase.

The addition of calcium ions increased the cleavage activity of purified *S. aureus* SrtA *in vitro* with smaller increases in activity seen with the addition of magnesium and manganese (Ilangovan *et al.* 2001b). Given that the concentration of calcium (2mM) which stimulated sortase was similar to the levels *S. aureus* might encounter in host tissues ($\geq$1.5mM) this may be a mechanism to increase surface protein anchoring during interaction with the host (Ilangovan *et al.* 2001b). Comparison of the NMR spectra of SrtA with and without calcium showed that ion binding perturbs the wall of the groove leading to the active site and so calcium may increase sortase activity by facilitating substrate binding (Ilangovan *et al.* 2001b).

Multiple sortase and substrate specificity

Although most extensively characterised in *S. aureus*, sortase and cell wall anchoring appear conserved among Gram-positive bacteria (Schneewind *et al.* 1995; Pallen *et al.* 2001). Invariably sortase homologues contain the active site cysteine and histidine shown to be essential for activity in *S. aureus* SrtA. Recently many of these sortase homologues have been shown to be functional. Gene knockout mutants in sortase homologues in *Listeria monocytogenes* (Bierne *et al.* 2002; Garandeau *et al.* 2002), *Streptococcus gordonii* (Bolken *et al.* 2001), *Streptococcus pyogenes* (Barnett and Scott 2002), *Streptococcus mutans* (Lee and Boran 2003) and *Streptococcus suis* (Osaki *et al.* 2002) as

with *S. aureus*, (Mazmanian *et al.* 2000; Mazmanian *et al.* 2002) all show defective surface expression of various LPXTG motif proteins.

Importantly not all Gram-positive bacteria have the same pentaglycine cell wall crossbridge as *S. aureus*, yet it appears anchoring may occur in the same way as it has been shown that in *Listeria monocytogenes* LPXTG motif proteins are attached to the meso-diaminopimelic acid found in its cell wall crossbridges (Dhar *et al.* 2000). Diversity in crossbridge structure may be a contributing factor to the low overall homology that sortase show as a family (Pallen *et al.* 2001). Additionally, some sortase proteins may be involved not in cell wall anchoring of proteins but in fimbrial assembly possibly through covalent linkage of the subunits via their LPXTG motif as suggested in *Actinomyces* species (Yeung *et al.* 1998; Li *et al.* 2001).

A striking feature of genome screening for sortase homologues is that typically more than one sortase is present (Pallen *et al.* 2001). The number of sortase varies among organisms with up to ten in *Corynebacterium effciens* (www.tigr.org). The role of these additional sortase is becoming clear, initially from work with *S. aureus* and *S. pyogenes*. In the case of *S. aureus* SrtB was identified and characterised by Mazmanian and colleagues (Mazmanian *et al.* 2002). Examining the cell wall anchoring of sixteen LPXTG motif proteins they found that a *srt*A mutant was defective in cell wall anchoring of them all, while a *srt*B knockout had no effect. Instead, SrtB recognised and anchored a protein, IsdC, with a cell wall anchor where the LPXTG motif is replaced by NPQTN. Purified SrtB cleaved a NPQTN peptide *in vitro*. This reaction was sensitive to a sulphydryl-modifying chemical suggesting that, like SrtA, the single cysteine residue of SrtB was essential for activity. The SrtB cleavage position was not identified, but as with SrtA, it

may be after the threonine. Thus in *S. aureus*, SrtA recognizes LPXTG motifs and SrtB a

NPQTN motif with the relationship specific and mutually exclusive (Mazmanian *et al.*

2002). In the case of *S. pyogenes* both Srt A and B recognized LPXTG motif proteins, but

anchored different proteins. Differential recognition was suggested to come from residues

immediately downstream of the LPXTG motif (Barnett and Scott 2002). In the case of

SrtA, all the four proteins that it anchors had an acidic residue immediately following the

LPXTG motif. The T6 protein anchored by SrtB however, had a serine in that position.

As with the two *S. aureus* sortase, a knockout in one *S. pyogenes* sortase had no affect on

the anchoring of proteins recognized by the other and vice versa. Additionally, the

majority of putative and known cell wall proteins in *S. pyogenes* have an acidic residue

following the LPXTG motif suggesting that as with *S. aureus* the majority of cell wall

proteins are anchored by SrtA. From this data it appears that a single sortase is

responsible for the anchoring of most cell wall proteins in an organism and that additional

sortase are responsible for the sorting of a limited subset of other proteins. This is

supported by the finding in *S. suis* that a *srt*A knockout was deficient in the cell wall

anchoring of fifteen proteins as assessed by 2-D electrophoresis (Osaki *et al.* 2002).

However, knockouts in the additional four sortase identified had no effect on these

proteins suggesting they had no role in the anchoring of SrtA substrates. The specific

substrates of these sortase were not detected, perhaps because they were present in

concentrations below the detection limit (Osaki *et al.* 2002). Finally, a *srt*A knock-out in

L. monocytogenes lacked nearly all proteins found in purified peptidoglycan preparations

suggesting SrtA anchors the majority of cell wall proteins in the species, with the

implication that the second *L. monocytogenes* sortase, SrtB is involved in the anchoring of only a small subset of cell wall proteins (Bierne *et al.* 2002).

Interestingly while *srt*A homologues appear ubiquitously within a species, there appears to be an inequality in the strain distribution of additional sortase in certain organisms. In the case of *S. pyogenes*, *srt*A was present in all twelve strains examined by Southern blot, but *srt*B was only evident in eight. Presumably the presence or absence of these additional sortase will correlate with the presence or absence of their specific substrates and may be a contributing factor to differences in phenotype such as virulence or tissue tropism.

Secretion of proteins for cell wall anchoring

Cell wall anchoring of proteins is dependent on the secretory pathway to transport proteins for anchoring out of the cytoplasm (Schneewind *et al.* 1992). Although protein secretion is not as well characterised in Gram-positive bacteria as it is in *Escherichia coli*, this export probably involves the Gram-positive homologues of the Sec pathway shown to be pivotal in Gram-negative protein secretion (Mazmanian *et al.* 2001).

However, some of the processes involved in Gram-positive protein secretion, are becoming clear and these include the identification of specialised export systems for specific cell wall proteins. In *Streptococcus gordonii* a locus encoding two accessory *sec* homologues, *sec*A2 and *sec*Y2 was identified and found to mediate the transport of GspB, a large serine rich repeat LPXTG cell wall anchored protein involved in adherence to platelets (Bensing and Sullam 2002). In contrast to the canonical *sec* homologues of *S. gordonii* and other organisms these accessory homologues were not essential for viability

or the export of most surface proteins. They appeared to perform the selective transport of GspB whose gene is found nearby secA2 and secY2 in the genome. In line with its specialised transport, GspB possesses an unusually long signal sequence, which presumably is responsible for its recognition by this alternative secretion pathway. Loci in *S. aureus* and *S. pneumoniae* show homology and comparable genomic organisation to the *S. gordonii* version and include the presence of an adjacent serine rich repeat cell wall anchored protein, therefore suggesting a similar function (Bensing and Sullam 2002). Genome sequencing of *S. pneumoniae* found that many LPXTG motif proteins contained an YSIRK-G/S motif in the signal sequence leading to the suggestion that this motif might be involved in the secretion of cell wall anchored proteins (Tettelin *et al.* 2001). The role for this motif in efficient secretion of *S. aureus* cell wall proteins has recently been confirmed (Bae and Schneewind 2003). Various deletions of the YSIR sequence or substitution of the G or S residues significantly impaired secretion of protein A in *S. aureus*. Cell wall anchoring of functional protein was not affected although the total amount of anchored protein was reduced. Interestingly not all cell wall anchor proteins of *S. aureus* contain this motif and so proteins containing the YSIRK-G/S motif may represent a subset upon which hitherto unknown factor(s) act (Bae and Schneewind 2003). Additionally, this motif is absent in some organisms such as *Listeria* and *Clostridia* while accessory *sec* homologues are not universal (Bensing and Sullam 2002; Bae and Schneewind 2003). Therefore although the general process of sortase mediated cell wall anchoring is conserved it is now becoming clear that much variation also exists in terms of the number of sortase homologues, the specificity of different sortase and additional specialised secretory factors.

Lee *et al.* have recently described a novel and unusual component of the cell wall anchoring pathway (Lee *et al.* 2002). Searching for LPXTG cleavage activity in *S. pyogenes* cell extracts they identified a protein in membrane fractions that specifically cleaves this motif. Named LPXTGase, this factor was distinct from sortase because unlike sortase, it was not sensitive to sulphydryl reagents and was inhibited by hydroxylamine, whereas sortase activity is stimulated (Ton-That *et al.* 1999). Additionally LPXTGase cleaved the LPXTG motif with different specificity to that of sortase. In contrast to sortase, which cleaves between the threonine and glycine, LPXTGase cleaved after the serine and glutamate of free LPSTGE peptide and after the glutamate of the surface attached peptide. Further highlighting its difference from sortase, LPXTGase is singular for a number of reasons. It is heavily glycosylated and contains seven different carbohydrates moieties that represent ~30% of its molecular mass. This glycosylation is essential for its activity (Lee *et al.* 2002). Amino acid composition analysis shows it contains an unusually high number of alanine residues with unknown/uncommon amino acids contributing ~30% of the protein backbone. It was speculated that LPXTGase is synthesised non-ribosomally on amino acid-activating polyenzyme templates. An inhibitor of LPXTGase was also identified in membrane extracts. How these two molecules interact with each other and with sortase in cell wall anchoring is unclear. Data discussed by the authors suggest that LPXTGase aids cell wall anchoring because addition of the inhibitor to *S. pyogenes* cultures reduced surface display of LPXTG motif proteins (Lee *et al.* 2002). More direct evidence on the role of LPXTGase in cell wall anchoring will come once the gene(s) responsible for its synthesis have been identified and can be inactivated. Given that many features of the cell wall

anchoring pathway are conserved it was suggested that LPXTGase and its inhibitor might

be found widely in Gram-positive bacteria. In support of this, both appear to be present in

S. aureus (Lee *et al.* 2002).

Virulence in animal models

Given the important role of many LPXTG motif proteins in infection (reviewed in

(Navarre and Schneewind 1999), it is of interest to investigate the potential effect of

sortase deletion on the virulence of Gram-positive pathogens. This has received much

attention recently with the comparison of sortase knockout strains with their parent in

animal models of infection. Investigated first in *S. aureus*, *srt*A mutants had impaired

virulence as assessed by bacterial load in renal abscesses following intravenous infection

and increased LD_{50} after intraperitoneal infection of mice (Mazmanian *et al.* 2000).

Interestingly, when a *S. aureus srt*B mutant was used in the renal abscess model it

showed no difference from the wild type early in infection (Mazmanian *et al.* 2002).

However, attenuation of the *srt*B mutant was evident later in infection. SrtB and the IsdC

protein it anchors are regulated by Fur boxes and are involved in iron uptake (Mazmanian

et al. 2002; Mazmanian *et al.* 2003) and so were proposed to be required for persistence

in host tissue but not initial colonisation. This suggests that this iron uptake system is not

required for initial stages of attachment and infection. In contrast SrtA may have been

important both early and later in infection because the adhesins and immune evasion

proteins it anchors would be required at all stages. Therefore although both *S. aureus*

sortase were virulence factors their effect on infection was distinct and likely to relate to

the function of the cell wall proteins they anchor.

Organism	Model	Effect of sortase knock-out compared to wild type	Reference(s)
Staphylococcus aureus	Renal abscess model	*srt*A and *srt*B mutants show reduced bacterial counts in kidneys	(Mazmanian *et al.* 2000; Mazmanian *et al.* 2002)
Staphylococcus aureus	Intraperitoneal infection	Increased LD_{50} for *srt*A mutant	(Mazmanian *et al.* 2000)
Staphylococcus aureus	Septic arthritis	*srt*A mutant attenuated, *srt*B mutant also attenuated but to much lesser degree	(Mazmanian *et al.* 2000; Jonsson *et al.* 2002; Jonsson *et al.* 2003)
Streptococcus mutans	Oral colonisation and dental caries	*srt*A mutant showed reduced colonisation and reduced severity of dental caries	(Lee and Boran 2003)
Listeria monocytogenes	Oral infection	*srt*A mutant showed reduced bacterial counts in infected tissues	(Bierne *et al.* 2002)
Listeria monocytogenes	Intravenous infection	*srt*A mutant displayed increased LD_{50} and reduced bacterial counts in infected tissues	(Garandeau *et al.* 2002)
Streptococcus gordonii	Oral colonisation	Reduced colonisation of *srt*A mutant	(Bolken *et al.* 2001)
Streptococcus pneumoniae	Intraperitoneal infection	*srt*A mutant had no effect on survival of infected mice	(Kharat and Tomasz 2003)
Streptococcus pneumoniae	Pneumonia infection	No effect for *srt*B or *srt*C mutants but *srt*D mutant attenuated	(Kharat and Tomasz 2003)
Streptococcus pneumoniae	Nasopharyngeal carriage	*srt*B mutant attenuated, no effect seen with *srt*D mutant but hypercolonisation displayed by *srt*C mutant	(Hava and Camilli 2002)
Streptococcus pneumoniae	Bacteraemia infection	No effect seen with *srt*B or *srt*D mutants	(Hava and Camilli 2002)

Table 1: Effect of sortase knockout in animals

Thus it appears the importance of a particular sortase in infection is determined by the function of the cell wall proteins it anchors with various sortase showing different contribution to virulence according to the type and stage of infection. In addition to the species discussed above, work has shown sortase mutants in *L. monocytogenes* (Bierne *et al.* 2002; Garandeau *et al.* 2002), *S. mutans* (Lee and Boran 2003), and *S. gordonii* (Bolken *et al.* 2001) to be attenuated in animal models in line with their defective surface display of LPXTG motif proteins known to be important in infection (Table 1). Since the being of this work data has also emerged on the role of pneumococcal sortase in virulence (Table 1, and discussed below).

Table 1 demonstrates the universal importance of sortase to Gram-positive pathogenesis. It is for this reason that sortase are regarded as potential drug targets to counter Gram-positive infections (Cossart and Jonquieres 2000; Mazmanian *et al.* 2000). In particular it is suggested that because this targeting will disrupt pathogenesis without affecting viability, the selective pressure towards drug resistance may be lessened (Mazmanian *et al.* 2000). In this regard the ubiquitous strain distribution and sequence/functional conservation of *srt*A in the species most extensively examined, namely *S. suis* (Osaki *et al.* 2003), *S. pyogenes* (Barnett and Scott 2002) and *S. aureus* (www.tigr.org) is a valuable feature in the exploitation of sortase as therapeutic targets. The search for sortase antagonists has already begun with the screening of medicinal plants for the ability to inhibit sortase activity *in vitro* (Kim *et al.* 2002). Although not widespread, sortase homologues have also be identified in a limited number of Gram-negative and Archea. Their role in these bacteria has yet to be investigated.

Pneumococcal sortase and cell wall proteins

The TIGR4 strain of *Streptococcus pneumoniae* has four sortase genes, named *srtA-D*, while R6 contains only *srtA* (Hava and Camilli 2003 and www.tigr.org) At the outset of this work no data was available on any pneumococcal sortase. However, *srtD* was subsequently identified in the signature-tagged mutagenesis screen of TIRG4 as being a virulence factor in pneumonia. As *srtD* is located in a putative pathogenicity island along with two other sortase genes (*srtB* and *C*) and three uncharacterised cell wall proteins this prompted the workers to examine the affect of individual gene knock-outs in this region (Hava and Camilli 2002). Re-inforcing the result of the screen, a *srtD* knock-out was

attenuated in a pneumonia model while no effect was seen with mutants in the *srtB* or *C*. In a colonisation model, the *srtB* mutant was attenuated, the *srtC* mutant showed hypercolonisation while the *srtD* mutant had no affect on colonisation. Mutants in *srtB* and *D* where then examined in a bacteraemia model where neither had an affect on virulence. These differences presumably relate to the function of the specific substrates anchored by these sortase, but further explanation awaits their identification and analysis. Interestingly each cell wall protein within the pathogenicity island varies slightly within the LPXTG motif (YPRTG, IPQTG and VPDTG). The significance of this variation is as yet unclear, but hints at the possible specific pairing of sortase with substrate within the pathogenicity island (Pallen *et al.* 2001). This pathogenicity island has since been named the *rlrA* pathogenicity islet because expression of the sortase and cell wall protein genes within it are regulated by a transcriptional regulator within the islet, *rlrA* (Hava *et al.* 2003). During the course of this work, pneumococcal *srtA* was investigated by Kharat and Tomasz (Kharat and Tomasz 2003) and was found to be required for the cell association of the LPXTG motif proteins β-galactosidase and neuraminidase A. In the absence of *srtA* these proteins were secreted into the supernatant rather than being associated with the cell pellet. Additionally, the mutant was defective in adherence to a human pharyngeal cell line presumably due to the defective surface localisation of cell wall proteins although the particular proteins responsible were not defined. In a study of virulence there was no difference in mouse survival following intraperitoneal infection with various doses of bacteria. Supporting the proposal of sortase specificity was the finding the mutants in *srtB* or *D* had no affect on the localisation of the SrtA substrate β-galactosidase (Kharat and Tomasz 2003).

An examination of sortase is incomplete without consideration of the proteins they anchor, as it is through these proteins that sortase influence bacterial biology. To identify likely cell wall anchored proteins the R6 and TIGR4 sequenced genomes (at TIGR Comprehensive Microbial Resource http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl) were analysied. Probable pneumococcal cell wall proteins were identified using the TIGRFAM facility, which employs the Hidden Markov Model to identify protein families based on conserved sequences and motifs. All pneumococcal proteins found to be members of TIGRFAM01167, representing an LPXTG-motif cell wall anchor domain were deemed likely to be putative cell wall anchored proteins see Table 2. Although a contribution to virulence has been demonstrated in animal models for several of these proteins their roles are poorly characterised. The best exemption is neuramidase A where it has been shown to contribute to colonisation nasopharynx and invade the middle ear in an otitis media model in chinchillas (Tong *et al.* 2000). This affect was by altering carbohydrate moieties in the tracheal epithelium to promote bacterial adherence (Tong *et al.* 2002).

Protein	R6 (Spr)	TIGR4 (Sp)	Demonstrated role in virulence	Reference(s)
β–galactosidase (β-gal)	0565	0648	Yes	(Hava and Camilli 2002; Robertson et al. 2002)
Neuraminidase A (Nan A)*	1536	1693	Yes	[Polissi et al. 199]; [ong et al. 2000; Hava and Camilli 200]; [aton et al. 1997]
Hypothetical protein	1345	1492	No	-
Hypothetical protein	1652	1833	No	-
Hypothetical protein	1806	1992	No	-
Alkaline pullulanase (SpuA)	0247	0268	Yes	(Hava and Camilli 2002)
Hyaluronidase (Hyl)	0286	0314	Yes	(Polissi et al. 1998) and (Hava and Camilli 2002) (Berry and Paton 2000; Chapuy-Regaud et al. 2003)
N-endo- β -N-acetylglucosamide	0440	0498	Yes	
Conserved Hypothetical protein	0075	0082	No	-
β-N-acetylglucosaminidase (StrH)	0057	0057	Yes	(Hava and Camilli 2002)
Hypothetical protein	1403	-	No	-
PrtA	0561	0641	Yes	(Bethe et al. 2001; Hava and Camilli 2002; Marra et al. 2002)
Conserved Hypothetical protein	0328	0368	No	-
RrgA (stands for RlrA-regulated gene)**	-	0463	Yes	(Hava and Camilli 2002)
RrgB	-	0464	No	-
RrgC	-	0465	No	-
serine rich repeat protein	-	1772	No	-
IgA1 protease***	1042	1154	Yes	[Polissi et al. 1998; Hava and Camilli 200]; [hiavolini et al. 2003]
Zinc metalloprotease B (ZmpB)***	0581	0664	Yes	Blue et al. Polissi et al [ava and Camilli 200]; [hiavolini et al. 2003]
Zinc metalloprotease C (ZmpC)***	-	0071	Yes	[Polissi et al. 199]; [ava and Camilli 200]; [hiavolini et al. 2003]

Table 2: Likely pneumococcal cell wall anchored proteins

Based on TIGRFAM facility at TIGR CMR using TIGRFAM01167. * Nan A in TIGR4 contains a deletion upstream of the LPXTG motif which likely results in the secretion of a truncated protein (Pericone et al. 2002) (Shakhnovich et al. 2002). ** Not included in TIGRFAM01167 but included here as contains likely cell wall sorting signal. *** Contain likely cell wall sorting signal but at the N-terminus unlike the other proteins and all characterised sortase substrates to date. Despite this, all were identified as belonging to TIGRFAM01167 and this was confirmed by manual checking of the sequence. For the other proteins a manual check of all sequences revealed a C-terminal LPXTG motif followed by hydrophobic domain supporting their inclusion in this family. Two exceptions were Sp0998 and Sp0931 (both excluded from the table) which were included in TIGRFAM01167 for only one strain despite having a weak score. Checking of their sequences determined they were unlikely to be cell wall anchored proteins based on the absence of a likely LPXTG motif and a C-terminal hydrophobic domain. All proteins in the table have a N-terminal hydrophobic domain likely to be involved in secretion. This domain was found upstream of the LPXTG motif in proteins with a N-terminal cell wall

anchoring sequence. The only exceptions are Spr1345 and 1806 both of which lack a N-terminal hydrophobic domain. Although only speculative this may suggest a distinct secretion mechanism for these proteins. Interestingly, both PavA and alpha-enolase are extracellular pneumococcal proteins but contain no apparent signal sequence and so a subset of secreted proteins may exist transported by a novel mechanism. (Holmes *et al.* 2001) (Bergmann *et al.* 2001). – denotes absence of homologue.

Pneumococcal surface protein C (PspC) is a highly polymorphic protein acting in adhesion, invasion and immune evasion (Iannelli *et al.* 2002). In the case of R6 and TIGR4 it is surface located by virtue of choline binding motifs that attach it to the cell surface via choline in teichoic acids. Sequence analysis of numerous strains suggests PspC is anchored to the cell wall by sortase in some strains, but is surface displayed in others by virtue of choline binding motifs. Furthermore, some strains posses two copies of PspC, one anchored to the cell wall and the other anchored to the choline of the cell wall (Iannelli *et al.* 2002). Thus the same protein can be surface anchored in different ways, although it is unclear what could be the functional significance of this.

Except for the polysaccharide capsule and the toxin pneumolysin most pneumococcal virulence factors are poorly characterised. In order to better understand pneumococcal pathogenesis we have used the sequenced genomes to identify putative virulence factors based on their homology to known factors. Sortase enzymes where identified as likely candidate virulence factors and were chosen for further investigation. Directly related to the investigation of sortase, several LPXTG motif proteins were also examined. Finally, two genes annotated as macrophage infectivity potentiator and exfoliative toxin A where also investigated given that these activities are important virulence factors for *Legionella*

pneumophila and *S. aureus* respectfully. The aim of this work was to investigate the role

in pneumococcal virulence of these putative virulence factors.

Chapter Two Materials and Methods

Composition of buffers and reagents are shown in the Appendix.

Bacterial cell culture

The pneumococcal strain D39, used throughout this work, is a virulent serotype 2 pneumococcus (MLST type 128) (Avery *et al.* 1944). It was received from the National Collection of Type Cultures and Pathogenic Fungi (NCTC) at the Central Public Health Laboratory, London. NCTC catalogue number 7466. R6 is an acapsular, avirulent derivative of D39 with improved transformation efficacy (Hoskins *et al.* 2001).

All *S. pneumoniae* strains were grown on Blood Agar Base, number 2 (BAB, Oxoid Ltd. Basingstoke) supplemented with 5 % defibrinated horse blood (E&O Laboratories, Bonnybridge, Scotland), or maintained in Brain Heart Infusion media (BHI, Oxoid Ltd., Basingstoke). Mutants were maintained with appropriate antibiotic selection. All cultures were incubated statically at 37 ° C under atmospheric conditions unless stated otherwise. Culture purity and optochin sensitivity were routinely confirmed throughout this work.

General confirmation of strains

All wild type strains used within this work were confirmed by colony morphology, production of α-haemolysis on blood agar and optochin sensitivity (5 μg optochin/ethylhydrocupreine hydrochloride, MAST Diagnostics). Multilocus sequence typing (MLST) was also performed to confirm strain identity. MLST analysis was performed at the Scottish Meningococcal and Pneumococcal Reference Laboratory (SMPRL), Stobhill Hospital, Glasgow, Scotland.

Glycerol stock preparation

Glycerol stocks of all *S. pneumoniae* strains were prepared by growing a culture from a single, colony to mid-exponential phase (OD_{600} nm 0.5-0.6). Glycerol was added to a final concentration of 12 % (v/v). 1 ml aliquots were prepared and frozen at - 70 °C in cryotubes. Viability and purity of stocks was determined 24 hr after freezing by viable counts (see below). Optochin sensitivity was also confirmed.

Transformation of *S. pneumoniae* (strains R6 and D39)

50 µl of thawed pneumococcal frozen glycerol stock was used to inoculate 30 ml BHI supplemented with 1 mM $CaCl_2$ (Sigma-Aldrich, UK). Cultures were grown to an OD_{600} nm of 0.1-0.3. For each transformation 1 ml of culture was added to a sterile Eppendorf tube. Competence-stimulating peptide-1 (CSP-1) (Havarstein *et al.* 1995) (Sigma Genosys, UK) was added to a final concentration of 100 ng ml^{-1}, and samples were incubated at 37 °C for 15 min. Donor DNA was added to a final concentration of 8-10 µg ml^{-1}. Total volume of DNA added to a maximum of 50 µl. For each transformation, a positive (plasmid pVA838 encoding erythromycin resistance) and a negative (no donor DNA) control were included. Samples were incubated for 75 min at 37 °C. Following incubation, aliquots of the transformation mixture were plated onto blood agar plates containing appropriate antibiotic selection. Plates were air-dried under a Bunsen burner and incubated overnight at 37 °C. Potential transformants were identified by their ability to grow on antibiotic-selective plates. Single colonies were sub-cultured into BHI from which glycerol stocks were prepared as described above.

Preparation of pneumococcal genomic DNA

A sweep of 4-5 bacterial colonies from a blood plate was used to inoculate 10 ml BHI. Following overnight incubation at 37 ° C, cultures were centrifuged for 15 min, 5,000 rpm at room temperature (RT) (Hereus Sepatech medifuge). The purity and optochin sensitivity of cultures for DNA preparations was always confirmed. The pellet was re-suspended in 1 ml extraction buffer and incubated for 1 hr at 37 ° C. Proteinase K (Sigma-Aldrich, UK) was added to a final concentration of 100 μg ml^{-1} and the mixture was incubated for 3 hr at 50 ° C. RNase (Sigma-Aldrich, UK) was added to a final concentration of 20 μg ml^{-1} and incubated at 37 ° C for 30 min. Resultant mixture was gently mixed with an equal volume of phenol/chloroform/isoamylalcohol (25:24:1 v/v, BDH & Fisher Scientific UK. Ltd.) and centrifuged at 13,000 rpm for 3 min (Sigma 113 microfuge). The aqueous layer was removed and added to 0.2 volumes of 10 M ammonium acetate (Sigma-Aldrich, UK). Ethanol (100 %) was added to precipitate the DNA and the mixture was centrifuged at 13,000 rpm for 30 min. The DNA pellet was left to air dry for 10 min, re-suspended in TE buffer, pH 7.4, and stored at 4 ° C.

Polymerase Chain Reaction (PCR)

PCR was optimised for each particular reaction. This optimisation was largely with regards annelling temperature. Approxiamtely 200-500 ng of DNA was used in each reaction. Typical reaction conditions were 95°C for 5 min then 30-35 cycles of 95°C for 30 sec, annealing at 50-65°C for 30 sec, extension at 72°C for1 min upwards (1 minutes for each 1Kb of expected product) with a final extension at 72°C for 5-10min. Taq polymerase (Promeaga) was used throughout. When a higher fidelity was required for

example the cloning of *srt*A, Vent polymease (New England Biolabs) was used. Primer

sequences are shown in the Appendix. When the presence or absence of genes were

investigated by PCR a genomic DNA preparation was made from sterile BHI and used in

each PCR to confirm there likely to be no cross-contamination of cultures and that

reagents were not contaminated. In addition, a positive control reaction was always

included with each PCR to ensure the reaction reagents (including the DNA preparation)

were working. Positive-displacement pipettes or filter tips were used throughout PCR

work to prevent contamination of reagents.

Molecular Biology

Restriction digest, ligations, sub-cloning and other common molecular biology methods

were performed using standard protocols (Sambrook *et al.* 1989).

Allelic Replacement of *srtA*

The strategy for allelic replacement of *srtA* is outlined in Figure 3.

Based on the R6 genome sequence, *srtA* and the surrounding region were amplified from

D39 genomic DNA with primers 19H and 19I and this product was cloned into pCR-4®-

TOPO (Invitrogen) following the manufacturers instructions. Inverse PCR was performed

on this clone with primers 19W and 19X. This deleted an internal portion of *srtA* and

introduced an AscI site within the pneumococcal sequences. This product was digested

with AscI and self-ligated and confirmed by sequencing. The AscI site was then used to

introduce an erythromycin resistance cassette (amplified previously in this laboratory

from the plasmid pGhost5) between the upstream and downstream pneumococcal

sequence. This plasmid was then used to transform D39 with selection for transformants on erythromycin (1µg/ml). Potential transformants were confirmed as the desired mutants using the primers 20o and 20N located immediately outwith the manipulated region. These primers gave an expected product of 1227 base pairs (bp) with wild type genomic DNA and a product of 1983bp with the *srtA* mutant allowing their differentiation.
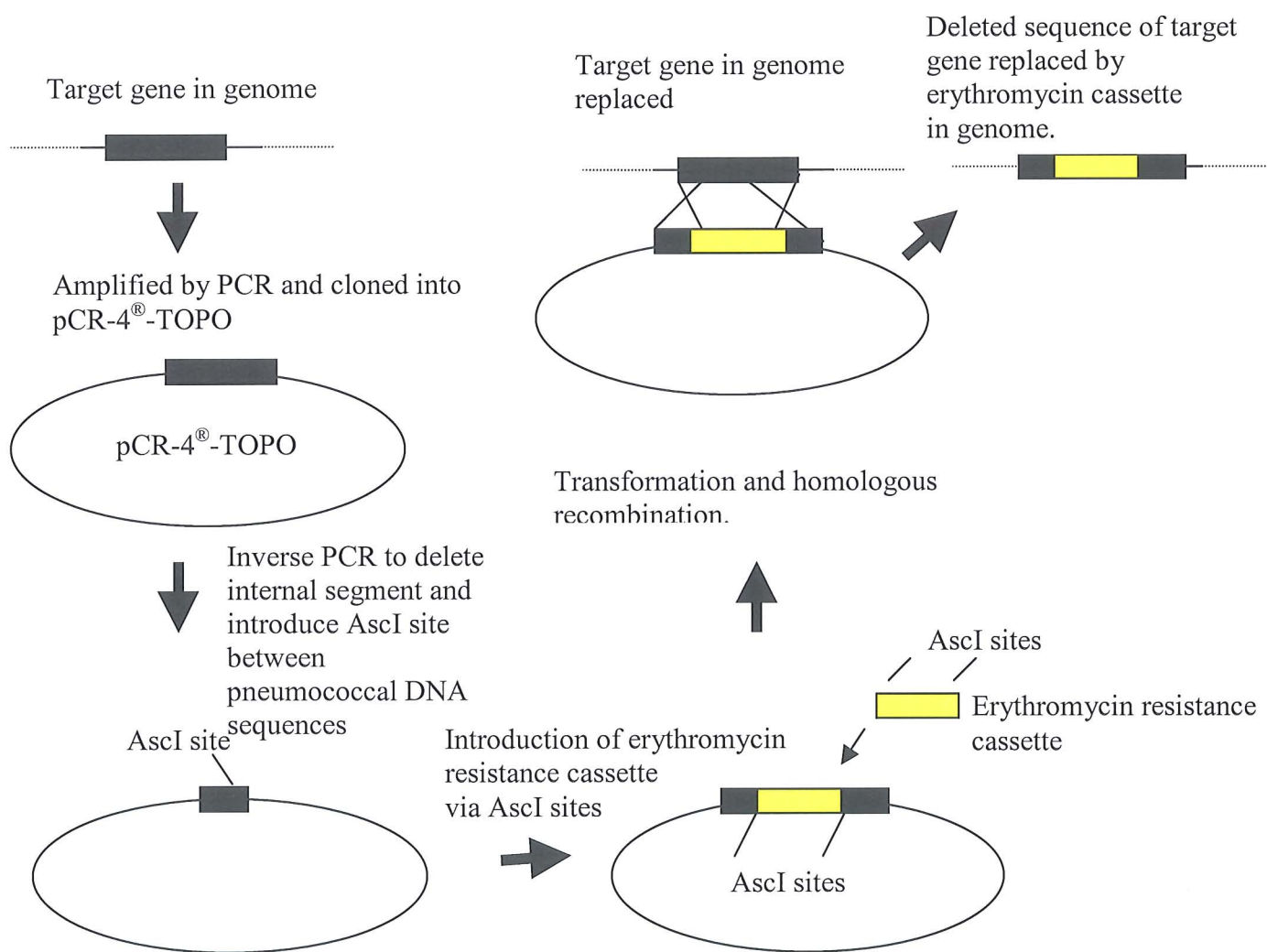
Figure 3: Strategy for allelic replacement of pneumococcal genes

Strain distribution of pneumococcal sortase genes

The strain distribution of the four known putative pneumococcal sortase genes among clinical isolates was examined by PCR. The clinical strains were isolated from invasive disease in Scotland and deposited at the Scottish Meningococcus and Pneumococcus Reference Laboratory (Stobhill Hospital, Glasgow, UK), (Table 5). Primers pairs for each gene were as follows; *srtA*, 25F and 25G, *srtB* 19Ia and 19J, *srtC* 19K and 19L and *srtD* 19M and 19A.

DNA sequencing

Sequencing of PCR products or plasmids was performed in-house (Molecular Biology Support Unit, DNA Sequencing Service, University of Glasgow, UK) with primers appropriate to the sample. Sequences were analysed using Edit View ASI Automated DNA Sequence Viewer (ABI Prism™) and DNASTAR software.

Cloning of *srtA*

srtA was amplified from D39 genomic DNA using the primers 19U and 19V, digested with the restriction enzymes BamHI and SacI (introduced by primers) and ligated into the expression plasmid pET-33b+ (Novagen). The construct was confirmed by sequencing. Primer design meant that SrtA would be expressed with a poly-histidine tag from the vector replacing residues 2-35. The absence of this section was intended to improve protein yield by removing the hydrophobic leader sequence and membrane anchor. This was performed previously for the purification of *S. aureus* SrtA and did not affect activity (Ton-That *et al.* 1999).

Allelic replacement of LPXTG protein genes

Due to the large size of some of the LPXTG motif protein genes targeted it was predicted that allelic replacement using the strategy employed for *srtA* might prove difficult with problems amplifying and cloning large PCR products. To bypass this potential difficulty an alternative method was used to generate constructs equivalent to that generated in *srtA* allelic replacement following inverse PCR (Figure 4). The regions for homologous recombination were amplified individually as the upstream and downstream regions and fused together by overlap PCR. This was then cloned into pCR-4®-TOPO and sequenced, thereby avoiding the need to amplify and clone the whole target region. After this point, mutagenesis continued as for *srtA* (Figure 2). Transformants were confirmed by PCR. Primers internal to the target gene were used to confirm its presence in wild type and absence in mutant strains. Primers located outside the manipulated region were not used on wild type genomic DNA as they would likely require much optimisation due to the large expected product. However, they were used to confirm the genotype of mutant strains because their expected product size was much smaller due to the deletion of an internal portion of the targeted gene. The primer pairs used for each of the four LPXTG motif protein genes targeted are shown in Table 3.

	Primer Pairs			
Gene	**Upstream**	**Downstream**	**Internal**	**Outside**
B-gal	27F and 27G	27P and 27I	35I and 35J	32M and 32N
nanA	27J and 25X	27K and 25Y	35E and 35F	31R and 31S
prtA	27C and 27Q	27E and 25Q	35C and 35D	32O and 32P
spuA	27L and 25L	27M and 25M	35G and 35H	32Q and 32R

Table 3: Primers used to construct and verify LPXTG mutants

Upstream region   Downstream region

```
         ╲                        ╲
  ────────┤IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII├────────
         ╲        Target gene        ╲
```

        ↓   Regions for homologous   ↓
   ████   Recombination amplified   ███
             by PCR

        ████
         │   These regions are fused
       AscI   together by overlap PCR and
       site   cloned into pCR-4®-TOPO

              ████
      ⟨   ⟩
          pCR-4®-TOPO
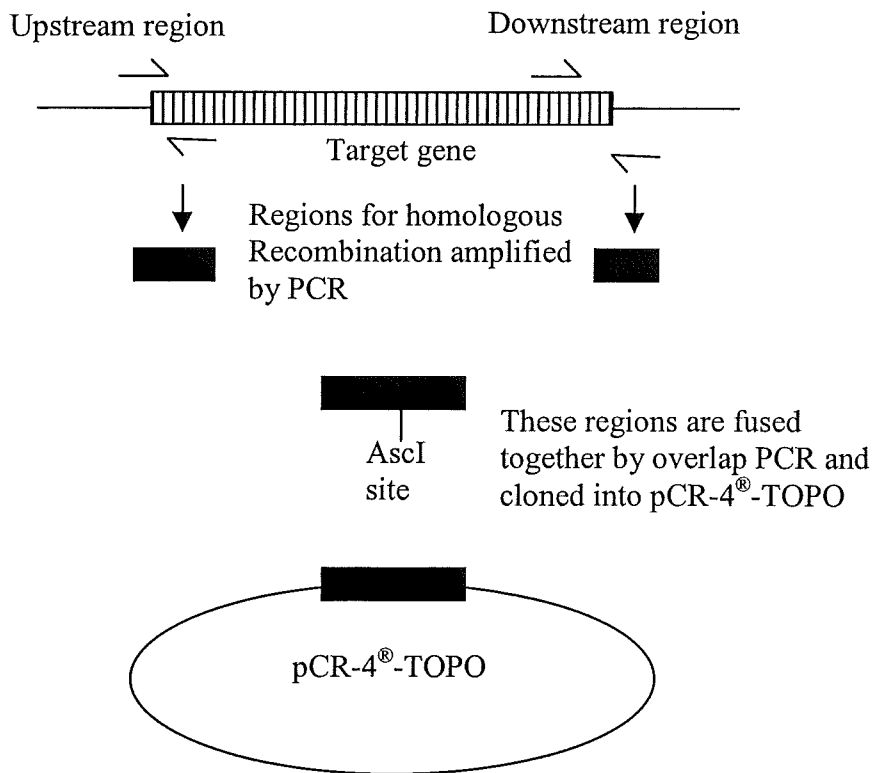
Figure 4: Schematic of strategy for allelic replacement of LPXTG motif protein genes
Due to the large size of the target genes involved the initial PCR and cloning as used in allelic
replacement of *srtA* might have proved problematic. Instead the two regions for homologous
recombination were amplified from the genome, named upstream and downstream in reference
to the positions relative to the target gene. These products were fused together in a second PCR
using the outermost primers from the initial reactions (relative to the target gene).
This fusion was aided by the introduction of complementary sequences via the primers used in
the original PCR. Also present due to primer design is the addition of an AscI between the two
pneumococcal sequences (upstream and downstream). This product is now cloned into pCR-
4®-TOPO confirmed by sequencing. The construct is then completed and mutagenesis
preformed as detailed in the latter stages of Figure 3.

Viable count of *S. pneumoniae*

To enumerate number of colony forming units (CFU) in frozen stocks, an aliquot of

culture was rapidly thawed in a 37 ° C water bath. 900 µl was removed and cells

collected by centrifugation for 3 min at 13,000 rpm. (Sigma 113 microfuge). Bacterial

pellets were re-suspended in 900 µl phosphate buffered saline (PBS). Serial 10-fold dilutions were carried out in a 96-well round-bottomed tissue culture plate (Nunclon™ surface, Nalge Nunc International™) to a final dilution of 10^{-6} in PBS. 6 x 20 µl of each dilution were plated onto blood agar and plates were incubated for at least 16 hr, 37 ° C. Following incubation, the dilution giving counts of 50-150 colonies per 20 µl spot were counted. From this the CFU ml^{-1} was calculated. This method is based on that of Miles and Misra (1938).

In vivo experiments

All *S. pneumoniae* strains used were mouse passaged (by intraperitoneal infection, see below) prior to use in the animal model, unless stated otherwise. This enhances virulence of the bacterial strains and is a technique commonly used in *S. pneumoniae* studies (Gingles *et al.* 2001; Tong *et al.* 2002; Blue and Mitchell 2003; Chiavolini *et al.* 2003).

Animal models

Adult female MF1 mice aged 9-13 wk and weighing 28-35 grams were used for the models of pneumococcal pneumonia, bacteraemia and colonisation. All mice were provided with pellet food (B&K Universal, North Humberside, England) and water *ad libitum* throughout the experiments. Mice were housed at the Central Research Facility, University of Glasgow, UK. All animal work was carried out under appropriate licensing and approval from the Home Office and the University of Glasgow.

Preparation of standard inoculum by intraperitoneal infection

Bacterial strains were re-suspended in sterile PBS to give approximately 1.0×10^{6} CFU ml^{-1}. 200 µl of this suspension was injected into the peritoneal cavity of MF1 mice using a 1 ml insulin syringe (Micro-fine, 12,7 mm, Becton Dickinson). 24 hr following injection, mice were sacrificed by cervical dislocation and the chest cavity opened. Blood was collected from the heart using a 23 gauge needle inserted into the right ventricle. 50 µl of recovered blood was inoculated into 10 ml BHI and grown overnight, statically at 37 ° C. Bacteria were harvested by centrifugation at 3,000 rpm. (Heraeus Sepatech medifuge) and re-suspended in 1 ml BHI. 100 µl of this suspension was used to inoculate 20 ml BHI containing 20 % (vol./vol.) heat inactivated foetal bovine serum (Gibco BRL, Life Technologies, U.K.) and then incubated at 37 ° C until mid-exponential phase was reached (OD_{600} nm 0.6 – 0.7). Culture was stored at – 70 ° C in 1 ml aliquots until required.

Intranasal challenge for pneumonia

900 µl of a frozen standard inoculum was rapidly thawed at 37 °C and centrifuged at 13,000 rpm. for 3 min (Sigma 113 microfuge). The pellet was re-suspended in 900 µl PBS and diluted to give 2.0×10^{7} CFU ml^{-1} (unless stated otherwise). Mice were lightly anaesthetised with 2.5 % (v/v) halothane (Zeneca Pharmaceuticals, Macclesfield, UK.) over oxygen (1.5 L min $^{-1}$) using a calibrated vapouriser. Confirmation of anaesthesia was determined when mice failed to produce a reflex reaction when pinched gently between the toes. 50 µl of bacterial suspension was administered, drop-wise, into the nares. Individual mice were thus given 1.0×10^{6} CFU. Mice were laid on their backs until they

recovered. Viable counts of the bacterial suspensions were determined prior to and immediately after challenge, by viable cell enumeration, as described to confirm the dose administered.

Intranasal challenge for colonisation

This was performed as for intranasal challenge for pneumonia with the exception that the dose of 1.0×10^6 CFU was administered in a 10 μl volume rather than 50 μl.

Intravenous challenge for bacteraemia

900 μl of standard inoculum was rapidly thawed and centrifuged at 13,000 rpm. for 3 min (Sigma 113 microfuge). The pellet was re-suspended in 900 μl PBS and diluted to give 2.0×10^6 CFU ml^{-1} (unless stated otherwise). Mice were placed in a ventilated, heated Perspex box for 5 min and restrained in appropriate apparatus. Tails were swabbed with ethanol and 100 μl of bacterial suspension was administered directly into the bloodstream via the lateral tail vein using a 1 ml insulin syringe (Micro-fine, 12,7 mm, Becton Dickinson). Blood was sampled from a different tail vein 1 min following challenge to ensure all mice were administered equal bacterial counts. At pre-determined time points, blood was sampled from the lateral tail vein as described above and used for viable cell enumeration, as described. Viable counts of the bacterial inoculum were determined prior to and immediately after challenge to confirm the dose administered.

Competitive infections

Challenges for competitive infections were performed as described above for intranasal, intravenous and colonisation infections with the exception that the dose was made of a 1:1 mix of two bacterial strains. Final dose administered was the same as challenges with a pure inoculum. For post-infection sampling dilutions of the sample were spread onto BAB plates and following overnight incubation, ~100 colonies were double picked onto BAB plates with and without antibiotic selection. Wild type colonies grew only without selection with the mutant able to grow on both. This gave the output ratio of mutant to wild type. The input ratio was determined in the same way and the competitive index calculated as by dividing the output ratio by the input ratio. Where no mutant colonies were recovered then the number one was used instead and CI expressed as less than (<) the value given. The ability of selective plates to permit mutant and retard wild type growth was confirmed before each experiment. It has previously been determined that insertion of an antibiotic cassette into strain D39 has no inherent affect on virulence (unpublished work in our laboratory).

Determining levels of bacteraemia

Blood was sampled at pre-determined intervals following challenge via removal of blood from peripheral veins. Mice were placed in a ventilated, heated Perspex box for 5 min and restrained in appropriate apparatus. Tails were swabbed with ethanol and a 1 ml insulin syringe (Micro-fine, 12,7 mm, Becton Dickinson) was used to remove a maximum of 100 µl blood from the lateral tail vein. Blood was immediately diluted 1:10 in PBS and stored on ice until viability counts were carried out, as described.

Determining bacterial loads in lung tissue

At pre-determined time points following intranasal challenge, mice were culled by cervical dislocation and the thoracic cavity opened up to expose the lungs. Lungs were aseptically removed and placed in 5 ml sterile PBS. Weight of the lungs was recorded. Lungs were homogenised for 15 sec using an electric hand-held tissue homogeniser set at high speed (IKA®-Ultra-Turrax® T25, Janke Kunkel Laboratories). Samples of homogenised lung tissue were used for viable cell enumeration as described.

Determining levels of colonisation

At pre-determined time points following intranasal challenge for colonisation mice were culled by cervical dislocation and the trachea exposed and clamped with Spencer-Wells forceps (Fisher Scientific Ltd. UK) to prevent flow of liquid into the lungs. A small incision was made in the trachea and the nasal cavity lavaged with 2 x 1ml PBS. Sutra was used to seal the month shut and fluid was collected as it passed out of the nostrils.

Murine survival and pain scoring

Mice were monitored frequently throughout the infection by a pain scoring system devised within this laboratory (Blue and Mitchell 2003). Symptoms progressed from normal to severe with mice showing starry coat/piloerection, hunched posture and lethargy. As infection proceeded to the terminal stages mice become moribund. Mice were sacrificed, by cervical dislocation, when showing signs of severe lethargy or upon becoming moribund. Where animals were actively culled, the time of reaching a moribund state was taken as the time of death ('mortality') / experimental end point for

survival data. Mice that survived the course of infection (7 days, unless stated otherwise) were assigned an arbitrary survival time of 168 hr for statistical analysis.

Localisation of LPXTG motif proteins

Preparation of samples

Wild type and *srtA* mutant D39 were both grown to mid-log in BHI at 37 °C. Cell pellet and supernatant were collected by centrifugation of the culture at 4500 rpm for 25 min at 4°C. The supernatant was collected and filtered (0.2 μm) to remove any remaining cells and concentrated ~180-fold using Amicon B15 concentrator (Millipore). Cell pellets were resuspended in lysis buffer and sonicated on ice to lyse the cells. Cell pellet samples were then concentrated ~2.5-fold using Amicon Ultra Centrafugal Filter Device 10000MWCO (Millipore).

Western Blot Analysis

Volume of sample loaded represented approximately 5×10^8 CFU as determined by viable count when the culture was taken for sampling. This represented the highest concetration of sample that could be loaded. Samples were mixed with 5x SDS-PAGE loading buffer and heated at 95 °C for 5 min before being loaded onto a precast SDS-PAGE gel (Cambrex Bio Science, Rockland ME USA) and electrophoresised for 2 hours at 125 volts. For transfer to nitrocellulose for Western blot transfer buffer was cooled to 4 ° C prior to use. The gel was placed in the cooled transfer buffer for 30 minutes and blotted onto Hybond™-C super membrane (Nitrocellulose, 0.45 μM, Amersham Life Science) with filter paper and fibre pads in a standard Western blot gel tank (BioRad Mini Protein

N™). The tank was filled with cold transfer buffer and the gel blotted at 30 V 80 mA overnight at 4 °C. The nitrocellulose membrane was subsequently blocked with Tris NaCl, pH 7.4, containing skimmed milk (3 % w/v, Marvel) for 1 hr at room temperature. Primary antibody (rabbit or mouse sera) was diluted 1:1000 in Tris NaCl, pH 7.4, containing 3 % skimmed milk, and incubated with the membrane with gentle shaking at 37 ° C for 3 hr. The membrane was washed several times in Tris NaCl, pH 7.4 containing 0.1% Tween 20, and transferred to Tris NaCl, pH 7.4 with 3 % skimmed milk and secondary antibody (anti-rabbit or anti-mouse, horseradish peroxidase-linked whole antibody, Amersham Biosciences) diluted 1:1000. Following incubation with gentle shaking at 37 ° C for 1 hr, the membrane was washed thoroughly in Tris NaCl, pH7.4 plus 0.1% Tween 20. Blots were visualised using ECL™ Western Blotting Detection Reagents (Amersham Biosciences) and developed on X-Mat UV film (Kodak).

Reverse transcriptase PCR (RT-PCR)

All glassware used for RNA / microarray work was washed thoroughly with diethyl pyrocarbonate (DEPC)-treated dH_2O prior to use and filter tips were used for pipetting all liquids during this work. Gloves were worn and frequently changed during RNA work.

Culture of bacterial strains for RNA isolation

Bacterial strains were grown in BHI, in the absence of antibiotic selection, to mid-log phase (OD $_{600 nm}$ 0.6-0.7) and 2 ml aliquots were removed and rapidly frozen in liquid nitrogen. This was to preserve the cellular RNA profile as quickly as possible. Frozen cultures were stored at - 70 ° C until required.

Isolation of bacterial RNA

Bacterial RNA was isolated using a commercially available kit (Rneasy® Mini Kit, Qiagen) with minor modifications to manufacturer's instructions. All reagents described in this section were provided with the kit, unless stated otherwise.

Aliquots of frozen bacterial cultures (described above) were removed and 8 ml sterile, RNase-free dH$_2$O was added to each. Once the bacterial stock had thawed in the dH$_2$O, suspensions were centrifuged at 5000 rpm for 10 min (Sigma 4K15 centrifuge). The supernatant was gently removed and tubes were inverted on a paper towel to remove all final traces of supernatant. 200 µl of freshly prepared TE buffer, pH 8.0 containing 15 mg ml^{-1} lysozyme (Sigma-Aldrich, UK) was used to re-suspend the bacterial pellet. The suspension was gently mixed by pipetting for 1 min. The resultant mixture was incubated at RT for 10 min with brief vortexing every 2 min. 700 µl of Buffer RLT containing β-mercaptoethanol (10 µl β-mercaptoethanol per 1 ml buffer, added immediately prior to use) was added to lyse the bacteria and mixed by vortexing vigorously. 500 µl of 100 % ethanol was added and mixed by pipetting. Solutions were applied to RNeasy® mini spin columns and centrifuged at 10,000 rpm for 15 sec (Sigma 113 microfuge). 350 µl wash Buffer RW1 was applied to each column and tubes were centrifuged at 10,000 rpm for 15 sec. The flow-through was discarded. DNA was removed by DNase treatment (70 µl Buffer RDD with 10 µl DnaseI stock per column Qiagen). DNase was applied directly onto the RNeasy® membrane and left to incubate at RT for 15 min. The membrane was washed by the addition of 350 µl Buffer RW1 followed by centrifugation of columns at 10,000 rpm for 15 sec. The RNeasy® columns were transferred to a new collection tube. 500 µl Buffer RPE was added and columns centrifuged at 10,000 rpm for 15 sec. The

flow-through was discarded. 500 µl Buffer RPE was added and columns centrifuged at 10,000 rpm for 2 min. The flow-through was discarded and columns were centrifuged at 10,000 rpm for an extra 1-minute. Columns were transferred to new RNase-free collection tubes and RNA was eluted in of 50 µl sterile, RNase-free dH$_2$O by centrifugation at 10,000 for 1 min.

cDNA synthesis

Reverse transcription reactions were carried out using the ThermoScript™ RT-PCR System (Invitrogen Life Technologies). This system utilises a reverse transcriptase that has high thermal stability and can increase specificity of reactions. Prior to use, RNA was treated with DNase (RQ1, Promega) according to manufacturer's specifications. cDNA synthesis using random hexamers was performed using the recommended Thermoscript ™ RT-PCR System protocol. Control reactions were performed in the absence of reverse transcriptase to confirm the absence of contaminating genomic DNA. Resulting cDNA was treated with 1 µl RNase H (*E. coli*, 2 U µl^{-1}, Invitrogen Life Technologies) at 37 ° C for 20 min, and stored at - 20 ° C until required.

PCR reactions were performed using 2 µl of the cDNA from the reverse transcription reaction described above. This was incorporated into a general PCR reaction as described using PCR conditions optimised to suit individual primer pairs. Negative control reactions were set up using cDNA synthesis reactions carried out in the absence of reverse transcriptase and PCR reactions where cDNA was substituted with dH$_2$O. A genomic DNA PCR was included as a positive control.

Susceptibility to killing to hydrogen peroxide

Wild type and mutant D39 were compared with regards hydrogen peroxide killing using the method of [ericone et al. 200]. Cells were growth in BHI at 37 °C to OD_{600} nm of 0.3-0.4 harvested by centrifugation 10 000g for 2 minutes at 4 °C. The cell pellet was resuspended in cold PBS, centrifuged again and resuspended in BHI. 100 μl of bacterial suspension was added in duplicate to wells of a round bottom 96-well plate containing 100 μl BHI plus different doses of hydrogen peroxide (Sigma-Aldrich, UK). Plates were incubated at 37°C for 30 minutes then viable counts performed on BHI agar plates containing 200U/ ml bovine liver catalase (Sigma-Aldrich, UK). Plates were incubated at 37°C for 18 hr in a candle jar to create a microaerobic environment. The percentage survival was calculated as a percentage relative to the addition of no hydrogen peroxide.

Statistical analysis

Statistical analysis was performed using Minitab™ Statistical Software (Minitab Inc.). Data from animal experiments using infection with pure cultures of strains were analysed by Mann-Whitney test. In all cases where no statistical difference was seen the data was re-evaluated using the less stringent student t-test and no difference was seen. Competitive index data was assessed by student t-test. Hydrogen peroxide killing was assessed by Mann-Whitney test. Unless stated otherwise data represent the geometric mean $\pm$ 1 standard error of the mean (SE). A P value of <0.05 was considered significant and statistically significant data is highlighted as * p<0.05 and ** p<0.025.

Chapter 3 Analysis of Pneumococcal Sortase, Results

Introduction

The important contribution of sortase enzymes to virulence in other organisms (reviewed by Paterson and Mitchell 2004) prompted our investigation of pneumococcal sortase homologues as candidate virulence factors.

Summary

Sortase are bacterial enzymes responsible for the anchoring of specific proteins to the cell wall. The strain distribution of the four known putative pneumococcal sortase genes was examined in a collection of clinical isolates. Unlike the other three sortase genes, *srtA* was found in all 57 strains examined. For this reason *srtA* was examined further in terms of; sequence conservation, role in pneumonia, bacteraemia, colonisation and localisation of cell wall proteins. In addition, SrtA has been purified allowing crystallisation trials to begin with the aim of providing structural data on this potential antimicrobial target.

The R6 genome has a single sortase gene annotated, *spr1098*. This gene is also present in the sequenced TIGR4 strain but is annotated as a conserved hypothetical protein, *sp1218*. TIGR4 has three additional annotated sortase genes *sp0466*, *sp0467* and *sp0468* all of which are absent in R6 and D39 (Tettelin *et al.* 2001). The basic features of the four

68

described pneumococcal sortase genes and their designation as *srt*A-D (Hava and Camilli 2002) are shown in Table 4.

Gene	*spr*1098[1]	*sp*0466[2]	*sp*0467[2]	*sp*0468[2]
Name	*sortase*A (*srt*A)	*sortase*B (*srt*B)	*sortase*C (*srt*C)	*sortase*D (*srt*D)
Present R6	Yes	No	No	No
Present TIGR4	Yes	Yes	Yes	Yes
Present D39	Yes	No	No	No
NCBI Accession	AAK99901	AAK74625	AAK74626	AAK74627
Coordinates	1095463 to 1094723	444857 to 445693	445791 to 446573	446563 to 447411
Gene Length (bp)	741	837	783	849
Protein Length (aa)	247	279	261	283
Molecular Weight (Da)	28132.94	31704.62	30091.57	31895.81
Percentage GC	32.66%	44.68%	46.10%	43.11

Table 4: Basic features of the four putative pneumococcal sortase genes and proteins
Identified in the sequenced genomes (strains R6 and TIGR4). Data from TIGR CMR.
[1]Data relates to R6 annotation. [2] Data relates to TIGR4 annotation.

Identification of sortase encoding genes

When this project begun there was no experimental evidence confirming the existence of pneumococcal sortase enzymes. To corroborate their annotation it was confirmed they all showed strong homology to experimentally proven sortase (data not shown). Importantly, this includes the active site cysteine and histidine residues, which are essential for the activity of SrtA from *S. aureus* (Ton-That *et al.* 1999; Ton-That *et al.* 2002). Additionally, using the TIGRFAM and Pfam Hits facility at the TIGR Comprehensive Microbial Resource http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl they all had strong positive matches to sortase families (data not shown).

Strain distribution of pneumococcal sortase genes

The strain distribution of the pneumococcal sortase genes *srt*B, C and D in the two sequenced strains (Table 4) showed that they are not present in all strains and prompted examination of sortase distribution in a larger population of strains. gDNA was purified from 57 clinical isolates isolated from invasive disease in Scotland and sortase gene content investigated by PCR.

Primers for *srt*B, C and D were initially tested against R6 and TIGR4 gDNA and found to produce the expected product with only the latter as expected based on genome sequencing. Primers for *srt*A failed to yield any product when used on gDNA from a *srt*A knock out (data not shown). Strains were selected to represent a variety of serotypes and sequence types as assessed by MLST. The strains examined and the distribution of pneumococcal sortase genes are summarised in Table 5.

Strain	MLST	Year	Specimen	Diagnosis	Serotype	srtA	srtB	srtC	srtD
3	227	01	blood	na	1	+	-	-	-
5	74	96	blood	na	2	+	-	-	-
11	180	01	ear	na	3	+	-	-	-
14	584	01	blood	pneumonia	4	+	+	+	+
15	246	01	blood	na	4	+	+	+	+
16	569	97	na	na	5	+	-	-	-
17	96	01	nasal	na	6	+	+	+	+
19	65	01	blood	na	6	+	-	-	-
20	65	01	blood	septicaemia	6	+	-	-	-
21	65	00	CSF	meningitis	6	+	-	-	-
22	581	00	blood	na	7	+	-	-	-
125	392	02	blood	pneumonia	7	+	-	-	-
25	577	01	blood	bacteraemia	8	+	-	-	-
26	578	01	blood	bacteraemia	8	+	-	-	-
29	156	00	eye	eye infection	9	+	+	+	+
30	156	00	tissue	na	9	+	+	+	+
31	156	00	ear	ear infection	9	+	+	+	+
34	NEW	01	blood	na	10	+	-	-	-
39	408	01	blood	na	11	+	-	-	-
41	62	00	serum	septicaemia	11	+	-	-	-
43	586	01	blood	bacteraemia	12	+	-	-	-
44	218	00	blood	pneumonia	12	+	-	-	-
45	218	00	blood	pneumonia	12	+	-	-	-
46	218	01	blood	bacteraemia	12	+	+	+	+
2968	na	01	na	na	13	+	-	-	-
47	70	01	blood	bacteraemia	13	+	+	+	+
119	574	96	na	na	14	+	-	-	-
48	124	00	blood	bacteraemia	14	+	-	-	-
49	156	00	ear	otitis	14	+	+	+	+

Strain	MLST	Year	Specimen	Diagnosis	Serotype	srtA	srtB	srtC	srtD
50	124	01	blood	pneumonia	14	+	-	-	-
51	NEW	00	blood	cellulitis	14	+	-	-	-
53	583	00	blood	bacteraemia	15	+	-	-	-
62	392	00	blood	pneumonia	17	+	-	-	-
63	392	00	blood	na	17	+	-	-	-
65	113	00	blood	bacteraemia	18	+	-	-	-
69	199	00	na	chest infection	19	+	-	-	-
70	96	00	serum	chest infection	19	+	+	+	+
71	587	01	blood	septicaemia	19	+	-	-	-
73	285	01	ear	na	19	+	-	-	-
74	567	00	blood	chest infection	20	+	-	-	-
75	235	00	blood	appendicitis	20	+	-	-	-
76	235	01	blood	pneumonia	20	+	-	-	-
79	433	01	blood	septicaemia	22	+	-	-	-
80	433	01	blood	pneumonia	22	+	-	-	-
84	573	01	ear	sudden death	23	+	-	-	-
90	571	01	blood	meningitis, fever	27	+	-	-	-
91	NEW	01	trachea	na	29	+	+	+	+
92	696	01	blood	pneumonia	31	+	-	-	-
115	444	01	blood	na	31	+	-	-	-
98	60	00	blood	pneumonia	33	+	-	-	-
99	100	01	blood	na	33	+	-	-	-
100	582	00	eye	na	35	+	-	-	-
101	NEW	01	pleural	na	35	+	+	+	+
102	572	01	hip aspirate	na	35	+	-	-	-
103	593	01	sputum	pneumonia	35	+	+	+	+
104	589	00	blood	chest infection	41	+	-	-	-
3653	na	00	na	na	NT	+	-	-	-

Table 5: Strain distribution of pneumococcal sortase genes in Scottish invasive disease isolates

Strain refers to laboratory reference number, year is year of isolation. + and – indicate presence or absence respectively of each gene. NT, non-typable, na data not available, new is new sequence profile yet to be assigned reference number.

srtA was found in all 57 strains examined, while srtB, C and D were present in only 13

(22.8%). These latter three sortase showed a mutually exclusive pattern whereby they

were either all present or all absent. Partial sequencing of srt B, C and D PCR products

confirmed their identity. This was seen as important because they share a high degree of

similarity   with   each   other.   Using   the   NCBI   facility

(http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) a BlastN search for these sortase

genes was performed against the unfinished pneumococcal genomes from strains 23F-1

and 670-6B (performed 2/10/03). srtA was present in both strains while srtB, C and D

were all absent in 23F-1 but all present in strain 670-6B.

Investigation of srtA

srtA, unlike the other sortase genes, was present in all pneumococcal strains tested and so

was chosen for further study as the results may be of relevance to the greatest number of

pneumococcal strains. In addition this meant srtA was present in the routinely used D39

strain that is well characterised in the mouse infection model to be used.

Sequence diversity of srtA

The published sequence of srtA in R6 and TIGR4 is highly conserved and was confirmed

by independent sequencing in this work. However, it must be noted that the TIGR4

annotation is truncated (123 amino acids in length) compared to that of R6 (247 amino acids). Despite this difference, the 'missing' sequence at the amino terminus of the TIGR4 protein is present in its genome based on examination of the published sequence and sequencing in this work (Figure 5 and 6). This disparity therefore appears due to differences in the recognition and annotation of the genome into genes with the designation of different ATG start codons. It is perhaps most likely that the R6 annotation is correct given it results in a protein similar in size to other sortase and includes a hydrophobic N-terminus seen in sortase enzymes which is involved in membrane association (Ton-That *et al.* 1999) (Figure 7). Western blot analysis if SrtA antisera were available could be used to confirm the size its size and therefore the likely start codon. To investigate if this sequence conservation extended to a larger range of strains, *srt*A was sequenced from a panel of twenty clinical isolates from Table 5. Strains were selected to represent diverse serotypes and sequence types.

*srt*A was amplified from these strains using primers 20N and 20o. These primers are sited in the upstream (formate-nitrate transporter) and downstream (DNA gyrase subunit A) genes. All strains yielded the expected product showing that the position of *srt*A is conserved relative to these genes (data not shown). From this PCR product *srt*A was sequenced using the primers 23U, 24D, 24E and 24F. Sequencing showed that *srt*A was highly conserved among all strains examined (Figure 8 and 9).

Figure 5: Nucleotide alignment of *srt*A from R6 and TIGR4
Name on right refers to strain. TIGR4 is the annotated version from that genome while TIGR4′ represents the gene sequence from TIGR4 based on the R6 annotation. Numbers refer to position relativeto gene start. Performed using DNAstar Megalign using ClustalW (Identitiy). Highlighted bases mismatch with R6 sequence.

```
1                                                                                          TIGR4
1    M S Q K N N K K K N K R K N L L T N I L A G F L I L L S L A L I F N T Q I R N I F I V W N T N K Y Q   R6
1    M S Q K N N K K K N K R K N L L T N I L A G F L I L L S L A L I F N T Q I R N I F I V W N T N K Y Q   TIGR4'

1                                                                                          TIGR4
51   V S Q V S K E K L E E N Q D T E G N F D F D S V K A I S S E A V L T S Q W D A Q K L P V I G G I A I   R6
51   V S Q V S K E K L E E N Q D T E G N F D F D S V K A I S S E A V L T S Q W D A Q K L P V I G G I A I   TIGR4'

1                                                        M K R E Q V M G E G N Y S L A S H H I F G V D N A N   TIGR4
101  P E L E M N L P I F K G L D N V N L F Y G A G T M K R E Q V M G E G N Y S L A S H H I F G V D N A N   R6
101  P E L E M N L P I F K G L D N V N L F Y G A G T M K R E Q V M G E G N Y S L A S H H I F G V D N A N   TIGR4'

27   K M L F S P L D N A K N G M K I Y L T D K N K V Y T Y E I R E V K R V T P D R V D E V D D R D G V N   TIGR4
151  K M L F S P L D N A K N G M K I Y L T D K N K V Y T Y E I R E V K R V T P D R V D E V D D R D G V N   R6
151  K M L F S P L D N A K N G M K I Y L T D K N K V Y T Y E I R E V K R V T P D R V D E V D D R D G V N   TIGR4'

77   E I T L V T C E D L A A T E R I I V K G D L K E T K D Y S Q T S [N] E I L T A F N Q P Y K Q F Y        TIGR4
201  E I T L V T C E D L A A T E R I I V K G D L K E T K D Y S Q T S D E I L T A F N Q P Y K Q F Y         R6
201  E I T L V T C E D L A A T E R I I V K G D L K E T K D Y S Q T S [N] E I L T A F N Q P Y K Q F Y        TIGR4'
```

Figure 6: Amino acid alignment of SrtA from R6 and TIGR4
Name on right refers to strain. TIGR4 is the annotated version from that genome while TIGR4′ represents the sequence from TIGR4 genome based on the R6 annotation. Numbers on left refer to position relative to amino terminal. Performed using DNAstar Megalign using ClustalW. Highlighted residues mismatch with R6 sequence.
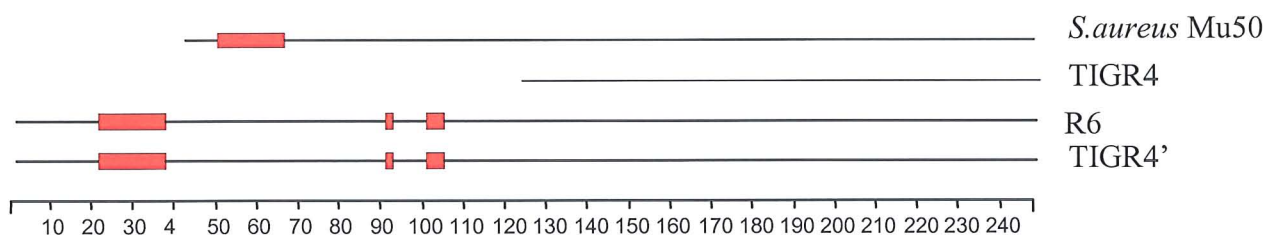
Figure 7: Hydropathy analysis of SrtA using Goldman-Engleman-Steitz
Performed using DNAStar Protean program with default window settings of 20. Scale bar indicates residue position relative to amino terminus. The Goldman-Engleman-Steitz plot shows predicted transmembrane domains indicated in red. TIGR4' represents the sequence from TIGR4 genome based on the R6 annotation.

Figure 8: Alignment of nucleotide sequence of *srt*A

Data from this work from various pneumococcal strains. Figures on left relate to nucleotide position from gene start. Strain shown on right. Bases in disagreement from consensus (R6) boxed. Performed using DNAstar Megalign using ClustalV (Identity)

Figure 9: Alignment of deduced amino acid sequence of SrtA
From this work from various pneumococcal strains. Figures on left relate to residue position from amino terminal. Strain name given on right. Only residues in disagreement with consensus are shown. Active site motif underlined in red.

Mutant construction and verification

To investigate the role(s) of *srt*A in pneumococcal biology a null mutant was made by allelic replacement in D39 as described in the Material and Materials (Figure 3). Potential transformants were confirmed by PCR using the primers 20o and 20N, located outside the manipulated region. These primers give different band sizes for wild type and *srtA* mutant (Figure 10). In additional confirmation, the PCR product from the mutant was sequenced and matched the expected sequence (data not shown). Transformants were readily obtaining showing that *srt*A was not essential for viability.
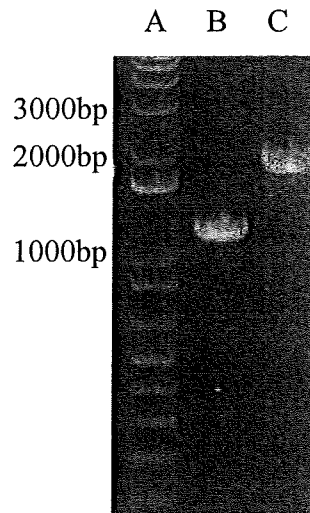
Figure 10: PCR confirmation of *srt*A knock out
Primers 20N and 20o located upstream and downstream of the manipulated region were used to confirm the allelic replacement of *srt*A. Lane A ladder marker, B wild type gDNA and C *srt*A knock out gDNA. The expected product sizes are wild type 1227bp and mutant 1983bp.

Analysis of virulence

The outbred MF-1 strain of mouse is widely employed as an animal model to study the pneumococcus and has been used to identify several virulence factors (Yesilkaya *et al.* 2000; Blue *et al.* 2003; Chiavolini *et al.* 2003). Wild type and *srtA* mutant D39 were therefore compared in this model to assess the contribution of *srtA* to pneumococcal virulence.

Following intranasal infection with 10^6 CFU survival time, bacterial blood (24 and 36 hr post infection) and lung counts (36 hr) were measured (Figure 11). No significant difference was seen between the two strains for any of the parameters measured.
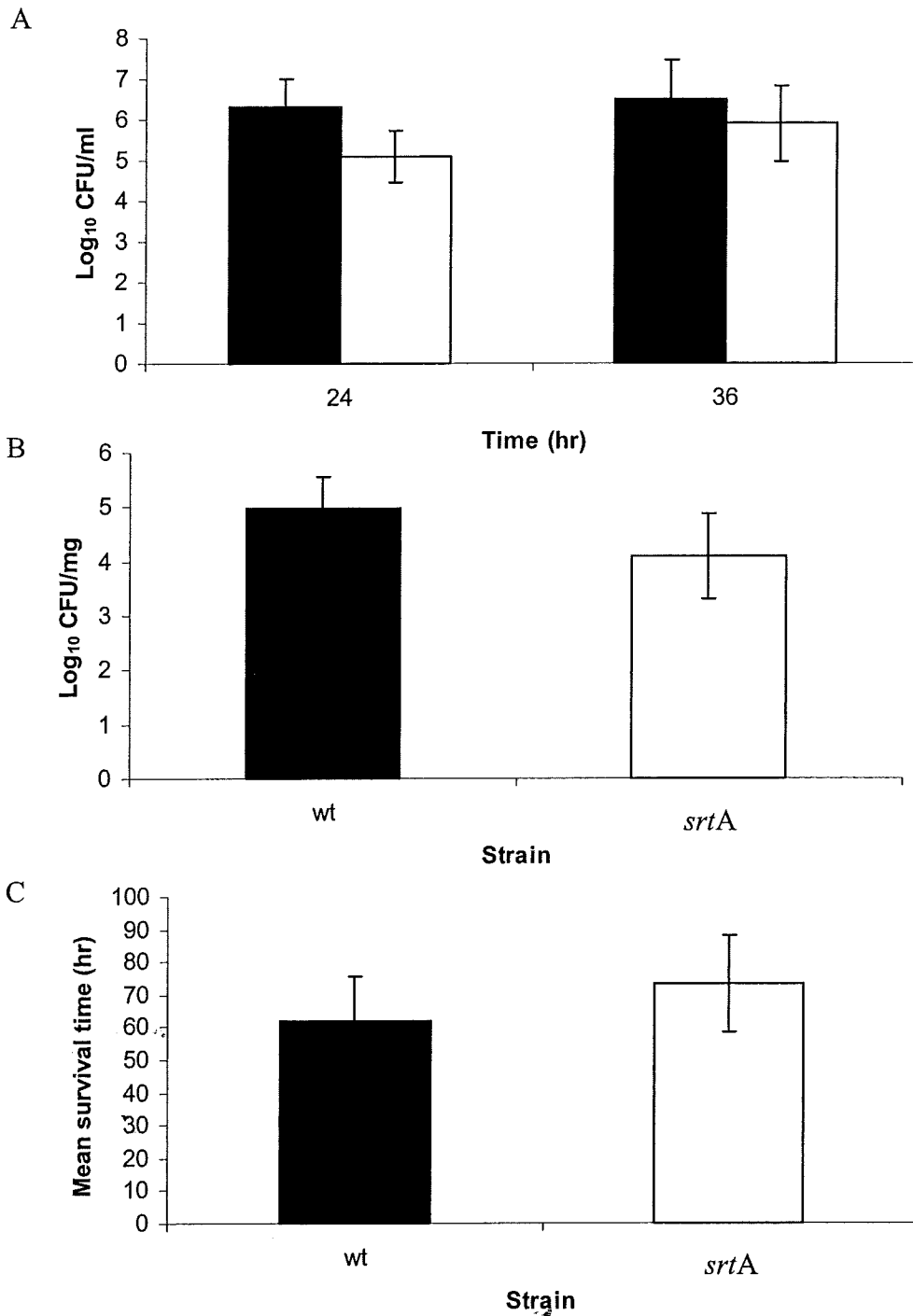
Figure 11: Comparison of virulence of D39 wt (black) and *srt*A knock out (white)
Following intranasal infection with 10^6 CFU strains were compared for A) blood and
B) lung (36 hr) bacterial counts and C) survival time. Data represent mean +/- 1SE
from 6-11 mice.

Comparison of virulence using competitive infections

Although there was no statistical difference between the virulence of wild type and *srt*A mutant there was a trend suggestive of slight attenuation of the mutant. For example, blood and lung counts were lower and mean survival slightly longer in mice infected with the *srtA* mutant D39 in comparison to wild type. This suggested there might be a subtle attenuation of the *srt*A mutant. To investigate this, competitive infections were used, which are sensitive to small differences in virulence. In this infection a 1:1 mixture of wild type and mutant are administered together and the ratio of mutant to wild type calculated from bacteria recovered following infection (output ratio). This ratio is then corrected by the input ratio of mutant to wild type to give the competitive index (CI) (Equation 1).

Equation 1

CI= <u>output ratio of mutant to wild type</u>
        input ratio mutant to wild type

The CI indicates if one strain has out-competed the other and is therefore presumably more virulent because it is better able to survive in the infection system. Interpretation of the possible CI results is shown in Table 6.

Competitive index	1	<1	>1
Interpretation	No difference in virulence	Wild type has out-competed mutant and so mutant is attenuated compared to wild type	Mutant has out-competed wild type and so mutant more is virulent than wild type

Table 6: Interpretation of CI

It was intended to differentiate wild type and *srtA* mutant bacteria in competitive infections by virtue of the erythromycin resistance of the mutant. However, following growth without erythromycin *in vitro* and *in vivo* (mouse infection) the *srt*A and *spr0370-0371* double mutant (Chapter 7 and 8) had lower counts on non-selective plates compared to those containing erythromycin. It therefore appeared that the mutations might be unstable and the erythromycin resistance cassette was being lost in the absence of selection. Although this might not affect any phenotype seen with the mutants (because they had a portion of the target gene deleted and therefore cassette lost would not restore the wild type genotype) it did have implications for the use of antibiotic resistance to identify strains in a competitive infection. When 12 *srtA* mutant colonies grown without selection were examined by the PCR used to confirm mutants (Figure 10) they all appeared to contain the erythromycin cassette. This was unexpected as based on the viable counts on non-selective versus selective plate only 1 in ~500 colonies were erythromycin resistant. This suggested the mutant genotype was stable but that erythromycin resistance was not being expressed in the absence of the antibiotic.

Furthermore, when mutants were grown on BAB plates without selection and subsequently stabbed picked onto plates with and without selection all colonies grew confirming that the contained the erythromycin cassette. Presumably the large inoculum used here (all the cells in a single colony) mean that at least some cells survive to form a colony either because they express resistance or are able to express it before being killed. In this way the apparent loss of resistance expression is masked. Such an affect will not be seen when colonies are grown from single cells explaining the original observation of resistance loss. It appears that the mutants are stable but the erythromycin resistance phenotype is not. Therefore for competitive infections bacteria were initially plated without selection then subsequently double picked onto plates with and without erythromycin in order to calculate the ratio of mutant to wild type bacteria.

Using this infection system D39 wild type and srtA mutant were compared following intranasal (pneumonia) and intravenous (bacteraemia) infections (Figure 12). At 0 hr post infection, following both intranasal and intravenous challenges the mean CI was not significantly different from 1. This control shows that the preparation and infection process themselves had no effect on the ratio of wild type and mutant. However, at all subsequent time points, in both infections, the CI was significantly below 1 indicating attenuation of the srtA mutant. When comparing between the infection types the CI was significantly lower following intravenous infection at 6 and 12 hours than after intranasal infection. This difference was not evident at 24 and 36 hr and suggests that attenuation of the mutant was greater in intravenous infection, at least at early time points.
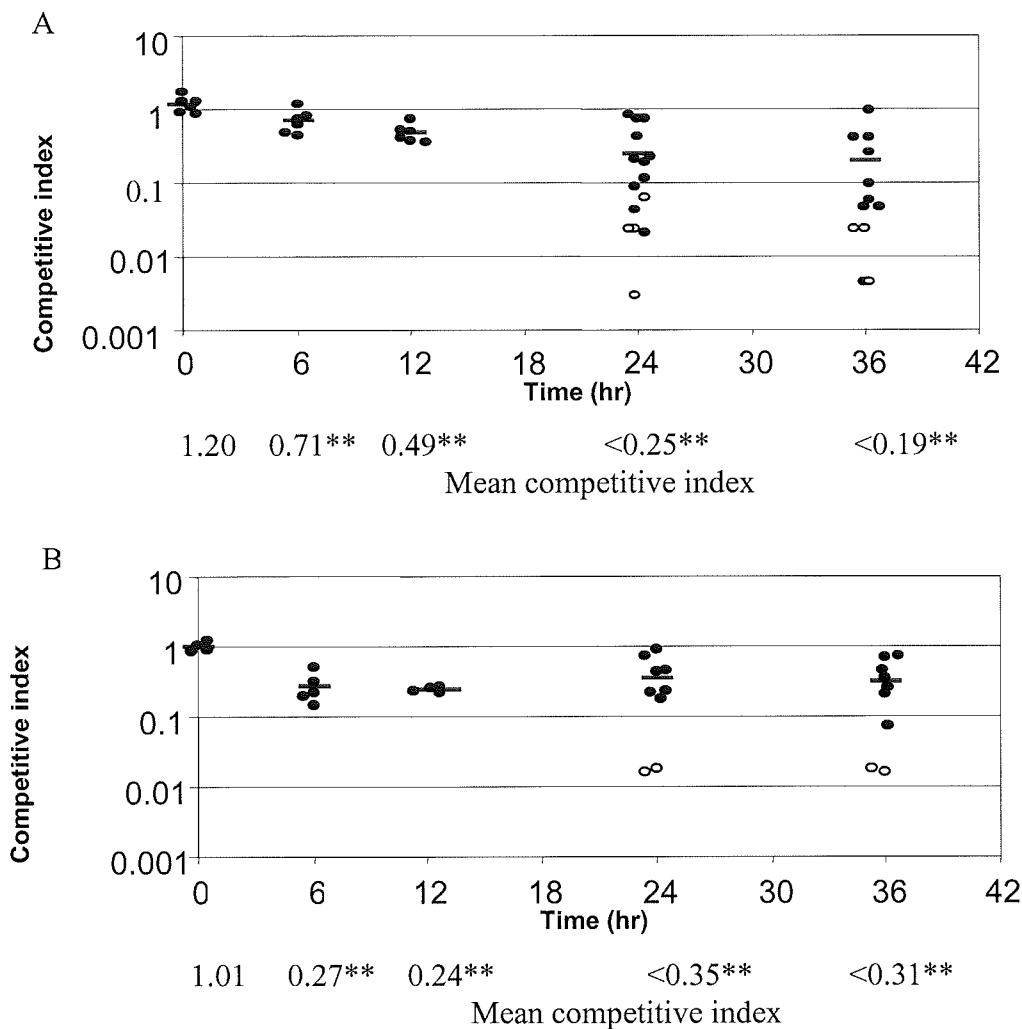
Figure 12: Virulence analysis of *srt*A mutant in competitive infections with parental wild type

Each point indicates the competitive index from an individual mouse following intranasal (A) and intravenous (B) challenges. Values <1 indicates attenuation of the mutant. Following intranasal challenge, bacteria were sampled from the lung. Following intravenous challenge bacteria were sampled from the blood. Red bar shows the mean CI and open circles indicate the recovery of no mutant bacteria. *p<0.05 **p<0.025 compared to one (one equals no attenuation).

Verification of competitive infection

In order to check that the genotype ascribed to colonies in competitive infections on the basis of antibiotic resistance was correct, it was verified by PCR genotyping of colonies. Genomic DNA was extracted from a total of eighty colonies (40 wild type and 40 mutant according to erythromycin resistance) from various time points from the competitive

infections. The *srt*A genotype (wild type versus *srt*A mutant) was confirmed by PCR with the primers 20o and 20N. These primers were used to confirm the original mutant based on the different product sizes (Figure 10). In all eighty cases the colony genotype as determined by PCR agreed with that based on erythromycin resistance. Thus the method of ascribing genotype based on erythromycin resistance appears to be accurate.

Comparison of colonisation using competitive infections

Having shown that the *srt*A mutant is attenuated in both pneumonia and bacteraemia infection the behaviour of the strain in a colonisation model was investigated. In this model MF-1 mice are infected with 10^6 CFU administered in a smaller volume of 10 µL as opposed to the 50 µL used in the pneumonia model. This results in nasal carriage but not invasive disease. Wild type and *srt*A were compared in this model using the competitive infection method used above. At both time points, 4 and 8 days post infection the CI was significantly less than one showing the *srt*A mutant to be attenuated with regards nasal colonisation (Figure 13). Although the attenuation seen at day 8 was more significant compared to 1 than that seen at day 4, comparison of the CI at the two time points showed no significant difference. In line with this being a model of asymptomatic colonisation and not invasive disease mice showed no clinical signs of infection throughout and bacterial blood counts obtained by tail bleeds at 24, 48 and 72 hr post infection were all below the detection limit (50 CFU/mL), in contrast to what occurs following invasive disease seen with intranasal infection using 10^6 CFU given in 50 µL. A 0 hr time point was not performed for the colonisation experiment as bacteria were

prepared in the same way as for the pneumonia model. The 0 hr CI therefore likely is the

same as for that infection (Figure 12) although this should still be confirmed in future.

Mean competitive index   0.32 *   0.37**
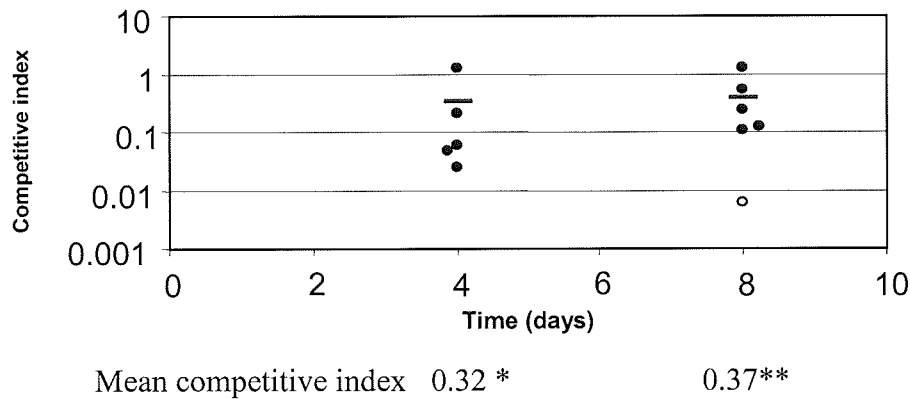
Figure 13: Competitive colonisation with wild type and *srt*A mutant D39
Each point indicates the competitive index from an individual mouse.
Values <1 indicates out-competition of the mutant. Following infection
bacteria were sampled from the upper respiratory tract. Red bar shows
mean and the open circle indicates recovery of no mutant bacteria.
*p<0.05 **p<0.025 compared to one (one equals no attenuation)

Growth *in vitro*

To determine if the out-competition of the *srt*A mutant *in vivo* was due to a general

growth defect the CI was calculated following growth in BHI at 37 °C. At mid-log and

late log stages of growth the mean CI were 0.95 and 1.16 (n=2) showing that under these

conditions the mutant is not out-competed by the wild type and suggesting it does not

have a general growth defect *in vitro*.

Surface anchoring of proteins by *srt*A

Previously it has been shown that proteins anchored to the cell wall by sortase are present at increased levels in the supernatant sortase mutants (Igarashi *et al.* 2003; Kharat and Tomasz 2003). Therefore, to investigate which proteins are anchored to the surface of D39 by SrtA, Western blots were preformed on cell lysates and supernatants from wild type and *srtA* mutant strains grown to mid-exponential phase. The four pneumococcal LPXTG motif proteins investigated were neuraminidase A (NanA), hyaluronidase (Hyl), zinc metalloprotease B (ZmpB) and protease A (PrtA) as antisera to these were available in the laboratory (Figure 14). In wild type D39 NanA and HYL were present in both the cell pellet and supernatant. In the *srt*A mutant there was a decrease in the cell-associated (pellet) signal for both these proteins. This was most evident with NanA where no cell-associated NanA was detected. The effect on Hyl distribution was less obvious with a strong pellet signal still present in the *srt*A mutant albeit at a lower intensity than in the wild type. This experiment needs to be repeated to confirm any subtle differences but was not performed due to limitations on time. Both ZmpB and PrtA were detected only in the supernatant with no cell-associated signal for either the wild type or *srt*A mutant. No difference was seen between the wild type and *srt*A mutant with regards to these proteins.

Cloning, over-expression and purification of SrtA

Pure SrtA would be a useful research tool to use in activity assays, structural studies and to produce antisera. With this in mind, *srt*A was amplified from D39 gDNA with primers 19U and 19V and cloned into the expression plasmid pET-33b+. Following over-expression in *E. coli* BL21 DE3λ cells, cell pellets were collected by centrifugation and

stored at –20°C until required. SrtA was purified from these cells by Ni-affinity column and gel filtration. Pure SrtA was confirmed by SDS-PAGE with Commassie Blue staining (Figure 15).
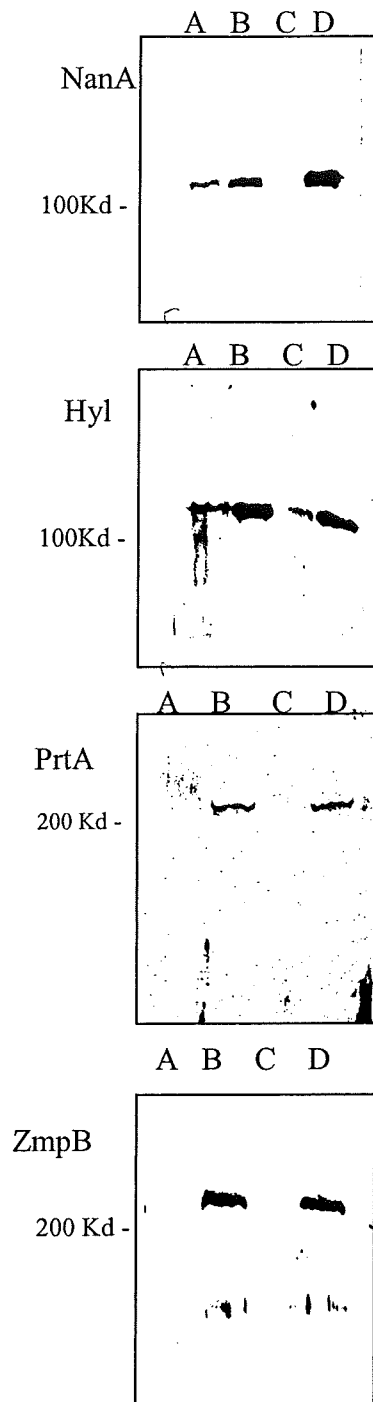
Figure 14: Localisation of LPXTG proteins in wild type and *srt*A mutant
Cell pellet and supernatant from the strains was probed with anti-sera as
indicated. Lane A wild type cell pellet, B wild type supernatant, C *srt*A
pellet and D *srt*A supernatant.

Figure 15: Purified SrtA
SDS-PAGE stained with Commassie Blue of
SrtA

Crystallisation of SrtA

To gain structural insight into the biology of pneumococcal SrtA crystallisation trails have been set up using the sitting drop vapour diffusion technique at 20 ºC. with commercially available sparse matrix screens (Hampton Research, Molecular Dimensions and Emerald Bioscience). Crystals formed but these appeared to be aggregates of several crystal and not likely to diffract satisfactory (Figure 16).

Figure 16: Protein crystals of SrtA

The crystals were taken to the synchrotron source Elettra (Trieste, Italy) however, diffraction experiments yielded little or no diffraction. While this is of little value it does suggests these crystals are formed by protein and are not salt artefacts. Further experiments are underway to improve the crystals produced.

The work on purified SrtA was performed in collaboration with Dr. Alan Riboldi-Tunnicliffe.

Findings

- The strain distribution of pneumococcal *srtA*, *B*, *C* and *D* genes was examined by PCR across 57 clinical isolates. *srtA* was present in all strains with *srtB*, *C* and *D* present in 13 (22.8%).

- Sequencing of *srtA* from 20 strains showed it was highly conserved.

- Survival, blood and lung bacterial counts from a *srtA* mutant following intranasal infection were not significantly different from wild type D39.

- A *srtA* mutant was attenuated in competitive infections with wild type D39 following intranasal and intravenous infections

- This *srtA* mutant was also out-competed by wild type D39 in a competitive colonisation model.

- NanA surface localisation was defective in the *srtA* mutant as measured by protein release into the supernatant.

- SrtA has been over-expressed and purified for further studies including crystallisation, which is currently underway.

Chapter 4 Analysis of Sortase, Discussion

The anchoring of specific proteins to the cell wall mean sortase enzymes have an important role in the pathogenesis of bacterial infections and are seen as a potential target for new antimicrobial drugs. Four sortase homologues were identified in the pneumococcal genomes and these investigated here as candidate virulence factors.

Strain distribution

To investigate the strain distribution of the four pneumococcal sortase genes, *srtA*, *B*, *C* and *D*, PCR with gene specific primers was performed on gDNA from a collection of 57 clinical isolates. Although we have taken the absence of a PCR product to suggest the absence of a gene this may also be due to sequence divergence in the primer binding sites. Other techniques such as microarray analysis or Southern blotting could be used to substantiate the likely absence of genes. These strains were selected from a collection of strains isolated from invasive disease in Scotland to represent a variety of serotypes and MLST sequence types. Sortase strain distribution was of interest because of the potential of sortase as antimicrobial or vaccine targets (Cossart and Jonquieres 2000; Mazmanian *et al.* 2000)). One criterion for any potential antimicrobial target is that it must be present in all or most virulent strains. At the start of this study no work was published looking at the strain distribution of any sortase genes. Strain distribution was also of interest as the presence or absence of particular genes between strains may be associated with certain disease types or outcomes. For example, in *S. pyogenes* the presence and expression of the fibronection-binding protein (FbaB) is associated with strains causing toxic shock-like syndrome (Terao *et al.* 2002). An investigation of gene distribution among different strains coupled with epidemiological/disease data from these strains may therefore

provide clues to such associations and offer important leads towards improving

understanding of pneumococcal carriage and pathogenesis. *srtA* was present in all 57

strains examined suggesting an important role in pneumococcal biology. This agrees with

work published during this study showing that *srtA* of *S. suis* was present in all 59 strains

investigated (Osaki *et al.* 2003). In the case of the three other pneumococcal sortase

genes, they were present in only 13 isolates representing 22.8% of the total examined.

These genes showed a mutually exclusive relationship whereby they were either all

present or all absent. A selected strain distribution for these sortase genes had already

been shown by their identification in TIRG4 but absence in R6 and D39 (Tettelin *et al.*

2001). Here, by examining a greater number of strains a more representative idea of their

distribution is given. Although these genes may contribute to pneumococcal pathogenesis

as suggested by the attenuation of *srt*B and D mutants in mouse infections they do not

appear essential for human disease (Hava and Camilli 2002). No obvious disease

association was seen with the strains positive for these three sortase genes. However,

given that only thirteen strains were positive, a larger sample size may be required for

any such association to be evident. Additionally, it should be noted that for many of the

57 isolates, clinical data is incomplete or absent, limiting the available data with which to

investigate potential associations. For example, no diagnosis was provided for over one

third of isolates (22 out of 57). Unfortunately strains were selected before this lack of

information became apparent otherwise it might have been useful to select strains for

which data was available. It would be of interest to compare the prevalence of these

genes between strains from a larger sample representing strains from various disease and

carriage states to try and identify any correlation with the presence or absence of these

sortase genes. In an effort to identify serotypes and MLST sequence types associated with carriage and invasive disease it was found that sequence types ST9, 124 and 113 had a propensity to cause invasive disease (Brueggemann *et al.* 2003). From the data presented here ST124 (2 strains tested) and ST113 (one strain) do not posses *srtB*, *C* and *D*. This might suggest that there is not a strong association between the ability to cause invasive disease and the presence of the sortase genes, *srtB*, *C* and *D*. However, a larger number of strains would need to be examined before a firm conclusion can be made. If these *srt* genes contribute to virulence then they may be more prevalent in the disease causing isolates. Alternatively, they may have a role in carriage without contributing to disease and so may be more prevalent in strains with a low propensity to cause disease. The linkage of *srt*B, C and D is unsurprising given they are found adjacent to each other in the *rlr*A pathogenicity islet and are transcribed together by the transcriptional regulator RlrA located in the same region (Hava *et al.* 2003). A negative regulator of this islet has also been identified as MgrA, a transcriptional regulator encoded in the genome, not the pathogenicity islet (Hemsley *et al.* 2003; Hava *et al.* 2003). Interestingly, these three sortase genes are found adjacent to three putative cell wall anchored proteins also encoded in this pathogenicity islet (Pallen *et al.* 2001; Hava and Camilli 2002). Each cell wall protein varies slightly within the LPXTG motif (YPRTG, IPQTG and VPDTG) the significance of which is as yet unclear but hints at the possible specific pairing of sortase with substrate within the pathogenicity island. Presumably the presence the three *srt* genes correlates with the presence of the *rlrA* pathogenicity islet and these cell wall proteins. No data is as yet available on the function(s) of these cell wall anchored proteins although one, RrgA (Sp0463) has been shown to contribute to pneumonia and carriage in

the mouse (Pallen *et al.* 2001; Hava and Camilli 2002). They all show weak homology to microbial surface components recognising adhesive matrix molecules (MSCRAMMs) (Pallen *et al.* 2001; Hava and Camilli 2002). Indirect evidence exists for a role of these three proteins in adhesion whereby null mutation or over-expression of the two transcriptional regulators of this islet increased or decreased binding to a lung epithelial cell line in accordance with their expected affects on islet gene transcription (Hemsley *et al.* 2003). This *rlrA* pathogenicity islet is similar to the FCT genome region of *S. pyogenes*, so named because it encodes fibronectin and collagen-binding proteins and the T antigen (Bessen and Kalia 2002). Interestingly, this region in *S. pyogenes* is highly variable between strains both in terms of its gene content and sequence conservation. It was described as a region with a high capacity for intergenomic recombination and therefore likely to contribute to the wide array of clinical and epidemiological phenotypes seen in different strains (Bessen and Kalia 2002). It would therefore be of interest to examine the *rlrA* region of different pneumococcal strains to see if such diversity occurs in this species also and furthermore if any association with particular clinical or epidemiological features exists. Preliminary PCR mapping experiments across this pathogenicity islet of several pneumococcal strains suggests variation may occur (data not shown). Of particular interest is the presence of a ~6Kb insert in two strains. Direct sequencing of this product produced only poor data suggestive of repetitive sequences within this insert. Time limitations meant this work was not pursued although further investigation is merited. The basis for the presence or absence of these three *srt* genes is of interest. The *rlrA* pathogenicity islet is possibility a mobile genetic element given it is flanked on either side by insertion sequences (IS1167). One of these in the TIGR4

96

sequenced genome is predicted to be inactive due to a frameshift mutation but this may not be the case in other strains. The presence of these insertion sequences and the fact that *srtB*, *C* and *D* are found in strains distantly related as judged by MLST sequence type would suggest this *rlrA* region is a mobile element. Its origin and distribution in other species is therefore of interest. Identification of a genomic region such as the *rlrA* pathogenicity islet as being from an exogenous source is possible by comparing several of its features with that of the genome average or housekeeping genes. These features include; G+C content, codon usage, dinucleotide bias and amino acid usage (Karlin 2001). The G+C content of the *rlrA* pathogenicity islet is 38.7% which is similar to that of the genome average of 39.7%. This offers no strong indication that the islet has been acquired by lateral gene transfer. However, the islet may have been acquired from a species with a G+C content close to the pneumococcus or that having been acquired from another species, the G+C content has been adapted towards the genome average as has been predicted following lateral gene transfer (Karlin 2001). In these circumstances it would be of value to investigate the other genome features mentioned above to determine the likelihood that the *rlrA* islet has been acquired from another species. Only one other probable pathogenicity island (PPI-1) has been described for *S. pneumoniae* and is involved in the uptake of iron (Brown *et al.* 2001).

Investigation of srtA

srtA was chosen for further study because it was present in all pneumococcal strains tested and so results may be of relevance to the greatest number of pneumococcal strains. At the beginning of this work no detailed examination of sortase genetic diversity had

been performed, although from the limited data from sequenced organisms it seemed that sortase genes were strongly conserved. As well as strain distribution, sequence diversity is another important consideration in the potential of sortase as antimicrobial or vaccine targets. Sequencing of srtA from twenty clinical strains selected to represent a variety of serotypes and MLST sequence types found the gene and deduced protein sequence to be highly conserved. The ubiquitous distribution and sequence conservation of srtA suggests it plays an important role in pneumococcal biology. The DNA sequence conservation of srtA between strains of 99.3-100% is much greater than the 70.0-99.5% shown between S. suis strains (Osaki et al. 2003). Despite this variation the different srtA alleles of S. suis where shown to be functionally conserved (Osaki et al. 2003).

To investigate further the function(s) of srtA in pneumococcal biology, a null mutant was created in strain D39 by allelic replacement. The mouse-virulent serotype 2 strain D39 was chosen as it has been well-characterised in our mouse infection model (Yesilkaya et al. 2000; Gingles et al. 2001; Blue and Mitchell 2003; Blue et al. 2003; Chiavolini et al. 2003; Kadioglu et al. 2003). Groups of MF-1 mice were infected intranasally with 10^6 CFU wild type or srtA mutant. To compare the virulence of the two bacterial strains mouse survival was monitored and bacterial counts taken from the lung (36 hr) and blood (24 and 36 hr). No statistically significant difference was seen between the two bacterial strains when comparing mouse survival or survival time. Nor was there a statistically significant difference in the bacterial counts in lung or blood. This result was somewhat unexpected because of the growing literature showing the importance of sortase in the virulence of various pathogens (Table 1). Additionally, several LPXTG motif proteins of the pneumococcus have previously been shown to be involved in virulence (Table 2).

However, for survival time and bacterial counts there was a trend toward attenuation of the $srtA$ mutant. For example, the mean blood counts at 24 hr post infection where $\log_{10}$ 5.1 CFU ml^{-1} for the $srtA$ mutant compared to 6.3 for the wild type. This suggested that the $srtA$ mutant was mildly attenuated. Although the MF-1 model is well established and has been used successfully in the past to identify virulence factors it may have the problem that use of a genetically diverse outbred strain creates variation that may mask differences between wild type and mutant. For example, following the standard intranasal infection with 10^6 CFU, as used throughout this work, the blood counts at 24 hr can vary from $\log_{10}$ 1.92 to $\log_{10}$ 9.00 (Blue 2002). Such variation may mask the statistical significance of small differences in virulence. The use of inbred mice might have helped overcome this problem. Being genetically identical they would be expected to respond more similarly to infection and thereby show less variation in the parameters measured. However, they have the disadvantage of costing more to purchase. Additionally, using an inbred strain is not guaranteed to limit variation. For example, following intranasal challenge with 10^6 CFU D39, several strains showed considerable variation between individuals with regards survival times (Gingles et $al.$ 2001). In addition the response between different inbred strains is quite different with some being highly resistant and other highly susceptible to pneumococcal infection (Gingles et $al.$ 2001). Each mouse strain may therefore have inherent disadvantages. For example, mouse survival or spread from lungs to blood could not be examined in the resistant BALB/c. While the difference between a wild type strain and an attenuated mutant might not be evident in a highly susceptible strain because both able to cause the same level of disease. To further study the possible role of $srtA$ in pneumococcal infection, wild type and the $srtA$ mutant were

compared in competitive infections. This type of infection has been used with various pathogens and is seen as potentially more sensitive to differences in virulence because the two strains are compared in the same host thereby reducing host-to-host variation. This allows the identification of small differences in virulence. For example, a knock-out in the ABC iron transporter *pit2* in type 3 pneumococcus had no affect on mouse survival following intranasal or intraperitoneal infections (Brown *et al.* 2001). However, this mutant was attenuated in competitive infections with the wild type strain in both infection types (Brown *et al.* 2001). It should be noted that a competitive infection measures virulence as a growth advantage/disadvantage between strains. The presumption therefore is that virulence/pathogenesis is directly related to bacterial load. This might not necessarily be the case. For example, a mutant made in *Mycobacterium tuberculosis* lacking the alternative sigma factor, *sigH*, produced reduced pathology and mortality compared to the wild type without a reduction in bacterial counts in a mouse infection (Kaushal *et al.* 2002). It appeared that the genes in the SigH regulon were dispensable for bacterial growth and persistence but were involved in triggering the harmful host response important in the development of pathology. Additionally, a *clpC* mutant pneumococcus showed reduced bacterial counts compared to the wild type without any reduction in mortality (Y. Ibraham, personal communication). In this situation it may be the case that some threshold may exist for bacterial counts above which mortality is assured. Any difference in bacterial counts above this threshold may have no impact on the infection outcome. Both these examples show that virulence need not be directly proportional to the ability of a pathogen to survive in the host. It therefore would be best

to measure virulence in terms of host damage, as this is possibly the most important measure of outcome (as discussed by Casadevall and Pirofski 2003].

This criticism also holds not only for competitive infections but also infections with pure inoculum followed by bacterial counts. The difficulty and challenge is therefore the ability to sensitively measure host damage. Another potential disadvantage with competitive infections is the inability to monitor the spread of bacteria from one site to another. This is an important feature of in our model and human infection where the spread of bacteria from the lung to blood may involve virulence factors distinct from those functioning at either site, and is associated with disease progression and mortality. For example, it is suggested that ZmpB is important in the dissemination of the pneumococcus from lung to blood (Blue et al. 2003). Hava and Camilli (Hava and Camilli 2002) found that in pneumococcal STM pools the bacteria spreading from the lungs to the blood were made of a random collection of a few clones. This will probably also apply to competitive infections using two strains and so such infections could not be used to investigate bacterial dissemination. In this case therefore the possible defective dissemination of the zmpB mutant would be overlooked. In addition, it is possible in a competitive infection for the wild type strain to complement a deficiency of the mutant. For example, a mutant lacking a secreted toxin might not be at a competitive disadvantage if the wild type is also present and producing that toxin.

One the other hand, competitive infections have the advantage of being sensitive to small differences between strains in their ability to survive in vivo and use less mice/host organisms and so are more ethical and economical. Bearing these points in mind, competitive infections were preformed between wild type D39 and srtA mutant following

intranasal and intravenous challenges. Both infection types were tested as the contribution to virulence of *srtA* may vary in different infections. The CI was calculated immediately following each infection (0hr). The value was not significantly different from 1 indicating that the preparation and infection processes themselves had no impact on the ratio of the two strains. At each other time point examined (6, 12, 24 and 36 hr) in both infections, the CI was significantly less than 1. This showed the wild type was out-competing the *srtA* mutant and was therefore better able to survive in the infection. *srtA* was therefore shown to be a virulence factor in pneumococcal pneumonia and bacteraemia. *srtA* was not identified as a virulence factor in the three previous pneumococcal STM screens (Polissi *et al.* 1998; Lau *et al.* 2001; Hava and Camilli 2002) and so this is the first demonstration of the role of *srtA* in pneumococcal pathogenesis.

During the course of this work *srtA* was deleted in pneumococcal strain R36ASIII, a type 3 encapsulated derivative of the acapsular R36A strain (Kharat and Tomasz 2003). This mutation caused no difference in mouse survival compared to wild type following intraperitoneal challenge. This experiment has the flaw that the role of capsule in virulence is dependant on the genetic background (Kadioglu *et al.* 2002) and so the expression of a heterologous serotype 3 capsule on a serotype 2 background may create artefacts. In addition, it can be argued that intraperitoneal challenge is not the most physiologically relevant route to pneumococcal infection. Indeed, the pneumococcus can be highly virulent following intraperitoneal challenge of mice and so this may mask any attenuation between wild type and mutant. However despite such caveats the lack of attenuation is in agreement with our first virulence studies following intranasal infection with wild type D39 and *srtA* mutant where no significant difference was seen between the

two strains. A significant attenuation of the *srtA* mutant was only detected using sensitive competitive infection. This suggests the contribution of *srtA* may be minor in pneumococcal virulence. This is in contrast to the apparently larger contribution seen in other organisms in animal models where attenuation of *srtA* mutants is more readily apparent, requiring straightforward infections of different groups of animals (Mazmanian *et al.* 2000; Bolken *et al.* 2001; Bierne *et al.* 2002; Garandeau *et al.* 2002; Mazmanian *et al.* 2002). There are several possible contributing factors to explain this. Neither R36A nor D39 have been sequenced and so may contain additional sortase genes complementing the loss of *srtA*. This seems unlikely given the dramatic affect of *srtA* deletion on the localisation of NanA and β-gal [harat and Tomasz 200] and our NanA Western blot, Figure 14). Additionally, both R36A and D39 are closely related to the sequenced R6 strain, which possesses only *srtA* and no other sortase (Hoskins *et al.* 2001). To overcome this potential problem it may have been advisable to use the virulent sequenced TIGR4 strain. However, not even using a sequenced strain removes this potential problem of redundancy or complementation masking a phenotype because the apparent absence of a gene in the genome does not confirm the absence of that particular activity as a non-homologous gene may encode it. An additional explanation is that the proper surface localisation of pneumococcal LPXTG motif proteins is not essential for their contribution to virulence. All of the characterised pneumococcal LPXTG motif proteins have enzymatic activity, it is therefore possible that their increased secretion into the extracellular environment in a *srtA* mutant allows them to perform their function almost as effectively as when found on the surface. This is in contrast to some of the LPXTG motif proteins acting as adhesins in other organisms such as fibronectin and

103

fibrogen binding proteins in *S. aureus* (Mazmanian *et al.* 2000; Mazmanian *et al.* 2002),

CshA fibronectin binding protein of *S. gordonii* (Bolken *et al.* 2001), P1 adhesin of *S. mutans* (Lee and Boran 2003), internalin A in *L. monocytogenes* (Bierne *et al.* 2002; Garandeau *et al.* 2002). Indeed, LPXTG proteins include important adhesins involved in the virulence of each of the organisms other than the pneumococcus examined for the role of *srtA* in pathogenesis. The requirement for the correct anchoring of these proteins to the cell wall may explain the large defect in virulence shown with *srtA* mutants. The enzyme LPXTG motif proteins of the pneumococcus may not have such an absolute requirement to be located on the cell surface and their increased secretion in a *srtA* mutant has little impact on virulence. A potential exception in the pneumococcus is the IgAse which has been shown to be involved in adherence (Weiser *et al.* 2003). However, it appears that enzymatic activity is central to this function rather than acting in a direct ligand-receptor binding interaction. It may therefore not be essential to be surface located. In support of this, a IgAse mutant could be partially complemented by addition of secreted IgAse from *Helicobacter pylori* to the medium. The binding of cationic cleaved antibodies fragments to the surface of the pneumococcus was thought to neutralise the inhibitory affect of the negativity charged capsule on adhesion leading to improved efficiency of the phosphorylcholine - PAFr mediated adherence. Additionally, as will be discussed with ZmpB, the cell wall anchoring signal of IgAse is found at the N-terminus not the C-terminus. IgAse may not therefore a substrate for sortase. The defect in pneumococcal host cell adherence shown by a *srtA* mutant by Kharat and Tomasz (Kharat and Tomasz 2003) suggest a role in adherence for LPXTG motif proteins. However, this effect was only shown in acapsular laboratory strains that as previously

discussed might not be representative of capsular clinical strains in such adherence assays (Brock *et al.* 2002). It would therefore be of interest to use our virulent capsular wild type and *srtA* mutant strains in a similar assay.

In contrast to other organisms it might be that pneumococcal ligands for adherence may be surface anchored by alternative methods such as choline binding proteins that are surface localised through binding of choline in teichoic acids. Or as yet uncharacterised mechanisms as suggested for PavA and alpha-enolase. Both these adhesins are surface proteins but lack any known motif for this localisation (Bergmann *et al.* 2001; Holmes *et al.* 2001). Finally, it may be possible that many of the LPXTG motif proteins in D39 do not contribute greatly to virulence in our model. Little detailed data exits on the role of LPXTG motif proteins in pneumococcal virulence. Additionally, it is difficult to compare what work has been done previously with our results given the use of different bacterial and host strains and infection types. For these reasons mutants were made in the pneumococcal LPXTG proteins *nanA*, *prtA*, *β-gal* and *spuA* to assess their role in virulence in a manner comparable with *srtA* (Chapters 5 and 6).

The affect of *srtA* deletion was also examined in a competitive colonisation model. This was performed in the same manner as the competitive pneumonia infections with the exception that the bacteria where administered in the smaller volume of 10µl not 50µl. The result is nasopharyngeal colonisation without invasive disease. In line with this, mice showed no clinical signs and had blood counts below the detectable level in contrast to the pneumonia model using a 50µl volume. Bacteria were recovered up to 8 days post challenge, later times were not tested but the bacterial levels at day 8 suggested persistence beyond that time point was likely. The decreased ability of *srtA* mutants to

adhere and invade the human pharyngeal tissue cell line Detroit 562 suggested a role for *srtA* in colonisation (Kharat and Tomasz 2003). Indeed, we found that the *srtA* was significantly out-competed by wild type D39 at both time points examined, 4 and 8 days. *srtA* appears therefore to contribute to pneumococcal nasopharyngeal colonisation in addition to pneumonia and bacteraemia.

When grown competitively in BHI *in vitro* there was no out-competition of *srtA* by wild type D39. The *srtA* mutant does not appear therefore to have a general growth defect that would cause it to be out-competed by the wild type. The *srtA* attenuation seen during competitive infection and colonisation studies may therefore be due to specific features of the *in vivo* environment rather than an inherent growth disability.

On the contribution of SrtA to pneumococcal virulence and carriage it may be the case that its role in the mouse is less important than in human colonisation and disease. Evidence for its importance might be taken for its ubiquitous presence in all strains tested and sequence conservation despite an apparently small contribution in mouse models. Although not feasible to test its role experimentally in human disease the recent use of human volunteers for experimental carriage (McCool *et al.* 2002; McCool *et al.* 2003) presents an interesting future experiment to examine the role of SrtA in pneumococcal biology. However, such studies are likely to be expensive and time-consuming.

It is also worth noting that the importance of different virulence factors varies with the pneumococcal strain being used (Blue and Mitchell 2003). For this reason it may be of interest to compare the affect of a *srtA* null mutation in other strains. In addition for a complete examination of virulence *srtA* could be examined in models of otitis media and meningitis to complement the work performed here. If our animal results with regards the

role of SrtA in virulence represent the situation in human infection it seems unlikely that SrtA would make a good antimicrobial target.

To assess which LPXTG motif protein are anchored to the cell wall by SrtA a series of Western blots were preformed. As has been shown in sortase mutants in other organisms there is an increased secretion of sortase substrate (Igarashi *et al.* 2003). This method was used by Kharat and Tomasz (Kharat and Tomasz 2003) to examine pneumococcal SrtA. Comparison of the cell pellet versus the supernatant showed this mutant had increased secretion of NanA and β-gal. We repeated and expanded this investigation by performing Western blots on cell pellet and supernatant from exponentially grown cells using antisera against NanA, PrtA, ZmpB and Hyl. Our NanA Western blot confirmed the results of Kharat showing that in wild type cells NanA is found in both the cell pellet and supernatant while in the *srtA* mutant NanA is largely absent in the cell pellet. This suggests that NanA is anchored by SrtA. Commensurate with its increased secretion in the *srtA* mutant, the signal from the supernatant was increased over that from the wild type supernatant. In the case of Hyl the results are less clear. A signal was present in all four lanes; wild type pellet and supernatant and *srtA* pellet and supernatant. The signal in the *srtA* mutant pellet was slightly less than that in the wild type pellet suggesting defective cell association. However, the difference was only slight and without an increased signal in the *srtA* supernatant. This would need to be repeated before making a firm conclusion on the cell association of Hyl by SrtA. Why there should be a difference between NanA and Hyl is unclear perhaps differences in their C-terminal hydrophobic tail mean that unanchored Hyl associates more with the membrane than does NanA. In the case of PrtA the entire signal in the wild type and *srtA* mutant was present in the

supernatant. No conclusion of the anchoring of PrtA by SrtA can therefore be made. The presence of PrtA in pneumococcal cell lysate and supernatant has previously been shown (Bethe *et al.* 2001) with the pellet signal greater than that from the supernatant. This current work used the same D39 strain that was employed previously so the reason for this discrepancy is unclear. We used cells from exponential growth phase and although not stated in the paper that may not have been the case for the previous work (Bethe *et al.* 2001). It may be that they used cells from an earlier stage of growth when the distribution of PrtA may be different. For example, in early growth PrtA may all be cell associated but is later secreted into the supernatant. A time course for the localisation of LPXTG motif proteins would be interest and has yet to be performed. The process by which these supposed cell wall anchored proteins are secreted into the supernatant is unclear. The band sizes of the three proteins described above were similar in the pellet and supernatant. This suggests that if cleavage of the LPXTG protein is involved in its secretion the cleavage site must be near the LPXTG motif. Additionally, it appears the proteins are present in the supernatant without being attached to a large portion of cell wall, which would cause an apparent increase in their size on the Western blots. It seems unlikely also that these proteins are present in the supernatant by virtue of cell lysis releasing cell wall fragments containing attached cell wall protein as during exponential growth such lysis should be limited (Balachandran *et al.* 2001). This could be examined further using a *lytA* null mutant, which is defective in autolysis (Tomasz *et al.* 1988). Finally, ZmpB was found only in the supernatant of wild type or *srtA* mutant with no apparent difference between the two strains. This confirms that ZmpB is secreted with no apparent role for SrtA. This does not however, exclude surface association or SrtA

involvement at another phase of growth again highlighting the potential value of a time course. It is worth mentioning that this approach to SrtA cell wall anchoring is limited by the availability of appropriate antibodies against the proteins of interest. This limitation could be overcome by the development of an appropriate reporter system for cell wall anchoring. Such a device has proved useful in studying *S. aureus* sortase. Here, the putative C-terminal cell wall sorting sequences from different proteins are fused to staphylococcal enterotoxin B, which can then be immunopreciptated from cells to identify their cellular location (Mazmanian *et al.* 2002). This allows identification of the proteins recognised by a sortase without the need for specific antibodies to each cell wall protein. To create a similar reporter system for the pneumococcus it was attempted to replace an internal fragment of NanA with EGFP as a reporter. NanA was chosen as it had already been confirmed to be surface located (Camara *et al.* 1994) while this work and that of Kharat and Tomasz (Kharat and Tomasz 2003) showed this to be SrtA mediated. Replacement of the original C-terminal cell wall sorting signal of NanA with that of other proteins could then be used to investigate if these were recognised and surface located by sortase. EGFP could be detected by microscopy or Western blot or measured by spectrophotometer thereby providing qualitative results. This work has not been finished and so has not been included here.

Purified SrtA would be a useful research resource for studying sortase biology. For example, it could be used in activity assays, structural studies and as vaccine antigen to generate antisera or to test its potential to provide protection from infection. With this in mind SrtA was overexpressed, purified and crystallisation trials begun in collaboration with Dr. A. Riboldi-Tunnicliffe. Crystals have been formed and it is hoped that in time

structural data may be generated by X-ray crystallography. The only available sortase

structure is that of *S. aureus* SrtA (Ilangovan *et al.* 2001a; Ilangovan *et al.* 2001b) there is

therefore value in solving the structure of an additional sortase. This is especially true

given the lack of sequence similarity between different sortase proteins that may extend

to structural and functional features. Attempts to model pneumococcal SrtA structure

based on the *S. aureus* protein using SWISS-MODEL Protein Modelling Server failed

due to lack of homology.

Chapter 5 Analysis of LPXTG Proteins, Results

Introduction

To complement the work on the role of *srtA* in pneumococcal virulence several LPXTG

motif proteins were also examined, as their roles in pathogenesis are poorly defined.

Summary

Null mutants were made in four pneumococcal LPXTG motif protein genes in order to

investigate their role(s) in virulence. Two of the four were attenuated in competitive lung

infections suggesting they contribute to pneumonia.

Janus mutagenesis is a technique allowing the introduction of mutations into the

pneumococcal genome that lack selectable markers. This can facilitate accessible

manipulation of the genome while limiting the chances of artefacts. This technique was

used to delete the LPXTG motif from ZmpB to investigate if this motif contributed to

virulence. This motif was of interest as unlike typical LPXTG motifs that are found at the

C-terminal, this was located at the N-terminal of ZmpB.

Mutant construction and verification

Four pneumococcal LPXTG motif proteins were selected for further study by creating

knock out mutants in D39. The four selected were neuraminidase A/NanA (spr1535), β-

galactosidase/β-gal (Spr0565), alkaline amylopullulanase/SpuA (Spr0247) and

pneumococcal protease A/PrtA (Spr0561). These were chosen because there is a lack of

information on their role in different pneumococcal infections. Null mutants were made

by allelic replacement with slight modification from the method used for *srt*A as

described in Material and Methods (Figures 3 and 4). Confirming the absence of the

target genes, primers internal to each failed to yield a product with gDNA from each

particular mutant but gave the expected product from wild type gDNA (Figure 17).

Furthermore, primers sited upstream and downstream of the manipulated region gave the

expected product for each mutant (Figure 17). These primers were not used on wild type

gDNA due to problems optimising the conditions needed to yield the large products

produced. This was not a problem with the mutants as the products were smaller due to

the deletion of the target gene. None of the four genes appeared essential for viability as

transformants were obtained for all.
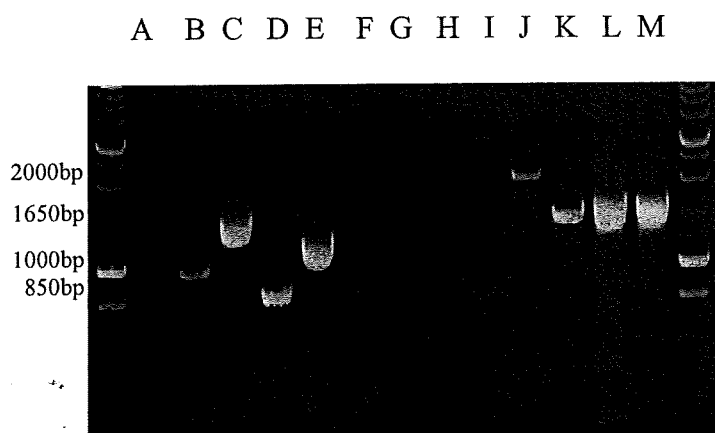
Figure 17: PCR confirmation of LPXTG mutants
Expected product size given in brackets expressed in base pairs. Lane A blank
gDNA prep checked with primers 20o and 20N (*srt*A) lanes B-E wild type
gDNA with internal primers against *prt*A (1050), *nan*A (1486), *pull* (835) and
β-*gal* (1252). Lanes F-I internal primers against *prt*A, *nan*A, *spuA* and β-*gal*
with gDNA from respective mutants. Lanes J-M outside primers against *prt*A
(2167), *nan*A (1560), *pull* (1631) and β-*gal* (1646) with gDNA from
respective mutants.
See Table 3 for primer pairs used.

Analysis of virulence

Each of the four LPXTG protein mutants was compared to wild type D39 with regards bacterial lung and blood counts 36 hr after intranasal infection of MF-1 mice with 10^6 CFU (Figure 18).
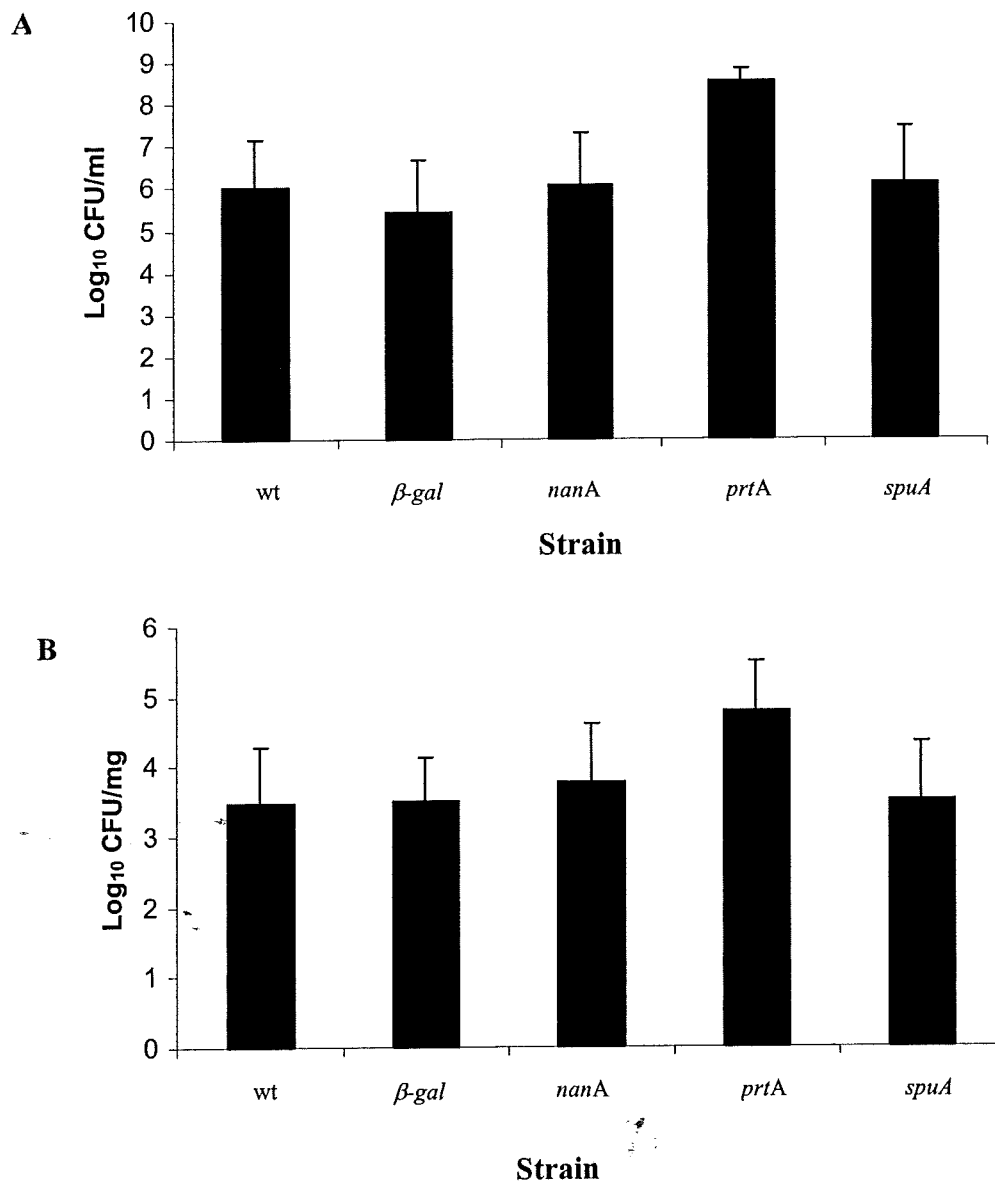
Figure 18: Virulence analysis of LPXTG protein null mutants
Bacterial blood (A) and lung counts (B) were measured 36 hr. after intranasal infection with 10^6CFU. n=6-8. Data are mean +/- 1 SE. No significant difference was seen compared to wild type D39 (wt).

No significant difference was evident between the bacterial counts with wild type D39 and any of the LPXTG protein mutants. In order to investigate any possible attenuation of these mutants a competitive infection was used as described for srtA. At 36 hr after intranasal infection the CI between each mutant and D39 wild type was calculated (Figure 19).
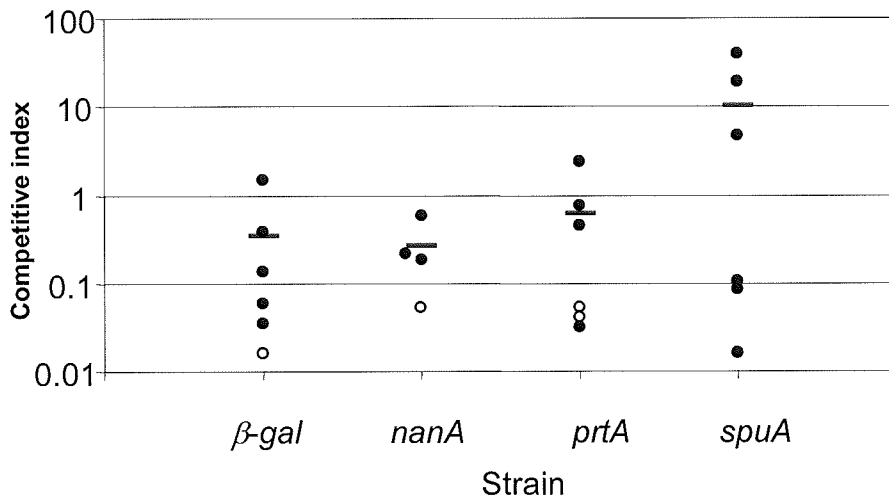
| Mean competitive index | <0.35* | <0.26** | <0.61 | 10.37 |

Figure 19: Virulence analysis of LPXTG protein mutants in competitive infections with parental wild type D39
Each point indicates the competitive index from an individual mouse following intranasal challenge. Values <1 indicates attenuation of the mutant. Bacteria were sampled from the lung at 36 hr post infection. Following intravenous challenge bacteria were sampled from the blood. Red bar shows the mean CI and open circles indicate the recovery of no mutant bacteria. *p<0.05 **p<0.025 compared to one (one equals no attenuation).

The mean CI for the β-gal and nanA mutants was significantly less than 1 indicating attenuation of these strains. In the case of prtA, the mean CI was not significantly less

than 1 but there was a trend towards attenuation, as 5 out of 6 mice had a CI less than 1 and 2 of these mice had no mutant recovered post-infection. It seems likely that repetition of this experiment might show attenuation when the results are pooled. Limitations on time precluded this experiment. The mean CI of the *spuA* mutant is not significantly different from 1 and on that basis this mutation appears not to affect virulence. However, it is interesting to note the large variation between individual mice with regard the CI. Three mice show a large attenuation of the mutant while the other three show increased virulence of the mutant. The reason for such variation is not easily apparent and it would be of value to repeat this infection if time had permitted.

To confirm the genotypes of colonies (wild type versus mutant) assigned by antibiotic resistance was correct, 4-5 wild type and 4-5 mutant colonies from each mixed infection were checked by PCR (same PCR as Figure 17). All genotypes by PCR agreed with those based on resistance to erythromycin.

Deletion of LPXTG motif of ZmpB

ZmpB has been shown to be an important pneumococcal virulence factor in both murine pneumonia and bacteraemia (Polissi *et al.* 1998; Hava and Camilli 2002; Blue *et al.* 2003; Chiavolini *et al.* 2003). It contains a signal sequence and an LPXTG (LPQTG) motif followed by a hydrophobic region and charged residues. ZmpB therefore contains the necessary features for sortase-mediated cell wall anchoring. Indeed, the TIGRFAM and Pfam Hits facility at the TIGR Comprehensive Microbial Resource http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl shows that ZmpB is classified as a cell wall anchored protein. However, the putative cell wall anchoring signal is found

at the N-terminus of ZmpB not the C- terminus as with typical cell wall anchored proteins. To investigate if this LPXTG motif is important in the localisation and virulence of ZmpB it was deleted using Janus mutagenesis. The strains used and created during this procedure are shown in Table 7.

Strain	Description/Use	Reference
CP1200	Soucre of rpsL1 allele	(Sung et al. 2001)
R1036	Source of Janus cassette	(Sung et al. 2001)
R6.1 Kan	R6 with Janus cassette replacing LPXTG motif coding sequence region of zmpB	This work
SmR1	D39 made Sm resistant by transfer of rpsL1 from CP1200 then backcrossed again onto D39	This work
SmR2	D39 made Sm resistant by transfer of rpsL1 from CP1200 then backcrossed again onto D39	This work
SmR3	D39 made Sm resistant by transfer of rpsL1 from CP1200 then backcrossed again onto D39	This work
JZ1	SmR1 with Janus cassette replacing LPXTG coding sequence region of zmpB made by transformation with R6.1 Kan gDNA then backcrossed again onto SmR1	This work
JM	Derivative of JZ1, Janus cassette removed and replaced by original sequence except with deletion of LPXTG coding sequence	This work

Table 7: Strains used or made during Janus mutagenesis of zmpB LPXTG motif sequence

SmR1 – 3 are three independent clones from the same transformation.

This mutagenesis strategy allows the introduction of mutations without a selectable marker into the pneumococcal genome. That is, mutations with no selection marker thus allowing more subtle and informative manipulations of the genome. The strategy begins with a pneumococcal strain resistant to streptomycin (Sm) due to the presence of a mutation within its genomic rpsL gene encoding the S12 ribosomal subunit. This resistance is conferred by a single base pair change converting lysine at position 56

117

(AAA) to threonine (ACA). In the next step the Janus cassette is inserted into the genome at the desired site of this Sm resistant strain. This is done through homologous recombination and so a construct is first made consisting of the Janus cassette flanked by the sequences required for this recombination. Depending on the purpose of mutagenesis and dictated by the sequences used for homologous recombination, this integration can involve replacement of host sequence with the cassette or the insertion of the cassette without any deletion. The Janus cassette is a bicistronic construct allowing selection for both its acquisition and its loss from the strain. It consists of three parts; the *ami*-A promoter, a 3'5''-aminoglycosidase phosphotransferase conferring kanamycin resistance and the wild type allele of *rps*L (*rps*L+). The promoter drives constitutive expression of the cassette independently of the site of integration. Kanamycin (Kan) resistance conferred by the 3'5''-aminoglycosidase phosphotransferase is the selective marker for insertion of the Janus cassette. The wild type *rps*L gene confers streptomycin susceptibility, as it is dominant over the resistant allele. This is possibly due to the higher transcription rate from the Janus cassette compared to the allele in the genome. Following integration of the Janus cassette the strain changes from Sm resistant/Kan susceptible to Sm susceptible/Kan resistant. Homologous recombination can then be used to remove the cassette and replace it with any DNA sequence of choice. Following this event the strain returns to being Sm resistant/Kan susceptible. This Sm resistance can be used to select for transformants and so the donor DNA used to for this transformation need not have a selectable marker (Figure 20).

	Procedure	Antibiotic resistance phenotype

- Streptomycin resistant host strain, *rpsL1* (S12 protein) Lys 56 (AAA) to Thr (ACA) ↓

     Kan Sus   Sm Res

     ↓

- Introduce Janus cassette into region of interest   **Kan Res   Sm Sus**

| amiA promoter | Kanamycin resistance | rpsL+ |

Janus casette

↓

- Replacement of Janus cassette with desired DNA

     Kan Sus   Sm Res
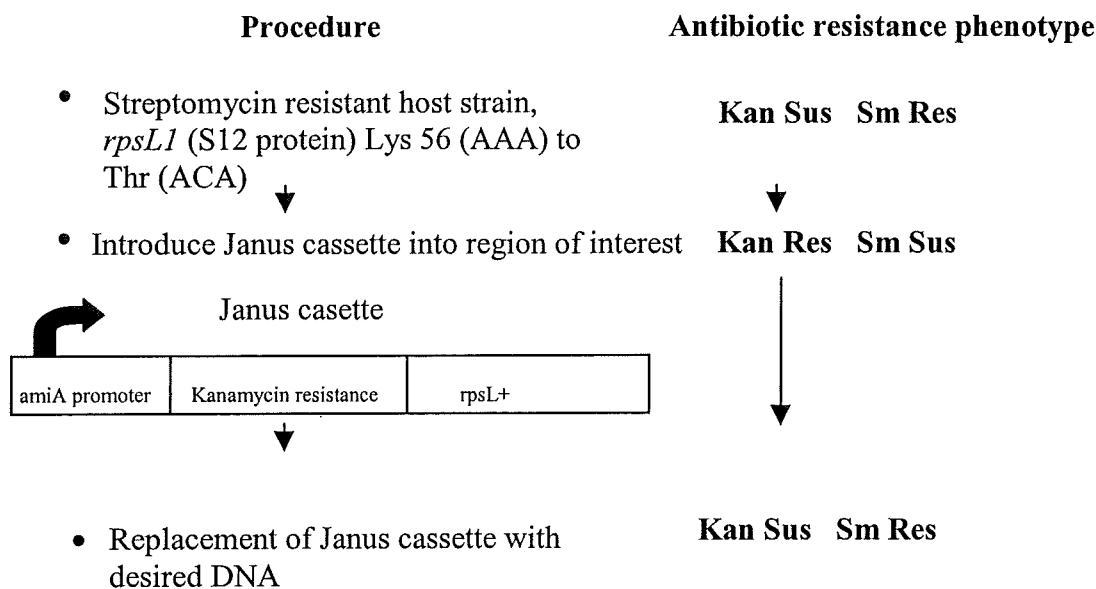
Figure 20: Outline of Janus mutagenesis
The starting strain is resistant (Res) to streptomycin (Sm) due to the presence of the resistant *rps*L1 allele in its genome. Integration of the Janus cassette into the region of interest is selected for by the acquisition of kanamycin (Kan) resistance. This intermediate strain also becomes Sm susceptible (Sus) due to the presence of the wild type *rps*L allele in the Janus cassette, which is dominant over the resistant genomic *rps*L1. Replacement of the Janus cassette by donor DNA is selected for by Sm resistance allowing the selection of mutations without any selectable marker.

The first step was to create a Sm resistant strain of D39. This was done by transformation of D39 with gDNA from strain CP1200 which possesses the *rps*L1 resistant allele, with selection on Sm ($150\mu g \ ml^{-1}$). Transformants were readily obtained. Genomic DNA from one of these D39 transformants was used to transform back into D39 to minimise the chances of CP1200 DNA integrating into D39 to give strain SmR1. Sequencing of the *rps*L gene of SmR1 confirmed the desired mutation. SmR1 and two other Sm resistant

D39 transformants (SmR2 and 3 produced in the same transformation) were compared to wild type with regards virulence. No difference in survival and blood counts following intranasal infection with 10^6 CFU (Figure 21) was seen between any of the SmR strains and wild type D39 showing that acquisition of Sm resistance did not alter virulence.
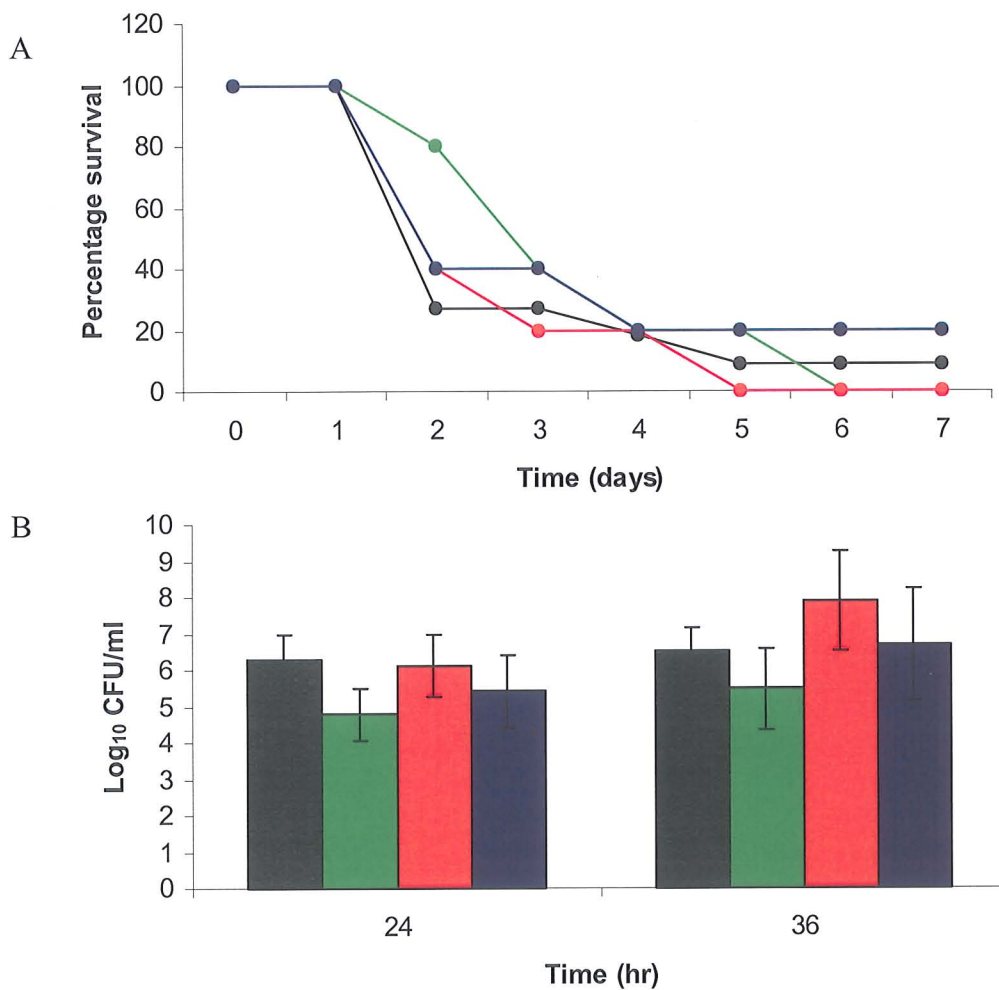
Figure 21: Virulence of D39 Sm resistance derivatives (SmR) compared to wild type
Survival was monitored (A) and blood counts (B) taken following intranasal challenge with 10^6 CFU. Black D39 wild type, green SmR1, red SmR2 and blue SmR3. No significant difference compared to wild type. n=5 for SmR 1-3 (n=11 for wild type). Blood counts show mean +/- 1 SE.

To integrate the Janus cassette into the LPXTG motif of the *zmp*B gene requires fusion of the sequences for the desired homologous recombination to the Janus cassette. To do this the desired sequences and the Janus cassette were amplified using primers that introduced restriction sites allowing ligation of the three products together. Using the primer pairs 21C & 21G and 21H & 21F the regions upstream and downstream respectively of the *zmp*B LPXTG motif were amplified from D39 gDNA. Similarly the Janus cassette was amplified from R1036 gDNA using the primers 21I and 21J. Each product was purified and digested with the appropriate restriction enzyme(s). These products were then purified again and all three added to the same ligation reaction. After ligation overnight this reaction was used to directly transform R6 with selection on kanamycin (200μg ml⁻¹) (Figure 22). R6 was used in preference to D39 due to its higher transformation rate. Integration of the Janus cassette was confirmed with the primers 21Q and 21R sited upstream and downstream respectively of the manipulated region (Figure 22).

(A) Target region in genome

21C 21G 21H 21F

*zmp*B LPXTG region

(B)

Upstream

PCR to amplify fragments and introduce restriction sites. Digestion with appropriate restriction enzyme(s)

Downstream

Janus cassette

HindIII NcoI

Ligation reaction with all three products and used directly for transformation leading to integration of Janus cassette into desired site.

(C)

Janus cassette

*zmp*B LPXTG region

21Q

Janus cassette

21R

Janus cassette insertion into genome in place of *zmp*B LPXTG coding region

Figure 22: Fusion of Janus cassette to flanking sequences for homologous recombination into LPXTG motif coding region of *zmpB*

(A) Using the primer pairs 21C & 21G and 21H & 21F the regions upstream and downstream respectively of the *zmp*B LPXTG motif coding sequence were amplified from D39 gDNA . Similarly the Janus cassette was amplified from R1036 gDNA using the primers 21I and 21J.
(B) Each product was purified and digested with the appropriate restriction enzyme(s). These products were then purified again and all three added to the same ligation reaction.
(C) After ligation overnight this reaction was used to directly transform R6 with selection on Kan (200µg ml^{-1}). This mutation can now be transferred to an Sm resistant strain to allow manipulation of the target region.

122

gDNA from this R6 derivative (R6.1 Kan) was then used to transform SmR1 with selection on kanamycin (300µg ml⁻¹) to transfer the Janus integration into this Sm resistant D39. Genomic DNA from a transformant from this transformation was purified and used to transform SmR1 again in order to restrict to co-transfer of R6 DNA to give strain JZ1. Janus cassette integration was confirmed with primers 21Q and 21R (Figure 22 C). This Janus cassette integration resulted in the deletion of a small region including the LPQTG coding sequence of *zmp*B. The next step was to replace the Janus cassette with the original sequence except with a deletion of the sequence for the LPQTG motif. Deletion of the coding sequence for this motif was performed by overlapping PCR (Figure 23). Primer pairs 21C & 21T amplified the 5' and 21S & 21F the 3' portion of the desired sequence respectively, minus the coding sequence for the LPXTG motif due to primer design. These two PCR products were purified and joined by a second PCR reaction with primers 21C and 21F. This was possible due to the inclusion of complementary sequences added in primers 21T and 21S during the initial PCR. The resulting product, the ZmpB LPQTG region minus the sequence for this motif plus the upstream and downstream regions for homologous recombination was cloned into TOPO-pCR4.
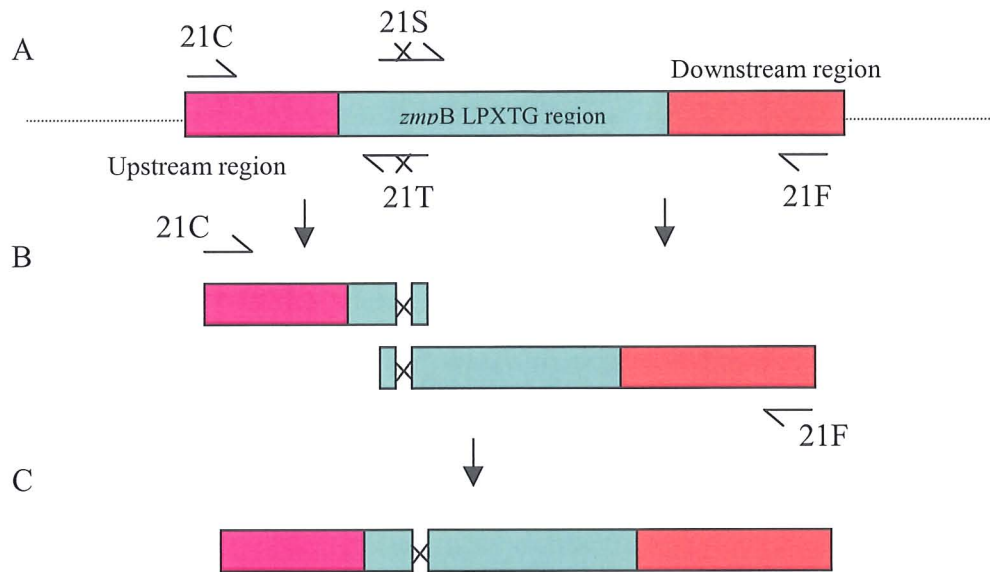
Figure 23: PCR deletion of the LPQTG motif coding sequence of *zmp*B
A) Primer pairs 21C & 21T were used to amplify the 5' region and 21S & 21F the 3' portion of the desired sequence minus the coding sequence for the LPXTG motif denoted by x. B) The two products were purified and joined together by a second PCR reaction with primers 21C and 21F due to complementary sequences added in primers 21T and 21S used the initial PCR. The resulting product, the LPQTG region minus the sequence coding this motif plus the upstream and downstream regions needed for homologous recombination. C) this product was subsequently cloned and sequenced, ready for subsequent use in the final Janus transformation to remove the Janus cassette from strain JZ1 and replace it with the original sequence with the LPQTG sequence deleted.

Sequencing of this plasmid confirmed the desired deletion and the absence of any other changes. The final plasmid construct was then used to transform JZ1 with homologous recombination replacing the Janus cassette with the original *zmpB* sequence containing a deletion of the LPQTG motif sequence. Selection was on Sm as the transformation should cause JZ1 to switch from Sm susceptible to Sm resistant due to loss of the

dominant susceptible allele in the Janus cassette (Figure 20). However, the results of initial transformations showed that the number of Sm resistant colonies produced by addition of the plasmid construct was the same as the number produced when no DNA was added. There was therefore a high level of background false positive colonies that made the identification of true transformants problematic. This was not unexpected, as it had been described in the original Janus mutagenesis description (Sung *et al.* 2001). The observation results from spontaneous recombination whereby one *rps*L allele replaces the other so the clone is no longer a *rps*L heterozygote but instead posses two identical copies of the gene. Where the Janus cassette wild type *rps*L allele is replaced by the genome allele the clone becomes Sm resistant but still contains the Janus cassette and represents a false positive during transformation. To limit the number of these spontaneous Sm resistant revertants several steps where taken. Increasing the concentration of Sm from 150 to 250 or 500 $\mu g\ ml^{-1}$ had no appreciable affect on the abundance of revertants. It then became apparent that the number of revertants was directly proportional to the length of incubation and so the period of incubation following a transformation was reduced. When this was done it was noticed that selective plates after transformation had a mixture of large and small colonies. In contrast, selective plates from no DNA controls had only small colonies. This suggested that the large colonies might be true transformants and the small ones Sm resistant revertants. Picking single large colonies without contamination from the smaller potential revertants was however a problem due to the great excess of the latter. This was overcome by plating out smaller volumes of transformation culture. For example 5-50μl was plated out instead of the more normal 300-500μl. Additionally, incubation of transformation plates in 5% CO_2

anaerobic bags seemed to enhance the size difference between the large and small colonies. When these steps were used it was possible to pick single colonies and it was found that of 20 large colonies picked, all appeared to be true transformants (Sm resistant, Kan susceptible) whereas all 20 small colonies appeared to be revertants (Sm resistant, Kan resistant). Replacement of the Janus cassette with the donor DNA in transformants was subsequently confirmed with PCR using the primer pair 21Q and 21R (Figure 24).
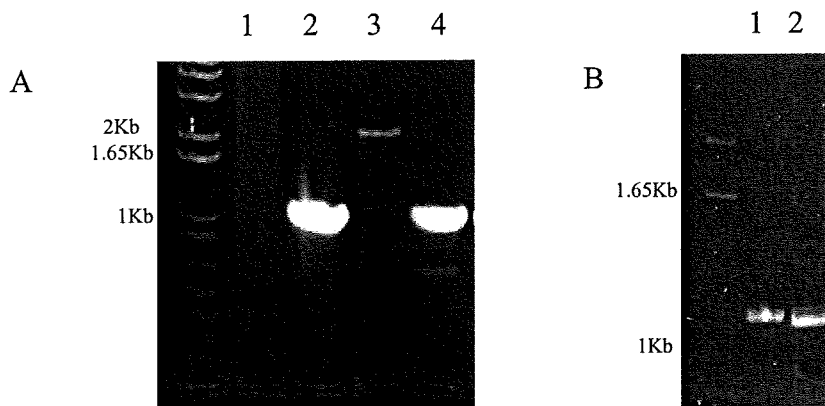
Figure 24: PCR confirmation of Janus mutagenesis strains
All reactions were performed with primer pair 21Q and 21R sited upstream and downstream respectively of the manipulated region. Expected product size in brackets expressed in base pairs. Gel A, lane 1) no template, 2) SmR1 (1097) gDNA, 3) JZ1 gDNA (~2100) and 4) JM gDNA (1082). Gel B, products from SmR1 (lane 1) and JM (lane 2) were run out together at low voltage for a prolonged period to show the small size difference (15bp) between them is possibly visible.

One transformant, JM was chosen for further work and the 21Q and 21R PCR product from this strain was sequenced and the desired mutation and the absence of others confirmed. This mutant was then compared to SmR1 with regards virulence following

intranasal infection. No difference was seen between the two strains with regards survival or bacterial blood counts suggesting the LPXTG motif of ZmpB is not essential for its contribution to virulence. (Figure 25).
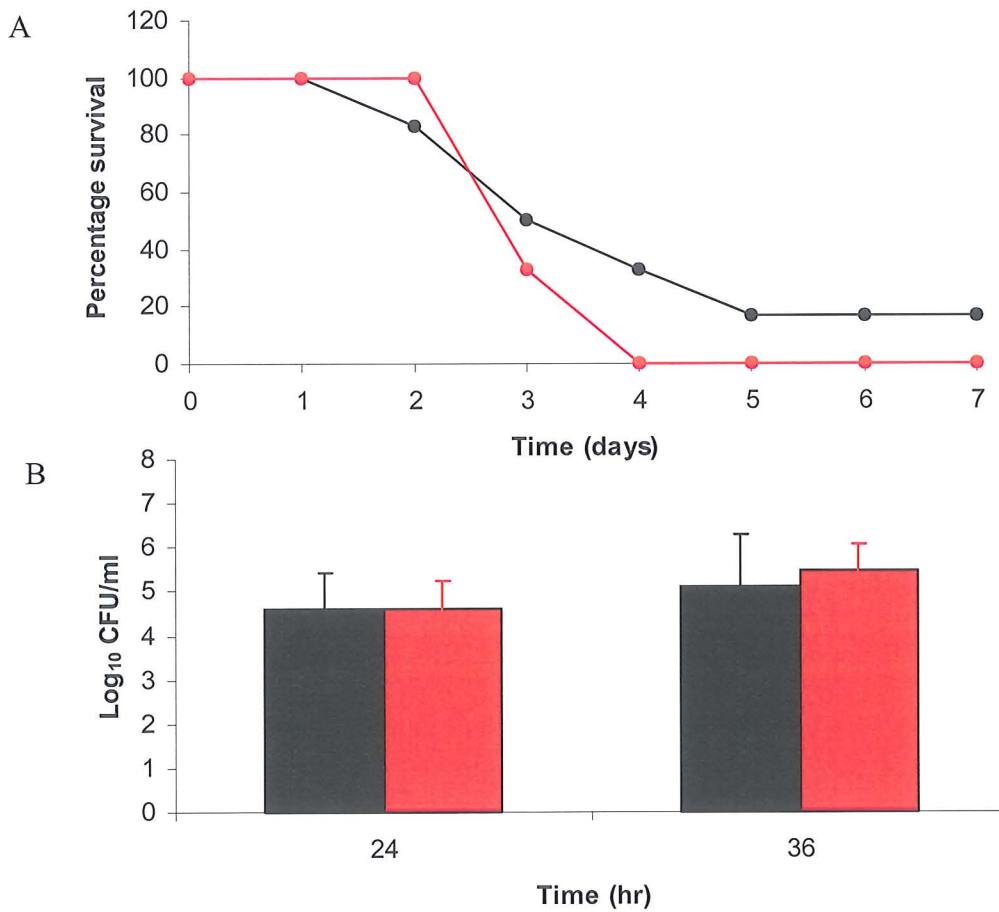
Figure 25: Analysis of virulence of Janus ZmpB LPXTG motif deletion mutant
Survival (A) and bacterial blood counts (B) where monitored following intranasal infection with 10^6 CFU. Wild type (SmR1) in black, mutant (strain JM) in red. Bacterial count data are mean +/- 1 SE. n=6. No significant difference was present between the strains.

Findings

- Allelic replacement mutants were made in the genes of four LPXTG proteins: *nanA*, *prtA*, *β-gal* and *spuA*.

- Blood and lung counts from these mutants following intranasal infection were not significantly different from wild type D39.

- In a competitive intranasal infection *nanA* and *β-gal* mutants were attenuated. p*rtA* and *pull* mutants were not attenuated.

- Janus mutagenesis was used to delete the N-terminal LPQTG motif of ZmpB.

- This deletion had no affect on virulence as determined by survival and bacterial blood counts.

128

Chapter 6 Analysis of LPXTG Proteins, Discussion

The limited effect on pneumococcal virulence seen with the *srtA* mutant in infections with single bacterial strains was surprising given the results from other organisms (discussed in Chpater 4). One of the possible explanations is that the pneumococcal LPXTG proteins anchored by SrtA do not have a major role in virulence in the mouse model used. Although previous studies suggests many of these proteins contribute to virulence (Table 2) this body of data has been accumulated from diverse studies using different bacterial and mouse strains, infection routes, doses and criteria for attenuation. This makes interpretation of the results with regards the *srtA* work difficult. Indeed, the only work looking at a LPXTG motif protein using the same model as this study for *srtA*, was examining *zmpB* mutants (Blue *et al.* 2003). Unlike the *srtA* mutant a *zmpB* mutant showed significant attenuation in pneumonia and bacteraemia infections. ZmpB carries a N-terminal cell wall sorting signal rather than the typical C-terminal and so it may not be processed by sortase. That work may therefore be irrelevant with regards to assessing the *srtA* results. Furthermore, there is also a lack of information on the role of LPXTG motif proteins in pneumococcal pathogenesis. For example, *spuA* (pullulanase) was identified as a virulence factor in pneumonia in the TIGR4 STM screen (Hava and Camilli 2002) but beyond that nothing is known about its influence and role(s) in pathogenesis. Indeed, little data exists on a role for pullulanase in other pathogens. For example, the degree of attenuation compared to the wild type or the contribution to other infection types or colonisation is unknown. So while STM has proved valuable in highlighting genes involved in virulence, a more thorough understanding of pathogenesis requires further detailed study. Together with this lack of information from animal virulence studies there is also little data from *in vitro* experiments such as host cell adhesion or phagocytic

killing assays that might illuminate the function of these proteins. Therefore, to be able to compare results from a *srtA* mutant with those from LPXTG motif protein mutants and the need for more comprehensive study of these proteins, null mutants were made in four pneumococcal LPXTG protein genes; *nanA, spuA, prtA* and *β-gal*. These four proteins were selected because they all contain typical C-terminal cell wall anchoring signals and are therefore likely to be anchored by SrtA. In the case of NanA and β-gal this has already been shown (Kharat and Tomasz 2003) and this work, Figure 14 NanA Western blot). Furthermore, all of these proteins identified as virulence factors in pneumococcal pneumonia in the TIGR4 STM screen and therefore may contribute to virulence in our model (Hava and Camilli 2002), although it is acknowledged that differences in bacterial strain, mouse strain and infection may invalidate this prediction. Antiserum is available three of these proteins, NanA (our laboratory), PrtA (Bethe *et al.* 2001) and SpuA (Bongaerts *et al.* 2000) while an activity assay exists for the remaining protein (β -gal) (Zahner and Hakenbeck 2000) which might be useful in the study of these proteins. Additionally, three have described activities (NanA, β-gal and SpuA), which may be valuable in elucidating their role in virulence by providing a rationale for any phenotypes shown by mutants and providing a platform for further work. For example, this might involve the deletion of an enzyme active site to determine if enzymatic activity is essential for a particular phenotype. Such work would be much more difficult in the study of conserved hypothetical or hypothetical proteins whereby any phenotype may be purely observational with no explanation or obvious route towards more detailed analysis. Finally, all four are known to be expressed by the pneumococcus in human as demonstrated by the development of convalescent sera against them (Zahner and

Hakenbeck 2000; Zysk *et al.* 2000). This supports the suggestion they play an important role in pneumococcal biology. Although it might be argued that all pneumococcal proteins are likely to be expressed during colonisation and this finding simply indicates these ones to be immunogenic. As with the *srtA* mutant, the LPXTG protein mutants were compared to wild type D39 with regards bacterial counts in the blood and lung following intranasal infection of MF-1 mice with 10^6 CFU at 36hr post infection. 36 hr was selected, as it is the latest time point under this infection at which samples can be taken from all mice. After this time point mice begin to die and so a full data set may not be obtainable. It was predicted that any attenuation would be evident at this time point. It may however, miss attenuation evident only early in infection. No statistical difference was seen for any of the four mutants compared to the wild type. This suggests that these proteins have a minor effect on virulence in this model. This may be a factor contributing to the limited affect seen with the *srtA* mutant (Chapters 3 and 4). Additionally, work with a hyaluronidase mutant (another pneumococal LPXTG motif protein)in our laboratory revealed no attenuation in terms of survival or blood counts in this infection model (A. Kerr personal communication). In this study only four LPXTG motif proteins were examined so it may be possible that others may have a much greater affect on virulence.

To further characterise the four LPXTG protein mutants made here, competitive infections were carried out between each individual mutant and wild type D39 as described previous for *srtA* (Chapter 3). Competitive indices were calculated at 36 hr. post infection. Both *nanA* and *β-gal* where out-competed by wild type with a mean CI of <0.35 and <0.26 respectively. This suggests they are attenuated compared to the wild

type. Although the mean CI of the *prtA* mutant was less than one (<0.61) this was not statistically significant suggesting that this mutation had no affect. However, as 5 out of 6 mice had a CI of less than one this mutant appears attenuated and it is likely repetition of this experiment would confirm this.

The results from the *spuA* mutant revealed that the mean CI was not significantly different from one, therefore revealing no attenuation of the mutant. However, there was considerable variation in the CI from individual mice with a strong polarisation whereby 50% showed decreased virulence and 50% showed increased virulence. The cause for such variation is unclear and repetition of this infection would have been of value if time had allowed. A possible contributing factor may be the presence of two additional pullulanase genes in the genome that may complement the loss of *spuA* (data not shown). Therefore, at least two of the LPXTG proteins examined, *nanA* and *β-gal* contribute to virulence in this model although attenuation could only be detected by the more sensitive competitive infections. No role for the other two, *spuA* or *prtA* could be confirmed. An area for future work would be to investigate these mutants further in terms of their contribution to pneumococcal biology. For example, examining the role for these LPXTG proteins in bacteraemia, carriage, meningitis and otitis media in addition to *in vitro* assays such as host cell adhesion and phagocytic killing. The importance of carbohydrate metabolism to the pneumococcus was shown in the genome sequence by the presence of numerous genes involved in sugar uptake and utilisation. As several LPXTG motif proteins such as *β-gal* and *spuA* have carbohydrate-degrading activity it may also be of value to investigate mutants in these with regards sugar utilisation and metabolism.

This study, demonstrating that *nanA* contributes to pneumococcal virulence, is in agreement with its identification as a virulence factor in the lung following STM screens with strains TIGR4 and G54 (Polissi *et al.* 1998; Hava and Camilli 2002). The work by Polissi *et al.* (1998) also compared the *nanA* mutant strain with wild type in a septicaemia model induced by intraperitoneal infection and found no difference as judged by mouse survival. This latter finding was later confirmed by Berry and Paton using strain D39 (Berry and Paton 2000). However, earlier work by this group, using the same experimental system found a significant, if small attenuation of a *nanA* mutant (Paton *et al.* 1997). This discrepancy was not explained in the subsequent study (Berry and Paton 2000). Additionally, a *nanA* mutant in D39, was shown to have reduced persistence in chinchilla nasopharyngeal carriage and was attenuated in an otitis media model in the same host (Tong *et al.* 2000). The ability of NanA to cleave N-acetylneuraminic acid from host cell glycoconjugates is proposed to expose surface receptors for the pneumococcus, thereby aiding mucosal colonisation and carriage (Andersson *et al.* 1983; Linder *et al.* 1994; Tong et al. 2001). Additionally, this activity may assist colonisation by decreasing the viscosity of mucous (Scanlon *et al.* 1989). The activity of NanA may also be directed against ecological competitors to aid nasopharyngeal colonisation (Shakhnovich *et al.* 2002). Desialylation of lipopolysaccharide from *H. pylori* and *N. meningitis* by NanA has been demonstrated and proposed to render these organisms more susceptible to complement clearance thereby giving the pneumococcus a competitive advantage (Shakhnovich *et al.* 2002). Intraperitoneal infection bypasses the requirement for adherence and colonisation steps and so this may explain the lack of phenotype seen with *nanA* mutants in this infection model. Supportive of a role for NanA in

pneumococcal biology is the finding that all clinical isolates tested were found to be positive for the production of NanA (O'Toole *et al.* 1971). A role for NanA has also been proposed in meningitis. In patients with pneumococcal meningitis, increased levels of N-acetylneuraminic acid, possibility released by NanA activity correlated with adverse outcome (O'Toole *et al.* 1971). Furthermore, intracerebal administration of purified pneumococcal neuraminidase in mice caused death (Kelly and Greiff 1970). However, a firm role for neuraminidase in pneumococcal meningitis remains to be established given some negative results reported using isogenic mutants in animal models (Winter *et al.* 1997; Wellmer *et al.* 2002). It should also be noted that some strains encode additional neuraminidase enzymes; NanB and NanC, which could mask the loss of NanA in mutants (Berry and Paton 2000; Shakhnovich et al. 2002). In the case of strain D39 used in this work, NanA and NanB are both present (Shakhnovich *et al.* 2002). Where and how these other enzymes are localised is as yet unclear, neither seems to be cell wall anchored due to the absence of a typical cell wall sorting signal.

This study did not support a conclusively a role for SpuA in pneumococcal virulence.

The only previous indication that SpuA might be a pneumococcal virulence factor came from the TIGR4 STM screen showing it to be involved in lung infection (Hava and Camilli 2002). The results presented here demonstrate no attenuation of a *spuA* mutant in D39 in intranasal challenges including a sensitive competitive infection. The reason for this discrepancy is unclear but may involve a combination of the use of different strains of mice or bacteria. This could be addressed by controlling for these factors in future experiments. Pneumococcal alkaline pullulanaseA has been postulated to contribute to colonisation or virulence by solublising mucus, altering the exposition of host

glycoconjugates, or scavenge potential carbon sources (Bongaerts *et al.* 2000). Its importance in pneumococcal biology is suggested by its serological conservation in all 41 strains examined (Bongaerts *et al.* 2000).

In the case of PrtA, the results presented here demonstrate no attenuation of a *prtA* mutant in D39 in intranasal challenges including in a sensitive competitive infection. However, there was a trend towards attenuation, which may become significant if the experiment was to be repeated. PrtA was found to be present in all 78 clinical isolates examined by PCR and sequencing (Bethe *et al.* 2001). It was also identified in the TIGR4 STM screen as a virulence factor in lung infection. In addition to that work, mice infected with a *prtA* mutant made in D39 showed improved survival after intraperitoneal infection (Bethe 2001). This result was confirmed by the work of Marra *et al.* (Marra *et al.* 2002) who discovered by DFI that *prtA* gene expression was up-regulated by iron starvation. Additionally, that study found no attenuation in virulence when comparing bacterial lung counts between wild type D39 and *prtA* mutant infected mice in agreement with our data (Marra *et al.* 2002) (Figure 18). The function of PrtA remains unclear. It is a serine protease belonging to the family of subtilisin-like protease called subtilase (Bethe *et al.* 2001). The highest NCBI Genbank BlastP hit against PrtA is an uncharacterised subtilisin-like serine proteases from *Thermoanaerobacter tengcongensis* (Accession NP_624131). Thereafter it shows limited local homology to *Streptococcal* C5a proteases involved in immune evasion, caseinases from *Lactobacillus* species and CspA from *S. agalactiae* which cleaves fibronection and protects the bacterium from phagocytosis (Harris *et al.* 2003). Investigation of these activities would therefore be of immediate interest in future studies with PrtA.

The results presented here confirm that *β-gal* contributes to pneumococcal virulence in the lung. *β-gal* had previously been identified as a pneumococcal virulence factor in the TIGR4 STM screen (Hava and Camilli 2002). Additionally, a D39 *β-gal* mutant showed transiently reduced bacterial counts in the lung after intranasal infection (Robertson *et al.* 2002). β-gal enzymes are often involved in lactose metabolism, however that does not appear to be the case for this pneumococcal enzyme and so it has been proposed it may act to alter host cell polysaccharides perhaps to expose bacteria binding sites or for nutrition (Zahner and Hakenbeck 2000).

The results in this work, showing no or limited contribution of these LPXTG proteins to virulence in this model may help explain why a *srtA* mutant had little affect on virulence. However, they are in contrast with the description of all four LPXTG motif proteins as being essential for lung infection following their identification as virulence factors in the STM screen of TIGR4 (Hava and Camilli 2002). Highly attenuated clones were identified in that work after two *in vivo* screens. In the first screen attenuated clones were identified as giving a reduced signal on the output blot compared to the input blot. These attenuated clones were then re-examined in a secondary STM screen and genes essential for infection identified as having no output signal. The STM results were validated when 16 out of a selection of 17 clones identified as attenuated were out competed by wild type in competitive lung infections. Examination of the mean CI from these competitive infections reveals that for some, the out competition by the wild type was small suggesting that some of these mutants would not be attenuated in straight single strain infections. In the MF-1 model used here such differences in bacterial counts would likely not be statistically significant. Indeed, as discussed previously, attenuation of a *pit2A*

mutant was not detected examining mouse death rates following intranasal infection but was apparent only in a competitive infections (Chapter 4). Notwithstanding the possible influence of different bacterial and mouse strains it appears the STM screen results do not represent exclusively genes essential to lung infection but includes also those with a more limited contribution. This sensitive detection of genes with a more subtle impact on virulence might be attributed to out-competition of attenuated clones by less attenuated ones in the secondary screen leading to the sensitive detection of virulence factors in a similar way to competitive infections with two strains.

The zinc metalloprotease ZmpB is an important pneumococcal virulence factor in mouse models of pneumonia and bacteraemia (Blue et al. 2003). Its function however, is so far unclear. The work of Blue *et al.* it may play a role in the spread of bacteria from the lung to blood or perhaps in the survival of bacteria following their dissemination from lung to blood (Blue *et al.* 2003). Additionally, a *zmpB* mutant was deficient in the induction of pulmonary TNF-α. The mechanism and significance of this is as yet unclear although ZmpB induced inflammation may contribute to pathology (Blue *et al.* 2003). ZmpB also contributes to bacterial survival and virulence following direct inoculation into the blood (Blue *et al.* 2003). As described, ZmpB possesses what appears to be a LPXTG cell wall sorting signal in the atypical position of the N-terminal region of the protein. Both ZmpC and IgA1 protease also have this feature, which suggests it may have functional importance. To investigate if this LPQTG sequence of the ZmpB N-terminus plays a role in the localisation of this protein and virulence, the motif was deleted using Janus mutagenesis. Deletion of the LPQTG motif coding sequence was successfully achieved in D39 SmR1 using Janus mutagenesis. This mutant, JM was then compared to the wild

138

type SmR1 with regards to mouse survival and blood counts following intranasal infection with 10^6 CFU given to separate groups of mice. No difference was apparent between the groups. The LPQTG sequence of ZmpB therefore does not appear to contribute significantly to the role of this protein in virulence. It would be of interest to see if this deletion alters the localisation of ZmpB. This may seem unlikely given that a *srtA* mutant had no apparent effect on ZmpB localisation together with the virulence data above suggest no role for SrtA in processing of ZmpB. Additionally, a *srtA* mutant (Chapter 3) showed none of the virulence attenuation that a *zmpB* mutant did (Blue *et al.* 2003). This too argues against a key role for SrtA in ZmpB biology.

Janus mutagenesis was used as this allows for the introduction of silent mutations into the genome. That is mutations such as this deletion, which lack a selectable marker. This feature makes Janus mutagenesis a potentially powerful technique to manipulate the pneumococcal genome, as there is little restriction with regards the alterations that can be made and it limits the generation of confounding artefacts. It is therefore likely to be used widely in the future. This is the first demonstration to our knowledge of Janus mutagenesis being utilised to examine pneumococcal virulence.

Chapter 7 Investigation of a putative macrophage infectivity potentiator protein and exfoliative toxin A, Results

Introduction

With the aim of identifying and investigating putative pneumococcal virulence factors two genes annotated in the genome sequences as macrophage infectivity potentiator protein and exfoliative toxin A were of interest as possible virulence factors. These genes were therefore investigated further.

Summary

The annotation of two adjacent pneumococcal genes, *spr0370* and *spr0371* as encoding a macrophage infectivity potentiator protein and exfoliative toxin A suggested these may be involved in virulence given the role of such genes in *Legionella pneumophila* and *S. aureus* respectively (Cianciotto and Fields 1992; Wintermeyer *et al.* 1995; Kohler *et al.* 2003; Gemmell 1995; Ladhani *et al.* 1999; Ladhani 2001). However, it appears these annotations may be incorrect and misleading with regards predicting the function of the pneumococcal genes. Sequence analysis revealed that *spr0370* may be an alkylhydroperoxidase (an enzyme acting to breakdown potentially toxic peroxide compounds). No probable function could be assigned to *spr0371*. RT-PCR analysis showed these genes were expressed as a bicistronic operon. PCR and sequencing found the two genes to be present and highly conserved in sequence among several pneumococcal clinical isolates. A null mutant lacking both genes was attenuated in mouse models of pneumonia and bacteraemia showing this operon to be a novel pneumococcal virulence determinate. Finally this double mutant had enhanced resistance to hydrogen peroxide supporting a role in the response to oxidative stress as suggested by the putative identification of *spr0370* as an alkylhydroperoxidase.

Basic features

For reference, the basic features of the pneumococcal genes investigated in this chapter, *spr0370* and *spr0371*, are shown in Table 8.

Locus Name	*spr0370*	*spr0371*
NCBI Accession	AAK99174	AAK99175
TIGR4 Locus Name	*sp*0409	*sp*0410
Putative Identification	Macrophage infectivity potentiator protein	Exfoliative toxin A
Coordinates	370629 to 371174	371275 to 372183
Gene Length (bp)	546	909
Protein length (aa)	182	303
Molecular Weight (Da)	19815.48	33303.29
Percent GC	43.77%	42.03%

Table 8: Basic characteristics of spr0370 and spr0371

Details relate to sequenced strain R6 as taken from the R6 annotation at the TIGR Comprehensive Microbial Resource CMR website.

Sequence Analysis of *spr0370-0371*

The probability that the annotations of *spr0370* and *spr0371* were indicative of their possible function was investigated. This was deemed especially important as no previous data suggested the presence of such genes and their associated activites in the pneumococcus. However, upon independent checking of these annotations it appears both may be incorrect and misleading with regards predicting the function of these pneumococcal genes.

The NCBI entry (AAB31024) used to annotate *spr0370* was named as macrophage infectivity potentiator (Mip) but the sequence for this entry is distinct from the well-characterised Mip (Accession S22665) and relates to a *L. pneumophila* protein of unknown function, identified due to its up-regulation within macrophages (Miyamoto *et*

al. 1993). This discrepancy was also noted by King *et al.* (King *et al.* 2001). In accordance with this inaccuracy the NCBI entry for this *L. pneumophila* protein has since been changed to a '24 kDa macrophage-induced major protein'. Therefore, despite *spr0370* showing limited homology to the genuine Mip, it and a number of other putative macrophage infectivity potentiator proteins remain in the databases due to annotation based on the flawed original NCBI entry. The proteins to which Spr0370 shows strongest homology are shown in Table 9, importantly none have a known or suggested function.

Organism	Accession[1]	Similarity	E Value	Length (aa[2])	Annotation
Streptococcus mutans	NP_721359	87% over 181 aa[2]	2e-75	183	Conserved hypothetical protein
Streptococcus suis	BAB85963	81% over 179 aa	2e-64	194	Conserved hypothetical protein
Actinobacillus pleuropneumoniae	ZP_00135195	75% over 182 aa	3e-52	180	Uncharacterized conserved protein
Burkholderia fungorum	ZP_00034893	69% over 181	3e-43	189	Hypothetical protein
Burkholderia fungorum	ZP_00034405	67% over 181 aa	7e-41	211	Hypothetical protein
Neisseria meningitis	NP_274031	64% over 183 aa	8e-41	227	Macrophage infectivity potentiator-related protein

Table 9: Highest BlastP hits from NCBI for R6 Spr0370 protein

[1] more than one entry from each organism may match enquiry due to multiple submissions and subtle variation in the sequence of these entries. Accession number given may therefore be representative of the best hits from a series of entries that appear to correspond to the same protein. [2] aa = amino acids. (Sept. 2003)

Using the TIGRFAM and Pfam Hits facility at the TIGR CMR, Spr0370 appears to contain an alkylhydroperoxidase AhpD family core domain. Additionally it may contain

carboxymuconolactone decarboxylase activity although, the match to this protein family is poor with the match score below the noise cutoff (Table 10).

TIGRFAM/Pfam Description	HMM Accession	Function	HMM Noise Cutoff	HMM Trusted Cutoff	Match Score	Match Evalue
Carboxymuconolactone decarboxylase	PF02627	Protocatechuate catabolism	-4.8	4	-14	0.91
Alkylhydroperoxidase AhpD family core domain	TIGR00778	Antioxidant activity	15	30	32.8	7.70E-06

Table 10 TIGRFAM/Pfam hits for Spr0370

The TIGRFAM/Pfam facility at TIGR CMR displays the protein families that proteins may belong to based on the Hidden Markov Model.

The alkylhydroperoxidase AhpD family core domain is characterised by a 20 residue sequence consisting of EXXXXXX[S/A]XXXXC[V/L]XCXXXH. This sequence is almost perfectly conserved in Spr0370 as EVVQITAAVTNGCA*FCVAGH except for an alanine instead of a valine or leucine (marked by asterisk). Only one protein containing this domain has been characterised, AhpD of *Mycobacterium tuberculosis* (Genbank Accession NP_216945) which has been shown to be involved in antioxidant protection (Hillas *et al.* 2000; Bryk *et al.* 2002; Nunn *et al.* 2002; Koshkin *et al.* 2003). Spr0370 and AhpD from *M. tuberculosis* show very little overall homology but do share the AhpD family core domain (Figure 26). Importantly, several residues within the AhpD family core domain shown to be involved in the activity of the *M. tuberculosis* protein are conserved in Spr0370. This suggests that Spr0370 may also possess alkylhydroperoxidase activity and therefore be involved in resistance to oxidative stress. Given that several other pneumococcal proteins involved in protection from oxidative

144

stress have been shown to be involved in virulence (discussed later in chapter 8) Spr0370

may also be a pneumococcal virulence factor.

A

1	M	T	T	F	T	H	T	V	E	S	A	-	P	A	E	V	K	E	I	L	E	T	V	E	K	D	N	N	G	Spr0370
1	M	S	I	E	K	L	K	A	A	L	P	E	Y	A	K	D	I	K	L	N	L	S	S	I	T	R	S	S	V	L	AhpD

30	Y	I	P	N	L	I	G	L	L	A	N	A	P	T	V	L	E	A	Y	Q	I	V	S	S	I	-	H	R	R	Spr0370
31	D	Q	E	Q	L	W	G	T	L	L	A	A	A	T	R	N	P	Q	V	L	A	D	H	G	A	E	A	T	AhpD

| 58 | N | S | L | T | P | V | E | R | E | V | V | Q | I | T | A | A | V | T | N | G | C | A | F | C | V | A | G | H | T | A | Spr0370 |
| 61 | D | H | L | S | A | A | A | R | H | A | A | L | G | A | A | A | I | M | G | M | N | N | V | F | Y | R | G | F | AhpD |

| 88 | F | S | I | K | Q | I | Q | M | N | D | D | L | I | Q | A | L | R | N | R | T | P | I | E | T | D | - | P | K | L | D | Spr0370 |
| 91 | L | E | G | R | - | - | I | Y | D | D | L | R | P | G | L | R | M | N | I | I | A | N | P | G | I | P | K | A | N | AhpD |

| 117 | - | T | L | A | K | F | T | L | A | V | I | N | T | K | G | R | V | G | D | E | A | L | S | E | F | L | E | A | G | Y | Spr0370 |
| 117 | F | E | L | W | S | F | A | V | S | A | I | N | G | C | S | H | C | L | V | A | H | E | H | T | L | R | T | V | G | V | AhpD |

| 146 | T | Q | N | A | L | D | V | V | F | G | V | S | L | A | I | L | C | N | Y | A | N | N | L | A | N | T | P | I | N | Spr0370 |
| 147 | D | R | E | A | I | F | E | A | L | - | - | K | A | A | A | I | V | S | G | V | A | Q | A | L | A | T | I | E | - | - | AhpD |

| 176 | P | E | L | Q | P | Y | A | Spr0370 |
| 173 | - | A | L | S | P | - | S | AhpD |

B

| Spr0370 | 66 | E | V | V | Q | I | T | A | A | V | T | N | G | C | A | F | C | V | A | G | H | 85 |
| AhpD | 118 | E | L | W | S | F | A | V | S | A | I | N | G | C | S | H | C | L | V | A | H | 137 |

Figure 26: Alignment of pneumococcal protein Spr0370 against *M. tuberculosis* AhpD
A) Full length alignment showing limited overall homology. B) Local alignment of AhpD active sites, residues highlighted by red arrows have been shown to be involved in activity of the *Mycobacterium* AhpD protein. The Spr0370 sequence used for the active site alignment is highlighted by a blue line in the full sequence alignment with the *Mycobacterium tuberculosis* AhpD active site sequence highlighted by a blue line. Alignments were performed using MegAlign Clustal IV (PAM250). Matching residues highlighted in black. Numbers refer to residue position relative to the amino terminus. NCBI Accession Numbers; Spr0370, AAK99174; AhpD NP_216945.

The best database homologues to Spr0371 are shown in Table 11. The presence of several putative exfoliative toxins in these matches would suggest that Spr0371 is indeed an exfoliative toxin. However, an alignment of Spr0371 against the best-studied exfoliative toxin, ETA from *Staphylococcus aureus* (NCBI Accession P09331) showed very little homology (data not shown). This prompted carefully analysis of the basis for the Spr0371 annotation. It was found to be annotated based on homology with *Staphylococcus hyicus* exfoliative toxin A (SHETA) (NCBI Accession BAB08178) (Table 11). This SHETA was subsequently shown to have very little sequence homology to other exfoliative toxins, which extended to the absence of the catalytic triad found in other exfoliative toxins (Yamaguchi *et al.* 2002). It was therefore important to verify this sequence did represent an exfoliative toxin. However, the literature reference for the NCBI Genbank sequence entry appears incorrect and relates to a paper from the same workers detailing sequencing of *S. hyicus* exfoliative toxin B not A. Therefore it appears no literature reference exists supporting this sequence as an exfoliative toxin. Attempts to clarify this with the group concerned were unsuccessful. There is therefore no sequence or experimental evidence that SHETA is an exfoliative toxin. Several bacterial open reading frames, including Spr0371 have been annotated against SHETA as exfoliative toxins yet there is no evidence that they are. Furthermore, the previously characterised exfoliative toxins are hydrophilic in nature in accord with their presence in culture supernatants (Gemmell 1995; Ladhani *et al.* 1999; Ladhani 2001). However, SHETA, Spr0371 as well as the other related proteins (Table 11) are all predominantly hydrophobic in nature and likely to membrane proteins (Figure 27). This would support the suggestion they are not exfoliative toxins. Additionally there is no indication as to

what the SHETA sequence is as a BlastP search gives the same hits as the Spr0371 search

(Table 11).

Organism	Accession[1]	Similarity	E Value	Length (aa[2])	Annotation
Streptococcus agalactiae	NP_688224	55% over 302 aa	3e-27	305	Putative exfoliative toxin A
Staphylococcus aureus	NP_371697	55% over 304 aa	1e-25	315	Conserved hypothetical protein
Streptococcus hyicus	BAB08178	56% over 305 aa	3e-25	306	Exfoliative toxin A
Streptococcus mutans	NP_722416	56% over 304 aa	2e-24	314	Conserved hypothetical protein
Streptococcus pyogenes	NP_664436	49% over 305 aa	3e-23	302	Putative exfoliative toxin A
Lactococcus lactis subsp. lactis	NP_266534	49% over 305 aa	2e-15	324	Permease

Table 11: Highest BlastP hits from NCBI against R6 Spr0371 protein

[1] more than one entry from each organism may match enquiry due to multiple submissions and subtle variation in the sequence of these entries. Accession given is therefore representative of best hits from a series of entries that appear to correspond to the same protein. aa= amino acids. (Sept. '03.).
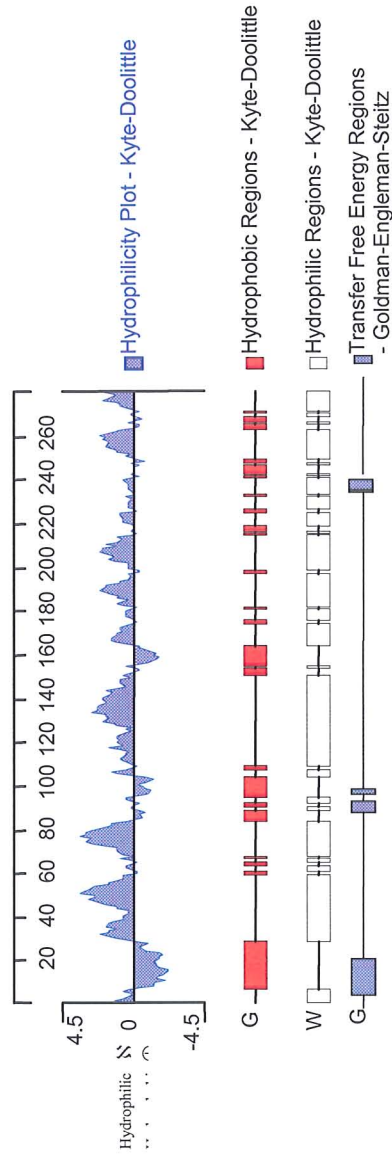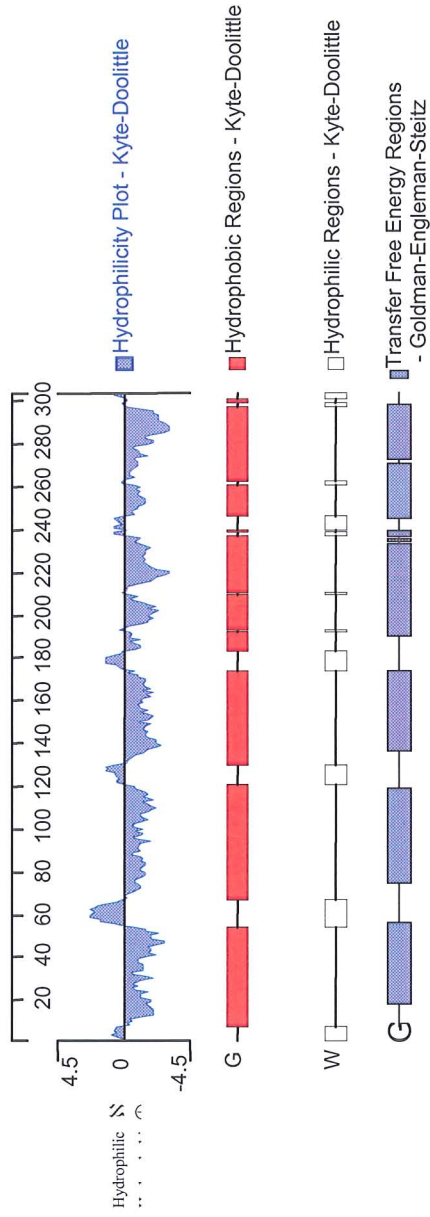
Figure 27: Hydrophilic/hydrophobic analysis of Spr0371 and *S. aureus* ETA. Performed using DNAStar Protean program with default window settings of 9 for Kyte-Doolittle and 20 for Goldman-Engleman-Steitz. Scale bar indicates residue position relative to amino terminus. The Kyte-Doolittle plots show predicted hydrophobic and hydrophilic regions. The Goldman-Engleman-Steitz plots shows predicted transmembrane domains.

The *Lactobacillus lactis* permease showing homology to Spr0371 is uncharacterised and seems to be annotated as such due to its hydrophobic nature and likelihood to be a membrane protein. All TIGRFAM and Pfam Hits at the TIGR CMR for Spr0371 were on or beneath the noise cutoff so it is hard to judge if these matches have biological significance (Table 12). Therefore any putative function of Spr0371 is unclear.

TIGRFAM/Pfam Description	HMM Accession	Function	HMM Noise Cutoff	HMM Trusted Cutoff	Match Score	Match Evalue
C4-dicarboxylate transporter/malic acid transport protein	PF03595	Dicarboxylate transporter/malic acid transport/tellurite resistance	-76.6	-65	-112.3	0.00036
Domain of unknown function (DUF389)	PF04087	Family of hypothetical bacterial proteins with an undetermined function	-54.7	204.5	-62.4	0.42
Got1-like family	PF04178	Membrane proteins involved in traffic to the Golgi complex	-47	185.2	-47	0.053
C4-dicarboxylate transporter/malic acid transport protein	TIGR00816	Dicarboxylate transporter/malic acid transport/tellurite resistance	40	300	-119.5	0.41

Table 12: TIGRFAM/Pfam hits for Spr0371

Transcript Analysis

Functionally related genes are often arranged together in the genome as operons and co-transcribed together in the same mRNA. RT-PCR was used to investigate if this was the case for *spr0370* and *spr0371*.

Between the stop codon of *spr0370* and the start codon of *spr0371* there are 96 base pairs. This close proximity and the apparent absence of a rho-independent potential terminator (as determined from TIGR CMR) between them suggested that they might be transcribed together in a single mRNA transcript. To test this, RNA was extracted from D39 grown to exponential phase and converted to cDNA using random primers. PCR

with various primer pairs was performed to analyse the transcript(s) (Figure 28). Both *spr0370* and *spr0371* were detected showing that they were expressed under the conditions used and are not pseudogenes. The results also showed that *spr0370-0371* were transcribed together given that primer pair B, spanning between the two genes gave the expected product with cDNA. It appears this transcript is bicistronic and does not include the upstream or downstream genes because in the case of the upstream gene *spr0369* (D-alanine glycine permease), no product was found using primer pair D, spanning from *spr0369* to *spr0370* on cDNA yet the expected product was amplified from gDNA. In the case of *spr0371*, the downstream gene seryl-tRNA synthetase is transcribed in the opposite direction to *spr0370-0371* so it is unlikely to be transcribed in the same transcript. Additionally both *spr0371* and *spr0372* poses downstream rho-independent transcription terminators (as determined from TIGR CMR). Importantly, a control in which reverse transcriptase was absent yielded no product with any primer pair confirming the absence of contaminating gDNA. Primers pairs specific for *spr0370* and *spr0371* failed to give a product (data not shown) when used on gDNA from a *spr0370-0371* deletion mutant (see below). Therefore *spr0370-0371* represents a pneumococcal operon that is expressed by D39 *in vitro* during exponential growth.
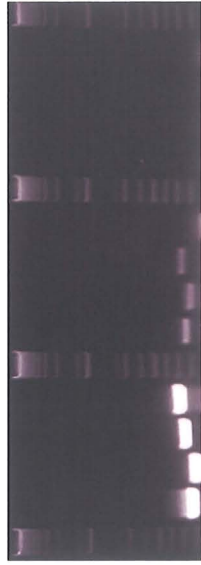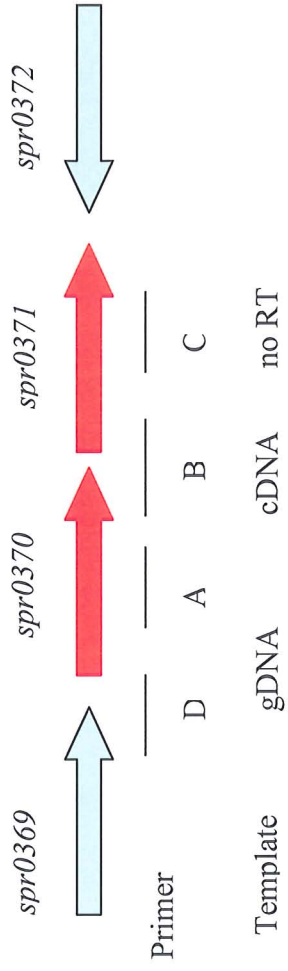
151

Figure 28: RT-PCR Analysis of *spr0370-0371* cDNA was prepared from D39 in exponential growth and subjected to PCR with the primer pairs as detailed. Templates; gDNA genomic DNA, cDNA complementary DNA and noRT cDNA prepared in the absence of reverse transcriptase.

Mutant construction and verification

To investigate the possible role(s) of the genes *spr0370* and *spr0371* in pneumococcal biology a double mutant strain lacking both of these genes was made by allelic replacement. This was done in D39 using the same strategy as for the *srtA* mutant (Figure 1). In this case, primers 17Q and 17R were used for the initial PCR and primers 17S and 17T for the inverse PCR. The final plasmid construct was used to transform D39 and potential transformants were selected for on erythromycin ($1 \mu g \ ml^{-1}$). Transformants were then checked by PCR using the primers 18I and 18Y which are located out with the manipulated region giving different sized bands from wild type and mutant (Figure 29). In addition confirmation this PCR product was sequenced and matched the expected sequence (data not shown). Mutants were readily obtaining showing that neither gene was essential for viability.

1650bp
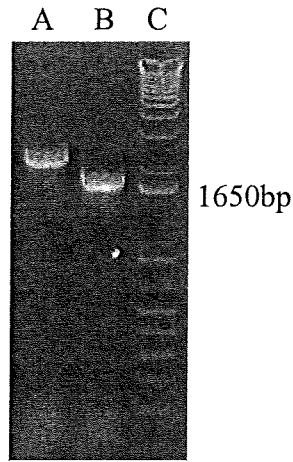
Figure 29: PCR confirmation of *spr0370-0371* double knock out
Primers 18I and 18Y located upstream and downstream respectively of the manipulated
region were used to confirm the allelic replacement of *spr0370-0371* as determined by the
different product sizes. Lane A wild type gDNA, B mutant gDNA C ladder. The expected
product sizes are wild type 2349bp and mutant 1844bp.

Analysis of virulence

The survival of MF-1 mice following intranasal challenge with 10^6 CFU was compared

when infected with D39 wild type and *spr0370-0371* double mutant strains. Although the

mutant infected mice showed a trend towards greater survival time (median 120 hours

post-infection n=11) compared to wild type infected mice (median 78 hours n=11) this

was not statistically significant (Figure 30). However, when blood counts were compared

from these mice the mutant was shown to be significantly attenuated (Figure 30). This

difference was only statistically significant at 36 hr. post infection and not at the earlier

time points of 24 and 30 hr, although the mean counts at these time points were still

lower for the mutant compared to the wild type.
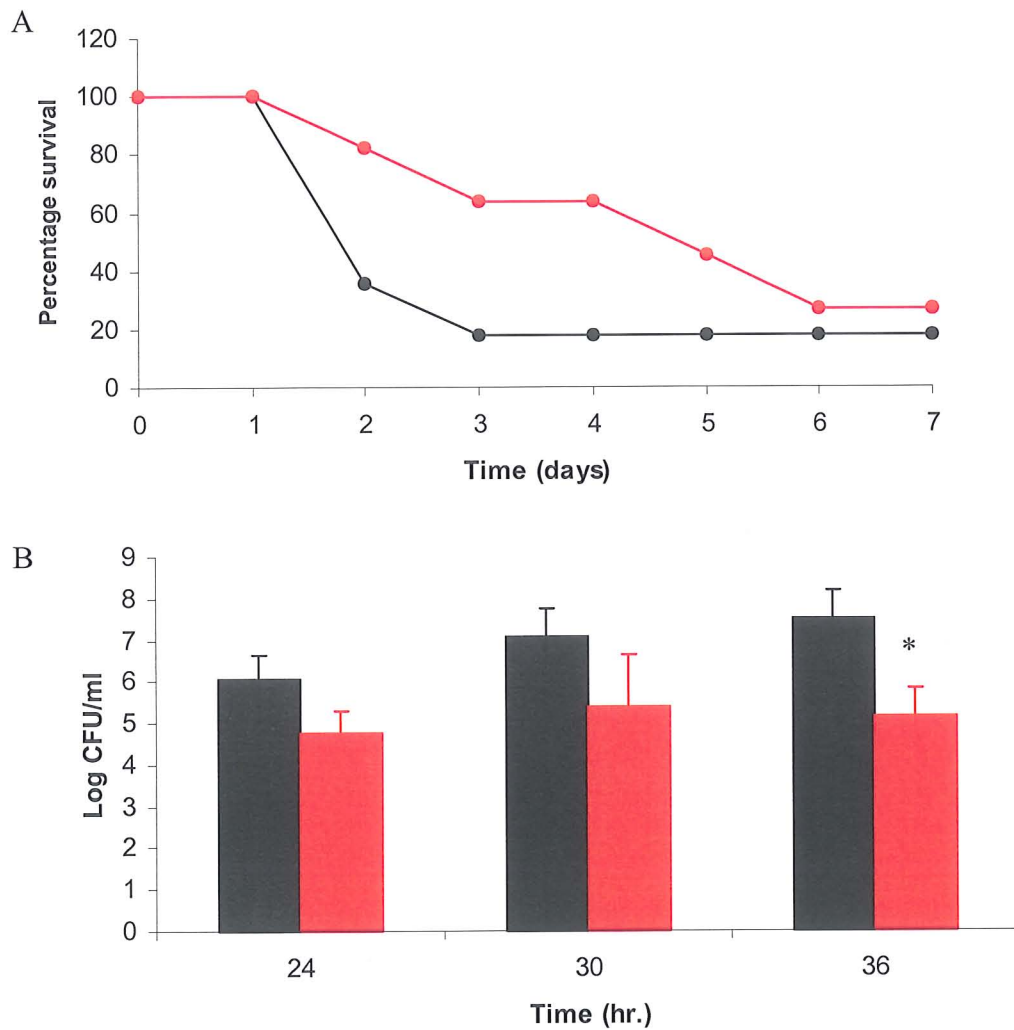
Figure 30: Survival (A) and bacteria blood counts (B) following intranasal challenge with wild type D39 (black) and *spr0370-0371* double mutant (red) Bacterial count data represent the mean +/- SE . * P<0.05 compared to wild type.

This showed that the *spr0370-0371* operon contributes to pneumococcal virulence in a murine model of pneumonia. To further characterise the virulence defect of the *spr0370-0371* mutant competitive infections were used as done for our *srt*A mutant (Chapter 3).

1:1 mixtures of wild type D39 and *spr0370-0371* double mutant were mixed and administered intranasally or intravenously to MF-1 mice. The ratio of recovered mutant and wild type was then calculated and corrected by the input ratio to give the competitive index. When this value is less than one the mutant is attenuated, where a value of greater than one indicates enhanced virulence of the mutant and a competitive index (CI) of one demonstrates no difference in virulence. Following intranasal challenge the CI at 0hr and 6hr was not significantly different from 1 (Figure 31). However, at 12, 24 and 36 hr the CI was significantly less than one showing attenuation of the mutant. At 36 hr the CI had fallen to <0.07 indicating that wild type out-competed the mutant by >14-fold. The CI after intravenous infection was not significantly less than one at 0, 6 or 12 hours post infection. Significant attenuation of the mutant was seen however, at 24 and 36 hr post infection. The CI at 0hr after both infections was not significantly different from one showing that the preparation and infection procedure themselves were not the cause of any out-competition of one strain by the other. The *spr0370-0371* operon therefore appears to contribute to the virulence of D39 in both experimental pneumonia and bacteraemia. Comparison of the CI from each infection at each time point found was no statistical difference at any time point between the CI found following intranasal or intravenous infection. This is despite the CI at 12 hr post intranasal infection being significantly less than one while the corresponding data following intravenous infection is not significantly different from one. *spr0370-0371* would therefore appear to contribute equally to both infection types, although that does not necessary mean the type of contribution is the same in both infections.

A

| 1.09 | 1.01 | 0.58* | <0.31** | <0.07** |

Mean competitive index

B

| 0.93 | 0.89 | 1.04 | <0.37** | <0.14** |

Mean competitive index
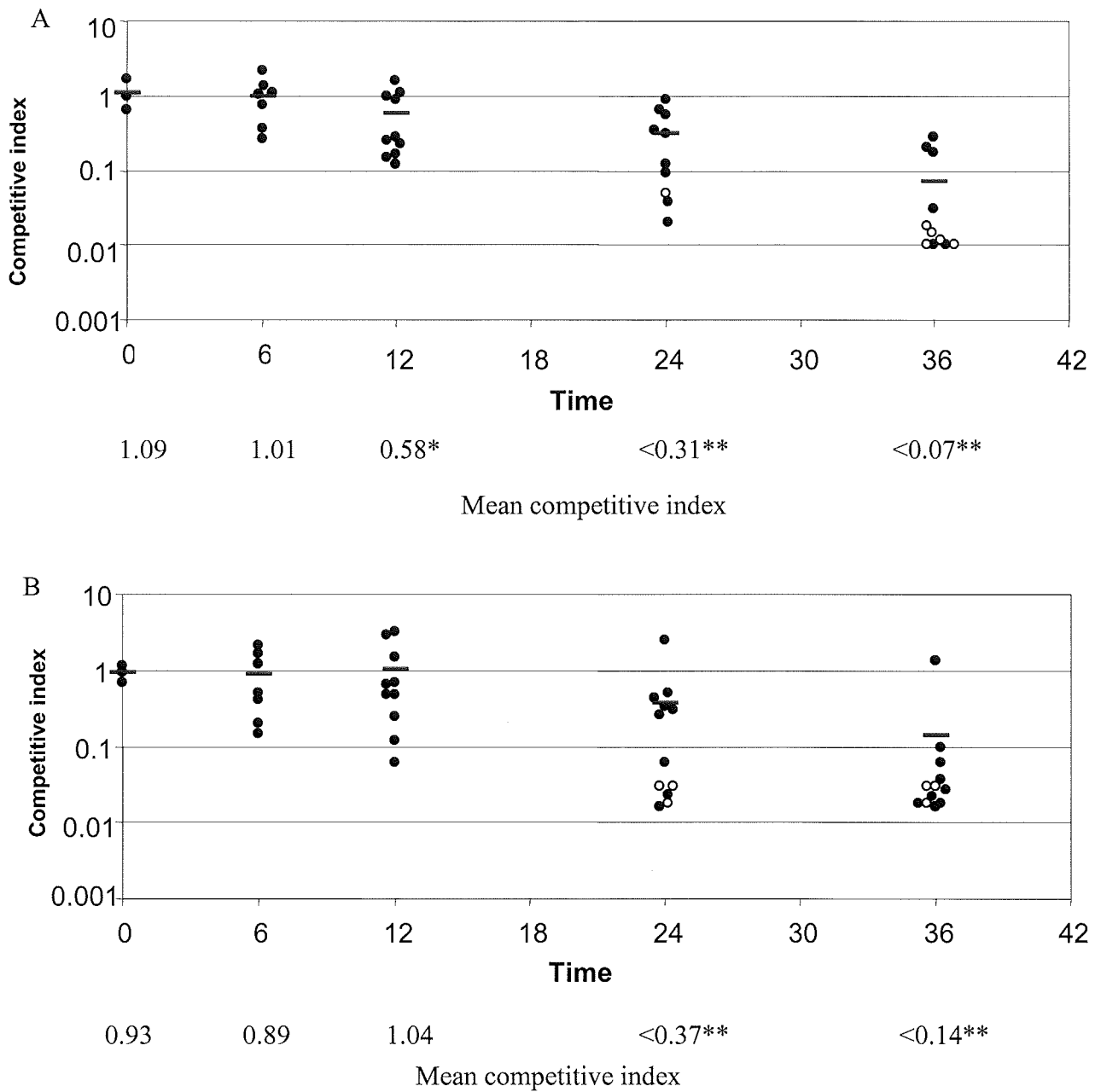
Figure 31: Virulence analysis of *spr0370-0371* double mutant in competitive infections with parental wild type D39
Each point indicates the competitive index from an individual mouse following intranasal (A) and intravenous (B) challenges. Values <1 indicates attenuation of the mutant. Following intranasal challenge, bacteria were sampled from the lung. Following intravenous challenge bacteria were sampled from the blood. Red bar shows the mean CI and open circles indicate the recovery of no mutant bacteria. *p<0.05 **p<0.025 compared to one (one equals no attenuation).

Growth in vitro

To determine if the out-competition of the mutant *in vivo* was due to a general growth defect the CI was calculated following growth in BHI at 37 °C. At mid-log and late log stages of growth the mean CI where 0.99 and 1.07 (n=2) showing that under these conditions the mutant is not out-competed and suggesting it does not have a general growth defect.

Strain distribution and sequence diversity

Both of the fully sequenced and annotated pneumococcal genomes contain *spr0370* and *spr0371*. To investigate whether this was true of a wider range of pneumococcal strains a selection of twenty clinical isolates from those used earlier to investigate sortase distribution where examined for the presence of *spr0370* and *spr0371* by PCR. The strains were selected to represent a variety of serotypes and sequence types, (Table 13).

Strain #	MLST	Year	Serotype	Specimen	Dianogsis
5	74	96	2	Blood	-
11	180	01	3	Ear swab	-
14	584	01	4	Blood	Pneumoina
16	569	97	5	-	-
17	96	01	6	Nasal	-
26	578	01	8	Blood	Bacteraemia
30	156	00	9	-	-
34	new	01	10	Blood	-
39	408	01	11	Blood	-
43	586	01	12	Blood	Bacteraemia
53	583	00	15	Blood	Bacteraemia
62	392	00	17	Blood	Pneumoina
65	113	00	18	Blood	Bacteraemia
69	199	00	19	-	Chest infection
74	567	00	20	Blood	Chest infection
79	433	01	22	Blood	Septicaemia
84	573	01	23	Ear isolate	-
90	571	01	27	Blood	Meningitis
91	new	01	29	Tracheal aspirate	-
119	574	96	13	-	-

Table 13: Strains examined for distribution of *spr0370-0371*

Strain # is laboratory reference. MLST is multisequence locus sequence type.
Year is year of isolation. – indicates data not available. New MLST is a new
sequence profile not yet assigned reference number

Using the primers 21X and 22C which lie within *spr0370* and *0371* respectively all 20
strains gave the expected product of ~800bp (Figure 32). Amplification using *spr0370-0371* mutant gDNA gave no product.

850b

L A B C   L
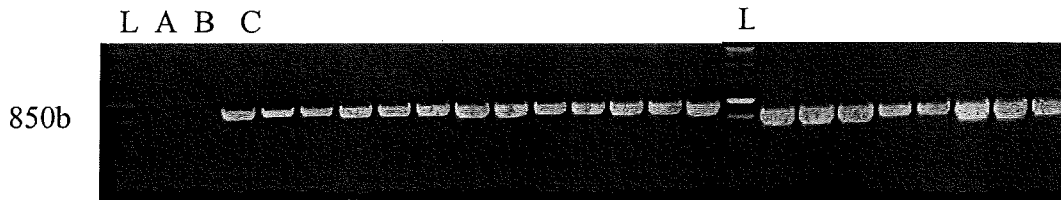
Figure 32: Strain distribution of *spr0370-0371* in various pneumococcal clinical isolates
Distribution of *spr0370-0371* was assessed by PCR with
primers 21X and 21Y which give an expected product of ~800bp.
Lane L, ladder marker, A gDNA preparation from sterile BHI, B gDNA
from *spr0370-0371* mutant C R6 gDNA as positive control. Other lanes are
the 20 clinical isolate as shown in Table 13 in numerical order.

In further confirmation, internal primer pairs for *spr0370* (21X & 21Y) and *spr0371* (22B

& 22C) gave the same expected product size from all strains (data not shown). Reactions

with these primers on gDNA from the *spr0370-0371* mutant yielded no product. Thus all

twenty strains examined possess *spr0370* and *spr0371* and they are found adjacent to

each other as seen in the sequenced strains R6 and TIGR4. Using the NCBI facility

(http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) a BlastN search for *spr0370* and

spr0371 was performed against the unfinished pneumococcal genomes from strains 23F

and 670 (done 24/9/03). Neither strain possessed either of the two genes. However, given

their ubiquitous distribution in our clinical isolates this is most likely due to the

unfinished status of these sequencing projects. To gain insight into the sequence diversity

on *spr0370-0371* the two genes were sequenced from five strains (strain numbers

5,17,30,69 and 79, Table 13). For this sequencing the primers 17Q and 17R (found

upstream and downstream of the operon respectively) were used to amplify the operon

spr0370-0371 from gDNA. The five strains selected for sequencing all gave the expected

product of ~1.9kb. Again strains were selected to represent different serotype and MLST sequence types. From this PCR product the two genes were sequenced using the primers 17Q, 21X, 28I, 28H. Sequencing showed that both *spr0370* and *spr0371* were highly conserved in these strains (Figures 33,34,35 and 36).

```
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  R6
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  TIGR4
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  5
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  17
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  30
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  69
  1  A T G A C A A C A T T T A C T A T C C A T A C A G T A G A G T C A G C A C C A G C A G A A G T G A A  79

 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  R6
 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  TIGR4
 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  5
 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  17
 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  30
 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  69
 51  A G A A A T T C T T G A A A C A G T A G A A A A A G A C A A C A A T G G C T A T A T T C C C A A C C  79

101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  R6
101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  TIGR4
101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  5
101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  17
101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  30
101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  69
101  T A A T C G G T C T C T T G G C C A A T G C C C C G A C T G T T T T A G A A G C C T A C C A A A T T  79

151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  R6
151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  TIGR4
151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  5
151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  17
151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  30
151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  69
151  G T C T C A T C T A T C C A C C G T C G C A A C A G C C T G A C A C C C G T T G A G C G T G A A G T  79

201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  R6
201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  TIGR4
201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  5
201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  17
201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  30
201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  69
201  G G T G C A A A T C A C G G C A G C C G T G A C C A A T G G T T G T G C C T T C T G T G T C G C A G  79

251  G T C A C A C A G C C T T T T T C C A T C A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  R6
251  G T C A C A C A G C C T T T T T C C A T C A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  TIGR4
251  G T C A C A C A G C C T T T T T C C A T C A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  5
251  G T C A C A C A G C C T T T T T C C A T C A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  17
251  G T C A C A C A G C C T T T T T C C A T ▉ A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  30
251  G T C A C A C A G C C T T T T T C C A T C A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  69
251  G T C A C A C A G C C T T T T T C C A T C A A A C A A A T C C A G A T G A A T G A T G A C T T G A T T  79

301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A C  R6
301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A C  TIGR4
301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A ▉  5
301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A C  17
301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A C  30
301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A C  69
301  C A A G C T C T T C G C A A T C G T A C T C C A A T T G A A A C A G A T C C T A A A T T G G A T A C  79

351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  R6
351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  TIGR4
351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  5
351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  17
351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  30
351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  69
351  C C T A G C T A A G T T T A C C T T G G C A G T T A T C A A T A C C A A G G G T C G T G T A G G A G  79

401  A T G A A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  R6
401  A T G A A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  TIGR4
401  A T G A₎A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  5
401  A T G A A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  17
401  A T G A A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  30
401  A T G A A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  69
401  A T G A A G C C T T G T C T G A G T T T T T A G A A G C T G G C T A C A C T C A A C A A A A T G C C  79

451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  R6
451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  TIGR4
451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  5
451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  17
451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  30
451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  69
451  T T G G A T G T G G T T T T T G G T G T C A G C C T A G C A A T C C T C T G T A A C T A T G C C A A  79

501  C A A C T T A G C T A A T A C A C C A A T T A A T C C A G A A T T G C A A C C T T A T G C C T A A   R6
501  C A A C T T A G C T A A T A C A C C A A T T A A T C C A G A A T T G C A A C C T T A T G C C T A A   TIGR4
501  C A A C T T A G C T A A T A C A C C A A T T A A T C C A G A A T T G C A A C C T T A T G C C T A A   5
501  C A A C T T A G C T A A T A C A C C A A T T A A T C C A G A₎A T T G C A A C C T T A T G C C T A A   17
501  C A A C T T A G C T A A T A C A C C A A T T A A T C C A G A₎A T T G C A A C C T T A T G C C T A A   30
501  C A A C T T A G C T A A T A C A C C A A T T A A T C C A G A₎A T T G C A A C C T T A T G C C T A A   69
501  C A A C T T A G C T A A T A C A C C A A T T A A T C C'A G A A T T G C A A C C T T A T G C C T A A   79
```

Figure 33: Alignment of nucleotide sequence of *spr0370* from various pneumococcal strains

Figures on left relate to nucleotide position from gene start, strain name given on right. Bases in disagreement with the consensus are highlighted in black. Performed with DNAstar MegAlign using ClustalV(Identity). R6 and TIGR4 sequences taken from TIGR CMR.

162

```
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  R6
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  TIGR4
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  5
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  17
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  30
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  69
  1  M T T F T I H T V E S A P A E V K E I L E T V E K D N N G Y I P N L I G L L A N A P T V L E A Y Q I  79

 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  R6
 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  TIGR4
 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  5
 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  17
 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  30
 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  69
 51  V S S I H R R N S L T P V E R E V V Q I T A A V T N G C A F C V A G H T A F S I K Q I Q M N D D L I  79

101  Q A L R N R T P I E T D P K L D T L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  R6
101  Q A L R N R T P I E T D P K L D T L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  TIGR4
101  Q A L R N R T P I E T D P K L D I L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  5
101  Q A L R N R T P I E T D P K L D T L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  17
101  Q A L R N R T P I E T D P K L D T L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  30
101  Q A L R N R T P I E T D P K L D T L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  69
101  Q A L R N R T P I E T D P K L D T L A K F T L A V I N T K G R V G D E A L S E F L E A G Y T Q Q N A  79

151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        R6
151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        TIGR4
151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        5
151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        17
151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        30
151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        69
151  L D V V F G V S L A I L C N Y A N N L A N T P I N P E L Q P Y A                                        79
```

Figure 34: Alignment of deduced amino acid sequence of Spr0370 from various pneumococcal strains

Figures on left relate to residues position from amino terminal, strain name given on right.
Residues in disagreement with the consensus are highlighted in black. Performed with DNAstar MegAlign using ClustalV(PAM250). R6 and TIGR4 sequences taken from TIGR CMR.

Figure 35: Alignment of nucleotide sequence of *spr0371* from various pneumococcal strains

Figures on left relate to nucleotide position from gene start, strain name given on right. Bases in disagreement with the consensus are highlighted in black. Performed with DNAtor MegAlign using ClustalV(Identity). R6 and TIGR4 sequences taken from TIGR CMR.

```
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L T G L V L W I Y F L I   R6
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L[I]G L V L W I Y F L I   TIGR4
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L T G L V L W I Y F L I   5
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L T G L V L W I Y F L I   17
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L T G L V L W I Y F L I   30
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L T G L V L W I Y F L I   69
  1   L E V R M K K L P L V F S G C L L G L A G A G N L I L D T L P V L S H L F S L T G L V L W I Y F L I   79

 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   R6
 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   TIGR4
 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   5
 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   17
 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   30
 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   69
 51   L H L F N W K E T K Q E L T K P P L L S G M A T F P M A G M I L S T Y V F R V F S Y L P L V A Q G I   79

101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   R6
101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   TIGR4
101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   5
101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   17
101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   30
101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   69
101   W W F S F L L D L T L I A G F T I K F A C P G R R V H A T P S W T V L Y V G I A V A A L T Y P L V G   79

151   I I E I A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   R6
151   I I E I A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   TIGR4
151   I I E I A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   5
151   I I E I A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   17
151   I I E I A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   30
151   I I E[M]A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   69
151   I I E I A Y A T L S F G F L L T F Y L Y P L I Y S D L K K H P L P L A L L G Q E G I Y C A P F S L L   79

201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   R6
201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   TIGR4
201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   5
201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   17
201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   30
201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   69
201   L A S L V R V G G T S L P T W V L I V M I L A S Q S F F F F V L T R L P N I L K Q G F Q P A F S A L   79

251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L[T]E T I I C L T I L F F V L G A Y L I W L R   R6
251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L A E T I I C L T I L F F V L G A Y L I W L R   TIGR4
251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L[T]E T I I C L T I L F F V L G A Y L I W L R   5
251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L A E T I I C L T I L F F V L G A Y L I W L R   17
251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L A E T I I C L T I L F F V L G A Y L I W L R   30
251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L A E T I I C L T I L F F V L G A Y L I W L R   69
251   T F P T I I T A T S L K M A Q G I L K L P F L D Y L V L A E T I I C L T I L F F V L G A Y L I W L R   79

301   K K V                                                                                                R6
301   K K V                                                                                                TIGR4
301   K K V                                                                                                5
301   K K V                                                                                                17
301   K K V                                                                                                30
301   K K V                                                                                                69
301   K K V                                                                                                79
```

Figure 36: Alignment of deduced amino acid sequence of Spr0371 from various pneumococcal strains

Figures on left relate to residues position from amino terminal, strain name given on right. Residues in disagreement with the consensus are highlighted in black. Performed with DNAstar MegAlign using ClustalV(PAM250). R6 and TIGR4 sequences taken from TIGR CMR.

Susceptibility to killing by hydrogen peroxide

The apparent presence of an AhpD domain in Spr0370 (Figure 26) suggests that this protein could be involved on resistance to oxidation stress. To test this, wild type and the *spr0370-0371* double mutant D39 were exposed to hydrogen peroxide at different doses and the percentage survival measured by viable count. The results showed that at the doses used the mutant was more resistant to killing by hydrogen peroxide (Figure 37). This difference was significant at both 18.75 and 37.5mM hydrogen peroxide. At the higher doses of 75 and 150mM no viable pneumococci were recovered (not shown).
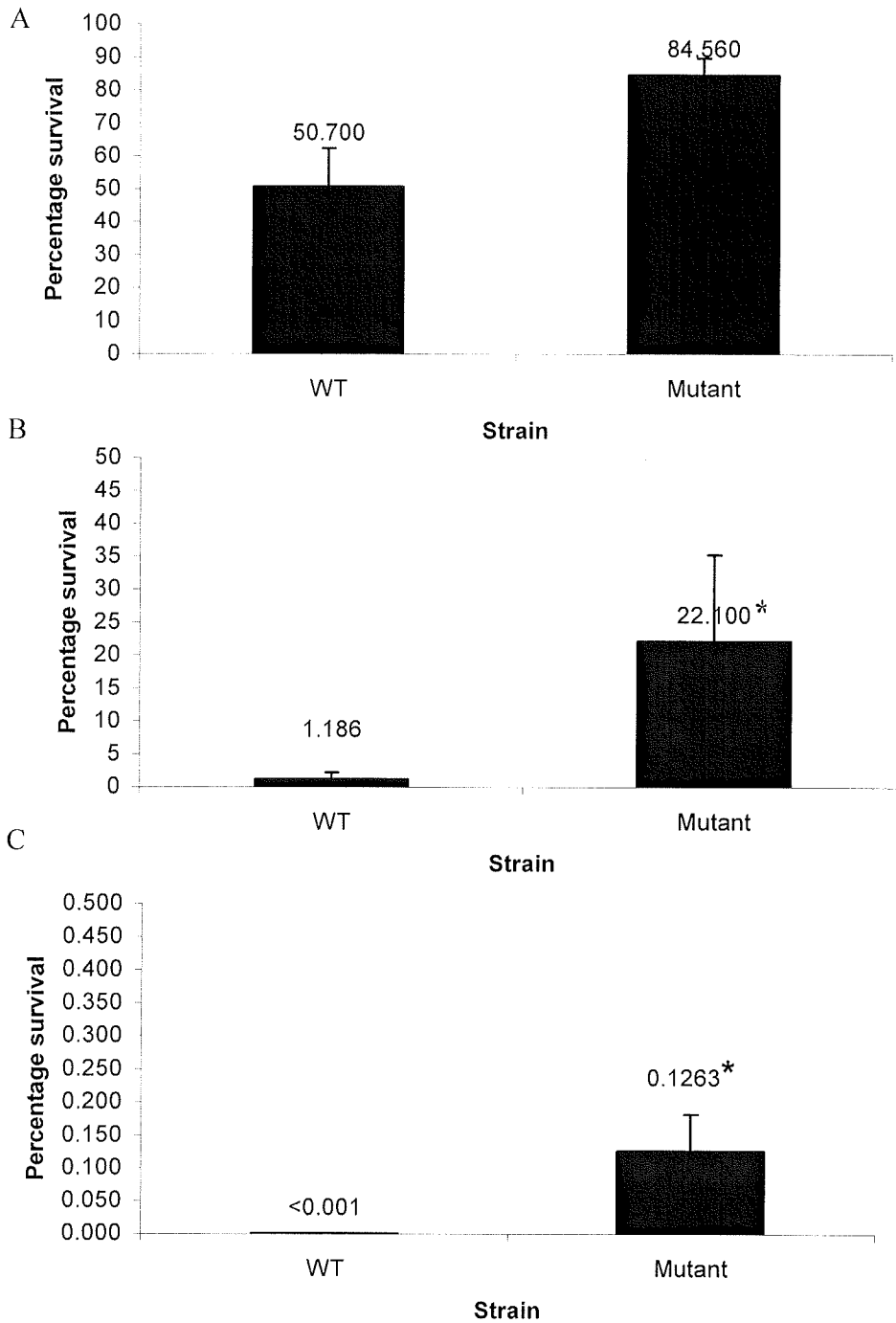
Figure 37: Susceptibility to killing by hydrogen peroxide

The percentage survival (relative to addition of no hydrogen peroxide) was compared between the *spr0370-0371* double mutant and the parental D39 wildtype (WT) at various doses of hydrogen peroxide A) 9.375mM B) 18.75mM and C) 37.5mM. Data represent mean +/- SE from two independent mutants each assayed in duplicate. * P<0.05. Figures give mean percentage survival.

Janus cassette integration in spr0370-0371

As the *spr0370-0371* mutant is a double mutant and lacks both genes it is unclear as to the relative contribution of each to the mutant phenotypes of reduced virulence and enhanced resistance to hydrogen peroxide. To address this would require single mutants. To create such strains Janus mutagenesis was used as described earlier (Chapter 5). The entire *spr0370-0371* open reading frames including the intragenic region were replaced with the Janus cassette. This was done first by amplifying the genome regions upstream (primers 17Q & 32T) and downstream of the target site (32F & 17R) and the Janus cassette amplified from strain R1036 (32S & 21J) (Table 7). These amplifications introduced a 3' BamHI site into the upstream sequence, a 5' NcoI site in the downstream region and a 5'BamHI and 3' NcoI site in the Janus cassette. Products were purified and digested with the appropriate restriction enzyme(s) As done for *zmpB* Janus mutagenesis Figure 22). All three were then purified again and mixed in a ligation reaction and used to directly transform R6. Potential mutants were selected on kanamycin ($200\mu g$ ml^{-1}). Several transformants were obtained and confirmed by PCR with the primers 18I and 18Y located outside the manipulated region. Wild type R6 gave the expected product size of ~2.3kb with the mutants giving the expected size of ~2.1kb indicating replacement of the *spr0370-0371* operon with the Janus cassette (data not shown). Time did not allow further manipulation where the negative selection procedure could be used to replace the Janus cassette with DNA of choice (Figure 20). For example, this could be used to introduce back the original DNA except with a stop codon to selectively knock out either *spr0370* or *0371*. Alternatively, the AphD domain of Spr0370 could be mutated to investigate what affect this might have on virulence or resistance to hydrogen peroxide.

Findings

- The annotation of *spr0370* and *spr0371* as a macrophage infectivity potentiator protein and exfoliative toxin respectively, appears incorrect.

- Spr0370 contains a putative AphD domain and may therefore be involved in resistance to oxidative stress.

- Sequence analysis of Spr0371 gives no firm indication to its possible functions, although it may be a membrane protein.

- The two genes are expressed *in vitro* during exponential growth and form a bicistronic operon.

- This operon contributes to virulence in both pneumonia and bacteraemia models.

- All twenty clinical isolates from a variety of serotypes and MLST sequence types contained the *spr0370-0371* operon as examined by PCR.

- Sequence comparison of the two genes from five clinical isolates and the two sequenced strains shows a high degree of conservation.

- A double mutant was more resistant to killing by hydrogen peroxide confirming that the operon is involved in the response to oxidative stress as suggested by the presence of a putative AphD domain in Spr0370.

- The Janus cassette has been introduced into the operon allowing future manipulation of the sequence, including the creation of single gene null mutants.

Chapter 8 Investigation of a putative macrophage infectivity potentiator protein and exfoliative toxin A, Discussion

Macrophage infectivity potentiator protein is an important virulence factor for *L. pneumophila* acting to facilitate infection of macrophages (Cianciotto and Fields 1992; Wintermeyer *et al.* 1995; Kohler *et al.* 2003) Likewise, exfoliative toxin A is a key virulence determinant in Staphylococcal scalded-skin syndrome caused by *S. aureus* (Gemmell 1995; Ladhani *et al.* 1999; Ladhani 2001; Ferrandiz and de la Campa 2002). The annotation of the pneumococcal genes *spr0370* and *spr0371* as a putative macrophage infectivity potentiator protein and putative exfoliative toxin A respectively, therefore attracted attention to them as potential virulence factors. This interest was further enhanced given that such activities had previously not been associated with *S. pneumoniae*. This novelty placed much significance on the proper annotation of these genes and subsequent functional confirmation. Considerable effort was therefore made to verify these annotations were likely to be accurate. However, unfortunately both identifications were shown to be incorrect and misleading with regards the potential functional of these genes. This clearly highlights the potential dangers of over-reliance on annotated functions without further sequence verification. Disregarding the original annotations, new database searches were made to try and ascertain possible functions to these genes and their products. Analysis of Spr0370 revealed the presence of a putative alkhydroperoxidase D (AhpD) domain suggesting a role in resistance to oxidative stress. Only a single protein containing this domain, AhpD from *M. tuberculosis* has been characterised to date (Hillas *et al.* 2000; Bryk *et al.* 2002; Nunn *et al.* 2002; Koshkin *et al.* 2003) with the best-studied alkylhydroperoxidase being the non-homologous AhpC from *Salmonella typhimurium* (Poole and Ellis 1996; Ellis and Poole 1997) (Figure 38).
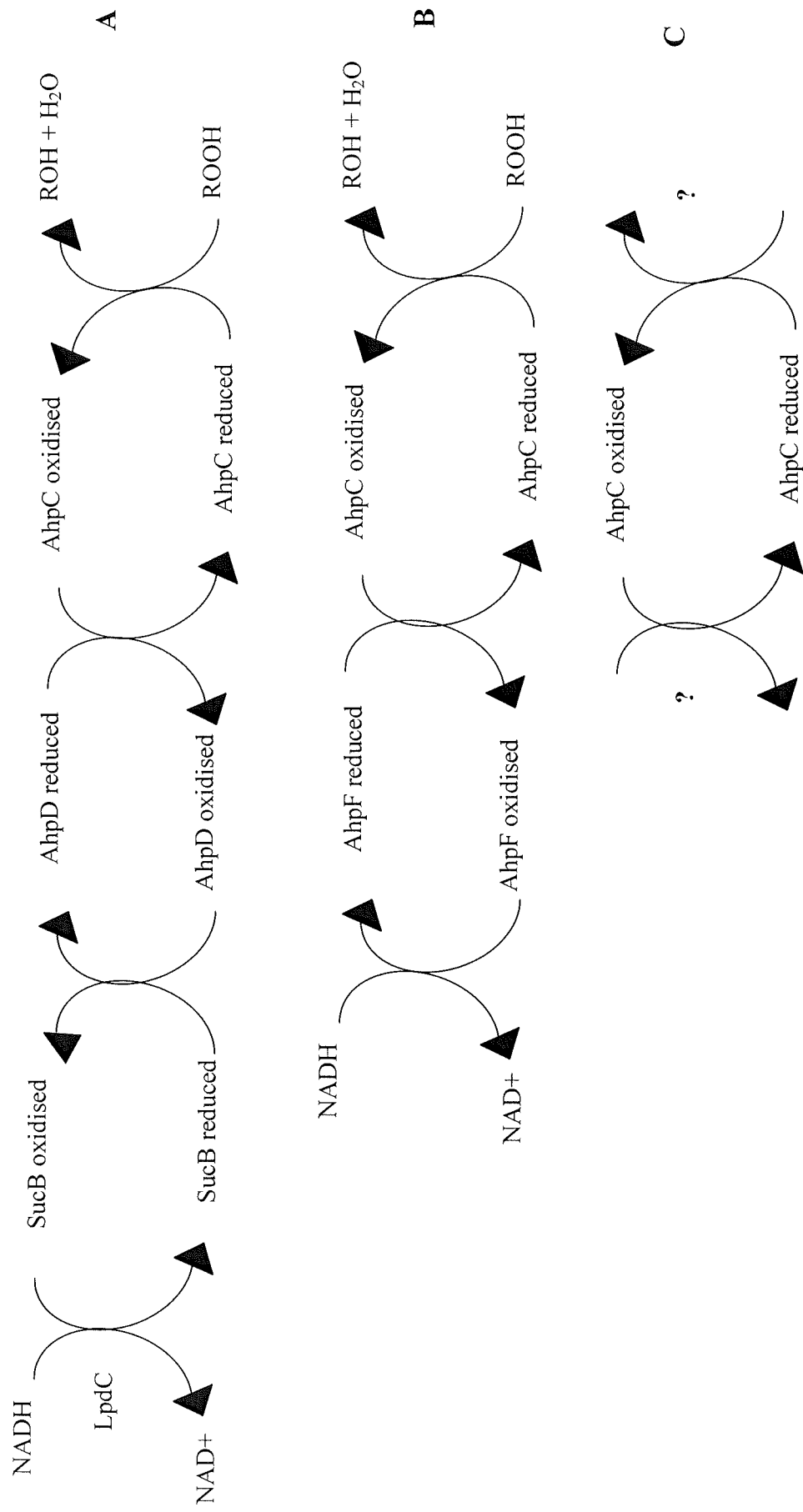
Figure 38: Schematic of electron flow in the AphC peroxidase pathways of *M. tuberculosis* (A), *Salmonella typhimurium* (B) and *Streptococcus pneumonia* (C)

In addition to reducing AhpD, AhpC from *M. tuberculosis* has independent peroxidase activity. The substrate(s) or reducing mechanism for the pneumococcal putative AhpC (Sp0370) are unknown although a role in the response to oxidative stress is supported by the heightened resistance of the *spr0370-0371* mutant to hydrogen peroxide.

172

This *S. typhimurium* AhpC protein contains two cysteine residues that catalyse the reduction of peroxides to the corresponding alcohols and water with concomitant oxidation of these amino acids to a disulphide bond. The catalytic cycle is completed and the active site thiol groups restored by the reduction of this disulphide bond by AhpF, a flavoprotein reductase located immediately downstream of AhpC using NADH as the reducing factor. *M. tuberculosis* possesses an AhpC homologue but AhpF is absent. Instead AhpD, which is not homologous to either AhpC or AhpF, is located immediately downstream of AhpC and is able to reduce AhpC (Bryk *et al.* 2002). Following the reduction of AhpC by AhpD, this latter protein must also be reduced to restore its thiol groups. This function is performed by a novel system consisting of dihydrolipoamide succinyltransferase (SucB), dihydrolipoamide dehydrogenase (Lpd) and NADH (Bryk *et al.* 2002), (Figure 38). In addition to its role in reducing AhpC thereby ensuring its catalytic turnover, AhpD has independent alkylhydroperoxidase activity of its own when AhpF from *S. typhimurium* was added as a reducing partner along with NADH (Hillas *et al.* 2000). Critical to the activity of AphD in reduction of AphC and its own alkylhydroperoxidase activity are its two cysteine residues. Substitution of either of these to a serine effectively abolished both activities (Hillas 2000 and Bryk 2002). The *M. tuberculosis* AhpD crystal structure has been solved independently by two groups (Bryk *et al.* 2002; Nunn *et al.* 2002). It forms a cloverleaf shaped homotrimer with monomers tightly associated through hydrophobic interactions and hydrogen bonds. Both the monomer, made entirely of α-helices and the trimer structures were novel. The trimer has three active sites, which appear to be structurally and functionally independent because all the residues surrounding the critical cysteine residues are from the same monomer. It

was proposed therefore that the individual subunits may be unstable and intersubunit interactions might be important in preserving the active site folding of the individual monomers (Nunn *et al.* 2002). A role for several active site residues in catalysis has been confirmed whereby their substitution decreased both the independent alkylhydroperoxidase ability of AhpC and its ability to reduce AhpC (Koshkin *et al.* 2003). Overall the homology between the pneumococcal Spr0370 protein and *M. tuberculosis* AhpD is low, however, there is a good match across the active site. Four of the five active site residues shown to be involved in the activity of AhpD are conserved in Spr0370 (Glu118, Cys130, Cys133 and His137). The exception to this conservation is His132, which is replaced by Phe in the pneumococcal protein. While substitution of His132 to Phe in AhpC from *M. tuberculosis* significantly reduced activity compared to wild type AhpD, this residue is not absolutely conserved in closely related proteins (Nunn *et al.* 2002). This suggests there may be subtle difference in catalysis where the role of His132 can be played by alternative residues. Therefore, the absence of His132 does not strongly preclude Spr0370 from being an alkylhydroperoxidase. This activity could be confirmed using purified Spr0370 and measuring its ability to breakdown various peroxides *in vitro* using for example HPLC to analysis the products as done for AhpC and D of *M. tuberculosis* (Hillas *et al.* 2000). This may prove difficult to do without the addition of a reducing partner for Spr0370 the identity of which, if any, is unknown. In this case AphF, from *S. typhimurium* could be tried, again as done for AhpC and D of *M. tuberculosis* when their reducing partners were unknown (Hillas *et al.* 2000). Alternatively, the addition of crude pneumococcal cell lysate could be added to provide the necessary factors for activity of Spr0370. This again proved useful in investigating

AhpC and D from *M. tuberculosis* with activity measured by consumption of NADH as determined by absorbance (Bryk *et al.* 2002).

The presence of a putative AhpD domain in Spr0370 suggested this protein might be an alkylhydroperoxidase involved in resistance to oxidative stress. To test if *spr0370-0371* was involved in defence against oxidative stress wild type and the double mutant were exposed to various doses of hydrogen peroxide during early logarithmic growth. The decrease in viable count was then assessed after exposure. In experiments with two independent mutants the *spr0370-0371* mutant showed greater survival compared to wild type D39. This confirmed this operon was involved in the response to oxidative stress but was surprising because it was suspected that the absence of a putative alkylhydroperoxidase (Spr0370) would cause enhanced susceptibility to hydrogen peroxide. However, a similar phenotype is seen with an *ahpC* mutant in *Bacillus subtilis* and single mutants in *ahpC* and glutathione peroxidase in *S. pyogenes* (Antelmann *et al.* 1996; Bsat *et al.* 1996; King *et al.* 2000). Although all these genes are involved in defence against oxidative stress, however, mutants displayed enhanced resistance to hydrogen peroxide. Based on the *ahpC B. subtilis* mutant this was postulated to occur due to low-level oxidative stress leading to the accumulation of an organic hydroperoxide. This did not occur in the presence of AhpC and promoted an adaptive response to peroxide whereby there was induction of the peroxide response genes including *katA* (catalase) and *mgrA* (DNA binding protein and protection protein) leading to increased resistance to subsequent exposure to hydrogen peroxide (Bsat *et al.* 1996). This may also explain the enhanced resistance seen in the *spr0370-0371* mutant. Microarray or 2-D gel proteome analysis could be used to identify genes up-regulated by hydrogen peroxide and

if their expression is enhanced in the mutant compared to the wild type. If *spr0370* is responsible for this, it is not clear yet if it is acting directly as an alkylhydroperoxidase itself or if its role, like that of AhpD in *M. tuberculosis* is in the reduction and catalytic turnover of an additional enzyme. No AhpD homologue exists in *S. pneumoniae* on the basis of BlastP searches against the R6 and TIGR4 genome sequences.

In contrast to Spr0370, no putative activity/function for Spr0371 is suggested other than it may be an integral membrane protein based on the presence of several putative transmembrane spanning regions.

RT-PCR analysis revealed that *spr0370-0371* are transcribed together in a bicistronic message. Genes with a similar function are often transcribed together and so the possible role for *spr0370* in resistance to oxidative stress may therefore hint at an undefined function for *spr0371* in this also. The finding that these genes form an operon and the fact no prior data existed to suggest a role for these genes in the virulence of the pneumococcus prompted the construction of a double mutant lacking both. This made mutagenesis easier and increased the chances of identifying an altered phenotype. During competitive growth *in vitro* this mutant was not out-competed by the wild type showing it did not have an inherent growth defect. When this double mutant was assessed in a model of pneumonia it was attenuated as assessed by reduced bacterial counts in the lung. In competitive infections it was attenuated in both pneumonia and bacteraemia models. These competitive infections were used for the majority of virulence work with this mutant due to the success in identifying small differences in virulence seen with *srtA*, *nanA* and β-gal mutants (Chapter 3 and 5) and the requirement for fewer mice. The *spr0370-0371* operon therefore codes for pneumococcal virulence factors contributing to

both pneumonia and bacteraemia in mouse models. The level of attenuation between the two infection types appeared similar but this does not necessitate the function/role of the operon was the same in the two infection types. The contribution to virulence of different pneumococcal virulence factors varies with the infection type (for example, Hava and Camilli 2002; Marra *et al.* 2002). This was the reason two infection types were analysed and why it would be of interest to examine other situations such as colonisation, meningitis and otitis media models. Additionally, the contribution to virulence of a particular factor may vary among different pneumococcal strains as demonstrated by Blue and Mitchell (2003). It will therefore be of interest to examine the affect of this *spr0370-0371* null mutant in other pneumococcal strains.

A drawback with the above work is that it relates to a double mutant lacking both genes and so it is unclear which are responsible for the phenotypes seen. For example, it is likely based on the presence of an AhpD domain that Spr0370 causes the different response to hydrogen peroxide but an influence for Spr0371 cannot be excluded. To overcome this single mutants can be made in either gene taking care not to disrupt the other given that they are transcribed together. In order to do this the whole *spr0370-0371* operon was replaced with the Janus cassette. As described in Chapter 5 this cassette can now be replaced with DNA of choice. This could include the original wild type DNA with non-polar deletions of one or other of the genes or the addition of a stop codon at the being or either gene resulting single mutants for further study. Additionally, it could be used to probe structure-function relationships, for example substitutions within the putative AhpD active site of Spr0370.

The *spr0370-0371* operon was shown to be ubiquitous among all twenty clinical isolates tested by PCR while sequencing of the two genes from five of these strains found them both to be highly conserved. These features would suggest an important role in pneumococcal biology with a selective advantage in the persistence and sequence conservation of these genes in the population. On the topic of strain distribution it is worth noting the situation for the *spr0370* homologue in *S. suis*. This pig pathogen encodes a cholesterol-binding cytotoxin, suilysin, homologous to the pneumococcal virulence factor pneumolysin. However, suilysin is not found in all strains and when it is absent it is found to be replaced by a *spr0370* homologue named *orfC* or *orf102* (King *et al.* 2001; Takamatsu *et al.* 2002). All other genes in this region of the genome were conserved with the exception of the 3' end of the gene immediately upstream of *suilysin/orfC* and the 5' end of the gene immediately downstream. These regions showed some divergence between *suilysin* positive and *orfC* positive strains. This divergence however, does not appear to alter their putative function. These regions where divergent between the suilysin positive and negative strains. This mutually exclusive relationship was conserved among 94 out of 100 strains examined in two studies (King *et al.* 2001; Takamatsu *et al.* 2002), the exception being six isolates whose suilysin gene region had undergone rearrangements leaving their genomic organisation in this area undefined (Takamatsu *et al.* 2002). The finding is enigmatic with no indication of the mechanism of this switch for example, repetitive structures or translocatable elements and no function suggested for *orfC* or *orf102* (King *et al.* 2001; Takamatsu *et al.* 2002). The *spr0370* homologue does not appear to be a haemolysin however, as that property correlated with the presence of *suilysin* and was absent in *suilysin* negative *orfC* / *orf102* positive strains.

Although the role of *orfC* / *orf102* is unknown its replacement of *suilysin* may suggest it performs some of the functions of this toxin although the role for *suilysin* remains undefined. Like pneumolysin, suilysin may inhibit the respiratory burst of phagocytes (Paton and Ferrante 1983; Nandoskar *et al.* 1986). OrfC / Orf102, like Spr0371 contains an AhpD domain and so it may function to enhance resistance to oxidative stress in the absence of suilysin.

Interestingly, although *spr0370-0371* forms an operon in the pneumococcus, this is not the situation in organisms containing homologues of these genes. For example, *N. meningitisis* encodes a *spr0370* homologue but not *spr0371*. Likewise *S. agalactiae*, *S. aureus*, and *L. lactis* subsp. *lactis* all contain a homologue of *spr0371* but not *spr0370*. *S. mutans* encodes homologues to both but these are found in separate locations in the genome. Furthermore, the homologues in other organisms appear to be located in regions distinct from that in the pneumococcus. No organism for which the genome has been sequenced therefore has the same genomic organisation of *spr0370-0371* shown in the pneumococcus. The reason and functional consequences of this are unclear.

If Spr0370 is acting as an alkylhydroperoxidase there may be a reducing partner ensuring its catalytic turnover (restoring the active site thiol groups allowing fresh attack on substrate). This is preformed in *Mycobacterium tuberculosis* by a novel pathway consisting of dihydrolipoamide succinyltransferase (SucB), and dihydrolipoamide dehydrogenase (Lpd) (Figure 26) (Bryk *et al.* 2002). Pneumococcal dihydrolipoamide dehydrogenase (Spr1049) shows strong homology to both Lpd (similarity 55% over 465 amino acids, E= 8.7 e-75) and SucB (similarity 50% over 303 amino acids, E= 1.2 e-38). Dihydrolipoamide dehydrogenase is therefore a candidate reducing partner for Spr0370.

Smith *et al.* 2002 investigated a dihydrolipoamide dehydrogenase mutant and found it to be defective in carbohydrate metabolism, capsule production and virulence (Smith *et al.* 2002). However, the response to oxidative stress was not investigated.

Alternatively, AhpF supported the direct alkylhydroperoxidase activity of AhpD from *S. typhimurium* (Hillas *et al.* 2000). BlastP searches against the R6 and TIGR4 genome sequences revealed several proteins with homology to AhpF from *S. typhimurium* (Spr1312 thioredoxin reductase, Spr1421 thioredoxin reductase, Spr1323 NADH oxidase, Spr1442 oxidoreductase and Spr0692 glutathione oxidoreductase, R6 gene name given). Any of these proteins may therefore be a reducing partner for Spr0370.

The pneumococcus is a facultative anaerobe able to grow in the presence or absence of oxygen although it does not require oxygen for its respiration, which is fermentative regardless of the growth conditions (Poolman 1993). The ability to grow in aerobic environments has the advantage of allowing a species to survive in an increased range of ecological niches. However, it does mean exposure to toxic products such as peroxide, superoxide and hydroxyl radicals that have many deleterious affects such as DNA strand damage and peroxidation of membrane lipids (reviewed by Miller and Britigan 1997). Several pneumococcal strategies to resist oxidative stress have already been described. The pneumococcus contains two superoxide dismutase enzymes, which convert superoxide to less toxic hydrogen peroxide and molecular oxygen. A mutant in the major superoxide dismutase (*sodA*) was found to have increased susceptibility to oxidative stress and reduced virulence (Yesilkaya *et al.* 2000). Likewise, NADH oxidase (Nox) converts oxygen to water using NADH. A pneumococcal mutant in *nox* was defective in growth under vigorous aeration but not under limited aeration (Yu *et al.* 2001). This

agrees with results relating to NADH oxidase in *S. pyogenes* and *S. mutans* (Gibson *et al.*

2000; Higuchi 1992). As with the *sodA* pneumococcal gene *nox* contributes to virulence

in animal models (Auzat *et al.* 1999; Lau *et al.* 2001; Yu *et al.* 2001; Hava and Camilli

2002). In addition to resistance to oxidative stress it is suggested that *nox* has a role in

improving the efficiency of glucose catabolism and regulating competence (Auzat *et al.*

1999). The pneumococcal surface antigen (*psa*) operon is also involved in oxidative

resistance. The genes *psaABC* encode an ATP-binding cassette manganese permease and

psaD a putative thiol peroxidase (Claverys *et al.* 1999). In addition to the uptake of

managense PsaA may contribute to adhesion with a mutant attenuated in mouse models

(Berry and Paton 1996). Mutants in *psaA* and *psaD* were hypersensitive to oxidative

stress which was suggested to be independent of their role in manganese uptake and so it

was suggested that *psaA* and *D* may function in a signal transduction system involved in

the response to oxidative stress (Tseng *et al.* 2002). ClpP an ATP-dependent protease that

contributes to the normal physiology, stress response and virulence of the pneumococcus

also has a role in oxidative stress. A mutant showed enhanced killing by hydrogen

peroxide that may have been due to down-regulation of *sodA* and a putative thioredoxin-

linked thiol peroxidase (Robertson *et al.* 2002). This was deemed an unexpected result

given that it had not previously been suggested that ClpP proteases where involved in the

degradation proteins damaged by oxidation (Robertson *et al.* 2002). In addition to the

genes discussed above BlastP searches of the R6 and TIGR4 genomes reveals

pneumococcal homologues of glutathione peroxidase from *S. pyogenes* and peroxide

resistance protein (Dpr) from *S. mutans* both of which contribute to oxidative resistance

in these species (Yamamoto *et al.* 2000; King *et al.* 2000). The response to hydrogen

181

peroxide in *B. subtilis* and *S. pyogenes* is negatively regulated by PerR (Bsat *et al.* 1996; King *et al.* 2000). BlastP searches of the R6 and TIGR4 genomes revealed no homologues to either the *B. subtilis* or *S. pyogenes* PerR regulator. OxyR is an important transcriptional activator in the response of some bacteria to oxidative stress (reviewed in Mongkolsuk and Helmann 2002). Spr0828, a LysR family transcriptional regulator, displays limited homology to *E. coli* OxyR but importantly appears to lack the two cysteine residues (at positions 199 and 208) that are critical for OxyR function. Likewise, OhrR is a negative regulator of the antioxidant protein OhrA in *B. subtilis* but is seemingly absent in *S. pneumoniae* based on a BlastP search of the R6 and TIGR4 genomes. Catalase is a key enzyme in resistance to peroxide among many bacterial species (Deisseroth and Dounce 1970; Farr and Kogoma 1991). However, as with other lactic acid bacteria, *S. pneumoniae* lacks this enzyme (Hardie 1986), which is of particular note given the ability of the pneumococcus to produce and be resistant to high levels of hydrogen peroxide (Pericone *et al.* 2000). Indeed this level of resistance exceeded that of several catalase positive species. The production of hydrogen peroxide by the pneumococcus is derived largely from pyruvate oxidase (SpxB) and may provide a competitive advantage to the pneumococcus during colonisation by inhibiting growth of competing organisms (Spellerberg *et al.* 1996; Pericone *et al.* 2000). SpxB generated hydrogen peroxide appears also to have a contributing role in pathogenesis via host tissue damage (Braun *et al.* 2002; Hirst *et al.* 2000). Interestingly, SpxB appears also to be required for resistance to hydrogen peroxide (Pericone *et al.* 2003). The mechanism may relate to the metabolic well being of cells given the important role of SpxB in generating ATP under aerobic conditions whereby the acetyl-P formed by SpxB can subsequently

donate it phosphate group to the formation of ATP via acetate kinase. In line with this, *spxB* mutants had a defect in generating ATP with levels more rapidly depleted upon hydrogen peroxide exposure compared to wild type cells (Pericone *et al.* 2003).

How the many factors that appear involved in the pneumococcal response to oxidative stress interact together is unclear and would appear to be complex. However, given the role of oxidative stress in pneumococcal viability, metabolism, virulence and competence it appears a pivotal issue in pneumococcal biology and worthy of further investigation.

The results presented suggest Spr0370 may be a novel pneumococcal virulence factor contributing to the resistance to oxidative stress. As discussed above it has been shown that pneumococcal mutants with enhanced susceptibility to oxidative stress are less virulent. Although these genes may have additional functions it is presumed that this enhanced susceptibility to oxidative stress contributes to their attenuation. Therefore the attenuation in mouse infections of the *spr0370-0371* mutant is in disagreement with this pattern given it was more resistant to hydrogen peroxide. However, as only resistance to hydrogen peroxide was assessed the mutant may be more susceptible to other oxidative stresses. For example, although a *S. pyogenes* mutant in *AhpC* was more resistant to hydrogen peroxide it showed increased susceptibility to oxidative stress induced by exposure to cumene hydroperoxide and methyl viologen (King *et al.* 2000). The attenuation of the *spr0370-0371* mutant may then relate to increased susceptibility to oxidative stresses other than hydrogen peroxide experienced *in vivo*. A potentially important source of oxidative stress in pneumococcal infections is phagocytes. It may therefore be of interest to compare the *spr0370-0371* mutant with wild type D39 with a wider range and combination of oxidative stresses in addition to phagocytic killing *in*

vitro. A further explanation for the apparent disagreement between hydrogen peroxide resistance and virulence is that the operon performs another additional, as yet unidentified function contributing to virulence.

Here we have investigated two pneumococcal genes whose annotation suggested they might be candidate virulence factors. Those annotations proved incorrect but the two genes were investigated as one appeared to be involved in the response to oxidative stress. Our work shows they form an operon found among all twenty pneumococcal strains examined. The sequence of both genes was conserved among sevens strains examined. Mouse infection studies show that this operon contributes to virulence in pneumonia and bacteraemia models. The increased resistance to killing by hydrogen peroxide of a null mutant suggests a role in the response to oxidative stress. Work is underway to create single gene null mutants to investigate their function further.

Concluding remarks

This thesis exploited the available pneumococcal genome sequences to identify putative pneumococcal virulence factors. The important role of sortase enzymes in the virulence of other organisms meant the four pneumococcal sortase homologues were likely candidates. *srtA*, the only pneumococcal sortase found in all strains examined, was chosen for further examination. A *srtA* null mutant was attenuated in mouse models of pneumonia, bacteraemia and colonisation. This is the first demonstration of *srtA* as a pneumococcal virulence factor. However, the defects in virulence appeared small and were only detected in sensitive competitive infections. This was somewhat surprising given the attenuation seen with sortase mutants in other organisms. Part of the reason for this appears to be the small contribution of sortase anchored proteins in this mouse model of infection. Another factor may be that the defective surface localisation of these proteins has a nominal affect on their contribution to virulence due to their activity as enzymes rather in direct ligands in adherence. If the limited role of *srtA* in mouse virulence extends to human infection, the development of novel antimicrobials targeting sortase will not be as effective in pneumococcal disease as suggested for other pathogens. In addition to sortase, two genes annotated as a macrophage infectivity potentiator protein and an exfoliative toxin A were investigated. The importance of checking the basis of genome annotations was highlighted here as both genes are incorrectly identified. One gene (*spr0370*) appears to encode an alkylhydroperoxidase involved in oxidative resistance while no putative function for the other could be suggested based on its sequence. These genes formed a novel pneumococcal operon conserved among clinical isolates and contributing to virulence. The role for this operon in oxidative stress appears

complex as unexpectedly a null mutant lacking both genes showed enhanced resistance rather than susceptibility to hydrogen peroxide. Finally, the ability to introduce mutations without a selectable marker into the pneumococcal genome using Janus mutagenesis is a potentially powerful technique. Here we provide the first demonstration of it use to study pneumococcal virulence.

Appendix

Reagents and buffers

Extraction Buffer (for genomic DNA preparation)
10 mM Tris pH 8.0 (Sigma-Aldrich, UK), 100 mM EDTA pH 8.0 (Fisher Scientific UK Ltd.), 0.5 % SDS (w/v)(Fisher Scientific Ltd.UK).

TE (Tris-EDTA) Buffer pH7.4
10 mM Tris-Cl (pH 7.4, Sigma-Aldrich , U.K.)
1 mM EDTA (pH 8.0, Fisher Scientific U.K. Ltd.)

PBS, pH 7.4
One tablet (phosphate buffered saline tablets, Dulbecco A, Oxoid Ltd. Basingstoke) was dissolved in 100 ml dH$_2$O and autoclaved.

DNA Gel Loading Buffer (6x)
0.25 % Bromophenol Blue (Sigma-Aldrich)
0.25 % Xylene cyanole FF
40 % (w/v) sucrose in dH$_2$O

Protein Gel Sample Buffer (5x)
60 mM Tris-HCl (pH 6.8)
25 % glycerol
2 % SDS
14.4 mM 2-mercaptoethanol
0.1 % bromophenol blue

Protein Gel Running Buffer (1x)
25 mM Tris
192 mM glycine
0.1 % SDS

Lysis buffer for Western blot samples
Solution 1
10 mM SDS
200 mM DTT
28 mM Tris base
20 mM Tris-HCl

Solution 2
24 mM Tris base
0.47 mM Tris-HCl
50mM MgCl$_2$
Added before use:
10mg/ml RNAse
5mM DNAse in 5mM CaCl$_2$

4.67 ml solution 1 and 330 µl solution 2 mixed and used.

1 % Agarose
1 g agarose (GibcoBRL)
100 ml TAE buffer

BHI contents

Calf brain infusion solids, beef heart infusion solids, protease peptone, glucose, NaCl, disodium phosphate

Transfer Buffer, pH 8.1-8.4 for Western blot

25 mM Tris (Sigma-Aldrich, U.K.)
192 mM glycine
20 % (v/v) methanol

Useful Web Addresses

BLAST	http://www.ncbi.nlm.nih.gov:80/BLAST/
PubMed	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi
The Institute of Genomic Research	http://www.tigr.org/

Primers

Chapter 3		
Name	**Sequence**	**Use**
25F	tacttgctcgcgtttcattgtacc	PCR detection of srtA
25G	atttctgatattactgtcactggc	PCR detection of srtA
19Ia	gtgtctcgcttgtattatcg	PCR detection of srtB
19J	atcacaatcaaaccaactgc	PCR detection of srtB
19K	tcgtattgagtcaaacgagg	PCR detection of srtC
19L	tctagtcctttgacaatccg	PCR detection of srtC
19M	acttgttgatgttggtagcc	PCR detection of srtD
19A	gagctctagtaaacgtgctagcttcc	PCR detection of srtD
23U	ttaataaaattgtttatatgg	Sequencing of srtA
24D	gaaggtgaagcataatgtctc	Sequencing of srtA
24E	cgtttcattgtaccagctcc	Sequencing of srtA
24F	ggagctggtacaatgaaacg	Sequencing of srtA
20o	ggtgtcatgattcgaacc	Confirmation of allelic replacement of srtA
20N	ctgcagcagtactgaagg	Confirmation of allelic replacement of srtA
19H	gaacatcgaacgaatcgc	Allelic replacement of srtA (initial PCR)
19I	ttcacaaacaggacgtgc	Allelic replacement of srtA (initial PCR)
19W	ggcgcgccctgctaggatatttgtcagc	Allelic replacement of srtA (inverse PCR)
19X	ggcgcgccctaacagctttcaatcaacc	Allelic replacement of srtA

		(inverse PCR)
Chapter 5		
27F	tgaaactaggttgtcatacc	Allelic replacement of β-*gal* (amplification of upstream region)
27G	gtatttggtaggcgcgccgattctccagctata tttcc	Allelic replacement of β-*gal* (amplification of upstream region)
27P	ggcgcgcctaccaaatacaggtagtgcg	Allelic replacement of β-*gal* (amplification of downstream region)
27I	tgaaatgaaataagaccagg	Allelic replacement of β-*gal* (amplification of downstream region)
35I	gttctcgtattcaagcgacc	Internal β-*gal*
35J	gaatctccaaacagaacacc	Internal β-*gal*
32M	actgacgaagtttgaattgg	Confirmation of β-*gal* allelic replacement (outside primer)
32N	ggtacttatggccaataacc	Confirmation of β-*gal* allelic replacement (outside primer)
27J	ggcgcgccctaacagctttcttccttgg	Allelic replacement of *nan*A (amplification of upstream region)
25X	tcttatgatttcgtatcagc	Allelic replacement of *nan*A (amplification of upstream region)
27K	agctgttagggcgcgcctgaaataagacattt attcc	Allelic replacement of *nan*A (amplification of downstream region)
25Y	aatagcataagctgatgtcc	Allelic replacement of *nan*A (amplification of downstream region)
35E	tgaccgcttctccagcatgc	Confirmation of *nan*A allelic replacement (outside primer)
35F	gaaaccaatgcttcaaatgg	Confirmation of *nan*A allelic replacement (outside primer)
31R	taacagtacaccttgactgc	Internal *nan*A
31S	gtgttcgataaggattgagc	Internal *nan*A
27C	acatctattctcagagtgcc	Allelic replacement of *prt*A (amplification of upstream region)
27Q	acattattccggcgcgccagttagtgacaatac tgtgc	Allelic replacement of *prt*A (amplification of upstream region)
27E	ggccgcgccggaataatgtttatagtagg	Allelic replacement of *prt*A

		(amplification of downstream region)
25Q	tctcatcacctctattagcc	Allelic replacement of *prt*A (amplification of downstream region)
35C	caataataatcaagctaagg	Confirmation of *prt*A allelic replacement (outside primer)
35D	gaatctgtcacaagatttcc	Confirmation of *prt*A allelic replacement (outside primer)
32O	tttctcagctatggaatccg	Internal *prt*A
32P	ggaattaagcaaccaaacgc	Internal *prt*A
27L	ggcgcgccgctgaacttccaaatacagg	Allelic replacement of *spu*A (amplification of upstream region)
25L	tcctaatcaagtttattagg	Allelic replacement of *spu*A (amplification of upstream region)
27M	aagttccgaggcgcgcctttctcagtgtgagatggtg	Allelic replacement of *spu*A (amplification of downstream region)
25M	ttcctaataatacggcttgg	Allelic replacement of *spu*A (amplification of downstream region)
35G	ttgtagctaacagtataagg	Confirmation of *spu*A allelic replacement (outside primer)
35H	aggatgctcaaggactatgg	Confirmation of *spu*A allelic replacement (outside primer)
32Q	gttgcacccaaagtacgagc	Internal *spu*A
32R	caacatttgtcagcacgtcc	Internal *spu*A
21I	ggcaagcttccgtttgatttttaatggataatg	Amplification of Janus cassette
21J	catccatggagagacctgggcccctttcc	Amplification of Janus cassette
21C	gcaattggcatgcatgtcaagc	ZmpB LPQTG deletion by Janus mutagenesis
21G	gacaagctttgcgtccactacagaagg	ZmpB LPQTG deletion by Janus mutagenesis
21F	acttcttcaactggagctggttctgc	ZmpB LPQTG deletion by Janus mutagenesis
21H	ccaccatggtgttggttatctatcagg	ZmpB LPQTG deletion by Janus mutagenesis
21Q	ttcatgcaggaggttcgg	ZmpB LPQTG deletion by Janus mutagenesis
21R	agctggtttgtctgatgg	ZmpB LPQTG deletion by

		Janus mutagenesis
21S	ggtgtatcaaaaagatcagacagaagcatctg	ZmpB LPQTG deletion by Janus mutagenesis
21T	cagatgcttctgtctgatctttttgatacacc	ZmpB LPQTG deletion by Janus mutagenesis
	Chapter 7	
17Q	gga tcc taa tgg aat cac aca gct cg	Allelic replacement of *spr0370-0371* (initial PCR) Amplification of *spr0370-0371* for sequencing
17R	gag ctc atc cgt tac cgt gat gaa gc	Allelic replacement of *spr0370-0371* (initial PCR) Amplification of *spr0370-0371* for sequencing
17S	ggc gcg ccg tgc tga ctc tac tgt atg g	Allelic replacement of *spr0370-0371* (inverse PCR) also used for RT-PCR
17T	ggc gcg ccg ctt atc tga ttt ggt tac g	Allelic replacement of *spr0370-0371* (inverse PCR)
18I	ctt ggt ctt gtc gcc cag tcg	Confirmation of *spr0370-0371* allelic replacement also used for RT-PCR and confirmation of Janus integration
18Y	cgt tct gag gct ggt tct gc	Confirmation of *spr0370-0371* allelic replacement also used for RT-PCR and confirmation of Janus integration
21X	tgt cgc agg tca cac agc	*spr0370* internal used for RT-PCR sequencing of *spr0370-0371* operon
21Y	gat tgc tag gct gac acc	*spr0370* internal used for RT-PCR
21Z	cta tgc caa caa ctt agc	*spr0370* internal used for RT-PCR
22A	caa acc tag caa aca acc	*spr0371* internal used for RT-PCR
22B	acc aag caa gaa ttg acc	*spr0371* internal used for RT-PCR
22C	tac cta cca gag gat agg	*spr0371* internal used for RT-PCR
28H	atg aca aca ttt act atc cc	Sequencing of *spr0370-0371* operon

28I	tta ggc ata agg ttg caa ttc tgg	Sequencing of *spr0370-0371* operon
32F	ctgccatggctagaaatagctaaacc	*spr0370-0371* Janus integration
32S	cgcggatccccgtttgattttttaatg	*spr0370-0371* Janus integration
32T	cgcggatccattgacactaatattattgg	*spr0370-0371* Janus integration

References

AlonsoDeVelasco, E., A. F. Verheul, J. Verhoef and H. Snippe (1995). "*Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines." Microbiol Rev **59**(4): 591-603.

Ament, A., R. Baltussen, G. Duru, C. Rigaud-Bully, D. de Graeve, A. Ortqvist, B. Jonsson, J. Verhaegen, J. Gaillat, P. Christie, A. S. Cifre, D. Vivas, C. Loiseau and D. S. Fedson (2000). "Cost-effectiveness of pneumococcal vaccination of older people: a study in 5 western European countries." Clin Infect Dis **31**(2): 444-50.

Andersson, B., J. Dahmen, T. Frejd, H. Leffler, G. Magnusson, G. Noori and C. S. Eden (1983). "Identification of an active disaccharide unit of a glycoconjugate receptor for pneumococci attaching to human pharyngeal epithelial cells." J Exp Med **158**(2): 559-70.

Andersson, B. and C. Svanborg Eden (1988). "Mechanisms of pneumococcal attachment." Monogr Allergy **24**: 44-5.

Antelmann, H., S. Engelmann, R. Schmid and M. Hecker (1996). "General and oxidative stress responses in *Bacillus subtilis*: cloning, expression, and mutation of the alkyl hydroperoxide reductase operon." J Bacteriol **178**(22): 6571-8.

Auzat, I., S. Chapuy-Regaud, G. Le Bras, D. Dos Santos, A. D. Ogunniyi, I. Le Thomas, J. R. Garel, J. C. Paton and M. C. Trombe (1999). "The NADH oxidase of *Streptococcus pneumoniae*: its involvement in competence and virulence." Mol Microbiol **34**(5): 1018-28.

Avery, O. T., C. M. MacLeod and M. McCarty (1944). "Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III." Journal of Experimental Medicine **79**: 137-157.

Bae, T. and O. Schneewind (2003). "The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing." J Bacteriol **185**(9): 2910-9.

Balachandran, P., S. K. Hollingshead, J. C. Paton and D. E. Briles (2001). "The autolytic enzyme LytA of *Streptococcus pneumoniae* is not responsible for releasing pneumolysin." J Bacteriol **183**(10): 3108-16.

Barnett, T. C. and J. R. Scott (2002). "Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs." J Bacteriol **184**(8): 2181-91.

Bartilson, M., A. Marra, J. Christine, J. S. Asundi, W. P. Schneider and A. E. Hromockyj (2001). "Differential fluorescence induction reveals *Streptococcus pneumoniae* loci regulated by competence stimulatory peptide." Mol Microbiol **39**(1): 126-35.

Bensing, B. A. and P. M. Sullam (2002). "An accessory sec locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets." Mol Microbiol **44**(4): 1081-94.

Bergmann, S., M. Rohde, G. S. Chhatwal and S. Hammerschmidt (2001). "alpha-Enolase of *Streptococcus pneumoniae* is a plasmin(ogen)-binding protein displayed on the bacterial cell surface." Mol Microbiol **40**(6): 1273-87.

Berry, A. M. and J. C. Paton (1996). "Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*." Infect Immun **64**(12): 5255-62.

194

Berry, A. M. and J. C. Paton (2000). "Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins." Infect Immun **68**(1): 133-40.

Bessen, D. E. and A. Kalia (2002). "Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in *Streptococcus pyogenes*." Infect Immun **70**(3): 1159-67.

Bethe, G., R. Nau, A. Wellmer, R. Hakenbeck, R. R. Reinert, H. P. Heinz and G. Zysk (2001). "The cell wall-associated serine protease PrtA: a highly conserved virulence factor of *Streptococcus pneumoniae*." FEMS Microbiol Lett **205**(1): 99-104.

Bierne, H., S. K. Mazmanian, M. Trost, M. G. Pucciarelli, G. Liu, P. Dehoux, L. Jansch, F. Garcia-del Portillo, O. Schneewind and P. Cossart (2002). "Inactivation of the srtA gene in *Listeria monocytogenes* inhibits anchoring of surface proteins and affects virulence." Mol Microbiol **43**(4): 869-81.

Black, S., H. Shinefield, B. Fireman, E. Lewis, P. Ray, J. R. Hansen, L. Elvin, K. M. Ensor, J. Hackell, G. Siber, F. Malinoski, D. Madore, I. Chang, R. Kohberger, W. Watson, R. Austrian and K. Edwards (2000). "Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group." Pediatr Infect Dis J **19**(3): 187-95.

Blue, C. E. (2002). Characterisation of the role of two two-component signal transduction systems and a putative zinc metalloprotease in the virulence of *Streptococcus pneumoniae*. A thesis submitted for the degree of Ph.D. Division of Infection and Immunity,, University of Glasgow.

Blue, C. E. and T. J. Mitchell (2003). "Contribution of a response regulator to the virulence of *Streptococcus pneumoniae* is strain dependent." Infect Immun **71**(8): 4405-13.

Blue, C. E., G. K. Paterson, A. R. Kerr, M. Berge, J. P. Claverys and T. J. Mitchell (2003). "ZmpB, a novel virulence factor of *Streptococcus pneumoniae* that induces tumor necrosis factor alpha production in the respiratory tract." Infect Immun **71**(9): 4925-35.

Bolken, T. C., C. A. Franke, K. F. Jones, G. O. Zeller, C. H. Jones, E. K. Dutton and D. E. Hruby (2001). "Inactivation of the srtA gene in *Streptococcus gordonii* inhibits cell wall anchoring of surface proteins and decreases in vitro and in vivo adhesion." Infect Immun **69**(1): 75-80.

Bongaerts, R. J., H. P. Heinz, U. Hadding and G. Zysk (2000). "Antigenicity, expression, and molecular characterization of surface-located pullulanase of *Streptococcus pneumoniae*." Infect Immun **68**(12): 7141-3.

Braun, J. S., J. E. Sublett, D. Freyer, T. J. Mitchell, J. L. Cleveland, E. I. Tuomanen and J. R. Weber (2002). "Pneumococcal pneumolysin and H_2O_2 mediate brain cell apoptosis during meningitis." J Clin Invest **109**(1): 19-27.

Brock, S. C., P. A. McGraw, P. F. Wright and J. E. Crowe, Jr. (2002). "The human polymeric immunoglobulin receptor facilitates invasion of epithelial cells by *Streptococcus pneumoniae* in a strain-specific and cell type-specific manner." Infect Immun **70**(9): 5091-5.

Brown, J. S., S. M. Gilliland and D. W. Holden (2001). "A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence." Mol Microbiol **40**(3): 572-85.

Brueggemann, A. B., D. T. Griffiths, E. Meats, T. Peto, D. W. Crook and B. G. Spratt (2003). "Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential." J Infect Dis **187**(9): 1424-32.

Bryk, R., C. D. Lima, H. Erdjument-Bromage, P. Tempst and C. Nathan (2002). "Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein." Science **295**(5557): 1073-7.

Bsat, N., L. Chen and J. D. Helmann (1996). "Mutation of the *Bacillus subtilis* alkyl hydroperoxide reductase (ahpCF) operon reveals compensatory interactions among hydrogen peroxide stress genes." J Bacteriol **178**(22): 6579-86.

Burrell, M. H., M. E. Mackintosh and C. E. Taylor (1986). "Isolation of *Streptococcus pneumoniae* from the respiratory tract of horses." Equine Vet J **18**(3): 183-6.

Camara, M., G. J. Boulnois, P. W. Andrew and T. J. Mitchell (1994). "A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein." Infect Immun **62**(9): 3688-95.

Casadevall, A. and L.-A. Pirofski (2003). "The Damage-Response Framework of Microbial Pathogenesis." Nature Reviews Microbiology **1**: 17-24.

Chapuy-Regaud, S., A. D. Ogunniyi, N. Diallo, Y. Huet, J. F. Desnottes, J. C. Paton, S. Escaich and M. C. Trombe (2003). "RegR, a global LacI/GalR family regulator, modulates virulence and competence in *Streptococcus pneumoniae.*" Infect Immun **71**(5): 2615-25.

Chiavolini, D., G. Memmi, T. Maggi, F. Iannelli, G. Pozzi and M. R. Oggioni (2003). "The three extra-cellular zinc metalloproteinases of *Streptococcus pneumoniae* have a different impact on virulence in mice." BMC Microbiol **3**(1): 14.

Cianciotto, N. P. and B. S. Fields (1992). "*Legionella pneumophila* mip gene potentiates intracellular infection of protozoa and human macrophages." Proc Natl Acad Sci U S A **89**(11): 5188-91.

Claverys, J. P., C. Granadel, A. M. Berry and J. C. Paton (1999). "Penicillin tolerance in *Streptococcus pneumoniae*, autolysis and the Psa ATP-binding cassette (ABC) manganese permease." Mol Microbiol **32**(4): 881-3.

Cockeran, R., R. Anderson and C. Feldman (2002). "The role of pneumolysin in the pathogenesis of *Streptococcus pneumoniae* infection." Curr Opin Infect Dis **15**(3): 235-9.

Connolly, K. M., B. T. Smith, R. Pilpa, U. Ilangovan, M. E. Jung and R. T. Clubb (2003). "Sortase from *Staphylococcus aureus* does not contain a thiolate-imidazolium ion pair in its active site." J Biol Chem **278**(36): 34061-5.

Cossart, P. and R. Jonquieres (2000). "Sortase, a universal target for therapeutic agents against gram-positive bacteria?" Proc Natl Acad Sci U S A **97**(10): 5013-5.

Cundell, D. R., N. P. Gerard, C. Gerard, I. Idanpaan-Heikkila and E. I. Tuomanen (1995a). "*Streptococcus pneumoniae* anchor to activated human cells by the receptor for platelet-activating factor." Nature **377**(6548): 435-8.

Cundell, D. R., B. J. Pearce, J. Sandros, A. M. Naughton and H. R. Masure (1995b). "Peptide permeases from *Streptococcus pneumoniae* affect adherence to eucaryotic cells." Infect Immun **63**(7): 2493-8.

Cunningham, A. S. (1979). "Morbidity in breast-fed and artificially fed infants. II." J Pediatr **95**(5 Pt 1): 685-9.

Davis BD, D. R., Eisen HN and Ginsberg HS (1990). Microbiology. Philadelphia, USA, J.B Lippincot.

de Saizieu, A., C. Gardes, N. Flint, C. Wagner, M. Kamber, T. J. Mitchell, W. Keck, K. E. Amrein and R. Lange (2000). "Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide." J Bacteriol **182**(17): 4696-703.

Deisseroth, A. and A. L. Dounce (1970). "Catalase: Physical and chemical properties, mechanism of catalysis, and physiological role." Physiol Rev **50**(3): 319-75.

Denny, F. W. and F. A. Loda (1986). "Acute respiratory infections are the leading cause of death in children in developing countries." Am J Trop Med Hyg **35**(1): 1-2.

Dhar, G., K. F. Faull and O. Schneewind (2000). "Anchor structure of cell wall surface proteins in *Listeria monocytogenes*." Biochemistry **39**(13): 3725-33.

Di Guilmi, A. M. and A. Dessen (2002). "New approaches towards the identification of antibiotic and vaccine targets in *Streptococcus pneumoniae*." EMBO Rep **3**(8): 728-34.

Dubos, R. J. (1937). "Mechanism of the lysis of pneumococci by freezing and thawing, bile and other agents." Journal of Experimental Medicine **66**: 101-112.

Ellis, H. R. and L. B. Poole (1997). "Roles for the two cysteine residues of AhpC in catalysis of peroxide reduction by alkyl hydroperoxide reductase from *Salmonella typhimurium*." Biochemistry **36**(43): 13349-56.

Enright, M. C. and B. G. Spratt (1998). "A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease." Microbiology **144**(Pt 11): 3049-60.

Eskola, J., T. Kilpi, A. Palmu, J. Jokinen, J. Haapakoski, E. Herva, A. Takala, H. Kayhty, P. Karma, R. Kohberger, G. Siber and P. H. Makela (2001). "Efficacy of a pneumococcal conjugate vaccine against acute otitis media." N Engl J Med **344**(6): 403-9.

Farr, S. B. and T. Kogoma (1991). "Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*." Microbiol Rev **55**(4): 561-85.

Ferrandiz, M. J. and A. G. de la Campa (2002). "The membrane-associated F_0F_1 ATPase is essential for the viability of *Streptococcus pneumoniae*." FEMS Microbiol Lett **212**(1): 133-8.

Finland, M. (1969). "[Letter to the editor]." N Engl J Med **284**: 212.

Finn, A., Q. Zhang, L. Seymour, C. Fasching, E. Pettitt and E. N. Janoff (2002). "Induction of functional secretory IgA responses in breast milk, by pneumococcal capsular polysaccharides." J Infect Dis **186**(10): 1422-9.

Garandeau, C., H. Reglier-Poupet, I. Dubail, J. L. Beretti, P. Berche and A. Charbit (2002). "The sortase SrtA of *Listeria monocytogenes* is involved in processing of internalin and in virulence." Infect Immun **70**(3): 1382-90.

Gemmell, C. G. (1995). "Staphylococcal scalded skin syndrome." J Med Microbiol **43**(5): 318-27.

Gibson, C. M., T. C. Mallett, A. Claiborne and M. G. Caparon (2000). "Contribution of NADH oxidase to aerobic metabolism of *Streptococcus pyogenes*." J Bacteriol **182**(2): 448-55.

Gillespie, S. H. (1989). "Aspects of pneumococcal infection including bacterial virulence, host response and vaccination." J Med Microbiol **28**(4): 237-48.

Gingles, N. A., J. E. Alexander, A. Kadioglu, P. W. Andrew, A. Kerr, T. J. Mitchell, E. Hopes, P. Denny, S. Brown, H. B. Jones, S. Little, G. C. Booth and W. L. McPheat (2001). "Role of genetic resistance in invasive pneumococcal infection: identification and study of susceptibility and resistance in inbred mouse strains." Infect Immun **69**(1): 426-34.

Goebel, W. F. and M. H. Adams (1943). "The immunological properties of the heterophile antigen and somatic polysaccharide of the pneumococcus." Journal of Experimental Medicine **77**: 435-449.

Gray, B. M. and H. C. Dillon, Jr. (1989). "Natural history of pneumococcal infections." Pediatr Infect Dis J **8**(1 Suppl): S23-5.

Griffith, D. (1928). "The significance of pneumococcal types." Journal of Hygiene **27**: 113-159.

Hardie, J. M. (1986). Genus *Streptococcus*. Bergy's Manual of Systematic Bacteriology. P. H. A. Sneath, N. S. Mair, M. E. Sharpe and J. G. Holt. Baltimore, MD., William and Wilkins: 1043-1071.

Harris, T. O., D. W. Shelver, J. F. Bohnsack and C. E. Rubens (2003). "A novel streptococcal surface protease promotes virulence, resistance to opsonophagocytosis, and cleavage of human fibrinogen." J Clin Invest **111**(1): 61-70.

Hava, D. L. and A. Camilli (2002). "Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors." Mol Microbiol **45**(5): 1389-406.

Hava, D. L., C. J. Hemsley and A. Camilli (2003). "Transcriptional regulation in the *Streptococcus pneumoniae* rlrA pathogenicity islet by RlrA." J Bacteriol **185**(2): 413-21.

Havarstein, L. S., G. Coomaraswamy and D. A. Morrison (1995). "An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*." Proc Natl Acad Sci U S A **92**(24): 11140-4.

Hemsley, C., E. Joyce, D. L. Hava, A. Kawale and A. Camilli (2003). "MgrA, an orthologue of Mga, acts as a transcriptional repressor of the genes within the rlrA pathogenicity islet in *Streptococcus pneumoniae*." J Bacteriol **185**(22): 6640-7.

Higuchi, M. (1992). "Reduced nicotinamide adenine dinucleotide oxidase involvement in defense against oxygen toxicity of *Streptococcus mutans*." Oral Microbiol Immunol **7**(5): 309-14.

Hillas, P. J., F. S. del Alba, J. Oyarzabal, A. Wilks and P. R. Ortiz De Montellano (2000). "The AhpC and AhpD antioxidant defense system of *Mycobacterium tuberculosis*." J Biol Chem **275**(25): 18801-9.

Hirst, R. A., K. S. Sikand, A. Rutman, T. J. Mitchell, P. W. Andrew and C. O'Callaghan (2000). "Relative roles of pneumolysin and hydrogen peroxide from *Streptococcus pneumoniae* in inhibition of ependymal ciliary beat frequency." Infect Immun **68**(3): 1557-62.

Holmes, A. R., R. McNab, K. W. Millsap, M. Rohde, S. Hammerschmidt, J. L. Mawdsley and H. F. Jenkinson (2001). "The pavA gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence." Mol Microbiol **41**(6): 1395-408.

Hook, E. W., 3rd, C. A. Horton and D. R. Schaberg (1983). "Failure of intensive care unit support to influence mortality from pneumococcal bacteremia." JAMA **249**(8): 1055-7.

Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud and J. I. Glass (2001). "Genome of the bacterium *Streptococcus pneumoniae* strain R6." J Bacteriol **183**(19): 5709-17.

Iannelli, F., M. R. Oggioni and G. Pozzi (2002). "Allelic variation in the highly polymorphic locus pspC of *Streptococcus pneumoniae*." Gene **284**(1-2): 63-71.

Igarashi, T., E. Asaga and N. Goto (2003). "The sortase of *Streptococcus mutans* mediates cell wall anchoring of a surface protein antigen." Oral Microbiol Immunol **18**(4): 266-9.

Ilangovan, U., J. Iwahara, H. Ton-That, O. Schneewind and R. T. Clubb (2001a). "Assignment of the 1H, 13C and 15N signals of Sortase." J Biomol NMR **19**(4): 379-80.

Ilangovan, U., H. Ton-That, J. Iwahara, O. Schneewind and R. T. Clubb (2001b). "Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*." Proc Natl Acad Sci U S A **98**(11): 6056-61.

Ispahani, P., V. C. Weston, D. P. Turner and F. E. Donald (1999). "Septic arthritis due to *Streptococcus pneumoniae* in Nottingham, United Kingdom, 1985-1998." Clin Infect Dis **29**(6): 1450-4.

James, P. A. and M. G. Thomas (2000). "*Streptococcus pneumoniae* septic arthritis in adults." Scand J Infect Dis **32**(5): 491-4.

Jonsson, I. M., S. K. Mazmanian, O. Schneewind, T. Bremell and A. Tarkowski (2003). "The role of *Staphylococcus aureus* sortase A and sortase B in murine arthritis." Microbes Infect **5**(9): 775-80.

Jonsson, I. M., S. K. Mazmanian, O. Schneewind, M. Verdrengh, T. Bremell and A. Tarkowski (2002). "On the role of *Staphylococcus aureus* sortase and sortase-catalyzed surface protein anchoring in murine septic arthritis." J Infect Dis **185**(10): 1417-24.

Kadioglu, A., J. Echenique, S. Manco, M. C. Trombe and P. W. Andrew (2003). "The MicAB two-component signaling system is involved in virulence of *Streptococcus pneumoniae*." Infect Immun **71**(11): 6676-9.

Kadioglu, A., S. Taylor, F. Iannelli, G. Pozzi, T. J. Mitchell and P. W. Andrew (2002). "Upper and lower respiratory tract infection by *Streptococcus pneumoniae* is affected by pneumolysin deficiency and differences in capsule type." Infect Immun **70**(6): 2886-90.

Kalin, M. (1998). "Pneumococcal serotypes and their clinical relevance." Thorax **53**(3): 159-62.

Karlin, S. (2001). "Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes." Trends Microbiol **9**(7): 335-43.

Kaushal, D., B. G. Schroeder, S. Tyagi, T. Yoshimatsu, C. Scott, C. Ko, L. Carpenter, J. Mehrotra, Y. C. Manabe, R. D. Fleischmann and W. R. Bishai (2002). "Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative sigma factor, SigH." Proc Natl Acad Sci U S A **99**(12): 8330-5.

Kellogg, J. A., D. A. Bankert, C. J. Elder, J. L. Gibbs and M. C. Smith (2001). "Identification of *Streptococcus pneumoniae* revisited." J Clin Microbiol **39**(9): 3373-5.

Kelly, R. T., S. Farmer and D. Greiff (1967)." Neuraminidase activities of clinical isolates of *Diplococcus pneumoniae*." J Bacteriol 94(1): 272-73.

Kelly, R. B. and D. Greiff (1970). "Toxicity of pneumococcal neuraminidase." Infect Immun **2**: 115-117.

Kharat, A. S. and A. Tomasz (2003). "Inactivation of the srtA gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells in vitro." Infect Immun **71**(5): 2758-65.

Kim, J. O., S. Romero-Steiner, U. B. Sorensen, J. Blom, M. Carvalho, S. Barnard, G. Carlone and J. N. Weiser (1999). "Relationship between cell surface carbohydrates and intrastrain variation on opsonophagocytosis of *Streptococcus pneumoniae*." Infect Immun **67**(5): 2327-33.

Kim, S. W., I. M. Chang and K. B. Oh (2002). "Inhibition of the bacterial surface protein anchoring transpeptidase sortase by medicinal plants." Biosci Biotechnol Biochem **66**(12): 2751-4.

King, K. Y., J. A. Horenstein and M. G. Caparon (2000). "Aerotolerance and peroxide resistance in peroxidase and PerR mutants of *Streptococcus pyogenes*." J Bacteriol **182**(19): 5290-9.

King, S. J., P. J. Heath, I. Luque, C. Tarradas, C. G. Dowson and A. M. Whatmore (2001). "Distribution and genetic diversity of suilysin in *Streptococcus suis* isolated from different diseases of pigs and characterization of the genetic basis of suilysin absence." Infect Immun **69**(12): 7572-82.

Kohler, R., J. Fanghanel, B. Konig, E. Luneberg, M. Frosch, J. U. Rahfeld, R. Hilgenfeld, G. Fischer, J. Hacker and M. Steinert (2003). "Biochemical and functional analyses of the Mip protein: influence of the N-terminal half and of peptidylprolyl isomerase activity on the virulence of *Legionella pneumophila*." Infect Immun **71**(8): 4389-97.

Koshkin, A., C. M. Nunn, S. Djordjevic and P. R. Ortiz de Montellano (2003). "The mechanism of *Mycobacterium tuberculosis* alkylhydroperoxidase AhpD as defined by mutagenesis, crystallography, and kinetics." J Biol Chem **278**(32): 29502-8.

Kramer, M. R., B. Rudensky, I. Hadas-Halperin, M. Isacsohn and E. Melzer (1987). "Pneumococcal bacteremia--no change in mortality in 30 years: analysis of 104 cases and review of the literature." Isr J Med Sci **23**(3): 174-80.

Ladhani, S. (2001). "Recent developments in staphylococcal scalded skin syndrome." Clin Microbiol Infect 7(6): 301-7.

Ladhani, S., C. L. Joannou, D. P. Lochrie, R. W. Evans and S. M. Poston (1999). "Clinical, microbial, and biochemical aspects of the exfoliative toxins causing staphylococcal scalded-skin syndrome." Clin Microbiol Rev 12(2): 224-42.

Lange, R., C. Wagner, A. de Saizieu, N. Flint, J. Molnos, M. Stieger, P. Caspers, M. Kamber, W. Keck and K. E. Amrein (1999). "Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of Streptococcus pneumoniae." Gene 237(1): 223-34.

Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison and D. W. Holden (2001). "A functional genomic analysis of type 3 Streptococcus pneumoniae virulence." Mol Microbiol 40(3): 555-71.

Lee, S. F. and T. L. Boran (2003). "Roles of sortase in surface expression of the major protein adhesin P1, saliva-induced aggregation and adherence, and cariogenicity of Streptococcus mutans." Infect Immun 71(2): 676-81.

Lee, S. G., V. Pancholi and V. A. Fischetti (2002). "Characterization of a unique glycosylated anchor endopeptidase that cleaves the LPXTG sequence motif of cell surface proteins of Gram-positive bacteria." J Biol Chem 277(49): 46912-22.

Li, T., M. K. Khah, S. Slavnic, I. Johansson and N. Stromberg (2001). "Different type 1 fimbrial genes and tropisms of commensal and potentially pathogenic Actinomyces spp. with different salivary acidic proline-rich protein and statherin ligand specificities." Infect Immun 69(12): 7224-33.

Linder, T. E., R. L. Daniels, D. J. Lim and T. F. DeMaria (1994). "Effect of intranasal inoculation of Streptococcus pneumoniae on the structure of the surface carbohydrates of the chinchilla eustachian tube and middle ear mucosa." Microb Pathog 16(6): 435-41.

Loeffler, J. M., S. Djurkovic and V. A. Fischetti (2003). "Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia." Infect Immun 71(11): 6199-204.

Loeffler, J. M., D. Nelson and V. A. Fischetti (2001). "Rapid killing of Streptococcus pneumoniae with a bacteriophage cell wall hydrolase." Science 294(5549): 2170-2.

Marra, A., J. Asundi, M. Bartilson, S. Lawson, F. Fang, J. Christine, C. Wiesner, D. Brigham, W. P. Schneider and A. E. Hromockyj (2002). "Differential fluorescence induction analysis of Streptococcus pneumoniae identifies genes involved in pathogenesis." Infect Immun 70(3): 1422-33.

Mazmanian, S. K., G. Liu, E. R. Jensen, E. Lenoy and O. Schneewind (2000). "Staphylococcus aureus sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections." Proc Natl Acad Sci U S A 97(10): 5510-5.

Mazmanian, S. K., E. P. Skaar, A. H. Gaspar, M. Humayun, P. Gornicki, J. Jelenska, A. Joachmiak, D. M. Missiakas and O. Schneewind (2003). "Passage of heme-iron across the envelope of Staphylococcus aureus." Science 299(5608): 906-9.

Mazmanian, S. K., H. Ton-That and O. Schneewind (2001). "Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus." Mol Microbiol 40(5): 1049-57.

Mazmanian, S. K., H. Ton-That, K. Su and O. Schneewind (2002). "An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis." Proc Natl Acad Sci U S A **99**(4): 2293-8.

McCool, T. L., T. R. Cate, G. Moy and J. N. Weiser (2002). "The immune response to pneumococcal proteins during experimental human carriage." J Exp Med **195**(3): 359-65.

McCool, T. L., T. R. Cate, E. I. Tuomanen, P. Adrian, T. J. Mitchell and J. N. Weiser (2003). "Serum immunoglobulin G response to candidate vaccine antigens during experimental human pneumococcal colonization." Infect Immun **71**(10): 5724-32.

McCullers, J. A. and K. C. Bartmess (2003). "Role of neuraminidase in lethal synergism between influenza virus and *Streptococcus pneumoniae*." J Infect Dis **187**(6): 1000-9.

Miles, A. A. and S. S. Misra (1938). "The estimation of the bactericidal power of blood." Journal of Hygiene **38**: 732-749.

Miller, R. A. and B. E. Britigan (1997). "Role of oxidants in microbial pathophysiology." Clin Microbiol Rev **10**(1): 1-18.

Miyamoto, H., S. Yoshida, H. Taniguchi, M. H. Qin, H. Fujio and Y. Mizuguchi (1993). "Protein profiles of *Legionella pneumophila* Philadelphia-1 grown in macrophages and characterization of a gene encoding a novel 24 kDa Legionella protein." Microb Pathog **15**(6): 469-84.

Mongkolsuk, S. and J. D. Helmann (2002). "Regulation of inducible peroxide stress responses." Mol Microbiol **45**(1): 9-15.

Nandoskar, M., A. Ferrante, E. J. Bates, N. Hurst and J. C. Paton (1986). "Inhibition of human monocyte respiratory burst, degranulation, phospholipid methylation and bactericidal activity by pneumolysin." Immunology **59**(4): 515-20.

Navarre, W. W. and O. Schneewind (1999). "Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope." Microbiol Mol Biol Rev **63**(1): 174-229.

Ng, W. L., K. M. Kazmierczak, G. T. Robertson, R. Gilmour and M. E. Winkler (2003). "Transcriptional regulation and signature patterns revealed by microarray analyses of *Streptococcus pneumoniae* R6 challenged with sublethal concentrations of translation inhibitors." J Bacteriol **185**(1): 359-70.

Nunn, C. M., S. Djordjevic, P. J. Hillas, C. R. Nishida and P. R. Ortiz de Montellano (2002). "The crystal structure of *Mycobacterium tuberculosis* alkylhydroperoxidase AhpD, a potential target for antitubercular drug design." J Biol Chem **277**(22): 20033-40.

Obaro, S. (2001). "Differences in invasive pneumococcal serotypes." Lancet **357**(9270): 1800-1.

Obaro, S. and R. Adegbola (2002). "The pneumococcus: carriage, disease and conjugate vaccines." J Med Microbiol **51**(2): 98-104.

Obaro, S. K. (2000). "Confronting the pneumococcus: a target shift or bullet change?" Vaccine **19**(9-10): 1211-7.

Osaki, M., D. Takamatsu, Y. Shimoji and T. Sekizaki (2002). "Characterization of *Streptococcus suis* genes encoding proteins homologous to sortase of gram-positive bacteria." J Bacteriol **184**(4): 971-82.

Osaki, M., D. Takamatsu, Y. Shimoji and T. Sekizaki (2003). "Allelic variation in srtAs of *Streptococcus suis* strains." FEMS Microbiol Lett **219**(2): 195-201.

O'Toole, R. D., L. Goode and C. Howe (1971). "Neuraminidase activity in bacterial meningitis." J Clin Invest **50**(5): 979-85.

Pallen, M. J., A. C. Lam, M. Antonio and K. Dunbar (2001). "An embarrassment of sortases - a richness of substrates?" Trends Microbiol **9**(3): 97-102.

Paterson, G. K. and T. J. Mitchell (2004). "The biology of Gram-positive sortase enzymes." Trends Microbiol 12 (2): 89-95.

Paton, J. C., A. M. Berry and R. A. Lock (1997). "Molecular analysis of putative pneumococcal virulence proteins." Microb Drug Resist **3**(1): 1-10.

Paton, J. C. and A. Ferrante (1983). "Inhibition of human polymorphonuclear leukocyte respiratory burst, bactericidal activity, and migration by pneumolysin." Infect Immun **41**(3): 1212-6.

Paton, J. C. and P. Giammarinaro (2001). "Genome-based analysis of pneumococcal virulence factors: the quest for novel vaccine antigens and drug targets." Trends Microbiol **9**(11): 515-8.

Pearce, B. J., Y. B. Yin and H. R. Masure (1993). "Genetic identification of exported proteins in *Streptococcus pneumoniae*." Mol Microbiol **9**(5): 1037-50.

Pericone, C. D., D. Bae, M. Shchepetov, T. McCool and J. N. Weiser (2002). "Short-sequence tandem and nontandem DNA repeats and endogenous hydrogen peroxide production contribute to genetic instability of *Streptococcus pneumoniae*." J Bacteriol **184**(16): 4392-9.

Pericone, C. D., K. Overweg, P. W. Hermans and J. N. Weiser (2000). "Inhibitory and bactericidal effects of hydrogen peroxide production by *Streptococcus pneumoniae* on other inhabitants of the upper respiratory tract." Infect Immun **68**(7): 3990-7.

Pericone, C. D., S. Park, J. A. Imlay and J. N. Weiser (2003). "Factors contributing to hydrogen peroxide resistance in *Streptococcus pneumoniae* include pyruvate oxidase (SpxB) and avoidance of the toxic effects of the fenton reaction." J Bacteriol **185**(23): 6815-25.

Perry, A. M., H. Ton-That, S. K. Mazmanian and O. Schneewind (2002). "Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. III. Lipid II is an in vivo peptidoglycan substrate for sortase-catalyzed surface protein anchoring." J Biol Chem **277**(18): 16241-8.

Polissi, A., A. Pontiggia, G. Feger, M. Altieri, H. Mottl, L. Ferrari and D. Simon (1998). "Large-scale identification of virulence genes from *Streptococcus pneumoniae*." Infect Immun **66**(12): 5620-9.

Poole, L. B. and H. R. Ellis (1996). "Flavin-dependent alkyl hydroperoxide reductase from *Salmonella typhimurium*. 1. Purification and enzymatic activities of overexpressed AhpF and AhpC proteins." Biochemistry **35**(1): 56-64.

Poolman, B. (1993). "Energy transduction in lactic acid bacteria." FEMS Microbiol Rev **12**(1-3): 125-47.

Robertson, G. T., W. L. Ng, J. Foley, R. Gilmour and M. E. Winkler (2002). "Global transcriptional analysis of clpP mutations of type 2 *Streptococcus pneumoniae* and their effects on physiology and virulence." J Bacteriol **184**(13): 3508-20.

Rubins, J. B. and E. N. Janoff (1998). "Pneumolysin: a multifunctional pneumococcal virulence factor." J Lab Clin Med **131**(1): 21-7.

Ruzin, A., A. Severin, F. Ritacco, K. Tabei, G. Singh, P. A. Bradford, M. M. Siegel, S. J. Projan and D. M. Shlaes (2002). "Further evidence that a cell wall precursor [C(55)-MurNAc-(peptide)-GlcNAc] serves as an acceptor in a sorting reaction." J Bacteriol **184**(8): 2141-7.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Scanlon, K. L., W. F. Diven and R. H. Glew (1989). "Purification and properties of *Streptococcus pneumoniae* neuraminidase." Enzyme **41**(3): 143-50.

Schneewind, O., A. Fowler and K. F. Faull (1995). "Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*." Science **268**(5207): 103-6.

Schneewind, O., P. Model and V. A. Fischetti (1992). "Sorting of protein A to the staphylococcal cell wall." Cell **70**(2): 267-81.

Sebert, M. E., L. M. Palmer, M. Rosenberg and J. N. Weiser (2002). "Microarray-based identification of htrA, a *Streptococcus pneumoniae* gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization." Infect Immun **70**(8): 4059-67.

Sethi, S. and T. F. Murphy (2000). "Chronic obstructive pulmonary disease." N Engl J Med **343**(26): 1969-70; author reply 1970-1.

Shakhnovich, E. A., S. J. King and J. N. Weiser (2002). "Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of Neisseria meningitidis and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract." Infect Immun **70**(12): 7161-4.

Smith, A. W., H. Roche, M. C. Trombe, D. E. Briles and A. Hakansson (2002). "Characterization of the dihydrolipoamide dehydrogenase from *Streptococcus pneumoniae* and its role in pneumococcal infection." Mol Microbiol **44**(2): 431-48.

Spellerberg, B., D. R. Cundell, J. Sandros, B. J. Pearce, I. Idanpaan-Heikkila, C. Rosenow and H. R. Masure (1996). "Pyruvate oxidase, as a determinant of virulence in *Streptococcus pneumoniae*." Mol Microbiol **19**(4): 803-13.

Sung, C. K., H. Li, J. P. Claverys and D. A. Morrison (2001). "An rpsL cassette, Janus, for gene replacement through negative selection in *Streptococcus pneumoniae*." Appl Environ Microbiol **67**(11): 5190-6.

Swiatlo, E. and D. Ware (2003). "Novel vaccine strategies with protein antigens of *Streptococcus pneumoniae*." FEMS Immunol Med Microbiol **38**(1): 1-7.

Takamatsu, D., M. Osaki and T. Sekizaki (2002). "Evidence for lateral transfer of the Suilysin gene region of *Streptococcus suis*." J Bacteriol **184**(7): 2050-7.

Tan, T. Q. (2000). "Pneumococcal conjugate vaccines - implications for community antibiotic prescribing." Curr Opin Microbiol **3**(5): 502-7.

Tan, T. Q. (2003). "Antibiotic resistant infections due to *Streptococcus pneumoniae*: impact on therapeutic options and clinical outcome." Curr Opin Infect Dis **16**(3): 271-7.

Terao, Y., S. Kawabata, M. Nakata, I. Nakagawa and S. Hamada (2002). "Molecular characterization of a novel fibronectin-binding protein of *Streptococcus pyogenes*

strains isolated from toxic shock-like syndrome patients." J Biol Chem 277(49): 47428-35.

Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead and C. M. Fraser (2001). "Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*." Science 293(5529): 498-506.

Throup, J. P., K. K. Koretke, A. P. Bryant, K. A. Ingraham, A. F. Chalker, Y. Ge, A. Marra, N. G. Wallis, J. R. Brown, D. J. Holmes, M. Rosenberg and M. K. Burnham (2000). "A genomic analysis of two-component signal transduction in Streptococcus pneumoniae." Mol Microbiol 35(3): 566-76.

Tomasz, A. (1965). "Control of the competent state in Pneumococcus by a hormone-like cell product: an example for a new type of regulatory mechanism in bacteria." Nature 208(6): 155-9.

Tomasz, A. (1999). "New faces of an old pathogen: emergence and spread of multidrug-resistant *Streptococcus pneumoniae*." Am J Med 107(1A): 55S-62S.

Tomasz, A., P. Moreillon and G. Pozzi (1988). "Insertional inactivation of the major autolysin gene of *Streptococcus pneumoniae*." J Bacteriol 170(12): 5931-4.

Tong, H. H., L. E. Blue, M. A. James and T. F. DeMaria (2000). "Evaluation of the virulence of a *Streptococcus pneumoniae* neuraminidase-deficient mutant in nasopharyngeal colonization and development of otitis media in the chinchilla model." Infect Immun 68(2): 921-4.

Tong, H. H., M. James, I. Grants, X. Liu, G. Shi and T. F. DeMaria (2001). "Comparison of structural changes of cell surface carbohydrates in the eustachian tube epithelium of chinchillas infected with a *Streptococcus pneumoniae* neuraminidase-deficient mutant or its isogenic parent strain." Microb Pathog 31(6): 309-17.

Tong, H. H., X. Liu, Y. Chen, M. James and T. Demaria (2002). "Effect of neuraminidase on receptor-mediated adherence of *Streptococcus pneumoniae* to chinchilla tracheal epithelium." Acta Otolaryngol 122(4): 413-9.

Ton-That, H., G. Liu, S. K. Mazmanian, K. F. Faull and O. Schneewind (1999). "Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif." Proc Natl Acad Sci U S A 96(22): 12424-9.

Ton-That, H., S. K. Mazmanian, L. Alksne and O. Schneewind (2002). "Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Cysteine 184 and histidine 120 of sortase form a thiolate-imidazolium ion pair for catalysis." J Biol Chem 277(9): 7447-52.

Ton-That, H., S. K. Mazmanian, K. F. Faull and O. Schneewind (2000). "Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates." J Biol Chem 275(13): 9876-81.

Ton-That, H. and O. Schneewind (1999). "Anchor structure of staphylococcal surface proteins. IV. Inhibitors of the cell wall sorting reaction." J Biol Chem **274**(34): 24316-20.

Tseng, H. J., A. G. McEwan, J. C. Paton and M. P. Jennings (2002). "Virulence of *Streptococcus pneumoniae*: PsaA mutants are hypersensitive to oxidative stress." Infect Immun **70**(3): 1635-9.

Van Ginkel, F. W., J. R. McGhee, J. M. Watt, A. Campos-Torres, L. A. Parish and D. E. Briles (2003). "Pneumococcal carriage results in ganglioside- mediated olfactory tissue infection." Proc Natl Acad Sci U S A **10**: 10.

Watson, D. A. and D. M. Musher (1990). "Interruption of capsule production in Streptococcus pneumonia serotype 3 by insertion of transposon Tn916." Infect Immun **58**(9): 3135-8.

Weiser, J. N. (2000). Phase variation in colony opacity by *Streptococcus pneumoniae*. Streptococcus pneumoniae: Molecular biology and mechanisms of disease. A. Tomasz. New York, Mary Ann Liebert, Inc: 245-251.

Weiser, J. N., D. Bae, C. Fasching, R. W. Scamurra, A. J. Ratner and E. N. Janoff (2003). "Antibody-enhanced pneumococcal adherence requires IgA1 protease." Proc Natl Acad Sci U S A **100**(7): 4215-20.

Wellmer, A., G. Zysk, J. Gerber, T. Kunst, M. Von Mering, S. Bunkowski, H. Eiffert and R. Nau (2002). "Decreased virulence of a pneumolysin-deficient strain of *Streptococcus pneumoniae* in murine meningitis." Infect Immun **70**(11): 6504-8.

Whatmore, A. M., S. J. King, N. C. Doherty, D. Sturgeon, N. Chanter and C. G. Dowson (1999). "Molecular characterization of equine isolates of *Streptococcus pneumoniae*: natural disruption of genes encoding the virulence factors pneumolysin and autolysin." Infect Immun **67**(6): 2776-82.

Winter, A. J., S. D. Comis, M. P. Osborne, M. J. Tarlow, J. Stephen, P. W. Andrew, J. Hill and T. J. Mitchell (1997). "A role for pneumolysin but not neuraminidase in the hearing loss and cochlear damage induced by experimental pneumococcal meningitis in guinea pigs." Infect Immun **65**(11): 4411-8.

Wintermeyer, E., B. Ludwig, M. Steinert, B. Schmidt, G. Fischer and J. Hacker (1995). "Influence of site specifically altered Mip proteins on intracellular survival of *Legionella pneumophila* in eukaryotic cells." Infect Immun **63**(12): 4576-83.

Wizemann, T. M., J. H. Heinrichs, J. E. Adamou, A. L. Erwin, C. Kunsch, G. H. Choi, S. C. Barash, C. A. Rosen, H. R. Masure, E. Tuomanen, A. Gayle, Y. A. Brewah, W. Walsh, P. Barren, R. Lathigra, M. Hanson, S. Langermann, S. Johnson and S. Koenig (2001). "Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection." Infect Immun **69**(3): 1593-8.

Wizemann, T. M., J. Moskovitz, B. J. Pearce, D. Cundell, C. G. Arvidson, M. So, H. Weissbach, N. Brot and H. R. Masure (1996). "Peptide methionine sulfoxide reductase contributes to the maintenance of adhesins in three major pathogens." Proc Natl Acad Sci U S A **93**(15): 7985-90.

Wuorimaa, T. and H. Kayhty (2002). "Current state of pneumococcal vaccines." Scand J Immunol **56**(2): 111-29.

Yamaguchi, T., K. Nishifuji, M. Sasaki, Y. Fudaba, M. Aepfelbacher, T. Takata, M. Ohara, H. Komatsuzawa, M. Amagai and M. Sugai (2002). "Identification of the

Staphylococcus aureus etd pathogenicity island which encodes a novel exfoliative toxin, ETD, and EDIN-B." Infect Immun **70**(10): 5835-45.

Yamamoto, Y., M. Higuchi, L. B. Poole and Y. Kamio (2000). "Role of the dpr product in oxygen tolerance in *Streptococcus mutans*." J Bacteriol **182**(13): 3740-7.

Yesilkaya, H., A. Kadioglu, N. Gingles, J. E. Alexander, T. J. Mitchell and P. W. Andrew (2000). "Role of manganese-containing superoxide dismutase in oxidative stress and virulence of *Streptococcus pneumoniae*." Infect Immun **68**(5): 2819-26.

Yeung, M. K., J. A. Donkersloot, J. O. Cisar and P. A. Ragsdale (1998). "Identification of a gene involved in assembly of *Actinomyces naeslundii* T14V type 2 fimbriae." Infect Immun **66**(4): 1482-91.

Yu, J., A. P. Bryant, A. Marra, M. A. Lonetto, K. A. Ingraham, A. F. Chalker, D. J. Holmes, D. Holden, M. Rosenberg and D. McDevitt (2001). "Characterization of the *Streptococcus pneumoniae* NADH oxidase that is required for infection." Microbiology **147**(Pt 2): 431-8.

Zahner, D. and R. Hakenbeck (2000). "The *Streptococcus pneumoniae* beta-galactosidase is a surface protein." J Bacteriol **182**(20): 5919-21.

Zhang, J. R., K. E. Mostov, M. E. Lamm, M. Nanno, S. Shimida, M. Ohwaki and E. Tuomanen (2000). "The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells." Cell **102**(6): 827-37.

Zysk, G., R. J. Bongaerts, E. ten Thoren, G. Bethe, R. Hakenbeck and H. P. Heinz (2000). "Detection of 23 immunogenic pneumococcal proteins using convalescent-phase serum." Infect Immun **68**(6): 3740-3.

Additional Papers

Additional papers published from work undertaken in this PhD but not forming part of the thesis itself are included in support:

Wei Xq, Orchardson M, Gracie JA, Leung BP, Gao Bm, Guan H, Niedbala W, **Paterson GK**, McInnes IB and Liew FY. 2001. The Sushi domain of soluble IL-15 receptor alpha is essential for binding IL-15 and inhibiting inflammatory and allogenic responses in vitro and in vivo. *Journal of Immunology* 167:277-82.

Blue CE, **Paterson GK**, Kerr AR, Berge M, Claverys JP and Mitchell TJ. 2003. ZmpB, a novel virulence factor of Streptococcus pneumoniae that induces tumor necrosis factor alpha production in the respiratory tract. *Infection and Immunity* 71:4925-35.

Paterson, GK and T. J. Mitchell. 2004.
The biology of Gram-positive sortase enzymes. *Trends in Microbiology* 12: 89-95.