Ecology of virulence genes in the human pathogen *Streptococcus pneumoniae*

A thesis submitted to the University of Glasgow for the degree of

Doctor of Philosophy

By

Calum Harvey Gustave Johnston B.Sc.

Submitted September 2008

Division of Infection and Immunity
Faculty of Biomedical and Life Sciences
Glasgow Biomedical Research Centre
Glasgow University
120 University Place
Glasgow G12 8TA

Acknowledgements

I would firstly like to thank Prof. Tim Mitchell and Dr. Andrew Smith, for the opportunity to work with them and for their support and input, which has been invaluable. Thanks to my assessor Dr. Roger Parton for assuring smooth progress of my PhD. Thanks also to many members of the Mitchell lab for help and support, both scientific and emotional. In particular, thanks to Ali Saville, Paul Hughes, Lea-Ann Kirkham, Patricia Romero, Jo Jefferies and Ashleigh Holmes for making the lab an enjoyable and fun place to work. Thanks to Denise Candlish, Donald Inverarity, Gill Douce and Kirsty Ross for invaluable help in the lab.

Many thanks go to my mum and dad for believing in me, and for their constant support, I would not have got this far without them.

Finally, thanks to Jen for keeping me going for the last 3 years, and for all her love and support, without which I would be lost.

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

Calum Harvey Gustave Johnston

September 2008

Abstract

Streptococcus pneumoniae, also known as the pneumococcus, is an important human pathogen, with high burdens of disease and mortality worldwide. There are over 90 serotypes of this pathogen, demonstrating the vast amounts of diversity present. Currently, there are two pneumococcal vaccines, both targeting the polysaccharide capsule. However, one vaccine is ineffective in the paediatric population, whilst the other only targets a minority of disease-causing serotypes, and has increased disease caused by serotypes not present in the vaccine. One solution is a new pneumococcal vaccine targeting a protein virulence factor possessed by all pneumococci, which would afford cross-serotype protection. As a result, it is important to assess the diversity of pneumococcal virulence factors in order to determine their potential as vaccine candidates, as excess diversity present may prevent full serotype-independent protection of a vaccine. Furthermore, diversity studies offer important insight on pneumococcal biology, epidemiology and pathogenesis.

The diversity in the toxin pneumolysin (Ply) was greater than previously thought, with 14 protein alleles discovered. However, diversity remained significantly lower than surface-exposed virulence factors, indicating this toxin may be more suitable as a vaccine candidate. Furthermore, all 14 alleles were recognised by polyclonal antibodies, showing the potential cross-serotype protection of a vaccine targeting this toxin. A novel non-haemolytic Ply allele was associated with clones recently expanding in the pneumococcal population, as well as serotypes associated with outbreaks of pneumococcal disease. The non-haemolytic toxin may therefore play a role in driving clonal expansion in certain genetic backgrounds, or be involved in establishing outbreaks of pneumococcal disease.

The diversity in the neuraminidase A (NanA) enzyme was significantly higher than in Ply, with many point mutations, mosaic blocks and insertions regions present in 18 divergent alleles. This level of diversity should not be prohibitive to the use of this protein as a vaccine candidate, as polyclonal antibodies recognised 4 NanA alleles of significant diversity, indicating the possibility of cross-serotype protection. The role of NanA in pathogenesis of pneumococcal

haemolytic uraemic syndrome (p-HUS) was investigated, although there was no correlation between p-HUS and NanA allele or activity.

The novel discovery that pneumococcal NanA was inhibited by viral neuraminidase inhibitors of influenza allowed insight into the synergistic relationship between these two deadly pathogens, and showed for the first time that treatment with these drugs acts on both the primary and secondary pathogen. One of these inhibitors, Oseltamivir, was found to have potential in treating secondary pneumococcal pneumonia, which may help decrease the significant burden of this disease, as well as reducing the over-reliance on antibiotics for treating pneumococcal diseases.

Homologues of Ply and NanA were identified and characterised in the related species *Streptococcus mitis* and *Streptococcus pseudopneumoniae*, giving insight into the evolutionary relationships between these species. Furthermore, the presence of these homologues in related species gives rise to the possibility of pneumococci acquiring altered genes through horizontal gene transfer. The selective pressure of a vaccine targeting these proteins may give evolutionary advantage to these pneumococci, resulting in evasion of a vaccine.

Microarray studies have been used to assess pneumococcal diversity at a genome-wide level. Gene expression studies identified candidate genes which may play a role in p-HUS pathogenesis. Further studies into this area will improve the diagnosis and treatment of this disease. Furthermore, a large number of established pneumococcal virulence factors, many of which are vaccine candidates, were found to have homologues in closely related commensal species. These results must be taken into consideration for future protein-based pneumococcal vaccine design.

Contents

List of Tables

List of Figures

Abbreviations

α-	Anti
ß-	Beta
°C	Degrees Celsius
Δ	Deletion
Δ6	Ply with AA146-147 deletion/*S. pneumoniae* D39 with this toxin
ΔPly stop	Ply with stop codon at start/S. *pneumoniae* D39 with this toxin
ΔNanA	S. *pneumoniae* D39 strain with NanA gene deleted
%	Percent
µl	micro litre
µm	micrometer
µM	micro molar
A	Absorbance
AA	Amino acid
Ab	Antibody
ABC	ATP-binding cassette
aHUS	Atypical haemolytic uraemic syndrome
APS	Ammonium persulphate
BAB	Blood Agar Base
BBB	Blood Brain Barrier
BHI	Brain Heart Infusion
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
C	Complement
Cbp	Choline Binding Protein
CDC	Cholesterol-dependent Cytolysin
cDNA	Complementary DNA
CFU	Colony Forming Unit
Clp	Caseinolytic protease
CPS	Capsule Polysaccharide
CSF	Cerebrospinal Fluid
CSP	Competence Stimulating Peptide
cv	Column Volumes
D39	S. *pneumoniae* serotype 2 strain D39 (NCTC number: 7466)
DEL	Deletion
dH_20	Distilled water

DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DLV	Double-locus variant
DTT	Dithiothreitol
ELISA	Enzyme Linked Immunosorbent Assay
EU	Enzymatic Units
FBS	Foetal Bovine Serum
g	Gram
g	centrifugal force
gDNA	Genomic DNA
h	Hour
H_2O_2	Hydrogen peroxide
His-Tag	Histidine Affinity tag
HIV	Human immunodeficiency virus
HPA	Health Protection Agency, Colindale, London
HRP	Horseradish peroxidase
HU	Haemolytic Units
HUS	Haemolytic uraemic syndrome
Hyl	Hyaluronidase
IAA	Isoamyl alcohol
IC_{50}	Concentration required to cause 50% inhibition
ICAM-1	Intercellular adhesion molecule 1
IFN-γ	Interferon gamma
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IL	Interleukin
ILY	Intermedilysin
INS	Insertion
i.n.	Intranasal
i.p.	Intraperitoneal
IPD	Invasive pneumococcal disease
IPTG	Isopropyl-β-D-Thiogalactopyranoside
Kb	Kilobase
kDa	Kilodalton
L	Litre

LB	Luria Broth
LPS	Lipopolysaccharide
LytA	Autolysin A
M	Molar
mAb	Monoclonal Antibody
mAbPLY-7	Monoclonal Antibody PLY-7
MEGA	Molecular Evolutionary Genetics Analysis
mg	Milligram(s)
min	Minute(s)
ml	Millilitre
MLST	Multi-Locus Sequence Typing
Mly	Mitilysin
mM	Millimolar
MW	Molecular weight
n	Group size
NA	Influenza neuraminidase
NanA	Neuraminidase A
NanB	Neuraminidase B
NanC	Neuraminidase C
Neu5ac	N-acetylneuraminic acid
nm	Nanometre
nM	Nanomole
OD	Optical Density
OC	Oseltamivir carboxylate
PAFr	Platelet activating factor receptor
PBP	Penicillin-binding protein
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PCV	Pneumococcal Conjugate Vaccine
PdT	Ply carrying triple mutations: D385N, C428G, W433F
PFO	Perfringolysin O
Pht	Pneumococcal histidine triad protein
p-HUS	Pneumococcal haemolytic syndrome
Pia	Pneumococcal iron acquisition protein
Piu	Pneumococcal iron uptake protein
Ply	Pneumolysin
pNP	p-nitrophenol

pNP-NANA	2-*O*-(*p*-nitrophenyl)-α-D-*N*-acetylneuraminic acid
pPly	Pseudopneumolysin
PpmA	Proteinase maturation protein A
PppA	Pneumococcal protective protein A
PPV	Pneumococcal Polysaccharide Vaccine
PS	Polysaccharide
Psa	Pneumococcal surface adhesion protein
Psp	Pneumococcal surface protein
RD	Region of diversity
RNA	Ribonucleic acid
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
RT	Room Temperature (~20°C)
RT-PCR	Reverse-transcriptase PCR
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
sec	Second(s)
SLO	Streptolysin O
SLV	Single-locus variant
SMPRL	Scottish Meningococcal and Pneumococcal Reference Laboratory
SodA	Superoxide dismutase
ST	Sequence Type
T antigen	Thomsen-Friedrich antigen
TE	TRIS-EDTA
TIGR	The Institute for Genomic Research
TIGR4	Serotype 4 *S. pneumoniae* sequenced by TIGR (ATCC: BAA-334)
TLR	Toll-like Receptor
TLV	Triple-locus variant
TNF	Tumour Necrosis Factor
U	Units
UK	United Kingdom
USA	United states of America
V	Volts
VGS	viridans group *Streptococci*
v/v	Volume/volume
WHO	World Health Organisation
wk	Week

WT	Wild type
X-GAL	5-Bromo-4-chloro-3-indolyl ß-D-galactopyranoside
ZmpB	Zinc metalloprotease B

1 Introduction

1.1 *Streptococcus pneumoniae*

Streptococcus pneumoniae, otherwise known as the pneumococcus, is a Gram-positive facultative anaerobe, and an important human pathogen. It causes a variety of diseases ranging from fatal meningitis and pneumonia to the relatively benign otitis media (Cartwright *et al.*, 2002). Diseases are caused primarily in children and the elderly, as well as immuno-compromised patients. The bacterium was identified by both Louis Pasteur in France and George Sternberg in America in 1881 (Pasteur, 1881, Sternberg, 1881). There are currently 91 different pneumococcal serotypes identified, which are characterised by polysaccharide immunochemistry (Henrichsen, 1995; Park *et al.*, 2007). Initially, serotypes were numbered chronologically by order of discovery (Dochez and Gillespie, 1913), although cross-reaction of a number of these has resulted in reclassification of serotype nomenclature to incorporate serogrouping of closely related serotypes (Smart, 1986). Serotypes can be identified by the Quellung reaction, which allows determination of pneumococcal capsule identity by reaction with antibodies to homologous capsule polysaccharides. This technique is still employed in laboratories today, although it is hindered by cross-reactivity of a number of serotypes resulting in mistyping, as well as non-expression of capsule genes resulting in non-typable isolates. Recent studies have concentrated on developing more sensitive methods of pneumococcal serotyping, with reduced expense (Batt *et al.*, 2005; Pai *et al.*, 2005a). Furthermore, a method of multiplex PCR has recently been shown to differentiate between pneumococcal serotypes (Pai *et al.*, 2006).

Multi-locus sequence typing (MLST) is a method of further pneumococcal classification pioneered by Enright and Spratt (Enright and Spratt, 1998). This method allows genetic characterisation of pneumococcal isolates by comparing the sequences of seven housekeeping genes (*aroE, gdh, gki, recP, spi, xpt and ddl*), with each allele assigned a different number after comparison with sequences in the MLST database (www.mlst.net). This results in a "barcode" of seven allele numbers for each isolate, and this sequence of numbers is compared to barcodes for other isolates present in the database. From this, each isolate is assigned a Sequence Type (ST), one number to represent the barcode of seven allele numbers produced. Therefore, isolates with identical alleles of these 7 housekeeping genes share the same ST. MLST is a powerful tool for molecular

characterisation, as it provides a portable, electronic method of comparing relatedness of species, allowing comparison of isolates between laboratories. Results of MLST have shown that a large variety of STs are present in each serotype, as well as STs being present in more than one serotype, showing that pneumococcal isolates have the ability to switch their polysaccharide capsule (Coffey *et al.*, 1999; Coffey *et al.*, 1998a; Coffey *et al.*, 1998b; Jefferies *et al.*, 2003). This has most recently been observed in relation to serotype replacement due to vaccine pressure (Brueggemann *et al.*, 2007; Pai *et al.*, 2005b). More accurate comparison of pneumococcal isolates can be determined by use of microarray technology and full genome sequencing. Indeed, use of microarrays has recently identified genetic differences between isolates considered identical by serotype and MLST (Silva *et al.*, 2006), indicating that even serotype and MLST are not robust enough methods of pneumococcal classification. Sequencing of entire pneumococcal genomes obviously has the most merit in terms of pneumococcal classification, and previous sequencing projects have returned valuable information on the pneumococcus (Hiller *et al.*, 2007; Hoskins *et al.*, 2001; Lanie *et al.*, 2007; Tettelin *et al.*, 2001). Furthermore, this technique is becoming more affordable, and our laboratory is currently sequencing ~25 pneumococcal genomes of various serotypes and ST. However, these techniques are not affordable or time-efficient for use in a reference laboratory. Therefore, it has been proposed that sequencing of essential pneumococcal virulence factors, along with house-keeping genes, may increase the distinguishing potential of MLST (Hanage *et al.*, 2005).

1.2 Carriage of *S. pneumoniae*

S. pneumoniae is carried asymptomatically in the upper respiratory tract of up to 30% of the human population, particularly in children (Tuomanen *et al.*, 2006). The nasopharynx is a highly competitive environment, colonised by other species including *Neisseria meningitidis, Haemophilus influenzae, Staphylococcus aureus* and other *Streptococcal* species (Bogaert *et al.*, 2004). The pneumococcus produces H_2O_2, which has been shown to inhibit or kill other nasopharyngeal microflora (Pericone *et al.*, 2000). Furthermore, the NanA enzyme of *S. pneumoniae* has been shown to desialiate the lipopolysaccharide of *N. meningitidis* and *H. influenzae,* exposing them to the host immune system. This confers a competitive advantage to the pneumococcus (Shakhnovich *et al.*,

2002). These facets of the pneumococcus offer a competitive advantage over other colonisers. The nasopharynx is where the pneumococcus comes into close contact with related members of the viridans group *streptococci* such as *Streptococcus mitis* and *Streptococcus oralis*. As these species are naturally transformable, and can uptake DNA from their environment, exchange of genetic information between species can occur. This can have important consequences for treatment of pneumococcal diseases, as has been demonstrated by the transfer of penicillin resistance from *S. mitis* and *S. oralis* to *S. pneumoniae* via the *pbp2b* gene (Coffey *et al.*, 1993; Dowson *et al.*, 1993). It has also been demonstrated that regions of variation within the pneumococcal virulence factors NanA and immunoglobulin A protease are a result of homologous recombination with *S. oralis* (King *et al.*, 2005; Poulsen *et al.*, 1998). Exchange of genetic material can therefore occur between these species, although it is unclear in which direction the flow of genetic material generally occurs.

Pneumococcal carriage is thought to be required for progression to pneumococcal disease. There are a number of risk factors that have been shown to affect colonisation levels, including antibiotic use, smoking, asthma, ethnicity and overcrowding (Cardozo *et al.*, 2008; DeMaria *et al.*, 1980; Obaro, 2000). Carriage studies are important in order to establish epidemiological information for comparison with studies of invasive pneumococcal disease. Although extensive pneumococcal carriage studies are rare, a review showed that the serotypes most commonly isolated from carriage in Europe included 6A, 6B, 9V, 14, 18C, 19F and 23F (Bogaert *et al.*, 2004), and a recent study showed that particular clones of pneumococci have higher propensities for transmission and persistence in the setting of a day care centre (Sa-Leao *et al.*, 2008).

1.3 Diseases caused by *S. pneumoniae*

The pneumococcus can cause a wide range of diseases, ranging from the relatively benign in otitis media and sinusitis to the severe and potentially fatal in meningitis, pneumonia and bacteraemia (Cartwright *et al.*, 2002). Other complications of pneumococcal disease include empyema and haemolytic uraemic syndrome. Although there is not an exclusive relationship, some serotypes are more frequently isolated from invasive pneumococcal disease (IPD) than others, whilst others appear to be more frequently isolated from carriage.

Brueggemann et al. (2003) employed a method allowing comparison of "invasiveness" of particular clones or serotypes by comparing numbers found in carriage to those found in IPD in children <5 in Oxford, UK, allowing comparison of odds ratios. The results showed that serotypes 1, 4, 14 and 18C were highly invasive, whilst 6B, 9V, 19F and 23F were more likely to be associated with carriage (Brueggemann *et al.*, 2003). Meanwhile, a study in Sweden compared odds ratios of invasive isolates, describing serotype 1 and 7F isolates as "primary pathogens" as they primarily infect healthy individuals, are rarely carried, and cause less severe IPD. Other serotypes were described as opportunistic pathogens, causing more severe diseases in immuno-compromised patients (Sjostrom *et al.*, 2006). These studies show that there are major differences observed in invasive potential of different pneumococcal serotypes and clones. The transition from carriage to invasive disease is poorly understood, although it has been shown that IPD results from a pneumococcal serotype already colonising the host (Lloyd-Evans *et al.*, 1996). This implies that isolates of serotypes such as serotype 1, which are very rarely found in carriage, may be only carried for a shrot period of time before causing IPD, whilst other serotypes may establish themselves in the host nasopharynx before progressing to IPD, possibly as a result of increased risk factors such as immuno-suppression or predisposing diseases. Much work has been done into such risk factors predisposing to pneumococcal disease. Clearly, the vast majority of IPD occurs in children, the elderly and immuno-compromised patients, and so falling into these categories can be considered a major risk factor for IPD (Tuomanen *et al.*, 2006). The immuno-suppression associated with the HIV virus has been shown to predispose to pneumococcal bacteraemia, with differences in serotype and antibiotic susceptibilities also noted (Jones *et al.*, 1998). Further studies noted differences in clinical presentation of pneumococcal diseases in HIV-infected patients. Infection with Influenza has been shown to predispose to pneumococcal disease, with the pneumococcus being the leading cause of secondary bacterial pneumonia after influenza infection (McCullers and Rehg, 2002). This lethal synergism is discussed further in Section 1.9.

1.3.1 Pneumonia and bacteraemia

Pneumococcal pneumonia is characterised by build-up of fluid in the lungs, which prevents optimum transfer of oxygen to the bloodstream by restricting

access to the alveoli. Pneumonia is responsible for around 20% of all childhood deaths, mainly in developing countries (Black *et al.*, 2003; Mulholland, 1999). The pneumococcus accounts for a large proportion of this, with an estimated 1-4 million cases of pneumococcal pneumonia every year in Africa alone (Scott, 2007) and 1 million deaths from pneumococcal pneumonia annually, mostly in children in developing countries. In developed countries, the pneumococcus is the most common cause of community-acquired pneumonia, and has a 10-20% mortality rate (www.who.int/vaccine_research/diseases/ari/en/index5.html). This mortality rate increased in high-risk groups such as paediatric and geriatric populations, as well as HIV-infected individuals (Amdahl *et al.*, 1995; Johnston, 1991). The roles of many pneumococcal virulence factors have been investigated in pneumococcal pneumonia, with Ply, autolysin and pneumococcal surface protein C shown to be essential for progression of pneumococcal pneumonia (Canvin *et al.*, 1995; Jounblat *et al.*, 2003; Rubins *et al.*, 1996). Pneumococcal bacteraemia often occurs after the onset of pneumococcal pneumonia, and is characterised by circulation of pneumococci in the bloodstream. It is thought that access to the bloodstream is gained through the alveoli in the lungs. A recent study reported that bacteraemia was the most common type of IPD reported in children under 3 in Switzerland, representing 50% of cases (Myers and Gervaix, 2007), with similar results reported in the USA (Kaplan *et al.*, 1998).

1.3.2 Meningitis

Pneumococcal meningitis is characterised by inflammation of the protective membranes covering the brain and spinal cord, known as the meninges. Symptoms range from headaches and a stiff neck to hearing loss, seizures and coma. Mortality from pneumococcal meningitis ranges from 16-37%, with neurological damage in 32-50% (Kastenbauer and Pfister, 2003; Ostergaard *et al.*, 2005). The World Health Organisation reported that in developing countries, 40-75% of cases of pneumococcal meningitis result in mortality or disability (www.who.int/mediacentre/factsheets/fs289/en/). Disabilities and neurological damage can include sequelae in the brain, resulting in mental retardation or learning difficulties, as well as permanent hearing loss (Bohr *et al.*, 1985; Keane *et al.*, 1979). Pneumococcal meningitis may also be recurring in particularly at-risk patients, demonstrated recently in a patient with a mild head injury, who

suffered 12 recurrences of bacterial meningitis, with the pneumococcus the main protagonist (Paudyal, 2007). Meningitis occurs after initial pneumococcal infection of the host, with 30% of cases being preceded by acute otitis media, and 18% with pneumonia (Ostergaard *et al.*, 2005). Furthermore, bacteraemia has been associated with poorer outcome of pneumococcal meningitis (Brandt *et al.*, 2008; Carrol *et al.*, 2008). Despite the apparent need for an initial pneumococcal infection to establish meningitis, it has been shown that pneumococci can invade the central nervous system directly via the olfactory neurones. This should allow pneumococci to penetrate the blood-brain barrier (BBB), and colonise the cerebro-spinal fluid (CSF) (van Ginkel *et al.*, 2003). This results in a large immune response in the host, resulting in inflammation and leukocyte recruitment, and occasionally brain injury (Hirst *et al.*, 2004a). A number of pneumococcal virulence factors have been shown to play a role in pneumococcal meningitis. These include Ply and autolysin, as knock-outs of these virulence factors displayed reduced virulence in a murine meningitis model (Hirst *et al.*, 2004a; Hirst *et al.*, 2008). However, isolates deficient in NanA or NanB or hyaluronidase displayed no reduction in virulence in the murine meningitis model, implying these proteins are not important in meningitis (Wellmer *et al.*, 2002).

1.3.3 Acute otitis media (AOM)

The most common infection caused by *S. pneumoniae* is the relatively benign AOM, which is common in children and characterised by growth of pneumococci in the host middle ear. The pneumococcus is one of the major causes of this infection, with 30-50% of cases attributed to this pathogen, and 7 million cases reported every year in the USA alone (Luotonen *et al.*, 1981; Prellner *et al.*, 1999) (www.who.int/vaccine/en/pneumococcus.shtml). An even higher rate of infection is observed in aboriginals, with most children having AOM at least once, although symptoms were found to be mild (Gibney *et al.*, 2005; Morris *et al.*, 2005). Infections are thought to occur via the nasopharynx in individuals colonised by pneumococci. Despite the comparatively mild symptoms of this infection, the sheer volumes of this infection exert great burdens on healthcare systems worldwide, and prevention by vaccination could have a significant socio-economic benefit. Indeed, in the USA alone, the economic cost AOM is estimated at $5 billion per year (Bondy *et al.*, 2000). Several serotypes have been

associated with AOM, such as 3, 19F and 23F, although this may be partly due to increased carriage rates of these serotypes (Hanage *et al.*, 2004; Hausdorff *et al.*, 2000a). NanA is thought to play an important role in AOM, as a knock-out strain was cleared from the middle ear in a chinchilla model, and immunization with recombinant NanA afforded protection against otitis media in the same model (Tong *et al.*, 2000; Tong *et al.*, 2005).

1.3.4 Complications of IPD

Further pneumococcal diseases becoming more prominent recently include empyema and haemolytic uraemic syndrome (HUS), with cases of these complications increasing rapidly in the UK recently (Byington *et al.*, 2006; Waters *et al.*, 2007). Empyema is characterised by a three-stage progression from pneumococcal pneumonia. Firstly, fluid accumulates in the pleural cavity, followed by the accumulation of pus and fibrin deposition in the pleural cavity, resulting in full empyema (Light, 1995). In the USA, *S. pneumoniae* is the primary cause of empyema, with serotype 1 accounting for 24-50% of cases between 1993 and 2000 (Byington *et al.*, 2002). In contrast, serotype 14 isolates are historically associated with empyema in the UK (Miller *et al.*, 2000). However, a recent study showed that a recent increase in serotype 1-related empyema mirrors the results found in the USA, with serotype 1 isolates accounting for the majority of empyema cases (Eastham *et al.*, 2004). A full review of HUS can be found in Section 1.4.

1.4 Pneumococcal HUS

HUS is characterised by a triad of microangiopathic haemolytic anaemia, thrombocytopenia and acute renal failure (Kaplan *et al.*, 1990). These symptoms translate to decreased red blood cell and platelet counts, with kidney failure. HUS is more commonly associated (>90% cases) with diarrhoea following infection with verocytotoxin producing bacteria, such as *Eschericia coli* O157:H7 and *Shigella dysenteriae*, with the causative agent being the bacteriophage-associated Shiga toxin (O'Brien *et al.*, 1984). The non-diarrhoeal/atypical form of HUS (aHUS) is rare, affecting 10% of all HUS in children in the UK (Neuhaus *et al.*, 1997). The aHUS variant has a poorer prognosis, (25%) compared to typical HUS caused by verocyto-toxin producing *E. coli* (Fitzpatrick *et al.*, 1993). More recently, aHUS has been shown to be a disease of complement dysregulation,

endothelial cell injury and activation which may be familial. Interestingly, pneumococcal infection has been reported to predispose to the development of aHUS. *S. pneumoniae* is responsible for almost 40% of all aHUS cases, and the incidence of aHUS following invasive pneumococcal disease has been reported as 0.4-0.6% (Arditi *et al.*, 1998; Cabrera *et al.*, 1998). aHUS as a result of pneumococcal infection has been termed pneumococcal HUS (p-HUS) to distinguish from other forms of aHUS (Copelovitch and Kaplan, 2007). Worryingly, the incidence of p-HUS has greatly increased in the UK. In a recent study, pneumococci were identified as causative agents in 34 of 43 cases, with empyema present in 23 of 35 pneumonia cases (Waters *et al.*, 2007). Although a number of pneumococcal serotypes have been implicated in p-HUS (Vanderkooi *et al.*, 2003; Waters *et al.*, 2007), it has been noted recently that serotype 19A isolates are an emerging cause of p-HUS (Waters *et al.*, 2007). Indeed, Waters and colleagues noted that whilst sequence types (STs) in isolates causing p-HUS were generally diverse, almost half of the isolates in their study were serotype 19A, ST199.

Little is known about the progression of invasive disease to p-HUS, although a model for the mechanism of pathogenesis has been suggested, wth neuraminidase activity central to pathogenesis (Klein *et al.*, 1977). The Thomsen-Friedrich antigen (T Antigen) is a surface molecule present on erythrocytes and renal capillary endothelial cells, although it is obscured by neuraminic acid (Klein *et al.*, 1977; Seger *et al.*, 1980). The neuraminidase enzymes produced by *S. pneumoniae* (Berry *et al.*, 1988; Camara *et al.*, 1991) are able to cleave the neuraminic acid from host cells, exposing the T Antigen on the cell surface. This in turn binds naturally occurring T antigen-specific immunoglobulin M (IgM), which is present in the plasma of all adults, and accelerates agglutination and haemolysis of erythrocytes, platelet agglutination and glomerular capillary damage, causing the symptoms seen in HUS (Kim *et al.*, 1979; Klein *et al.*, 1977; Novak *et al.*, 1983; Seger *et al.*, 1980). It is therefore possible that pneumococci possessing over-expressed or over-active neuraminidases may be more likely to cause p-HUS (Copelovitch and Kaplan, 2007), and it has been suggested that risk of developing p-HUS is increased by a heavy bacterial load, based on the relatively high numbers of patients developing p-HUS after empyema-associated pneumonia (Brandt *et al.*, 2002).

Indeed, although more historic cases of p-HUS were associated with pneumococcal pneumonia and meningitis (Alon *et al.*, 1984; Begue *et al.*, 1991; Cabrera *et al.*, 1998; Constantinescu *et al.*, 2004; Feld *et al.*, 1987; Huang and Lin, 1998; Martinot *et al.*, 1989), a number of recent cases have been associated with empyema (Huang *et al.*, 2006; Lee *et al.*, 2006; Vanderkooi *et al.*, 2003). This implies an association between clones causing empyema and p-HUS. However, the predominance of serotype 1 isolates in empyema has yet to be observed in p-HUS.

1.5 Pneumococcal virulence factors

1.5.1 Overview

The pneumococcus possesses many virulence factors which allow it to colonise a host and go on to cause invasive pneumococcal disease. Well-characterised virulence factors include the polysaccharide capsule, against which current pneumococcal vaccines are targeted, and the toxin Ply. Molecular techniques such as signature-tagged mutagenesis (Hava and Camilli, 2002; Lau *et al.*, 2001; Polissi *et al.*, 1998) and lambda-display libraries (Beghetto *et al.*, 2006), along with full genome sequencing (Hiller *et al.*, 2007; Tettelin *et al.*, 2001), have allowed identification of new putative pneumococcal virulence factors. Indeed, a signature-tagged mutagenesis study identified 337 genes essential for pneumococcal virulence in a murine pneumonia model (Hava and Camilli, 2002). However, many of these putative virulence factors remain as hypothetical proteins with unknown functions, demonstrating that a large amount of research remains required in this area. Some of the best characterised pneumococcal virulence factors, along with their location and role in pathogenesis, if known, can be found in Table 1.1. Furthermore, the functions of the pneumococcal capsule, the cholesterol-dependent cytolysin Ply and the sialidase enzyme NanA are discussed in greater detail in sections 1.5.2, 1.6 and 1.7 respectively.

Virulence factor	Location	Role in pathogenesis
Capsule	Covering pneumococci	Protects pneumococci from phagocytosis (Austrian, 1981). Shields surface proteins from IgG and complement (Musher, 1992; Winkelstein, 1981). Protects from neutrophil extracellular traps (Wartha *et al.*, 2007).
Pneumolsyin (Ply)	Cytoplasm	Highly active cytolytic toxin Inhibits ciliary beating of respiratory epithelia (Feldman *et al.*, 1990). Induces brain cell apoptosis during meningitis (Braun *et al.*, 2002). Crucial for pathogenesis in pneumonia (Rubins *et al.*, 1996) and meningitis (Winter *et al.*, 1997). Essential for survival in both upper and lower RT (Kadioglu *et al.*, 2002; Orihuela *et al.*, 2004).
Neuraminidase A (NanA)	Surface-exposed and secreted	Cleaves terminal sialic acid residues from host glycoproteins, glycolipids and cell-surface oligosaccharides, revealing host receptors for pneumococcal adherence (Tong *et al.*, 1999; Tong *et al.*, 2000). Removes sialic acid from host proteins such as lactoferrin and IgA2 (King *et al.*, 2004), and competitor species (Shakhnovich *et al.*, 2002). Important for pneumococcal survival in RT and blood (Manco *et al.*, 2006) as well as for colonization and acute otitis media (Tong *et al.*, 2000). Not important in meningitis (Winter *et al.*, 1997).
Hyaluronidase (Hyl)	Surface-exposed and secreted	Degrades hyaluronic acid in host connective tissue and may promote dissemination (Paton *et al.*, 1993).
Autolysin A (LytA)	Cell wall	Degrades pneumococcal peptidoglycan (Howard and Gooder, 1974) resulting in pneumococcal lysis and release of Ply toxin. Has a role in pneumonia and bacteraemia

		in murine models (Berry *et al.*, 1989a; Berry and Paton, 2000; Canvin *et al.*, 1995). Involved in pneumococcal transition from nasopharynx to lower RT (Orihuela *et al.*, 2004). Related enzymes LytB and LytC involved in pneumococcal colonisation (Gosink *et al.*, 2000).
Cell wall	Beneath capsule	Provides base for attachment of surface-exposed proteins. Inflammatory effect when degraded and released by autolysin. Responsible for sepsis-associated neuronal damage, usually prevented by interactions by TLR2 and Nod2 (Orihuela *et al.*, 2006).
Pneumococcal surface protein C (PspC, CbpA, SpsA)	Attached to cell wall	Involved in pneumococcal colonization (Rosenow *et al.*, 1997). Mutants show reduced virulence in pneumonia and sepsis (Iannelli *et al.*, 2004; Kerr *et al.*, 2006). Protects against pneumococcal opsonization by binding host factor H, and binds secretory IgA (Dave *et al.*, 2004).
Pneumococcal surface protein A (PspA)	Surface-exposed	Inhibits complement binding (Jedrzejas, 2001) and activiation (Ren *et al.*, 2004). Protects from bactericidal effect of apolactoferrin (Shaper *et al.*, 2004). May play a role in colonization (LeMessurier *et al.*, 2006).
Pneumococcal surface antigen A (PsaA)	Surface-exposed	Essential for colonization, bacteraemia and pneumonia (Berry and Paton, 1996; Johnson *et al.*, 2002; Marra *et al.*, 2002). Metal-binding part of manganese ABC transporter (Dintilhac *et al.*, 1997). May play a role in protection from oxidative stress (Tseng *et al.*, 2002).
Pili	Attached to cell wall	Involved in adhesion and important for invasion of host (Barocchi *et al.*, 2006; Nelson *et al.*, 2007). Present in small subset of serotypes,

		therefore not essential for virulence (Aguiar *et al.*, 2008; Basset *et al.*, 2007). May be associated with penicillin-resistant isolates (Sjostrom *et al.*, 2007). A second pilus recently identified also plays a role in adhesion (Bagnoli *et al.*, 2008).
IgA1 protease	Cytoplasm	Fights host immunity by cleaving host IgA1 (Poulsen *et al.*, 1996). May play a role in allowing pneumococcal adhesion and persistence on mucosal surfaces (Weiser *et al.*, 2003).
Caseinolytic protease C (ClpC)	Cytoplasm	Strain dependent role in virulence. Appears to be required for growth in lungs and blood in a mouse model (Ibrahim *et al.*, 2005).
Hydrogen peroxide (H_2O_2)	Produced during aerobic growth	Involved in early stages of meningitis, mediating brain cell apoptosis (Braun *et al.*, 2002; Hoffmann *et al.*, 2007). Bactericidal towards *Staphylococcus aureus* and other nasopharyngeal competitor species (Pericone *et al.*, 2000; Regev-Yochay *et al.*, 2006).
Zinc metalloprotease (ZmpB)	Surface exposed	Induces production of TNF-α in the lower RT, resulting in inflammation (Blue *et al.*, 2003).

Table 1.1: Pneumococcal virulence factors

Pneumococcal virulence factors, their location in the pneumococcal cell, and roles in pathogenesis as described by previous research. RT = respiratory tract, IgA/G = Immunoglobulin A/G, ABC transporter = ATP-binding cassette transporter, TNF-α = Tumor necrosis factor alpha.

1.5.2 Capsule

The pneumococcal capsule is a layer of polysaccharide covering the bacteria. This is a major virulence factor of the organism, and enables resistance to phagocytosis (Jonsson *et al.*, 1985). The capsule is also thought to be essential for pneumococcal survival in blood, and although un-encapsulated strains do exist, these are generally thought to be avirulent. However, there have been exceptions, with un-encapsulated strains isolated from invasive disease in immuno-compromised patients, and commonly responsible for conjunctivitis, an eye infection (Buck *et al.*, 2006; Muller-Graf *et al.*, 1999; Porat *et al.*, 2006). There are 91 identified pneumococcal serotypes (Henrichsen, 1995; Park *et al.*, 2007), and recent sequencing of the capsular genes of 90 serotypes has shown the high levels of diversity present between the loci, with total genetic content of variants in this loci being almost equivalent to the total genetic content of a single pneumococcal isolate (Bentley *et al.*, 2006). The pneumococcal polysaccharide capsule is the target for current pneumococcal vaccines, with a polysaccharide vaccine targeting 23 serotypes, and a conjugate polysaccharide vaccine targeting 7 serotypes. These vaccines are described in greater detail in Sections 1.6.2 and 1.6.3 respectively. One of the main roles of the pneumococcal capsule is to protect pneumococci from phagocytosis (Austrian, 1981), and it has also been shown to prevent interaction of the Fc part of IgG or the complement component C3b with receptors on the pneumococcal cell wall (Musher, 1992; Winkelstein, 1981). Furthermore, the capsule has been shown to reduce overall levels of complement attached to the pneumococcal cell (Abeyta *et al.*, 2003), as well as reducing trapping in neutrophil extracellular traps (Wartha *et al.*, 2007). The opaque and transparent phenotypes of pneumococci result in changes in expression levels of the pneumococcal capsule, with the transparent phenotype expressing less capsule and adapted to colonization, and the opaque phenotype expressing higher levels of capsule and adapted to survival in blood (Weiser *et al.*, 1994). The transition between these phenotypes is thought to play a role in progression from carriage to invasive disease (Hammerschmidt *et al.*, 2005; Tuomanen, 1999). However, it has been shown that different serotypes vary greatly in their ability to cause disease, which is presumably related to differences in the ability to resist host defences. This is supported by the observation that despite the presence of 91 serotypes, the

majority of disease is caused by less than 20% of these serotypes (Hausdorff *et al.*, 2000b).

1.6 Ply

S. pneumoniae possesses a pore-forming toxin called Ply (Walker *et al.*, 1987). Ply is an important virulence factor for the pneumococcus, is present in all disease-causing isolates (Kalin *et al.*, 1987), and has many important roles in pneumococcal pathogenesis (Mitchell and Andrew, 1997). This toxin has been shown to have two main roles in pneumococcal pathogenesis: it forms pores in host cell membranes, causing lysis of host cells (Boulnois *et al.*, 1991), and has the ability to activate complement in the host (Paton *et al.*, 1984). To date, Ply has been shown to be cytotoxic to all known eukaryotic cells due to the presence of cholesterol in the membranes of these cells. These include cells at important sites of human contact with *S. pneumoniae* such as pulmonary epithelial and endothelial cells (Feldman *et al.*, 1990; Steinfort *et al.*, 1989), ciliated brain ependymal cells (Mohammed *et al.*, 1999) and cerebral endothelial cells (Zysk *et al.*, 2001). Further to this work, a study using microarrays showed that many host cell genes are differentially regulated upon exposure to Ply, showing the major effect of this toxin on the host cell (Rogers *et al.*, 2003). Interestingly, despite the defined role of Ply in pneumococcal diseases, a study showed there was no significant difference in expression of Ply between *S. pneumoniae* in the nasopharynx, lung or blood. The authors hypothesise that production of Ply may be controlled by an environmental factor, or that levels of the toxin released differ in different niches (LeMessurier *et al.*, 2006). Established roles of Ply in colonization and disease are discussed in sections 1.6.4 and 1.6.5.

The effect of Ply on a variety of host cells has been determined in many studies since the identification of this toxin. Low concentrations of Ply were shown to promote migration and lysosomal enzyme secretion by neutrophils, as well as inhibiting their ability to migrate, and to kill opsonized pneumococci (Johnson *et al.*, 1981; Paton and Ferrante, 1983). Furthermore, Ply induces necrosis of neutrophils, showing that as well as reducing their ability to kill pneumococci, Ply actively targets these cells (Zysk *et al.*, 2000). Ply was shown to inhibit the ciliary beat of respiratory epithelial cells (Feldman *et al.*, 1990; Steinfort *et al.*,

1989), as well as being toxic to pulmonary epithelial cells (Rubins *et al.*, 1993), showing that the toxin has a role in allowing pneumococcal survival and dissemination in the lungs. Furthermore, Ply was shown to disrupt the integrity of these cells, thought to facilitate dissemination of pneumococci from lungs to blood (Rubins *et al.*, 1992). Monocytes are also affected by Ply, which inhibits phagocytosis and stimulates release of TNF-α and IL-1ß (Houldsworth *et al.*, 1994; Nandoskar *et al.*, 1986), thought to prevent pneumococcal clearance by the host. Furthermore, Ply was shown to up-regulate ICAM-1 in monocytes, which facilitates migration of neutrophils, although the benefit of this to pneumococci remains unclear (Thornton and McDaniel, 2005). Ply has also been shown to subvert the immune system in a number of ways, including inhibition of antibody production, induction of apoptosis in dendritic cells and preventing proliferation of lymphocytes (Colino and Snapper, 2003; Ferrante *et al.*, 1984).

The Ply toxin is thought to be well conserved (Mitchell *et al.*, 1990) due to the cytoplasmic nature of the protein affording it protection from the selective pressure exerted on surface-exposed proteins by the immune system. However, diversity has been observed previously (See section 1.6.2) (Kirkham *et al.*, 2006a; Lock *et al.*, 1996). Despite this, Ply is a pneumococcal vaccine candidate for future protein-based vaccines (Tai, 2006), and studies have shown Ply affords protection against pneumonia and bacteraemia in a murine model (Ogunniyi *et al.*, 2001; Paton *et al.*, 1983). A Ply toxoid also afforded protection against challenge with 9 different pneumococcal serotypes, displaying its potential as a cross-serotype vaccine candidate (Alexander *et al.*, 1994). Furthermore, research from our laboratory has recently shown that a non-toxic Ply mutant retained immunogenicity comparable to wild-type, and that vaccination with this mutant protects from *S. pneumoniae* challenge in a murine model (Kirkham *et al.*, 2006b). Another study showed that a stronger immune response was elicited to the pneumococcal capsule when Ply rather than a Tetanus toxoid was used as a carrier (Michon *et al.*, 1998). In a similar vain, recent research in our laboratory has also indicated potential use of Ply as an adjuvant, as immunization with a fusion of PsaA to Ply increased antibody titres to PsaA compared to immunization with both individual proteins (Unpublished results).

1.6.1 Cholesterol-dependent cytolysins

Ply belongs to the cholesterol-dependent cytolysins (CDCs), a family of toxins widely present in several genera of Gram-positive bacteria, particularly *streptococci, bacilli* and *clostridia*. A list of the toxins in this family can be found in Table 1.2. These toxins form pores in membranes possessing cholesterol, and as a result are toxic to all mammalian cells (Palmer, 2001). It has been shown that cholesterol or a related sterol is essential for membrane binding of these toxins (Ohno-Iwashita *et al.*, 1992). CDCs have been shown to bind cholesterol in membranes in a perpendicular manner (Ramachandran *et al.*, 2005), followed by oligomerization of up to 50 monomers, resulting in the formation of large pores of up to 30nm in diameter (Morgan *et al.*, 1995) and subsequent lysis of the target cell. With one known exception, all CDCs possess a secretion signal, and as a result are secreted. The exception to this is Ply, which does not possess a secretion signal and is thought to rely on pneumococcal autolysis by the virulence factor LytA to allow escape from the cytoplasm (Walker *et al.*, 1987). However, it has been shown that Ply may be secreted in a LytA-independent manner (Balachandran *et al.*, 2001).

Species	CDC Identity
Arcanobacterium pyogenes	Pyolysin
Bacillus anthracis	Anthrolysin O
B. cereus	Cereolysin
B. thuringiensis	Thuringiolysin O
Brevibacillus laterosporous	Laterosporolysin
Clostridium perfringens	Perfringolysin O
C. bifermentas	Bifermentolysin
C. botulinum	Botulinolysin
C. chauvoei	Chauveolysin
C. histolyticum	Histolyticolysin
C. septicum	Septicolysin
C. tetani	Tetanolysin
C. novyi	Novyilisin
Listeria ivanovii	Ivanolysin O
L. monocytogenes	Listeriolysin O
L. seeligeri	Seeligeriolysin O
Paenibacillus alvei	Alveolysin
Streptococcus canis	Streptolysin O
S. equisimilis	Streptolysin O
S. intermedius	Intermedilysin
S. pneumoniae	Ply
S. pyogenes	Streptolysin O
S. suis	Suilysin

Table 1.2: Cholesterol-dependent cytolysins

Cholesterol-dependent cytolysins produced by Gram-positive bacterial species (Billington *et al.*, 2000; Shannon *et al.*, 2003).

The structures of two CDCs have been solved, namely perfringolysin O (Rossjohn *et al.*, 1997) and intermedilysin (Polekhina *et al.*, 2005), and structure/function studies of Ply have employed a homology model of perfringolysin O, as shown in Figure 1.1. CDCs are composed of 4 domains designated 1-4, each with a different function. Domain 4 has been shown to be involved in host cell binding, whilst domain 3 is responsible for oligomerization and pore formation. Domain 3 of perfringolysin O inserts into the host cell membrane after initial toxin binding (Heuck *et al.*, 2003). A conserved 11 amino acid region of domain 4 (ECTGLAWEWWR) is responsible for host cell binding (Jacobs *et al.*, 1999), and thought to form the "receptor" region of the toxin for cholesterol recognition and subsequent binding of eukaryotic cells (Alouf, 2000; Tweten *et al.*, 2001). Furthermore, recombinant domain 4 was shown to inhibit haemolytic activity of full-length Ply, streptolysin O, and listeriolysin O by competitive binding to cholesterol (Baba *et al.*, 2001; Kohda *et al.*, 2002; Weis and Palmer, 2001).

Despite the conserved nature of this region, diversity has been observed (Billington *et al.*, 2000) with the variations observed in this region of intermedilysin resulting in a specificity towards human cells (Nagamune *et al.*, 2004). The receptor for binding of intermedilysin has been show to be CD59, a membrane protein involved in inhibition of complement (Giddings *et al.*, 2004).

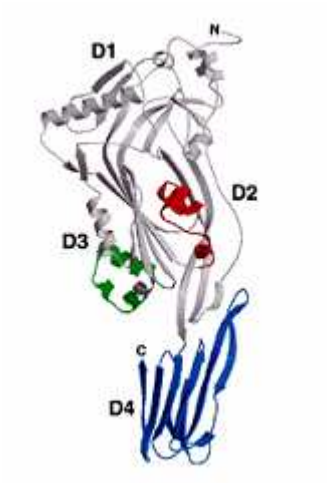


Figure 1.1: Structural model of perfringolysin O

Structure of the CDC perfringolysin O from Ramachandran et al. 2002. Domains represented by D1-4, and transmembrane helices highlighted in red and green. Domain 4, responsible for host cell-binding, highlighted in blue.

The oligomerization and insertion of CDCs into host membranes has been studied in detail. It has been shown that domain 2 is situated on the outside of the oligomers, with domains 3 and 4 on the inside, in both Ply and perfringolysin O (Dang *et al.*, 2005; Gilbert *et al.*, 1999a). Two mechanisms for the oligomerization and pore-formation in the CDCs have been suggested. The first proposed that CDC monomers bind and insert into membranes individually, and are able to move within the lipid bilayers, allowing oligomerization and formation of a pore which grows progressively in size (Figure 1.2B) (Palmer *et al.*, 1998; Weis and Palmer, 2001). This mechanism was proposed in streptolysin O, and accounts for the arc-like structures, essentially incomplete pores, observed on host cells. The second hypothesis, which has been more readily accepted, proposed that toxin monomers insert partially into the host membrane, and that this is followed by addition of further monomers, producing an oligomer as they insert (Figure 1.2A) (Gilbert *et al.*, 1999b; Shepard *et al.*, 2000). This mechanism, developed with perfringolysin O by the Tweten lab, is

known as the “prepore” theory, and has been strengthened by further study from their laboratory showing that formation of a prepore is the limiting step in pore formation (Hotze *et al.*, 2001) and that interactions between individual monomers drive the transition from prepore to pore (Hotze *et al.*, 2002). Another study showed that neither perfringolysin O nor streptolysin O formed small pores during oligomerization (Heuck *et al.*, 2003), with further research supporting the prepore hypothesis (Dang *et al.*, 2005; Ramachandran *et al.*, 2005). These studies were carried out on the toxins perfringolysin O and streptolysin O, and a recent study has shown that the prepore mechanism of pore formation is also employed by Ply (Tilley *et al.*, 2005). However, this theory does not account for the presence of arcs in membranes, and despite presence of arcs in figures from prepore studies, they are not mentioned in the research (Czajkowsky *et al.*, 2004).

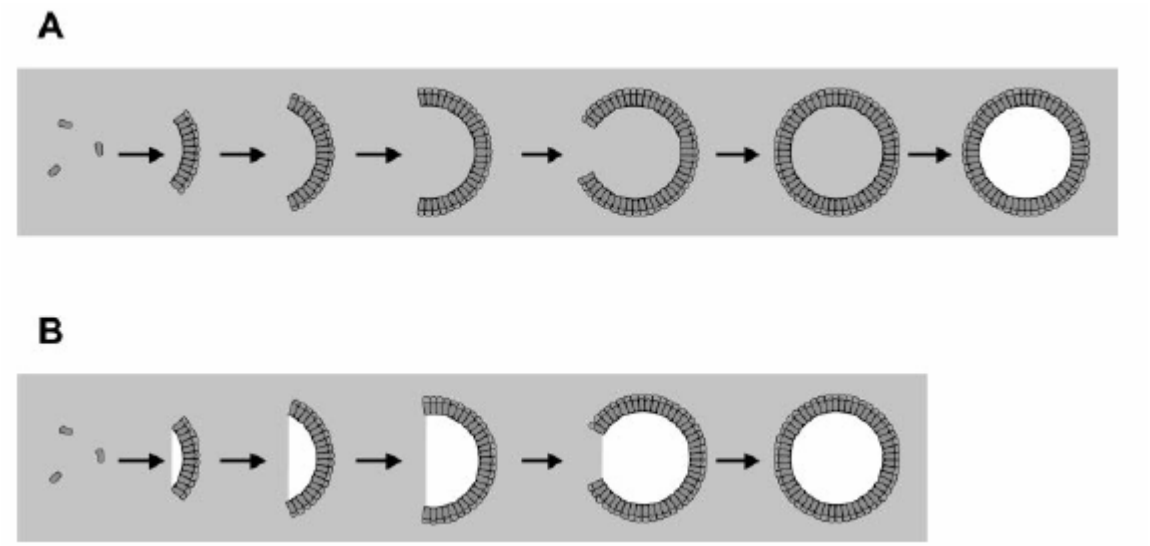


Figure 1.2: Two models of pore formation by CDCs

A: Prepore model suggested, where a complete oligomer is formed before insertion into the membrane (Shepard et al., 2000). B: Model suggested where the pore gradually increases in size through-out oligomerization (Palmer et al., 1998). Figure from (Heuck et al., 2003).

1.6.2 Ply diversity

Ply is a cytoplasmic protein thought to be released upon lysis of the pneumococcal cell by the autolytic enzyme LytA. As a result, Ply is thought to be highly conserved, (Mitchell *et al.*, 1990) as it is shielded from the selective pressure of the host immune response. However, a number of studies have previously demonstrated variation within the Ply protein. Lock and colleagues

(1996) explored the production of Ply by 30 pneumococcal isolates. They noted that whilst all isolates produced Ply, serotype 7F and 8 isolates produced Ply with reduced mobility on SDS-PAGE gel, and reduced haemolytic activity. The sequences of these toxin variants showed 3-4 amino acid substitutions and the deletion of valine and lysine residues at positions 270-271. Further study identified substitution of a threonine residue at position 172 with an isoleucine as responsible for reduction in both haemolytic activity and electrophoretic mobility (Lock *et al.*, 1996). Recent research from our laboratory has discovered a non-haemolytic Ply variant associated with serotype 1, ST306 pneumococci, with the mutations responsible present in the pore-forming region of the toxin. Epidemiological data also showed that the ST306 clone possessing the non-haemolytic toxin has undergone clonal expansion in the serotype 1 IPD population in Scotland since the year 2000. Interestingly, these variants were harboured by isolates causing IPD, and authors hypothesised that in some genetic backgrounds, the pneumococcus does not require the haemolytic activity of the toxin to progress to IPD (Kirkham *et al.*, 2006a). Furthermore, this study showed that whilst serotype 1 isolates possessing the wild-type toxin formed pores as described with the wild-type toxin (Tilley *et al.*, 2005), the non-haemolytic variant, although able to bind to the membrane, was not able to form pores and lyse erythrocytes, although formation of arcs was observed. This agreed with their findings that mutations in this toxin were present in the pore-forming domain, with the toxin retaining the ability to bind, but not form pores (Kirkham *et al.*, 2006a). An interesting subset of pneumococcal isolates were found to cause pneumonia in horses. These isolates possessed a genetic deletion causing a fusion between Ply and LytA, resulting in abrogation of the both protein activities (Whatmore et al. 1999c).

1.6.3 Effect of defined mutations on Ply activities

A number of Ply mutants have been constructed in order to determine amino acids essential to certain functions of the toxin. Mutation of amino acids at positions 428 or 433, part of an 11-amino acid conserved region of the toxin, as well as 387, was found to have a drastic effect on haemolytic activity of the toxin (Boulnois *et al.*, 1991; Saunders *et al.*, 1989). Meanwhile, mutation of amino acid 385 abolished the ability of Ply to activate the complement pathway (Mitchell *et al.*, 1991). A random mutagenesis study also identified the

tryptophan residue 433 as having importance in haemolytic activity, as well as the histidine residue 156 (Hill *et al.*, 1994). Further study of a mutant replacement of the tryptophan residue at position 433 with a phenalalanine showed that this residue was important in channel formation prior to oligomerization of Ply during pore formation (Korchev *et al.*, 1998). An *in vivo* study showed that mutation of amino acids 387 and 433, or 428 and 433 had a significant attenautive effect on pneumococci. Meanwhile addition of mutation at position 385, whilst abolishing complement activation, had no increase in attenuation, implying that the haemolytic activity of the toxin was more important than complement activation in systemic infection. The mutant with both haemolytic and complement activating ability abolished was termed PdT (Berry *et al.*, 1995). A second *in vivo* study showed that whilst the virulence of pneumococci was decreased when possessing the PdT toxoid, a significantly greater decrease was observed in a complete Ply knock-out, showing that Ply possesses another role in pathogenesis (Berry *et al.*, 1999). A recent study in our laboratory involved deletion of amino acids from a region involved in oligomerization (AA134-151) (Garcia-Suarez Mdel *et al.*, 2004), identifying a number of non-haemolytic mutants with deletions in this region. The Ply mutant Δ6, with deletion of amino acids 146-147, was recognised by a Ply antibody and retained immunogenicity comparable to wild-type Ply. This toxoid was suggested for use in future pneumococcal vaccines (Kirkham *et al.*, 2006b).

1.6.4 Role of Ply in pneumococcal colonization

The role of Ply in establishment of pneumococcal colonization is unclear, with several contrasting results uncovered. A Ply knock-out was shown to display reduced adherence and damage to respiratory epithelial cells *in vitro*, although colonization comparable to or exceeding wild-type isolates was observed with knock-outs *in vivo* (Rayner *et al.*, 1995; Rubins *et al.*, 1998). A different study noted that in a serotype 2 isolate, Ply was required for survival and proliferation in the nasopharynx, implying a role for Ply in colonization and carriage. However, a recent study identified Ply as being the causative agent of enhanced localised inflammatory responses resulting in clearance of colonizing pneumococci (van Rossum *et al.*, 2005) agreeing with the findings of Rubins et al (1998). Despite this, a Ply knock-out strain showed slightly reduced ability to colonize the nasopharynx in a murine model, although authors suggest Ply plays

a relatively minor role in colonization compared to PspC and PspA (Ogunniyi *et al.*, 2007a; Ogunniyi *et al.*, 2007b). Meanwhile, Malley and colleagues showed that the recognition of Ply by Toll-like receptor 4 (TLR4) was central to the innate immune response against the pneumococcus, and that mice lacking TLR4 were more susceptible to colonization by pneumococci possessing Ply (Malley *et al.*, 2003). A related study showed that the Ply-induced apoptosis of macrophages was dependent on TLR4, implying that apoptosis is involved in the interaction between Ply and TLR4 (Srivastava *et al.*, 2005). Authors suggest that Ply-induced apoptosis and interaction with TLR4 is an innate immune mechanism of fighting pneumococcal infection. These results support that hypothesis that Ply enhances the immune response to the pneumococcus upon colonization.

1.6.5 Role of Ply in pneumococcal diseases

Despite a proposed role in pneumococcal colonization, the main role of Ply appears to be as a virulence factor allowing the pneumococcus to cause IPD. Many studies have demonstrated the roles of Ply in pneumococcal pneumonia. A pair of studies in a rat model of pneumonia showed that whilst administration of Ply resulted in similar symptoms to pneumococcal pneumonia, immunization with Ply and Freunds adjuvant afforded protection (Feldman *et al.*, 1991; Roberts *et al.*, 1992). A mouse model of pneumonia was used to show that Ply was required for successful pneumococcal replication within alveoli as well as penetration of the bacteria into the interstitium of the lung (Rubins *et al.*, 1995), whilst both the haemolytic and complement-activating activities of Ply were shown to play distinct roles in progression of pneumonia (Rubins *et al.*, 1996). These roles have been determined, with the cytolytic activity of the toxin promoting neutrophil recruitment into lung tissue, whilst complement activation by Ply results in accumulation of T cells (Jounblat *et al.*, 2003). Despite this, the role of the cytolytic activity of Ply appears not necessary for certain clones pneumococci to cause pneumonia, since clones harbouring a recently discovered non-haemolytic toxin were isolated from cases of pneumonia (Kirkham *et al.*, 2006a). This finding is supported by recent data suggesting that lung injury in pneumonia is mediated by the pro-apoptotic and pro-inflammatory activities of Ply rather than the cytolytic activity, as Ply was present in the lungs at sublytic concentrations *in vivo,* and induced apoptosis and inflammation (Garcia-Suarez Mdel *et al.*, 2007). However, a clone producing no recogniseable Ply, due to

insertion of the mobile genetic element IS1515, has recently been identified in patients with pneumonia, showing that some pneumococcal clones can cause pneumonia without producing a fully functional Ply protein (Garnier *et al.*, 2007). A number of animal studies have shown that mutants of Ply have reduced virulence in *in vivo* pneumonia models (Berry and Paton, 2000; Ogunniyi *et al.*, 2000; Ogunniyi *et al.*, 2007a), whilst further studies have confirmed that immunization with Ply or monoclonal anti-Ply antibodies protect mice from pneumonia (Briles *et al.*, 2003; Garcia-Suarez Mdel *et al.*, 2004). These studies together offer clear indication of the role and importance of Ply in causing pneumococcal pneumonia.

Bacteraemia commonly results from progression of pneumococcal pneumonia, and Ply also plays an important role in this transition. Early studies showed that both cytolytic and complement-activating functions of Ply facilitated the transition to bacteraemia in mice (Rubins *et al.*, 1995), although another study reported different findings dependent on the strain of mouse used (Benton *et al.*, 1995), and both functions have been reported to affect the timing of bacteraemia progression (Jounblat *et al.*, 2003). Similar studies in a rat model showed that Ply contributes to bacteraemia by inhibiting pneumococcal clearance from the bloodstream and reducing the opsonising activity of serum (Alcantara *et al.*, 1999; Alcantara *et al.*, 2001). Furthermore, Ply has recently been demonstrated to play a role in evasion of the early innate immune response, as a knock-out displayed higher levels of C3 deposition than wild-type in a murine model (Quin *et al.*, 2007). Finally, a study by Musher and colleagues showed that hospitalised patients with pneumococcal pneumonia had higher levels of anti-Ply IgG than those with both pneumonia and bacteraemia. Authors hypothesised that higher IgG levels may neutralise Ply in the lungs and prevent or delay development of bacteraemia. Anti-Ply human IgG was also shown to offer protection against bacteraemic pneumococcal infection in a murine model, adding to the body of evidence that the toxin plays an important role in bacteraemia (Musher *et al.*, 2001).

Ply has been reported as one of the most important factors in the progression of pneumococcal meningitis (Kostyukova *et al.*, 1995). The role of Ply in the pathogenesis of pneumococcal meningitis has been explored using *in vivo* meningitis models. Studies in a rabbit model reported no difference in brain

inflammation in animals infected with Ply knock-out pneumococci compared to wild-type (Friedland *et al.*, 1995). However, a number of more recent studies have reported contrasting evidence. Virulence of pneumococci injected directly into murine brains was attenuated in a Ply-deficient strain (Wellmer *et al.*, 2002), whilst apoptosis of brain cells was shown to occur in a Ply-dependent manner (Braun *et al.*, 2002). These results imply a significant role for Ply in pneumococcal meningitis, and are congruent with studies in a rat model by Hirst and colleagues, who have demonstrated that Ply inhibits the ciliay beat of brain ependymal cells, as well as the attenuation of this inhibition by anti-Ply antibodies (Hirst *et al.*, 2000; Hirst *et al.*, 2004b). Furthermore, the authors have recently used the same model to show that a Ply mutant was almost avirulent in a meningitis model, and did not cause disruption to the brain ependyma, which was observed with the wild-type parent strain (Hirst *et al.*, 2008). Furthermore, Ply has also been shown to be the cause of hearing loss associated with pneumococcal meningitis (Winter *et al.*, 1997), and these data taken together show that Ply plays a number of important roles in the progression of pneumococcal meningitis.

The study of the importance of Ply in AOM has been less substantial, and indeed the toxin is thought to play a less significant role in this infection, with cell wall teichoic acids released by LytA thought to be more important for the inflammation observed. Firstly, Ply was shown to have cytotoxic effects on hair cells of guinea pig cochlea, and authors hypothesised that the toxin may play a role in deafness caused by acute otitis media, a result supported by further study (Comis *et al.*, 1993; Skinner *et al.*, 2004). The related toxin streptolysin O was also shown to be responsible for hearing loss in otitis media, related to changes in permeability of the round window membrane (Engel *et al.*, 1995). Furthermore, the presence of anti-Ply IgA antibodies was detected in the nasopharynx, middle ear fluid and serum of children with otitis media, showing that the host produced a response to this toxin upon progression of otitis media (Virolainen *et al.*, 1995a; Virolainen *et al.*, 1995b; Virolainen *et al.*, 1996). The development of an *in vivo* chinchilla otitis media model allowed further study into the role of Ply in acute otitis media (Giebink, 1999). As a result, Ply was shown to play a less significant role than LytA in otitis media pathogenesis in a serotype 3 isolate of pneumococci (Sato *et al.*, 1996).

Ply has also been shown to have a role in pneumococcal eye infections such as conjunctivitis and endophthalmitis, with an early study showing that Ply was responsible for the corneal damage opbserved upon intra-corneal injecton of crude pneumococcal lysate in a rabbit model (Johnson and Allen, 1975). Another study implicated Ply as a virulence factor in endophthalmitis by showing that injected Ply reproduced the main symptoms of this infection (Ng *et al.*, 1997; Ng *et al.*, 2002).

1.7 NanA

NanA is the main sialidase enzyme produced by the pneumococcus. The pneumococcus was first reported to possess neuraminidase activity in 1959, and the enzyme responsible was shown to be secreted soon after (Laurell, 1959; Lee and Howe, 1966). However, pneumococcal neuraminidase was not purified and characterised until much later, with reports of numerous neuraminidases of differing sizes attributed to degradation (Berry *et al.*, 1988; Lock *et al.*, 1988b; Scanlon *et al.*, 1989). However, this theory was disproved when a second pneumococcal neuraminidase was characterised. This neuraminidase enzyme, now known as NanB, was shown to be distinct from the main pneumococcal neuraminidase, known as NanA (Camara *et al.*, 1991). Sequencing of the *nanA* gene showed that this enzyme had all the characteristics of a neuraminidase gene, with 4 copies of the SXDXGXTW motif present in all bacterial neuraminidases (Camara *et al.*, 1994; Roggentin *et al.*, 1989). Furthermore, the protein was found to possess a putative signal sequence relating to export across the cellular membrane, and a C-terminal LPXTG motif implying that NanA may be anchored to the cell wall as well as secreted (Camara *et al.*, 1994).

The function of NanA is based on the cleaving of terminal sialic acid from host cell glycans such as glycolipids, glycoproteins and mucin. This function has resulted in the demonstration of numerous roles for NanA in pneumococcal pathogenesis. The main role of NanA is thought to be to expose receptors on the host cell surface, allowing pneumococcal adhesion, the first stage in colonization (Tong *et al.*, 1999; Tong *et al.*, 2000). Furthermore, NanA has been implicated in desialiation of molecules involved in bacterial clearance from blood, and interspecies competition (King *et al.*, 2004; Shakhnovich *et al.*, 2002). As a result, NanA has been implicated as playing an important role in

pneumococcal colonization as well as acute otitis media (Tong *et al.*, 2000; Tong *et al.*, 2002). The roles and importance of this enzyme in colonization and disease are described in more detail in Sections 1.7.2 and 1.7.3 respectively. NanA was found to be present in all pneumococci, whilst related enzyme NanB was present in 96% of isolates, and the poorly characterised homologue NanC in 51% (Pettigrew *et al.*, 2006).

A number of amino acids have been identified as important for the enzyme activity of NanA. Replacement of a glutamic acid residue with a glutamine at position 647 (E647Q) or of tyrosine with phenalalanine at position 752 (Y752F) resulted in complete abrogation of enzymatic activity, whilst an R633H mutation resulted in a 98% reduction in activity. This demonstrated the importance of these residues towards enzymatic activity of the protein (Yesilkaya *et al.*, 2006). A study of diversity present in the NanA gene reported overall amino acid diversity estimated at 14.8% (King *et al.*, 2005). However, this study did not cover the full NanA gene, suggesting this may be an underestimate of diversity. Four large mosaic blocks were identified in the gene, with one sharing homology with a sialidase gene from *S. oralis*, implying homologous recombination between these species. However, despite the high levels of variation, no differences were observed in the essential aspartic boxes or predicted active site, demonstrating the important of these regions to the activity of the protein (King *et al.*, 2005).

1.7.1 Role of sortase A in NanA surface exposure

Analysis of the NanA sequence has revealed an LPXTG motif at the C-terminal of the protein. This motif is absent from NanB, which is as a result presumed to be absent from the pneumococcal cell surface. Proteins possessing a C-terminal LPXTG motif have been shown to be covalently attached to the cell wall peptidoglycan of Gram-positive bacteria, resulting in surface-exposure on the bacteria. These include many virulence factors including the M proteins of *Streptococcus pyogenes*, and protein A of *Staphylococcus aureus* (Navarre and Schneewind, 1999). A number of pneumococcal virulence factors possess this motif, including NanA and hyaluronidase, another surface-exposed enzyme (Berry *et al.*, 1994; Camara *et al.*, 1994). The anchoring of these proteins to the cell wall in the pneumococcus is carried out by a sortase enzyme, known as

sortase A, and it has been shown that inactivation of this enzyme affects the localization of NanA, as well as decreasing adhesion of *S. pneumoniae* to human pharyngeal cells (Kharat and Tomasz, 2003). Furthermore, sortase A has a role in the pathogenesis of pneumococcal pneumonia and bacteraemia, as well as colonization, most likely due to sortase A-deficiency resulting in differences in presentation of virulence factors, including NanA, on the bacterial surface (Chen *et al.*, 2005; Paterson and Mitchell, 2006).

1.7.2 Role of NanA in pneumococcal colonization

The sialidase activity of NanA has implicated it strongly as having a role in pneumococcal colonization. It is thought that cleavage of sialic acid from host glycans modifies the cell surface of host cells and exposes receptors for interaction with the pneumococcus, possibly allowing pneumococcal adherence (Tong *et al.*, 1999). This hypothesis was supported by further data showing decreased ability of a NanA-deficient mutant to colonize and persist in the chinchilla nasopharynx, whilst immunization with recombinant NanA afforded protection against challenge in a chinchilla colonization model (Tong *et al.*, 2000; Tong *et al.*, 2005). The authors also showed that NanA was responsible for alteration of carbohydrates in the tracheal epithelium, corroborating the previous studies (Tong *et al.*, 2002). This hypothesis was initially disputed by King and colleagues, who observed no decrease in pneumococcal adherence to human epithelial cells or colonization in a rat model (King *et al.*, 2004). However, further study by this group showed that NanA works synergistically with β-galactosidase and β-N-acetylglucosaminidase to expose mannose receptors in the host nasopharynx, and that these three enzymes play a role in pneumococcal colonization (King *et al.*, 2006). Furthermore, both NanA and NanB were shown to have an important role in establishment of colonization in a murine model (Manco *et al.*, 2006; Orihuela *et al.*, 2004).

Microarray studies showed that NanA was up-regulated in pneumococci isolated from the nasopharynx compared to blood isolates, further implicating the enzyme in colonization (King *et al.*, 2004). The NanA enzyme of *S. pneumoniae* is also involved in interspecies competition with *Neisseria meningitis* and *Haemophilus influenzae*. Both of these bacteria employ sialic acid on their cell surfaces as a mechanism of molecular mimicry of host cells, to protect

themselves from opsonization by host complement. The pneumococcus utilises NanA to degrade this sialic acid, thus exposing the competing bacteria to the host immune system and conferring a competitive advantage to the pneumococcus during colonization (Shakhnovich *et al.*, 2002). Furthermore, the sialidase activity of NanA has also been shown to cleave human proteins including lactoferrin and IgA2, which are involved in clearance of pneumococci from the nasopharynx and blood (King *et al.*, 2004). The pneumococcus has also been shown to utilise mucin, which is plentiful in the nasopharynx, as an energy source (Yesilkaya *et al.*, 2008). NanA activity plays a role in this, as NanA transcription was increased in presence of mucin, whilst a NanA-deficient mutant was unable to utilise mucin as a sole energy source (Yesilkaya *et al.*, 2008).

1.7.3 Role of NanA in pneumococcal diseases

The roles of NanA in IPD have not been studied in as great detail as in Ply. However, a number of potential roles have been defined. An important role in pathogenesis of acute otitis media has been attributed to NanA, with the initial observation of destruction of chinchilla middle ear membrane cells (LaMarco *et al.*, 1986). Subsequent studies showed reduced colonization and persistence of a NanA-deficient mutant in the middle ear, whilst immunization with native or recombinant NanA afforded protection against acute otitis media in chinchillas (Long *et al.*, 2004; Tong *et al.*, 2000). Although NanA does not appear to be important in pneumococcal pneumonia, with no reduction in virulence of a NanA-deficient strain observed in a murine pneumonia model, it may play a role in the transition from nasopharynx to lung prior to establishment of pneumonia (Berry and Paton, 2000; Orihuela *et al.*, 2004).

The importance of NanA in pneumococcal bacteraemia is unclear. Whilst one study reported mutants deficient in NanA or NanB with an inability to persist in murine blood after intravenous injection (Manco *et al.*, 2006), two other studies showed no reduction of survival in blood of a NanA-deficient mutant after intravenous and intraperitoneal injection respectively (Grewal *et al.*, 2008; Orihuela *et al.*, 2004). It appears that NanA plays no role in pneumococcal meningitis. Early studies in an *in vitro* canine model found no role for NanA in neuronal damage involved in meningitis (Carruthers and Kanokvechayant, 1973).

Similarly, no role for NanA was found in relation to the hearing loss and cochlear damage associated with pneumococcal meningitis, and no decrease in virulence was observed with a NanA-deficient mutant in two separate studies in murine meningitis models (Orihuela *et al.*, 2004; Wellmer *et al.*, 2002; Winter *et al.*, 1997).

NanA has recently been implicated in the pathogenesis of conjunctivitis, as microarray analysis of conjunctivitis isolates showed up-regulation of neuraminidase genes, and exogenous neuraminidase increased adherence of pneumococci to human conjunctival epithelia *in vitro* (Williamson *et al.*, 2008). Furthermore, NanA appears to play a role in progression of p-HUS, a complication of IPD (Klein *et al.*, 1977). This hypothesis is discussed in detail in Section 1.4. Finally, neuraminidase activity has been implicated as important in the lethal synergism observed between the pneumococcus and Influenza virus (Peltola and McCullers, 2004; Peltola *et al.*, 2005), and this relationship, along with the potential role of neuraminidases, is discussed in Section 1.9.

1.7.4 NanA as a vaccine candidate

Due to its role in pneumococcal colonization and importance in several pneumococcal diseases, NanA is being considered as a pneumococcal vaccine candidate for future protein-based vaccines (Tai, 2006). An early study using formaldehyde-treated NanA to immunize mice before intranasal challenge showed only small levels of protection (Lock *et al.*, 1988a). However, recent studies in the chinchilla model have shown that immunization with NanA affords protection against pneumococcal colonization and otitis media, whilst immunization of mice with NanA toxoids significantly delayed the onset of pneumococcal pneumonia (Long *et al.*, 2004; Tong *et al.*, 2005; Yesilkaya *et al.*, 2006). However, these studies challenged chinchillas with pneumococci possessing either identical or similar NanA proteins to those used for immunization. This does not take into account the high levels of diversity in the NanA enzyme, which may restrict its use in a species-wide vaccine due to the high levels of diversity present across different serotypes (King *et al.*, 2005).

1.8 Pneumococcal vaccines

1.8.1 History

The first pneumococcal vaccine consisted of killed whole pneumococci, and success was reported in preventing pneumococcal pneumonia and death in South African miners (Wright *et al.*, 1914). Further research in the 1930s and 40s, based on the recognition of multiple pneumococcal serotypes, showed that anti-capsular polysaccharide antibodies played a central role in protection against the pneumococcus (Finland and Sutliff, 1932; Francis and Tillett, 1930), and demonstrated the importance of incorporating multiple serotypes into a vaccine (McLeod *et al.*, 1947). After further study in South Africa (Austrian *et al.*, 1976; Smit *et al.*, 1977), the first pneumococcal polysaccharide vaccine (PPV), a 14-valent vaccine, was licensed for use in the USA, and subsequently Europe in 1977. In the 1980s, the valency of this vaccine was increased to cover 23 serotypes, which remains the valency of this vaccine today (French, 2003). Further to this, and due to limitations of the 23-valent PPV, a 7-valent conjugate polysaccharide vaccine was developed by Wyeth, and licensed in the USA in 2000. In 2006, this vaccine was introduced into the UK vaccination schedule. In the near future, a number of increased-valency conjugate polysaccharide vaccines will become available. These are an 11-valent vaccine developed by GlaxoSmithKline, and a 13-valent vaccine developed by Wyeth. In the future, it is expected that protein-based vaccines will be licensed to protect against the pneumococcus. These are likely to include a number of protein virulence factors in order to afford cross-serotype protection (Ogunniyi *et al.*, 2000; Ogunniyi *et al.*, 2007a).

1.8.2 Polysaccharide vaccines

The first pneumococcal polysaccharide vaccine was a vaccine targeting the capsule of the 14 serotypes most commonly isolated from IPD. The next generation of this vaccine, a 23-valent PPV called PneumoVax, was launched by Merck. This vaccine was effective against 23 pneumococcal serotypes (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F, 33F), which caused 90% of IPD in the USA, and included the six serotypes most commonly causing invasive antibiotic resistant pneumococcal infections in the USA (Robbins and Schneerson, 1983). The 23-valent PPV vaccine is generally

administered to elderly patients, and patients with underlying predispositions such as sickle cell disease, HIV, or immuno-compromised patients where there is an increased risk of IPD (Gebo *et al.*, 1996; Pebody *et al.*, 2005). The efficacy of this vaccine has been researched in a number of studies. One study showed that HIV-infected children developed a significant immune response to the 23-valent PPV (Tangsinmankong *et al.*, 2004), whilst another study showed that the 23-valent PPV was ineffective in Ugandan HIV-infected adults, and noted increased rates of IPD in vaccine recipients (French *et al.*, 2000). Furthermore, although this vaccine protects immuno-competent adults from IPD, there have been differing conclusions on efficacy in the elderly population. The vaccine was initially shown to be inefficacious in the elderly population (Ortqvist *et al.*, 1998), although efficacy has since been observed (Melegaro and Edmunds, 2004).

One of the main drawbacks of the 23-valent PPV is that it is ineffective in children <2 years old. Use of this vaccine has been shown to produce anti-capsular polysaccharide antibodies by stimulating clonal expansion of B-cells, resulting in maturation of B-cells into plasma cells, which produce antibody. This is in a T-cell independent manner, and these cells have a short lifespan, with no immune memory established. The reason for the lack of efficacy in children is thought to be the poor expression of CD21, a complement receptor which acts to improve B-cell response (Weintraub, 2003). Since the infant population is one of the main risk groups for IPD, a new vaccine was developed.

1.8.3 Polysaccharide conjugate vaccines

The conjugation of capsular polysaccharides to selected proteins has been shown to increase their immunogenicity as well as being efficacious in infants. The pneumococcal polysaccharide conjugate vaccine (PCV7) was developed after the outstanding success of the *Haemophilus influenzae* serotype b conjugate vaccine, with dramatic decline in invasive disease (Obaro and Adegbola, 2002). Indeed, vaccination has resulted in a 90-100% decrease in serotype b *H. influenzae* invasive disease. This vaccine had such a high level of success as serotype b clones are almost exclusively responsible for invasive disease cases (Kelly *et al.*, 2004). The problem with the pneumococcus is that whilst 6 serotypes of *H. influenzae* exist, with one causing almost all invasive disease, high levels of diversity in the pneumococcus result in over 90 pneumococcal

serotypes, with many responsible for IPD. Due to the complexity and cost of the conjugation process, only a limited number of serotypes could be included in the vaccine. In 2000, Prevnar (PCV7), a heptavalent pneumococcal polysaccharide conjugate vaccine was launched. This vaccine targeted the serotypes 4, 6B, 9V, 14, 18C, 19F, and 23F based on epidemiological data from the USA, which showed these to be responsible for 78% of IPD in children (Butler *et al.*, 1995). The protein component of this vaccine was the Diphtheria toxoid CRM_{197}, which increased the immunogenicity of the capsular polysaccharides in infants (Black *et al.*, 2000). However, the coverage of this vaccine varies globally, ranging from 70-88% in Europe and Africa, to <65% in Latin America and Asia, were serotypes 1 and 5 cause most IPD (Hausdorff *et al.*, 2000b; Lloyd-Evans *et al.*, 1996; Obaro and Adegbola, 2002). This represents very low coverage, and demonstrates the need for either geographically tailored vaccines or non-serotype-specific vaccines.

The efficacy of this vaccine has been tested in several studies. The efficacy against invasive disease was tested by Black et al, who showed a 97.4% efficacy against vaccine serotypes causing invasive disease. Acceptable levels of efficacy were also observed for otitis media. However, results on the impact on pneumonia were inconclusive due to difficulties with diagnosis (Black *et al.*, 2000). Further to this study, it was concluded that the PCV7 vaccine was safe and efficacious for use in the general population, with particular efficacy in <5 year child populations in comparison to PPV (Black and Shinefield, 2002). However, this vaccine has been shown to be less efficacious in elderly populations, where the 23-valent PPV remains useful (Briles, 2004).

Since its introduction into the vaccination schedule in 2000, Prevnar has had a drastic effect on IPD in the USA. In particular, IPD has decreased in children, although decreases have also been observed in older groups, mostly attributed to herd immunity. However, whilst there has been a decrease of 75% in vaccine serotypes in the >65 age group, there has also been a significant 14% increase in non-vaccine serotypes (Lexau *et al.*, 2005). The success of the PCV7 vaccine in the USA has been so great that the coverage of this vaccine had decreased to 20% by 2005, due to replacement by non-vaccine serotypes (Whitney, 2005). This is the largest problem created by vaccinating against a sub-population of serotypes of this pathogen, which is evasion of the vaccine. This phenomenon

has been termed serotype replacement, and many studies have shown the replacement of vaccine serotypes by non-vaccine serotypes both in carriage and IPD in vaccinated individuals. Studies in the USA have shown that carriage of vaccine serotypes has decreased by 98% since the introduction of Prevnar, whilst vaccine-related serotypes have decreased by 79%. Conversely, non-vaccine serotypes have increased by 45% (Pai *et al.*, 2005b). In particular, the non-vaccine serotype 19A has increased 148%, and is the most significant of the replacement serotypes observed in the USA. Increases in this serotype have resulted in increase in penicillin non-susceptibility in the overall pneumococcal population due to over-representation of a serotype 19A, ST199 clone. Furthermore, several clones from vaccine serotypes have been discovered in the serotype 19A population, demonstrating their ability to switch capsule and evade the vaccine, and resulting in diversification of the serotype 19A population and emergence of this serotype as the predominant cause of IPD in the USA (Moore *et al.*, 2008; Pai *et al.*, 2005b). Similarly, a separate study demonstrated a reduction in IPD in HIV infected individuals, separated into a 62% decrease in IPD caused by vaccine serotypes, and a 44% increase in IPD caused by non-vaccine serotypes (Flannery *et al.*, 2006). Similar results have been observed in most studies (Frazao *et al.*, 2005; McEllistrem *et al.*, 2003), and recent studies have indicated that such changes in circulating clones may have an effect on disease presentation in populations including HIV-infected patients and Alaskan natives (Lexau, 2008). Another drawback has been the emergence of a multi-drug resistant 19A clone causing acute otitis media, which is resistant to all approved methods of treatment, demonstrates the risk of vaccination targeting a subset of the pneumococcal population (Pichichero and Casey, 2007).

Serotype replacement represents one of the major shortcomings of the PCV7 vaccine. This has resulted in development of polysaccharide conjugate vaccines with increased valency. In the near future, an 11-valent vaccine will be launched by GlaxoSmithKline, covering further serotypes 1, 3, 5 and 7F, and has been shown to increase coverage of IPD-causing serotypes (Nurkka *et al.*, 2004). Furthermore, 13-valent vaccine will be launched by Wyeth, covering further serotypes 1, 3, 5, 6A, 7F and 19A, which will further increase coverage. These vaccines will afford greater coverage of serotypes, and target the problematic serotypes, such as 19A and 1, which were absent from PCV7. However, the

serotype replacement observed upon the introduction of PCV7 in the USA has shown that this may not be the answer. Although these vaccines will increase coverage, there is an associated increase in cost, which will make these even more prohibitive for use in developing countries. Therefore, research into future pneumococcal vaccines is focussing on alternatives to serotype-specific vaccines, with the hope of producing a serotype-independent pneumococcal vaccine. This research has so far focussed on protein virulence factors shared by all pneumococci (Briles *et al.*, 2000a; Ogunniyi *et al.*, 2007a; Tai, 2006).

1.8.4 Protein-based vaccines

Many pneumococcal virulence factors have been studied as potential vaccine candidates for a protein-based pneumococcal vaccine. The benefits of such a vaccine would be to provide protection against all pneumococci irrespective of serotype, as well as being cheaper than the prohibitively expensive conjugate vaccines, thus becoming globally accessible. The ideal protein vaccine candidate would be highly conserved across all pneumococci and illicit a strong immune response. Surface-exposed pneumococcal virulence factors such as the sialidase enzyme NanA (Berry *et al.*, 1988) should promote a strong immune response, however, these proteins are exposed to a greater selective pressure from the host immune system, and therefore tend to display a greater level of diversity, which may prevent full coverage of a vaccine. Cytoplasmic proteins such as the cholesterol-dependent cytolysin Ply (Walker *et al.*, 1987) are highly conserved across pneumococci, due to their protection from the host immune system. Advances in technology, in particular those allowing rapid genome sequencing of the pneumococcus, have allowed identification of virulence factors with potential as pneumococcal protein vaccine candidates (Hava and Camilli, 2002; Hiller *et al.*, 2007; Tettelin *et al.*, 2001; Wizemann *et al.*, 2001). Indeed, over 30 pneumococcal genomes from a variety of serotypes will soon be publicly available, allowing for much more pertinent, cross-genome analysis of potential candidate proteins. The proteins that are being considered as vaccine candidates can be found in Table 1.3, along with information about their diversity across pneumococcal serotypes, and protection elicited in previous studies. Two of the most promising vaccine candidates are Ply and PsaA, since they have been shown to be highly conserved across all pneumococci, and have shown promising levels of protection in animal models (Briles *et al.*, 2000b; Briles *et al.*, 2003; Ogunniyi

et al., 2001; Talkington *et al.*, 1996). Furthermore, serum antibodies to these proteins were found in children, associated with carriage or infection, implying that these proteins are targeted by the host immune response to pneumococci (Rapola *et al.*, 2000). A further protein, PspA, was shown to be highly immunogenic and protect against both IPD and carriage (Briles *et al.*, 2003). This was despite high levels of diversity demonstrated in this pneumococcal protein (Hollingshead *et al.*, 2000). However, this protein was found to share 27% sequence identity with human cardiac myosin upon Phase I clinical trials, and research into its potential as a vaccine candidate has since ceased.

Protein	Diversity	Immunogenicity Studies
Ply	Thought to be highly conserved (Lock *et al.*, 1996; Mitchell *et al.*, 1990) although recent studies in our lab have shown diversity between and within serotypes (Kirkham *et al.*, 2006a).	Protects against pneumonia and bacteraemia (Ogunniyi *et al.*, 2001; Paton *et al.*, 1983). Toxoid offers protection to 9 pneumococcal serotypes (Alexander *et al.*, 1994). Shown to increase efficacy of CPS vaccines when used as a carrier protein (Lee *et al.*, 2001). A non-toxic mutant retains high levels of immunogenicity (Kirkham *et al.*, 2006b).
PsaA	High levels of variation between 2 isolates reported (Berry and Paton, 1996), however, larger study reported relative conservation (Sampson *et al.*, 1997).	Protects against pneumococcal colonisation and sepsis in a mouse model (Briles *et al.*, 2000b; Talkington *et al.*, 1996).
PspA	Diverse, with highly mosaic structure indicating many recombination events and resulting in >20% diversity between alleles. Split into distinct clades (Hollingshead *et al.*, 2000).	Intranasal immunization protects mice from carriage, pulmonary infection and sepsis (Wu *et al.*, 1997). Protects against sepsis in murine model (Swiatlo *et al.*, 2003). In humans, antibodies against PspA protect from carriage and invasive disease (Palaniappan *et al.*, 2005).
PiaA & PiuA	PuiA conserved within pneumococci, and related species. PiaA 100% conserved in pneumococci, and absent from related species (Whalan *et al.*, 2006).	Immunization with both or either protects mice from bacteraemia (Brown *et al.*, 2001). Antibodies promote pneumococcal opsonophagocytosis (Jomaa *et al.*, 2005a; Jomaa *et al.*, 2005b).
PspC (CbpA, SpsA)	Present in 75% of pneumococcal strains. Similar to PspA, and split into 2 distinct clades. One clades has a domain shared by PspA, the other clade lacks this domain (Brooks-Walter *et al.*, 1999).	Protects against pneumococcal bacteraemia (Brooks-Walter *et al.*, 1999), but no effect on virulence seen when deleted (Berry and Paton, 2000).
ClpP	Undetermined.	Protects from systemic

		challenge. Mutant shows reduced colonisation and virulence levels (Kwon *et al.*, 2004).
LytA	Highly conserved, showing less than 3% amino acid diversity (Whatmore and Dowson, 1999).	Immunization significantly increases survival of challenged mice (Lock *et al.*, 1992). Partially effective as carrier protein for CPS (Lee and Li, 2001).
NanA	Study showed high levels of diversity, although did not sequence full gene (King *et al.*, 2005).	Immunization confers slight increase in survival time of challenged mice (Lock *et al.*, 1988a). Immunization reduces colonisation and otitis media in chinchillas (Long *et al.*, 2004; Tong *et al.*, 2005).
PhtA-E	Antigenically conserved (Tai, 2006).	Protect mice against pneumococcal sepsis (Adamou *et al.*, 2001). PhtB and PhtE protect against pneumonia in mice (Hamel *et al.*, 2004; Zhang *et al.*, 2001).
PpmA	Relatively conserved, with most mutations discovered synonymous (Overweg *et al.*, 2000).	Elicits species-specific, cross-reactive antibodies, and a knock-out mutant shown reduces virulence in mouse pneumonia model (Overweg *et al.*, 2000).
PppA	Antigenically conserved across serotypes, although nothing mentioned regarding amino acid similarity (Green *et al.*, 2005).	Elicits cross-reactive antibodies and reduces colonisation in a murine model (Green *et al.*, 2005). Nasal immunization with Lactococcus lactis expressing PppA shown to elicit cross-protective immunity against both respiratory and systemic challenges (Medina *et al.*, 2008).

Table 1.3: Pneumococcal protein vaccine candidates

A list of pneumococcal protein vaccine candidates, diversity present and immunization studies. Ply = pneumolysin, PsaA = Pneumococcal surface antigen A, PspA = pneumococcal surface protein A, PiaA = pneumococcal iron acquisition protein A, PiuA = pneumococcal iron uptake protein A, PspC = pneumococcal surface protein C, ClpP = caseinolytic protease P, LytA = autolysin, NanA = neuraminidase A, PhtA-E = pneumococcal histidine triad proteins A-E, PpmA = proteinse maturation protein A, PppA = pneumococcal protective protein A.

The most successful pneumococcal vaccine may turn out to be a combination of pneumococcal proteins, and recently, a number of studies have investigated this area. In 2000, it was shown that while mutants deficient in Hyl or PspC had no effect on virulence, significant additive attenuation was observed in double mutants of each of these proteins with Ply (Berry and Paton, 2000). Furthermore, investigation of additive immunization with PspA, PspC and ClpP or their polyclonal antibodies showed that survival increased significantly in mice immunized with all three proteins compared to single or double combinations (Cao *et al.*, 2007). Similarly, another study showed that immunization with the Ply toxoid PbB, PspA and PspC together significantly increased survival times in mice compared to any other combination of proteins, showing that immunization with several pneumococcal proteins may have more success in protecting against pneumococcal disease (Ogunniyi *et al.*, 2007a).

1.9 The relationship between Influenza and the pneumococcus

There is an important relationship between Influenza infections and secondary pneumonia caused by bacterial pathogens, resulting in increased mortality during Influenza epidemics and pandemics (Brundage and Shanks, 2008; Glezen, 1982; Simonsen, 1999). The main bacterial pathogen involved in this is the pneumococcus, and together influenza and *S. pneumoniae* are responsible for most deaths from infectious diseases worldwide (World Health Organization, www.who.int). An example of the catastrophic cooperative interaction between these pathogens can be found in the 1918 influenza pandemic, where 40-50 million people are reported to have lost their lives, mostly to secondary pneumococcal pneumonia (Brundage and Shanks, 2008). However, the mechanisms behind the relationship between these organisms are poorly understood. It has been hypothesised that destruction or alteration of the human respiratory epithelium or inhibition of phagocytosis by the Influenza virus may predispose the host to pneumococcal colonization, leading to IPD (Abramson *et al.*, 1982; Plotkowski *et al.*, 1986). Indeed, whilst one of these studies showed that adherence of pneumococci to murine tracheal epithelium was increased following inoculation of influenza, another showed that neutrophil dysfunction caused by influenza resulted in increased levels of pneumococcal otitis media (Abramson *et al.*, 1982; Plotkowski *et al.*, 1986). The cytokine response to this

virus may also play a role in exposing pneumococcal adherence receptors and promoting invasion, and although the PAF-receptor was implicated, studies in mice did not confirm this hypothesis (McCullers *et al.*, 2008). Another factor recently implicated has been the viral accessory protein PB1-F2, with studies showing that a mutant possessing the variant of PB1-F2 present in the 1918 pandemic strain was more virulent than its parent strain, with an increase in secondary pneumococcal pneumonia noted. Authors report this result may explain the virulence of the 1918 strain as well as the high mortality levels as a result of pneumococcal pneumonia in this time (McAuley *et al.*, 2007). Furthermore, a study of cytokine expression in mice with secondary pneumococcal pneumonia showed high levels of both pro- and anti-inflammatory cytokines in the blood and lungs, with elevated neutrophil influx also noted. Detailed study of these cytokines and their roles in this synergistic infection may improve understanding of the pathogenic methods employed (Smith *et al.*, 2007).

Neuraminidase activity has been strongly implicated as being important in the synergism between these pathogens, with both possessing active neuraminidase enzymes, and several studies providing strong evidence for this hypothesis (Peltola and McCullers, 2004; Peltola *et al.*, 2005). Historically, in a chinchilla otitis media model, co-infection of Influenza with the pneumococcus caused significantly greater levels of disease than infection with either pathogen alone, confirming the synergistic relationship between pathogens (Giebink *et al.*, 1980). In retrospect, since neuraminidase activity has been suggested as important in this synergism, and NanA is important in otitis media, these studies agree with this hypothesis. More recently, a robust model of secondary pneumococcal pneumonia after influenza infection was developed by the group of Jonathan McCullers, allowing further research into the relationship between influenza and *S. pneumoniae,* as well as the role of neuraminidases in this relationship. Initial studies showed that infection with influenza followed by the pneumococcus resulted in 100% mortality from pneumococcal pneumonia (McCullers and Rehg, 2002). Further to this, specific inhibitors of influenza neuraminidase (NA) became available for clinical use, including Zanamivir and oseltamivir, facilitating studies in this area (Hayden *et al.*, 1999a; Hayden *et al.*, 1999b). Studies using this model showed that inhibition of influenza NA using

oseltamivir resulted in reduced mortality from secondary pneumonia, as well as showing that whilst pre-incubation with influenza increased pneumococcal adherence to human lung epithelial cells, inhibition of influenza neuraminidase activity by oseltamivir prevented this increase. The authors suggest that viral NA plays an important role in the synergism between these organisms, possibly by exposing pneumococcal adherence receptors, due to the reduced virulence observed when specifically inhibiting the viral NA. However, they state that oseltamivir does not have an effect on pneumococcal NanA (McCullers and Bartmess, 2003). The importance of NA activity ahead of other influenzal virulence factors was further strengthened by the finding that whilst treatment with oseltamivir increased survival after secondary pneumococcal pneumonia from 0% to 75%, whilst other antiviral drugs had no effect, and infection with an influenza strain deficient in NA produced a similar increase in survival (McCullers, 2004). Furthermore, Peltola and colleagues used recombinant NA from pandemic strains to show that the levels of neuraminidase activity in these strains correlated with their ability to promote secondary bacterial pneumonia, further strengthening the hypothesis that the activity of viral NA predisposes infected hosts to pneumococcal colonization and resulting pneumococcal pneumonia (Peltola *et al.*, 2005). However, all of these studies have focussed on the role of influenza NA in this relationship, without studying the possible role of pneumococcal NanA, or the effect that inhibitors may have on this enzyme.

1.10 Related viridans group *streptococci*

Members of the *streptococcus* genus are Gram-positive cocci, consisting of both commensal and pathogenic species. The earliest attempt at classification of *streptococci* was by Schottmuller in Germany in 1903, when bacteria were split into beta-haemolytic and non-beta-haemolytic groups. Since then, many methods of classification have been developed including differences in phenotypic traits such as fermentation and pH tolerance and characterisation of carbohydrate antigens (Lancefield, 1933). However, classification was hampered by inaccuracies in phenotypic tests, resulting in species classifications qualified by several exceptions. The advent of gene sequencing allowed further exploration of species classification, with differences in the 16S rRNA gene allowing production of a genetic tree to show the relatedness of streptococcal species (Facklam, 2002). *S. pneumoniae* belongs to the viridans *streptococci*,

which is split into 5 main groups, namely the mutans, salivarius, mitis, anginosus and sanguinis groups (Bruckner and Colonna, 1997). *S. pneumoniae* is a member of the mitis group, which includes other species such as *Streptococcus mitis, Streptococcus oralis* (Kilpper-Balz *et al.*, 1985), *Streptococcus cristatus* (Handley *et al.*, 1991) and *Streptococcus infantis* (Kawamura *et al.*, 1998). This group contains non-beta-haemolytic *streptococci*, which are mostly associated with commensal carriage in the human nasopharynx and oral cavity. *S. pneumoniae* is the notable exception to this rule, being a debilitating human pathogen. However, this species is closely related to other members of the mitis group, and indeed shares >99% sequence identity with its closest relations *S. oralis* and *S. mitis* by 16S rRNA analysis, which identifies 12 member species in this group (Figure 1.3) (Facklam, 2002; Kawamura *et al.*, 1995). Despite their normally commensal nature, *S. mitis* and *S. oralis* are able to cause opportunistic infections, particularly endocarditis in patients with replacement heart valves (Douglas *et al.*, 1993). Furthermore, these organisms can cause infections in immuno-compromised patients, particularly after transplants or in cancer patients (Beighton *et al.*, 1994; Lucas *et al.*, 1997). A number of diagnostic kits exist to distinguish the pneumococcus from related species, including the Rapid ID 32 STREP (Biomerieux, France) and the STREPTOGRAM (Wako Pure Chemicals, Japan). However, success of these kits has been shown to be as low as 50% in correct identification of species (Hoshino *et al.*, 2005).

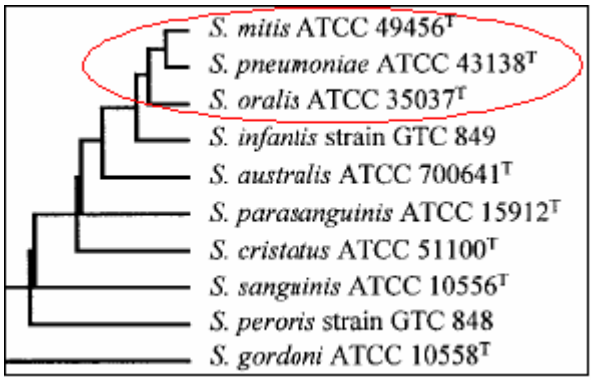


Figure 1.3: Relationship between *S. pneumoniae* and closest relatives

Part of a genetic tree of relatedness between *streptococci* based on 16S rRNA analysis taken from a previous study (Facklam, 2002). *S. pneumoniae* and two closest relations circled in red.

S. pneumoniae can be distinguished from related species by a number of phenotypic and genotypic tests including tests of bile solubility, optochin resistance and agglutination with specific pneumococcal anti-capsular antibodies. However, the classification of these species is blurred, and there are many exceptions when using these classification techniques, with resulting isolates termed “atypical” pneumococci, which may be bile insoluble or optochin resistant (Diaz *et al.*, 1992; Fenoll *et al.*, 1990; Kontiainen and Sivonen, 1987; Munoz *et al.*, 1990; Phillips *et al.*, 1988). A study of the house-keeping genes of these “atypical” isolates uncovered a group of isolates distinct from both *S. pneumoniae* and *S. mitis* (Whatmore *et al.*, 2000). It is likely that these isolates belonged to the species *S. pseudopneumoniae*, characterized in 2004. This species was similar to *S. mitis* by phenotypic tests, but was not distinguishable from *S. pneumoniae* by genotypic tests such as *ply* or *lytA* PCR (Arbique *et al.*, 2004). More recently, phylogenetic analysis of 4 housekeeping genes was shown to allow definitive classification of mitis group s*treptococci*. However, comparison of results for individual housekeeping genes showed that analysis of the *sodA* gene of isolates was highly discriminatory and allowed species differentiation. Authors noted than the exceptions to this were some *S. sanguinis* isolates which failed to PCR for *sodA*, and an inability to distinguish between *S. pneumoniae* and *S. pseudopneumoniae* isolates (Hoshino *et al.*, 2005).

S. pneumoniae, *S. mitis* and *S. oralis* are naturally transformable bacteria, meaning they can take up DNA from their environment and insert it into their genome by homologous recombination (Havarstein *et al.*, 1996; Morrison, 1997). The fact that these 3 organisms share a niche in the human nasopharynx suggests the possibility that they may share genetic material, with the ability to pass genes between species. Indeed, a number of studies have shown that regions of diversity identified in pneumococcal genes have resulted from homologous recombination with related species. This was initially demonstrated in penicillin-biding protein PBP2B, which was identical between species, and thought to be transferred between *S. pneumoniae* and *S. mitis*, *S. sanguinis* and *S. oralis* (Dowson *et al.*, 1989; Dowson *et al.*, 1993; Dowson *et al.*, 1994). A similar mechanism was suggested for the PBP2X gene, and study has shown that evolution of penicillin resistance by this gene originated in *S. oralis*, before

horizontal gene transfer to *S. pneumoniae* (Chi *et al.*, 2007; Laible *et al.*, 1991; Sibold *et al.*, 1994). These results demonstrated the presence of a "pool" of genes for penicillin resistance shared between these species (Reichmann *et al.*, 1997). The transfer of IgA protease genes between *S. pneumoniae*, *S. mitis* and *S. oralis* has also been demonstrated, as was homologous recombination in the *nanA* gene between *S. pneumoniae* and *S. oralis* (King et al., 2005; Poulsen et al., 1998). Other genetic regions of *S. pneumoniae* which have been shown to possess mosaic blocks, a sign of homologous recombination between species, include the competence regulation operon (Havarstein *et al.*, 1997; Whatmore *et al.*, 1999), fluoroquinolone resistance genes (Ip *et al.*, 2007; Janoir *et al.*, 1999) and topoisomerase IV (Balsalobre *et al.*, 2003). Furthermore, a number of studies have demonstrated that isolates characterized as *S. mitis* or *S. oralis* can possess homologues of other pneumococcal virulence genes, or homologues of these genes. This was demonstrated for the *lytA* and *ply* genes, as well as *nanA* (King *et al.*, 2005; Neeleman *et al.*, 2004; Obregon *et al.*, 2002; Whatmore *et al.*, 2000). Taken together, these results imply that the definition between the pathogenic organism *S. pneumoniae* and normally commensal relatives *S. mitis* and *S. oralis* is not simple. Indeed, research into the relationship described a "smooth transition" between species due to the overlap in genetic information observed between species (Hakenbeck *et al.*, 2001). However, a recent study has offered greater insight into the evolution of and relationships between these species. The results showed that *S. oralis* and *S. infantis* isolates cluster distinctly from *S. pneumoniae,* which forms a large cluster with *S. mitis* and *S. pseudopneumoniae*. This cluster was found to represent many evolutionary lineages claimed to represent individual species by taxonomy, with particular diversity observed in the *S. mitis* isolates. Authors suggest that the *S. mitis* "species" is evolving from a pneumococcus-like pathogenic ancestor towards commensality, with pneumococcal virulence genes being lost over time, which would account for their seemingly random distribution in *S. mitis* isolates (Kilian *et al.*, 2008). However, the possibility of evolution occurring in the other direction, with the pneumococcus evolving towards pathogenicity from a commensal ancestor, with *S. mitis* and *S. pseudopneumoniae* in the process of this evolution by acquisition of virulence genes, is not addressed by authors despite the plausibility of this hypothesis as an alternative.

1.11 Aims of the project

The aims of this project were to assess the diversity of two of the main virulence factors of the pneumococcus, and determine the importance of this diversity on virulence of pneumococcal clones. The two virulence factors examined were the cytolytic toxin Ply and the sialidase enzyme NanA. Furthermore, the distribution and diversity of homologues of these virulence factors in related commensal *streptococci* was investigated. Study of the diversity present in these virulence factors will not only further our understanding of pneumococcal pathogenesis, but also give insight into their suitability for use in future protein-based pneumococcal vaccines. Finally, pneumococcal diversity was explored at a more global level by use of pneumococcal microarrays, in order to allow relation of clonal differences to invasive pneumococcal disease, and identification of differences between this debilitating human pathogen and commensal relatives.

2 Materials and Methods

2.1 Bacterial strains, growth conditions and storage

S. pneumoniae strains were initially grown on BAB (Blood agar base, Oxoid, UK) supplemented with 5% horse blood (E&O Laboratories, Scotland), and single colonies inoculated into BHI (Brain heart infusion broth, Oxoid, UK) and grown statically at 37°C until mid-log growth phase (OD_{600nm} 0.6) was reached. These cultures were then stored in 1ml aliquots at -80°C with 10% glycerol (Riedel-de Häen, Germany). Strain purity and optochin sensitivity was tested prior to strain freezing by streaking on BAB + 5% horse blood, with optochin disc (Mast diagnostics, UK). *E. coli* strains were grown from single colonies in LB (Luria broth, Sigma-Aldrich, UK) supplemented with appropriate antibiotics, and stored in 1ml volumes with 10% glycerol.

2.2 Preparation of genomic DNA from *S. pneumoniae*

Strains were grown at 37ºC overnight in 20ml BHI, and then a BAB plate was aseptically streaked with culture to check purity, and identity using an optochin disc. The remaining culture was centrifuged at 3,500g at 4ºC for 15 minutes to pellet the cells, before the removal of the supernatant. The pellet was subsequently resuspended in 1ml lysis buffer (10mM Tris, 100mM EDTA, 0.5% SDS) before incubation at 37ºC for 1 hour. Proteinase K (Invitrogen, Scotland) was then added to the sample to a final concentration of 20µg/ml, before incubation at 50ºC for 3 hours. After this, RNase A (Invitrogen, Scotland) was added to each sample, to a final concentration of 20µg/ml, and the samples were incubated at 37ºC for 30 min, before addition of an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma-Aldrich, UK). The tubes were then inverted sharply several times to mix the samples, and centrifuged at 12,000g for 3 min. The upper phase of the solution was then removed and placed into a fresh 1.5ml tube, before addition of 0.2 volumes of 10M ammonium acetate (Sigma-Aldrich, UK) and ~600µl absolute ethanol (Fisher Scientific, UK, analytical reagent grade). The tubes were then gently inverted, and centrifuged at 12,000g for 30 min in a benchtop centrifuge (Centrifuge 5417C, Eppendorf, UK) to pellet the DNA, before careful removal of the supernatant. The pellets were then air dried for 15-20 minutes to remove any remaining ethanol, before being resuspended in ~200µl TE buffer (10mM Tris, 1mM EDTA) and incubated at

65ºC, with intermittent mixing, to resuspend the DNA fully. The samples were then stored at 4ºC until required.

2.3 Serotyping, Multi Locus Sequence Typing (MLST) and eBURST

Serotyping and MLST was carried out at the Scottish Meningococcal and Pneumococcal Reference Laboratory (SMPRL). Serotyping was done using an agglutination reaction. MLST was done using pre-defined primers (Enright and Spratt, 1998) using a semi-automated technique (Jefferies *et al.*, 2003). This involved DNA sequencing of 7 highly conserved housekeeping genes (*aroE, gdh, gki, recP, spi, xpt* and *ddl)* from genomic DNA of pneumococcal isolates. Each allele was assigned a number in reference to the MLST website (http://www.mlst.net) depending on sequence, which results in a seven-digit "barcode" for each particular isolate. This barcode can then be used to determine the ST of the isolate by comparison the isolates present in MLST database, resulting either in isolates being assigned an ST matching other isolates in the database, or an isolate being assigned a new, novel ST. The main advantage of this is the portability of the system, with global surveillance and comparison of isolates at a genetic level possible. E-Burst 2 (http://www.mlst.net) was used to construct eBURST diagrams illustrating relationships between isolates listed in the pneumococcal MLST database. These diagrams allow comparison of strains from a particular population (eg. All pneumoccocal isolates from a particular serotype, geographical location or time period present in the database) and relates the MLST profiles of these clones by linking single locus variants (clones sharing 6/7 housekeeping alleles) into clonal complexes. STs in clonal complexes which are blue are the predicted founders of those complexes, as predicted by the fact that they have the higher number of single locus variants. The size of the point for each ST represents the number of isolates from that ST present in the MLST database.

2.4 PCR and gene sequencing

Pneumococcal gene *ply* was amplified with external primers 27R and S for gene sequencing, giving a PCR product of 2411bp. The *nanA* gene was amplified with

external primers 31R and S for gene sequencing, giving a PCR product of 4847bp. A 440bp fragment of the streptococcal gene *sodA* was amplified using internal primers previously designed (Hoshino *et al.*, 2005). In each case, a Phusion High fidelity polymerase enzyme (New England Biosciences, UK) was used. In all cases, the following PCR conditions were used, with extension times noted:

PCR reaction conditions

- 98ºC for 30s
- 30x
    - 98ºC for 10s
    - 55ºC for 30s
    - 72ºC for 30s (*sodA*), 60s (*ply*) or 140s (*nanA*)
- 72ºC for 5 min
- Stored at 4ºC

Genes amplified, primers used for amplification and PCR product sizes can be found in Table 2.1.

PCR products were diluted 3:1 in DNA loading buffer (0.25% bromophenol blue, 40% (w/v) sucrose in dH_2O), run in 0.7% agarose (Roche, UK) with 0.1% SYBRSAFE (Invitrogen, UK) and viewed under ultraviolet light in a UVIpro Gold Gel-doc system (UVItec, UK) to confirm successful PCR. PCR products were sent for sequencing at the Molecular Biology Sequencing Unit at the University of Glasgow, where they were cleaned using the AMPure® PCR purification method (Agencourt Bioscience Corporation, Beverly, MA, USA). DNA sequencing was performed using Big Dye Terminator cycle sequencing reagents version 1.1 (Applied Biosystems, UK) and a MegaBace 1000 DNA sequencer (Amersham Biosciences, UK). Primers used for gene sequencing can be found in Table 2.1. Resulting sequence data were assembled aligned and translated using Vector NTI software (Invitrogen, Scotland). Minimum evolution trees showing relationships between nucleotide sequences were constructed using MEGA4 software (Molecular Evolutionary Genetics Analysis, www.megasoftware.net).

Primer	Name	DNA sequence 5'-3'	Use
sodA			
39I	*sodA* forward	TRCAYCATGAYAARCACCAT	PCR amplification and sequencing of *sodA* gene fragment
39J	*sodA* reverse	ARRTARTAMGCRTGYTCCCA RACRTC	PCR amplification and sequencing of *sodA* gene fragment
ply			
4T	*ply* internal forward	GTTGATCGTGCTCCGATGAC	Sequencing of *ply* gene
4V	*ply* internal forward	CAATACAGAAGTGAAGGCGG	Sequencing of *ply* gene
4W	*ply* internal forward	GATCATCAAGGTAAGGAAGT C	Sequencing of *ply* gene
9Y	*ply* internal forward	CGGGATCCGGCAAATAAAGC AGTAAATGACTTT	Sequencing of *ply* gene
27R	*ply* external forward	CTTGGCTACGATATTGGC	PCR amplification and sequencing of ply gene
27S	*ply* external reverse	TACTTAGTCCAACCACGG	PCR amplification and sequencing of *ply* gene
27T	*ply* internal reverse	ATAAGTCATCGGAGCACG	Sequencing of *ply* gene
nanA			
22X	*nanA* internal forward	GAGGAACAGTATGAATCGGA GTG	Sequencing of *nanA* gene
31R	*nanA* external forward	TAACAGTACACCTTGACTGC	PCR amplification of *nanA* gene
31S	*nanA* external reverse	GTGTTCGATAAGGATTGAGC	PCR amplification of *nanA* gene
35E	*nanA* internal forward	TGACCGCTTCTCCAGCATGC	Sequencing of *nanA* gene
35F	*nanA* internal reverse	GAAACCAATGCTTCAAATGG	Sequencing of *nanA* gene
38P	*nanA* internal forward	TCCTCAGTAATGCAGGTGGA C	Sequencing of *nanA* gene
38Q	*nanA* internal reverse	TGGAGCCTTGTTGACCAATA C	Sequencing of *nanA* gene
62F	*nanA* internal forward	CCAGAAGAGGTACAAAAACG TAGTCAAC	Sequencing of *nanA* gene
62G	*nanA* internal forward	AGGGTGATCTATACAAGGGT GACCA	Sequencing of *nanA* gene

62H	*nanA* external reverse	TGCCTTTGGTTTTCGGAACTTT	Sequencing of *nanA* gene
62I	*nanA* internal reverse	CCCCTTCTTGAGCTAAAACAGGAGA	Sequencing of *nanA* gene
62J	*nanA* internal forward	AATCTTTACAAAGGTCAGGAATTGATT	Sequencing of *nanA* gene
62K	*nanA* internal forward	GCCGTCTTCATTCTAGTGACTGGGG	Sequencing of *nanA* gene
62L	*nanA* internal forward	CAACAAACCAAATAAAGATGGAAT	Sequencing of *nanA* gene
64I	*nanA* internal forward	GAATAAATGTCTTATTTCAGAAATC	Sequencing of *nanA* gene
64J	*nanA* internal reverse	TTATTGTTCTCTTTTTTCCCTAGC	Sequencing of *nanA* gene

Table 2.1: Primers used for PCR amplification and sequencing of pneumococcal genes

2.5 Testing production and activity of virulence factors

2.5.1 Preparation of samples

Isolates were grown statically at 37^{o}C, in 20ml BHI until mid-log phase (OD_{600nm} 0.6), spun for 15 minutes at 3,500g, 4^{o}C in a 4K15 centrifuge (Sigma-Aldrich, UK). When required, induction of neuraminidase activity was accomplished by addition of 50µg/ml N-acetylneuraminic acid (Neu5ac) to growth medium. Pellets were resuspended in 1ml PBS (Phosphate-buffered saline) with protease inhibitor cocktail (Roche, UK). Samples were sonicated at 10 microns for eight repeats of 30 seconds with 30 seconds rest using a Soniprep 150 sonicator (MSE, UK). Supernatants, where required, were concentrated 20x to a final volume of 1ml, using an Amicon centrifugal concentrator with cut-off 30kDa (Millipore, UK). All samples were stored on ice until use.

2.5.2 SDS-PAGE and Western blotting

Samples were diluted 3:1 in NuPAGE sample buffer (Invitrogen, Scotland) and heated to 70^{o}C for 10 minutes before loading onto gel. Gels used were 10%

polyacrylamide gels throughout (Invitrogen, Scotland). Markers used were either Kaleidoscope protein markers (Bio-rad, UK) or SeeBlue +2 pre-stained markers (Invitrogen, Scotland). A sample volume of 20µl was loaded in each well, and run at 100V for 60-80 minutes. SDS-PAGE gels were stained overnight in Coomassie stain (500ml dH_2O, 400ml methanol, 100ml acetic acid, 0.5g Coomassie blue R250), and destained with destain solution (500ml dH_2O, 400ml methanol, 100ml acetic acid) until clear.

For Western blot, the proteins were transferred to Hybond-C nitrocellulose membrane (Amersham Biosciences, UK) by blotting at 100V for 1 hour. The membrane was the shaken at 37ºC for 2-3h in 3% skimmed milk (Marvel, UK) in Tris NaCl, pH 7.4 (10mM Tris, 150mM NaCl, 8mM HCl), with a 1/1000 dilution of primary antibody before washing x 4 in 40ml Tris NaCl pH7.4 for 5 minutes. The membrane was then transferred to 3% skimmed milk in Tris-NaCl pH 7.4 with 1/1000 dilution of HRP-labelled secondary antibody (α-rabbit IgG antibody raised in donkey, Amersham Biosciences, UK), and shaken for 1 h at 37ºC, before washing x 4 as described, incubated with developer (40ml Tris NaCl, pH7.4, 10ml Methanol, 30µl Hydrogen peroxide (30% w/v) and 30mg 4-chloro-1-naphthol, prepared just before use) at room temperature in the dark and stopped with dH_2O.

2.5.3 Haemolytic assay

The haemolytic activity of Ply in culture lysates was measured using a haemolytic assay previously described (Walker *et al.*, 1987), with modifications. Briefly, doubling dilutions of 50µl culture lysates, starting at a 1/5 dilution were made in PBS in duplicate in a round-bottomed 96-well plate (Costar, UK). A control double dilution of a 1/1000 dilution of purified Ply (0.7mg/ml) (Mitchell *et al.*, 1989) was included in duplicate. PBS was included as a negative control. 50µl of 10µM dithiothreitol (DTT) (Sigma-Aldrich, UK) was added to each well, to act as a reducing agent for any Ply oxidised in the lysate. The plate was then incubated at 37ºC for 15 min with a lid to prevent evaporation.

A 2% (vol/vol) solution of either sheep erythrocytes (E & O Laboratories, Scotland) or human erythrocytes was prepared and 50µl was added to each well. Unless stated, the assay was run with sheep erythrocytes. The plate was then

incubated at 37°C for 30 min before addition of a further 50µl PBS, and centrifugation at 500g for 1 minute in a 4K15 centrifuge (Sigma-Aldrich, UK) to pellet the intact cells. A volume of 100µl of supernatant was removed from each well and added to the corresponding well of a fresh flat-bottomed 96-well plate (Costar, UK), and a spectrophotometer reading at 540nm taken in a FLUOstar Optima plate-reader (BMG Labtech, UK) to measure the levels of haemoglobin released in each well. From this result, the percentage (%) lysis in each well was calculated using the mean negative control value as the 0% lysis value, and the mean value from the wells where 100% haemolysis has occurred as the 100% value. A curve of % lysis against well number was plotted for each strain, giving a typically sigmoid curve, using GraphPad Prism 4 software (GraphPad software, USA).

2.5.4 Ply sandwich ELISA

Ply expression levels in culture lysates were quantified using a Ply sandwich ELISA described previously (Cima-Cabal *et al.*, 2003) with modifications. Maxisorp 96-well plates (NUNC, UK) were coated with 100µl 2.5µg/ml mAb PLY-7 (de los Toyos *et al.*, 1996) in coating buffer (carbonate/bicarbonate buffer, 0.05M, pH9.6) by incubation at 4ºC overnight. Plates were blocked with blocking buffer (10% foetal calf serum [Invitrogen, Scotland] in PBS) and washes were made with PBS + 0.05% Tween20 (Sigma-Aldrich, UK). Blocking buffer with 0.05% Tween20 (assay buffer) was used to dilute samples. Cell extracts were diluted 1:7500 and 100µl added to wells in duplicate. A standard curve of purified Ply (2000pg/ml to 31.25pg/ml) was used to quantify Ply levels. A 1:2000 dilution of polyclonal α–Ply antibody (Mitchell *et al.*, 1989) in assay buffer was added to each well at a volume of 100µl, and plates were shaken at 37°C for 30 min. After 4 washes, 100µl of biotinylated α-rabbit IgG antibody (Amersham Biosciences, UK) diluted 1:500 in assay buffer was added to each well and plates were shaken at 37°C for 30 min. After a further 4 washes, 100µl streptavidin HRP (KPL, Maryland, USA) diluted 1:2000 in assay buffer was added to each well, and plates were incubated at room temperature for 30 min before being developed with with 3, 3′, 5, 5′-tetramethylbenzidine (TMB) substrate (KPL, Maryland, USA) according to the manufacturers instructions. 50µl of stop solution (10% H_2SO_4) was then added and the plate read at 450nm to allow construction of standard curve and calculation of unknown Ply concentrations.

2.5.5 Calculation of Ply specific activity

From the results of the haemolytic assay and Ply sandwich ELISA on culture lysates, the specific activity of Ply present was calculated. The Ply concentration in the cell lysates was calculated from the ELISA results, and from this, the concentration of Ply present in each well of the haemolytic assay was calculated for each cell lysate tested. From this, the concentration of Ply required to lyse 50% of erythrocytes in the assay was calculated, and reciprocal of this value gave the specific activity of the Ply in the sample, in haemolytic units/mg (HU/mg).

2.5.6 Neuraminidase activity

The neuraminidase activity present in culture lysates and concentrated supernatants was measured using a modified version of a previously described assay (Manco *et al.*, 2006). Total protein concentrations were calculated using a Bradford's assay as previously described (Bradford, 1976) and concentrations were standardized. In a round-bottomed 96-well plate, 25µl of each sample was added to 25 µl of 0.3mM pNP-NANA (2-*O*-(*p*-nitrophenyl)-α-D-*N*-acetylneuraminic acid (Sigma-Aldrich, UK) prepared in 50mM Sodium citrate pH6.0) in triplicate. 25µl of 40mM Tris-HCl (pH 7.6) was added as a negative control. Samples were incubated at 37ºC for 30min, and 100µl 0.5M ice cold Na_2CO_3 was added to each well. The plate was read at 405nm to measure the release of p-nitrophenol (pNP) from the substrate. For strain lysates, this was compared to a standard curve of pure pNP (Sigma-Aldrich, UK) to determine neuraminidase activity in the sample, with the units in nM pNP released/mg total protein/hour. The results for sonicate and supernate were averaged and combined for each isolate to determine overall neuraminidase activity. For purified NanA proteins, % release of pNP from the pNP-NAN substrate was plotted against protein concentration, and specific activity in enzymatic units/mg calculated by taking the reciprocal of the NanA concentration required to cause 50% pNP release in the assay. Statistical analyses of results were performed using GraphPad InStat 3 (Graphpad Software,)

2.6 Construction of *S. pneumoniae* Ply knockout strain

2.6.1 Rationale of Janus technique

A strain of wild-type D39 *S. pneumoniae* expressing no Ply (D39 ΔPly stop) was constructed using the Janus technique (Sung *et al.*, 2001). This method allowed production of unmarked mutations in *S. pneumoniae* by employing a method of negative selection. Firstly, D39 was transformed with a gene to confer streptomycin resistance (*rpsL*-). Secondly, the Janus cassette, conferring both kanamycin resistance and dominant streptomycin sensitivity (*rps*L+), was inserted into the *S. pneumoniae* genome in position of the *ply* gene by homologues recombination, resulting in clones with kanamycin resistant/streptomycin sensitive phenotype. A *ply* gene with the desired mutation was then transferred into these clones by homologues recombination, replacing the Janus cassette and rendering the clones kanamycin sensitive/streptomycin resistant. The result of this was a *S. pneumoniae* D39 isolate with the desired mutation but no antibiotic resistance cassette inserted into the mutated gene, as shown in Figure 2.1.

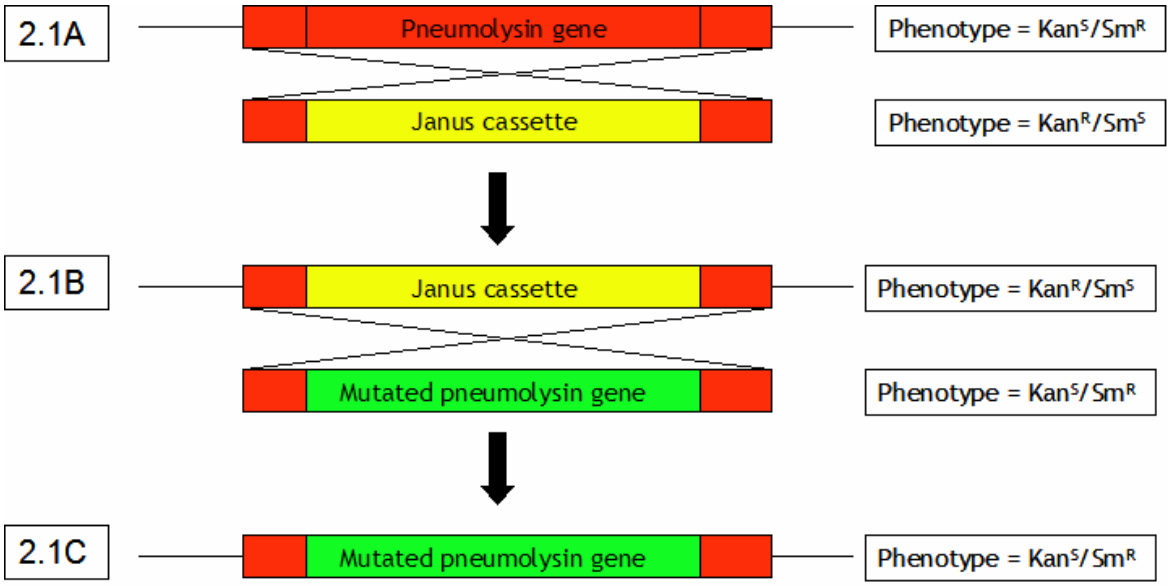


Figure 2.1: Rationale of the Janus technique

The process involved in replacing the *ply* gene with a mutated variant, using the Janus technique. Firstly, a streptomycin resistant wild-type D39 isolate is transformed with the Janus cassette with ligated *ply* flanks, which replaces the *ply* gene, and gives the transformants a kanamycin^R/streptomycin^S phenotype (2.1A). Then the resulting transformant is then transformed with the *ply* gene with flanking DNA and the desired mutation inserted, resulting in loss of the Janus cassette and a kanamycin^S/streptomcin^R phenotype (2.1B-C). The result is a mutant with the desired mutation in *ply*, and no resistance marker inserted in the gene.

2.6.2 Transformation of S. pneumoniae

This method involved three transformation steps. Firstly, the *rpsL*- gene from R6 CP1200 *S. pneumoniae* (Sung *et al.*, 2001) was transformed into D39. Secondly, the Janus cassette with *ply* flanking DNA (Provided by Dr Gavin Paterson) was transformed into the D39 Smr strain in place of the *ply* gene. Finally, the mutated *ply* gene, with a stop codon inserted at the beginning of the gene (Figure 2.2, provided by Dr Lea-Ann Kirkham), was transformed into the clones to replace the Janus cassette. The transformation method employed in all three cases was as follows. D39 pneumococci (and clones from various stages of the process) were grown in 10ml BHI with 1mM $CaCl_2$ until an OD_{600nm} of 0.1 was reached. Aliquots of 1ml of the culture were prepared, and 100ng/ml competence stimulating peptide 1 (CSP-1, Sigma-Aldrich, UK) was added to induce uptake of extracellular DNA. Samples were then incubated at 37^{o}C for 15 min before addition of 100ng test DNA (PCR products of either *rpsL*- gene, Janus cassette with *ply* flanks, or mutated *ply* gene with *ply* flanks) and a further incubation at 37^{o}C for 75 min.

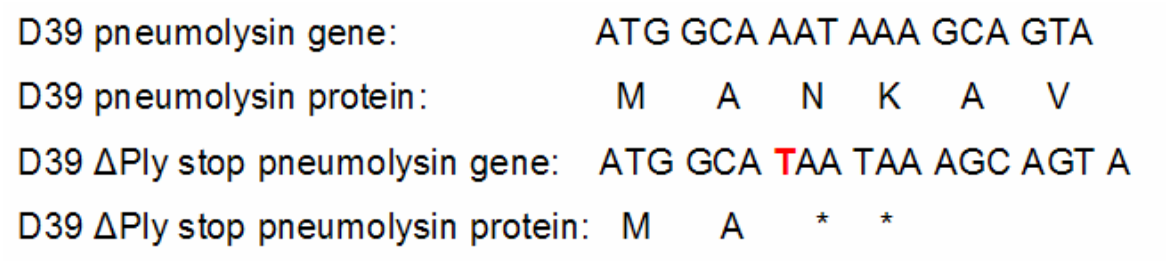


Figure 2.2: Sequence of D39 ΔPly stop mutant

Diagramatic representation of mutation inserted into D39 ΔPly stop using the Janus technique. Insertion of an extra T base after base 6, highlighted in red, results in a frameshift and, the presence of 2 TAA codons, underlined, at the start of the gene. When translated, this results in 2 stop codons after amino acid 2 in the *ply* gene, represented by asterisks. The result of this is the loss of Ply protein production by the mutant.

Samples were then plated on BAB with 5% horse blood and either 300µg/ml streptomycin in the first and third transformations, or 200µg/ml kanamycin in the second transformation. Plates were incubated for 9h at 37°C under anaerobic conditions using GasPak™ Pouches (Becton Dickenson, Oxford, UK). This short anaerobic incubation time minimised the chance of streptomycin resistant revertants occurring which do not have the desired phenotype. Successful transformants in the second and third transformation were then inoculated onto both streptomycin and kanamycin selective plates to confirm the correct resistance phenotype in each case.

2.7 Characterisation of D39 ΔPly stop

2.7.1 Confirmation of successful mutation

The successful mutation of the *ply* gene in D39 ΔPly stop, resulting in no expression of the *ply* protein, was confirmed by sequencing, haemolytic assay and Western blot as described previously.

2.7.2 Animal studies using D39 ΔPly stop

The effect of the loss of Ply on virulence of the *S. pneumoniae* strain D39 was investigated by studies in mice. Three strains were included in the study, namely D39, D39 ΔPly stop, and D39 Δ6, a Ply mutant with amino acids 146-147 deleted, created by Dr. Lea-Ann Kirkham. This mutant produces Ply protein, but has no haemolytic activity. These strains were passaged through three female MF1 mice

(Harlan, UK), and standard inoculums prepared for further infection. Ten 6 week old female MF1 mice per strain were inoculated intranasally with $1x10^6$ colony forming units (cfu) of the appropriate bacteria, and viable counts in blood were measured after 24 and 48h periods. Furthermore, the survival of the mice was monitored for a period of 2 weeks. Animal studies were conducted by Dr Gill Douce and Kirsty Ross. Results were analysed by Mann-Whitney U test using GraphPad InStat software (GraphPad software, UK).

2.7.3 Determination of viable counts

The method used to determine viable counts in a sample was based on the original method of viable counting (Miles and Misra, 1938). At 24 and 48h time points, ~50µl of blood was removed from mice using an insulin syringe. Viable counts of *S. pneumoniae* were calculated by addition of 20µl blood to 180µl PBS, followed by six 1:10 dilutions. Three 20µl samples from each dilution were plated out on BAB plates with 5% horse blood and incubated overnight at 37°C. The number of viable *S. pneumoniae* in the blood, in cfu/ml, was then determined by averaging the number of colonies from a chosen dilution, multiplying by the dilution factor, and multiplying by a factor of 50 to account for the 20µl sample tested.

2.8 Cloning, expression and purification of NanA fragments

2.8.1 Cloning method

PyMol software (DeLano Scientific, USA) was used to map mutations onto a structural model of a NanA fragment (AA318-792). This allowed selection of NanA alleles for further study. The method used for cloning NanA active fragments (AA318-792) was based on the high-throughput method developed previously (Berrow *et al.*, 2007) using In-Fusion cloning technology (Zhu *et al.*, 2007), with a number of modifications. The vector used was pOPINF, based on the three-promoter vector pTriEx2. The cloning method involves amplification of a target sequence with primers possessing overlaps matched to regions of the pOPINF vector, and cloning with this cut vector to insert the target sequence

into the vector in *E. coli* (Figure 2.3). The NanA fragment can then be purified from the *E. coli* cells, possessing a C-terminal His tag and a C3 protease site allowing purification.

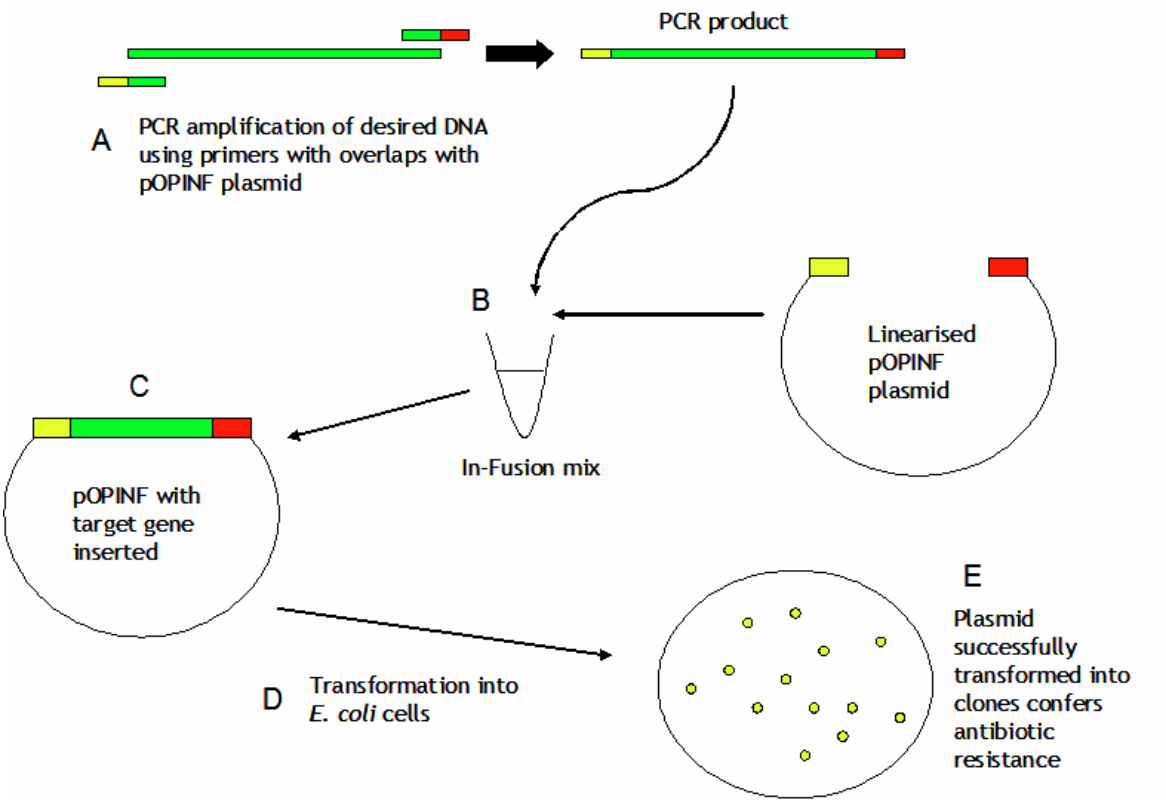


Figure 2.3: In-Fusion cloning technology

The method used to clone NanA fragments into *E. coli* cells. Firstly, the desired DNA was amplified using primers with overlap regions to the plasmid pOPINF (2.3A). The resulting PCR product was then mixed with the linearised pOPINF plasmid and the In-Fusion mix (2.3B), producing the pOPINF plasmid with the desired DNA inserted into the correct position to allow expression (2.3C). Transformation of this plasmid into *E. coli* cells (2.3D) results in clones of *E. coli* with the plasmid and associated antibiotic resistance (2.3E).

2.8.2 Primer design and PCR

Primers were designed to amplify the selected 1425bp region from different alleles of the *nanA* gene (AA318-792). Primers included overlap regions for the pOPINF vector, and can be found in Table 2.2. 1425bp fragments from NanA alleles 12, 16 and 18 were amplified by PCR using primers detailed in Table 2.2 and Phusion high fidelity DNA polymerase enzyme.

Primer	DNA sequence 5'-3'	Use
64U	AAGTTCTGTTTCAGGGCCCGCCTGAAGGAGCGG TTTTAACAGAGAAA	PCR of *nanA* alleles 12 and 16
64V	AAGTTCTGTTTCAGGGCCCGCCAGAAGGTGCGA AAATCTCAGAGAAA	PCR of *nanA* allele 18
64W	ATGGTCTAGAAAGCTTTAATCTTTGCTCAAAAAG TCCCAATTAAA	PCR of *nanA* allele 12
64X	ATGGTCTAGAAAGCTTTAATTTTTGCTCAAAAAG TCCCAATTAAA	PCR of *nanA* allele 16
64Y	ATGGTCTAGAAAGCTTTAATTTTTGCTCAAAAAT TCCCAATTAAA	PCR of *nanA* allele 18

Table 2.2: In-Fusion cloning primers

Primers used for PCR of *nanA* fragments for In-Fusion cloning into pOPINF vector. Bases in red represent those from *nanA* alleles, whilst bases in blue are from pOPINF vector and in black are inserted stop codons.

2.8.3 Cloning

100ng PCR product and 100ng linearised and purified vector pOPINF were diluted in nucleotide-free water (Sigma-Aldrich, UK) to a final volume of 10µl. This was added to resuspended In-Fusion dry-down mix and incubated at 42°C for 30 min. The mix was immediately diluted in 40µl TE buffer (10mM Tris, 1mM EDTA), 2µl added to 50µl *E. coli* DH5α cells, and incubated on ice for 30 min. The cells were then heat-shocked at 42°C for 30s and incubated on ice for 2 min. 450µl GS96 medium with 1% glycerol was added, and samples incubated for 1h at 37°C. Varying volumes of culture were then plated on LB agar with 1µM Isopropyl ß-D-1-thiogalactopyranoside (IPTG, Melford Laboratories, UK), 50µg/ml ampicillin and 20µg/ml 5-Bromo-4-chloro-3-indolyl ß-D-galactopyranoside (X-GAL, Sigma-Aldrich, UK), and plates incubated at 37°C overnight. White colonies were selected as successful tranformants, and plasmids purified using a Miniprep spin kit (QIAGEN, UK) as per manufacturer's instructions. Plasmids were then sequenced to confirm success of cloning reaction.

2.8.4 Transformation of pOPINF plasmid into E. coli

100ng of expression vector pOPINF was added to 50µl *E. coli* cells (either B834 or Rosetta, Invitrogen, Scotland) and incubated on ice for 30 min. The cells were then heat-shocked at 42°C for 30s and incubated on ice for 2 min. 450µl GS96 medium (Qbiogene, UK) with 1% glycerol was added, and samples incubated for 1h at 37°C. Varying volumes of culture were then plated on LB agar with 1%

glucose and either 54ug/ml carbenicillin (Sigma-Aldrich, UK) or 34ug/ml chloramphenicol (Sigma-Aldrich, UK) and 54ug/ml carbenicillin for B834 and Rosetta cells respectively. Colonies were picked and plasmids purified and tested for successful transformation as described previously.

2.8.5 Expression and purification of NanA fragments

A colony of Rosetta cells containing pOPINF plasmid with *nanA* allele fragment was inoculated into 100ml LB with 34µg/ml chloramphenicol and 54µg/ml carbenicillin and incubated at 37^{o}C overnight, with shaking. A volume of 5ml overnight culture was inoculated into 500ml TB overnight expression medium (Novagen, UK) and incubated shaking for 4h at 37ºC, followed by 20h at 25ºC. The resulting culture was centrifuged at 3,500g for 30 min at 4ºC in a J-B6 centrifuge (Beckman, UK), and the pellet resuspended in 30ml lysis buffer. A protease inhibitor tablet (Roche, UK), 400 Kunitz units DNase I (Invitrogen, UK) and 1µM $MgCl_2$ were added to the sample, which was sonicated at 8µm for 10 x 15s with 45s rest. The sample was then centrifuged at 6000g for 20 min (4K15, Sigma-Aldrigh, UK), supernatant removed and centrifuged at 15,000g for 30 min (Allegra™ 64R, Beckman, UK) and supernatant collected. A 1ml Histrap HP column (GE Healthcare, UK) was stripped with 5 column volumes (cv) of 500mM EDTA, primed with 5 cv Nickel sulphate ($NiSO_4$), 5 cv elution buffer (50mM Tris pH 7.5, 500mM NaCl, 500mM imidazole) was passed through, and the column was washed with 10 cv wash buffer (50mM Tris pH 7.5, 500mM NaCl, 20mM imidazole) using a peristaltic pump. The protein sample was passed through the nickel column, where the His-tagged protein bound the column and the rest of the proteins present passed through. The column was then washed with 10 cv wash buffer to remove residual contaminants, and 10 cv elution buffer to elute the bound protein.

2.8.6 C3 protease cleavage of purified protein

To remove residual contaminants from the purified NanA sample, the His tag was removed from the protein by cleavage at the C3 site by C3 protease (provided by Dr. Alan Riboldi-Tunnicliffe). 5µl C3 protease was added to the sample, along with 1.5mg EDTA and 1mg DTT, before incubation at 4^{o}C for 2h. The sample was then passed through the nickel column again, and whilst the purified NanA

fragment, with the His tag removed, passed through the column, previous contaminants remained bound to the column. Protein samples were run on SDS-PAGE gels and coomassie stains as described to confirm purification.

2.8.7 Gel filtration of purified protein

When necessary, proteins were further purified by gel filtration using an AKTAprime plus (GE Healthcare) with Superdex 75 300ml column (GE Healthcare, UK). A 2 ml volume of wash buffer was passed through the column to clean, before 10ml sample was added. After this, 450ml gel filtration buffer (50mM Tris pH 7.5, 500mM NaCl) was passed through the column with a flow rate of 0.2ml/min and collected in 5ml fractions. The fractions were tested for presence of NanA fragment by SDS-PAGE and coomassie stain as described, and fractions containing pure protein were pooled.

2.8.8 Dialysis and concentration of purified NanA fragments

Dialysis tubing with a molecular weight cut off point of 14,000Da (Medicell International Ltd, UK) was previously boiled in 2% sodium bicarbonate solution with 1mM EDTA for 10 min. The sample was sealed in a 20cm length of the pre-boiled dialysis membrane and dialysed overnight at 4^{o}C in 1L wash buffer (50mM Tris pH 7.5, 500mM NaCl, 20mM imidazole) to remove high concentrations of imidazole from the sample. The protein was then concentrated to a volume of ~1ml by centrifugation in an Amicon centrifugal filter (30kDA cut-off, Millipore, UK) at 4,100g, 4^{o}C. The absorbance at 280nm (A_{280}) of a 1:50 dilution of the protein sample was measured in a spectrophotometer, and extinction coefficient (ε) calculated using protparam software (www.expasy.ch/tools/protparam). The protein concentration in mg/ml was calculated by the formula ((A_{280}/ε) x dilution factor).

2.9 Study of effects of viral neuraminidase inhibitors on NanA *in vitro* and *in vivo*

2.9.1 Inhibition of purified NanA fragments by viral neuraminidase inhibitors in vitro

The inhibition of purified active NanA fragments (AA318-792) with inhibitors of Influenza neuraminidase (NA) was investigated. Three inhibitors were tested, namely Oseltamivir carboxylate (OC, the active form of Oseltamivir phosphate, marketed as Tamiflu, Roche, UK), Zanamivir (marketed as Relenza, GlaxoSmithKline, UK) and the competitive inhibitor 2,3-dideoxy-2,3-didehydro- N -acetyl-neuraminic acid (DANA, Sigma-Aldrich, UK). The inhibitory effect of these compounds was tested by preparing a round-bottomed 96-well plate with a concentration of NanA that conferred 100% activity in the NanA activity described, and adding doubling dilutions of inhibitor, starting from 3.15mg/ml, 35mg/ml and 8.75mg/ml respectively. Plates were then incubated at 37^{o}C for 15 min before addition of pNP as described, and continuation of NanA assay. Curves of inhibition and concentrations required to inhibit 50% of activity (IC_{50}) were calculated using GraphPad Prism 4 (GraphPad software, USA).

2.10 Microarray analysis

2.10.1 S. pneumoniae microarray

Microarray technology involves attachment of genetic probes onto slides for hybridization with test genomic DNA or RNA, allowing the presence or expression of genes to be determined, respectively. The microarray slides used were prepared by the Bacterial Microarray Group at St. Georges Hospital, London (BµG@S). The probes on the array represent all the genes of reference isolate TIGR4 (2240 genes) as well as unique genes present in R6 (117 genes). DNA or RNA from a test isolate is labeled with either Cy3 or Cy5 fluorescent dye (GE Healthcare, UK), with the reference isolate TIGR4 labeled with the opposite dye. Competitive hybridization to the probes on the microarray slide then occurs between the two labeled samples, allowing determination of genes present or

absent in test gDNA and up- or down-regulated in test RNA, all in comparison to the reference isolate TIGR4. A schematic of microarray technology is shown in Figure 2.4.

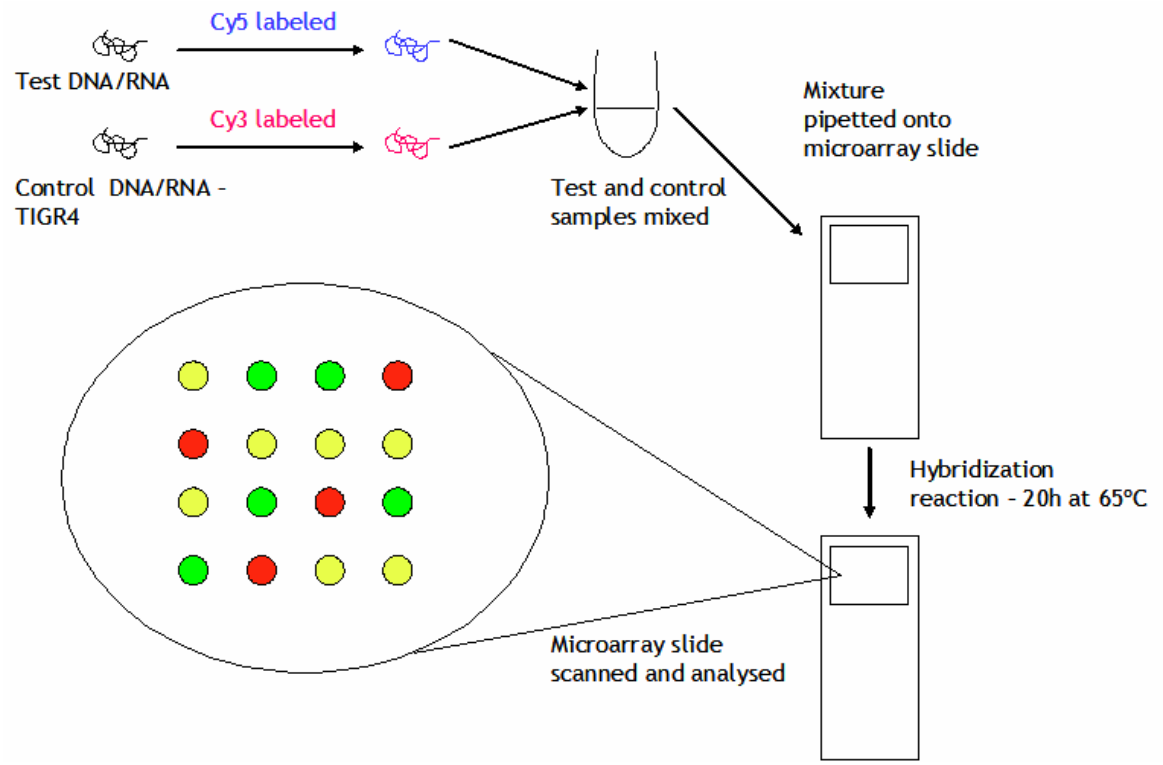


Figure 2.4: The basic methodology involved in microarray technology

Test gDNA and control gDNA were labelled with either Cy3 or Cy5 fluorescent dye, combined, and pipetted onto the microarray slide. The hybridization reaction then ran for 20h at 65ºC, after which the microarray slide was scanned, and the results interpreted. Each spot on the slide contained a probe complementary to DNA from a particular gene of the reference isolate, and a yellow spot on the slide indicated competitive hybridization, where the gene was present in both test and control isolates. Green and red spots represent non-competitive hybridization where one isolate possessed the gene whilst the other lacked it.

2.10.2 Preparation of genomic DNA and RNA for microarrays

Pneumococcal genomic DNA for use in microarrays was prepared as previously described (Section 1.2). Genomic DNA from other viridans *streptococci* was prepared using a DNeasy blood and tissue kit (QIAGEN, UK), as per manufacturer's instructions for purification of Gram positive gDNA.

Pneumococcal RNA was prepared by growing strains in 10ml BHI until mid-log phase (OD_{600nm} 0.6). The culture was then pelleted by centrifugation at 4,100g for 5 minutes (4K15 centrifuge, Sigma-Aldrich), supernatants discarded and pellets immediately frozen in liquid nitrogen. Pellets were then resuspended in

200µl lysozyme TE buffer (10mM Tris, 1mM EDTA, 15mg/ml lysozyme), vortexed and incubated at room temperature (RT) with vortexing every 2 min. A RNeasy mini kit (QIAGEN, UK) was used to purify RNA, using the following protocol. 700µl RTL buffer was added, and the sample was vortexed to 10s. The sample was then transferred to fresh 1.5ml centrifuge tube (Eppendorf, UK) containing 50mg 100µm glass beads (Sigma, UK), and cells disrupted using a ribolyser (Hybaid, UK) at speed 4 for 4 x 20s. The samples were then centrifuged at 12,000g for 10s (1K15 centrifuge, Sigma-Aldrich, UK), the supernatant transferred to a fresh tube and mixed with 500µl absolute ethanol (Fisher Scientific, UK, analytical reagent grade). 700 µl of this solution was applied to a RNeasy Mini column (QIAGEN, UK) and centrifuged at 12,000g for 30s. The flow-through was discarded, and the remaining 700µl added to the column before re-centrifugation, and flow-through was again discarded. 350µl of RW1 buffer was added, and the column was centrifuged at 12,000g for 5 min. A volume of 80µl DNase 1 (28 Kunitz units) (QIAGEN, UK, not in kit) was added to the column, and column was incubated at RT for 15 min. 350 µl Buffer RW1 was pipetted into the column, before centrifugation for 30s at 12,000g. The flow-through was discarded, and 700µl RW1 buffer was added before centrifugation at 12,000g for 30s. The flow-through was again discarded, and 500µl RPE buffer was added. The column was centrifuged at 12,000g for 30s, then the flow-through was discarded and an additional 500µl RPE buffer was added, and the column was centrifuged at 12,000g for 2 min to dry the membrane. The column was transferred to a fresh 1.5ml tube, and 50µl nuclease-free water was added directly onto the membrane to elute the RNA. The column was then left to stand for 1 min, before centrifugation at 12,000g for 1 min. RNA was stored at -80^{o}C until required.

2.10.3 Analysis of DNA and RNA samples for microarray

The concentration of DNA or RNA was measured using a nanodrop ND 1000 (Thermo Scientific, UK). The purity of RNA in the samples was confirmed by use of a Bioanalyser 2100 (Agilent, UK). Using this technique, the output shows that two clear peaks are present in the sample, representing 16S and 23S RNA, when RNA has been successfully purified (Figure 2.5). Presence of gDNA in desired samples was confirmed by PCR for either the *ply* or *sodA* genes as described.

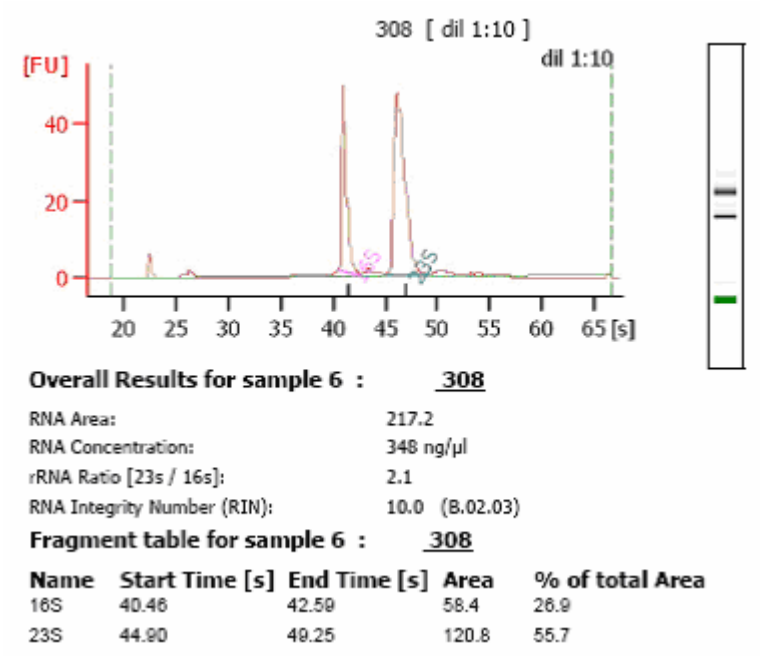


Figure 2.5: Sample result of RNA purity confirmation

Purity confirmation carried out using Bioanalyser 2100. The result showed that two clear peaks were present in the plot, representing 16S and 23S RNA. The lack of other peaks confirmed the purity of the sample, with no DNA contamination present. A gel of the two pure peaks can also be seen on the right hand side.

2.10.4 DNAvsDNA microarray hybridization

The hybridization reaction was carried out as per manufacturer's instructions. Briefly, 5µg test gDNA was mixed with 1µl random primers (Invitrogen, Scotland) and made up to 41.5µl with nuclease-free water. The samples were incubated at 95°C for 5 min, before snap cooling, and addition of 1µl Klenow polymerase (Invitrogen, Scotland), 5µl 10x REact buffer, 1µl dNTPs (5mM dA/G/TTP, 2mM dCTP, Invitrogen, Scotland) and 1.5µl Cy3/Cy5 dye (GE Healthcare, UK). The samples were incubated in the dark at 37°C for 90 min.

The microarray slides were prepared for the hybridization by incubation with pre-hybridization solution (3.5xSCC buffer (Ambion, UK), 0.1% SDS, 10mg/ml BSA) for 20 min, before washing in dH_2O for 1 min and isopropanol (Fisher Scientific, UK) for 1 min. The slides were then centrifuged in a 4K15 centrifuge (Sigma-Aldrich, UK) at 800g for 5 min to dry them, and stored in a dark, dust-free box until required.

The Cy3/Cy5-labelled test DNA was then mixed with TIGR4 DNA labeled with the opposite dye, and the resulting sample purified using a MinElute purification kit (QIAGEN, UK) as per manufacturer's instructions, with the final elution being in 15.9µl nuclease-free water (Ambion, UK). 14.9µl of sample was mixed with 4.6µl filtered 20xSCC buffer (Ambion, UK) and 3.5µl filtered 2% sodium dodecyl sulphate (SDS, Sigma-Aldrich, UK). The sample was heated to 95°C for 2 min, and allowed to cool, and briefly centrifuged. A lifter slip (Erie Scientific Company, Portsmouth, NH, USA) was then placed on the microarray slide, to cover the probes, and the sample pipetted under the lifter slip. The slides were carefully placed into a microarray hybridization cassette, which was sealed, submerged in water and incubated at 65°C for 16-20h.

The slides were washed in Wash A buffer (1xSCC buffer, 0.05% SDS, pre-heated to 65°C) for 2 min to remove the cover slip and wash the slide, followed by washing in Wash B buffer (0.06xSCC buffer) for 4 min. The slides were then centrifuged at 800g for 5 min before scanning.

2.10.5 RNAvsRNA microarray hybridization

The RNAvsRNA hybridizations were carried out as per manufacturer's instructions. These were identical to the DNAvsDNA hybridization method, with the following exceptions. 10µg RNA was initially added to 1µl random primers, and made up to 11µl in nuclease-free water. After incubation, 2.5µl Superscript II (Invitrogen, Scotland), 5µl 5x First strand buffer, 2.5µl 100mM DTT, 2.3µl dNTPs and 1.7µl Cy3/Cy5 dye was added, before incubation at 25°C in the dark for 10 min, followed by 42°C for 90 min.

2.10.6 Microarray normalization and analysis

Microarray slides were scanned using ScanArray Express™ (Packard Biosciences Biochip Technologies, Perkin Elmer), and resulting TIF images were entered into Bluefuse (BlueGnome Ltd, Cambridge, UK) for Microarrays 3.5 © BlueGnome Ltd with the Cy3 labeled image in Channel 1 and the Cy5 labeled image in Channel 2. The array gridmap files utilised were provided by the Bacterial Microarray Group at St George's (BµG@S). The post processing protocol was comprised of initial exclusion of unreliable results due to poor quality hybridizations with a confidence estimate of less than 0.1. Controls spots on the array were identified

using an array gridmap GAL file (SPv1_1_0_CGH_Gridmap.bcf) and data from control spot hybridizations was automatically removed from the analysis. To correct for spatial, intensity and dye related effects, normalization was performed using the option “Global Lowess excluding all with text.” Confidence flags were set at default settings. Replicates of each dye swap were combined by fusion.

Comparative genomic hybridization was performed in Bluefuse through the identification of a normal distribution of experimental variability and by identifying variability which was two standard deviations from the median of this normal distribution of all results for the TIGR4 and R6 genes represented on the microarray. Automated classification of regions of variability was performed by setting a ratio threshold for amplification as 1.0 and ratio threshold for deletion at -1.0 with the minimum number of clones included in the region in order for it to be classified as an amplification or a deletion set at 1. Dye swap processing was enabled.

After normalisation in Bluefuse, output_fused_CGH files were imported into Genespring GX 7.3.1 (Agilent Technologies, USA) and further normalization was performed after data transformation to account for dye swaps. This normalization was performed using the “Per spot and divided by control channel” protocol with a cross gene error model using the error model for one-colour data. The error model was based on deviation from 1. The generation of gene lists using Genespring GX 7.3.1 was accomplished by importing the Bluefuse generated output_fused.xls files to create an experiment whereby the microarray dye swaps for each strain could be analysed using the “Filter on data file” option. To generate a gene list a search was performed using the “Type” column employing search criteria, “Column values must be not equal to NO CHANGE,” and “Value must appear in at least 1 of the selected columns.” The resulting gene list could be saved or exported into Microsoft Excel for comparison with further strains.

For RNA expression experiments, normalization was performed as above by importing the output_fused.xls files into Genespring GX 7.3.1 for the 3 biological replicates of each isolate. However, no dye swap procedure was used for RNA expression experiments. Statistical analysis of RNA expression data generated by

Genespring GX 7.3.1 was performed using the statistical analysis (ANOVA) tool. This performs a 1-way parametric test without assuming variances are equal. The false discovery rate was set at 0.05 resulting in a false discovery rate of about 5% of genes. Multiple testing correction was performed using a Benjamini and Hochberg False Discovery Rate. No *post hoc* tests were used. Gene lists produced were then imported into Microsoft Excel for comparison to lists for other tested isolates.

2.10.7 Validation of microarray results

2.10.7.1 DNA microarray validation

Results of DNA microarrays were validated where required by PCR using primers used to design the probes of specific genes on the microarray slide. Genomic DNA from selected isolates was tested along with a positive control of TIGR4 gDNA to confirm the hybridization result generated by the DNA microarray. Genes validated and primers used can be found in Table 2.3. Phusion High fidelity polymerase enzyme (New England Biosciences, UK) was used in all cases, the following PCR conditions were used for every reaction:

PCR reaction conditions

- 98ºC for 30s
- 30x
    - 98ºC for 10s
    - 50ºC for 30s
    - 72ºC for 30s
- 72ºC for 5 min
- Stored at 4ºC

Gene number	Primer name	DNA sequence 5'-3'
SpTIGR4-0268	SpTIGR4-0268_f	GATGCCCATAAAGTGGCTAAAG
SpTIGR4-0268	SpTIGR4-0268_r	GCATAGGTTCTCCAACCTTCAC
SpTIGR4-0314	SpTIGR4-0314_f	TGAAACAGCACCAAATCGATAC
SpTIGR4-0314	SpTIGR4-0314_r	CCCAAACTCCTTGTTTAGCATC
SpTIGR4-0377	SpTIGR4-0377_f	TATGCTGTAAAGACAGGCTGGA
SpTIGR4-0377	SpTIGR4-0377_r	CTGTAACCATTGCTCCATTTGA
SpTIGR4-0461	SpTIGR4-0461_f	AATTTCCACTTGAGTTCCCTGA
SpTIGR4-0461	SpTIGR4-0461_r	AAGGCAATCGTCAAAAAGTGAT
SpTIGR4-0463	SpTIGR4-0463_f	AGAATGACTGAAGGTTTGGCAT
SpTIGR4-0463	SpTIGR4-0463_r	ATTACAAATTCTGCCCCAGCTA
SpTIGR4-0468	SpTIGR4-0468_f	CGGAGGGATATGAGGTCAATTA
SpTIGR4-0468	SpTIGR4-0468_r	TAAACGTGCTAGCTTCCACAAA
SpTIGR4-0530	SpTIGR4-0530_f	TGGCACAAAACCCTAATCTCTT
SpTIGR4-0530	SpTIGR4-0530_r	ATTAATATGACGGCGCAAGACT
SpTIGR4-0966	SpTIGR4-0966_f	AATCCTGCAAACCCAAGAACTA
SpTIGR4-0966	SpTIGR4-0966_r	TGTTTGAATCAATTCTTCACGG
SpTIGR4-0978	SpTIGR4-0978_f	ATGTTTGTTGCGAGAGATGCTA
SpTIGR4-0978	SpTIGR4-0978_r	GGTCGAATTTGTGGATACCATT
SpTIGR4-1154	SpTIGR4-1154_f	GCGTGATATTCGGAAACAATTT
SpTIGR4-1154	SpTIGR4-1154_r	CACGAAATTCTTTACTGAGGGG
SpTIGR4-1396	SpTIGR4-1396_f	TGGTTTTCCAACAACCTAATCC
SpTIGR4-1396	SpTIGR4-1396_r	AATTTTACCAGCTGAAATCGGA
SpTIGR4-1770	SpTIGR4-1770_f	TATTGCGTCAGAGTGGTTTTTG
SpTIGR4-1770	SpTIGR4-1770_r	CATGCTCCATCTCACAACTAGC
SpTIGR4-1771	SpTIGR4-1771_f	TTATGTGACCTTTGTGGACTCG
SpTIGR4-1771	SpTIGR4-1771_r	AATCCATTCATTTGGAAAATCG
SpTIGR4-1772	SpTIGR4-1772_f	CCTCAGCAAGTACAAGTGCATC
SpTIGR4-1772	SpTIGR4-1772_r	TAGCAGCGTAAGGGGTAAATGT
SpTIGR4-2190	SpTIGR4-2190_f	AGTCAGGCAGAACAAGGAGAAC
SpTIGR4-2190	SpTIGR4-2190_r	TGGAAGAGTCTGAACTTGACGA

Table 2.3: DNA microarray validation primers

Primers used to validate DNA microarray hybridizations by PCR. Gene numbers equate to the position of the gene in the annotated TIGR4 reference genome used in the arrays.

2.10.7.2 RNA microarray validation

Results of RNA microarrays were validated where required by real-time RT-PCR to confirm differences in gene expression levels between selected isolates. Firstly, cDNA was synthesised from the purified RNA as follows. Any residual DNA contamination was removed by incubation of sample with Turbo DNA-free™ (Ambion, UK) as per manufacturer's instructions. 2µg of RNA, 2µl random hexamers (Invitrogen, Scotland) and 1µl RNase OUT (Invitrogen, Scotland) were then mixed with RNase-free water to a final volume of 17.5µl before denaturation at 70°C for 10 min in a heat block and storage on ice. A 12.5µl mixture containing 6µl 5x first strand buffer, 3µl 0.1M DTT, 1.5µl 10mM dNTP and 2µl Superscript III reverse transcriptase (Invitrogen, Scotland) was added before incubation at 42°C for 16h. The reaction was then inactivated by denaturation at 70°C for 15 min before addition of 1µl *E. coli* RNase H (Invitrogen, Scotland) and a further incubation at 37°C for 20 min to remove residual RNA from the sample. The cDNA concentration was then measured using a nanodrop ND 1000 (Thermo Scientific, UK)

Real-time PCR was then carried out using SYBR green (Roche, UK) according to manufacturer's instructions in a LightCycler® 480 System (Roche, UK). Primers used to validate gene expression levels can be found in Table 4. Results were analysed using a previously described method (Muller *et al.*, 2002), and using Qgene 4.0 (Qgene, USA). The control gene, to which expression of test genes was compared, was *gyrA*, a pneumococcal house-keeping gene (Enright and Spratt, 1998).

Gene	Gene number	Primers	DNA sequence 5'-3'
gyrA	SpTIGR4-1219	SpTIGR4-1219_f	CAAGGTGCATGAGCATATTGTT
gyrA	SpTIGR4-1219	SpTIGR4-1219_r	GATCCAAACGCTTAATGTAGCC
gdhA	SpTIGR4-1306	SpTIGR4-1306_f	GGTATCGACTTCGATCTTTTGG
gdhA	SpTIGR4-1306	SpTIGR4-1306_r	TCATGGCATTTGCTACATTTTC
guaA	SpTIGR4-2072	SpTIGR4-2072_f	TCATTCCTGTTGGTGATGAGTC
guaA	SpTIGR4-2072	SpTIGR4-2072_r	TTAATGCTTTGACGATGGACTG

Table 2.4: RNA microarray validation primers

Primers used to validate RNA microarray hybridizations by real-time RT-PCR. Gene numbers equate to the position of the gene in the annotated TIGR4 reference genome used in the arrays.

2.10.8 Growth curves of S. pneumoniae in ammonium.

The growth of *S. pneumoniae* isolates in differing concentrations of ammonium (NH_4Cl_2, Sigma-Alrdich, UK) was investigated. BHI with ammonium concentrations ranging from 1-1000µg/ml (0, 1, 10, 50, 200, 500, 1000) was prepared in a 96-well plate (Costar, USA), and seeded with 2µl of overnight culture. Growth of bacteria was monitored by reading OD600 every 1800s for 24 reads in a FLUOstar Optima plate reader (BMG Labtech, UK), allowing plotting of growth curves in GraphPad Prism 4 (GraphPad Software, USA).

3 Diversity studies of the virulence factor Ply

3.1 Summary

Ply has long been known to be an important virulence factor of the pathogen *S. pneumoniae,* and is being considered as a potential component of next-generation protein vaccines. The aim of this chapter was to assess the diversity present in the Ply protein, and relate the uncovered differences to the ability of clones to cause invasive pneumococcal disease. Furthermore, the importance of Ply in invasive pneumococcal disease was explored. This study of diversity present in the Ply toxin will uncover important information in the context of vaccine design as well as giving insights into pneumococcal biology and pathogenesis.

Ply is a cytoplasmic toxin, thought to be released upon autolysis of the pneumococcal cell. As such, it was thought to be a highly conserved protein (Mitchell *et al.*, 1990), although a number of variants have previously been identified (Kirkham *et al.*, 2006a; Lock *et al.*, 1996; Tettelin *et al.*, 2001; Walker *et al.*, 1987). This work reports the presence of a further nine Ply protein alleles after a large-scale screen of clinical isolates. These Ply alleles display varying levels of haemolytic activity as a result of mutations present within the *ply* gene. Interestingly, a fully non-haemolytic variant first identified by Kirkham et al. (Kirkham *et al.*, 2006a) (allele 5) was found to be widespread within serotype 1 and serotype 8 clones, present in the founder clones of two main clonal complexes. These two clones, ST306 and ST53, were shown to be either the established dominant clone (ST53) or undergoing stark clonal expansion (ST306) (Kirkham *et al.*, 2006a) in the Scottish IPD population. Since these clones are unrelated, and both appear to be dominant within their serotypes, it was hypothesised that the presence of this non-haemolytic Ply variant, in certain genetic backgrounds, may play a role in driving expansion of these clones.

Further identification of a clinical isolate possessing an insertion sequence in the *ply* gene, resulting in no Ply functional activity, shows for the first time that pneumococci can cause IPD without expressing a functional Ply protein. This, together with the discovery of the non-haemolytic allele, implies that the haemolytic properties of Ply may be less important in establishing IPD in certain genetic backgrounds.

An unmarked Ply knock-out was created in a D39 background using the Janus technique (Sung *et al.*, 2001). The advantage of this technique is that the mutations are made without insertion of an antibiotic cassette into the target gene. Initial results show that this mutant is less virulent than the wild-type parent strain, in agreement with previous data (Berry *et al.*, 1989b; Wellmer *et al.*, 2002; Winter *et al.*, 1997), although further studies are required. Furthermore, this knock-out is a useful tool for further study, and is now routinely used by our laboratory in Ply studies.

3.2 Diversity of Ply in *S. pneumoniae* clinical isolates

3.2.1 Screen of S. pneumoniae clinical isolates for Ply diversity

The pneumococcal toxin Ply is an important virulence factor, and is being considered as a vaccine candidate for next generation protein vaccines. Uncovering the distribution and diversity of alleles of this toxin within the pneumococcal population can not only further understanding of pneumococcal biology and the relationships between clones, but also give insight into the suitability of this protein as a vaccine candidate. The *ply* gene from 121 clinical isolates was sequenced (work partly done by Dr Johanna Jefferies [~50%]) and aligned to the sequences from readily available sequenced *S. pneumoniae* genomes using Vector NTI™ software (See Appendix I). The wild-type *ply* gene was taken from the strain D39 (Walker *et al.*, 1987). From this alignment, 22 isolates were selected for further study, possessing protein alleles differing from the wild-type D39 sequence. The information on these strains can be found in Table 3.1. Variation in the DNA sequence occurred at 37 positions in the *ply* gene of these isolates. The DNA sequences were translated into amino acid sequence, and aligned using Vector NTI™ software. This showed that these 37 mutations resulted in 15 amino acid changes across the *ply* gene, indicating that a number of synonymous mutations were present. Furthermore, a number of alleles possessed a double amino acid deletion at position 270-271, whilst 2 Ply alleles possessed regions of insertion. A summary of mutations present in these alleles can be found in Figure 3.1. Five of the Ply alleles discovered had been documented previously, namely alleles 1, 2, 3, 4, and 5 (Kirkham *et al.*, 2006a; Lock *et al.*, 1996; Tettelin *et al.*, 2001; Walker *et al.*, 1987). Another allele, termed Ply7 by authors, was found previously in serotype 7F isolates but was not seen in this study (Lock *et al.*, 1996). The Ply assay developed was used to calculate the specific activity of the Ply proteins where possible. The plots of mean haemolytic activity produced can be found in Figure 3.2, with specific activities calculated from these plots noted in Figure 3.1.

Strain	Serotype	Sequence Type (ST)	Ply amino acid allele	Source
D39	2	128	1	(Walker *et al.*, 1987)
TIGR4	4	205	2	(Tettelin *et al.*, 2001)
01-2696	1	227	2	(Kirkham *et al.*, 2006a)
2PN00495	8	404	3	HPA
H040920498	8	944	3	HPA
00-3645	1	227	4	(Kirkham *et al.*, 2006a)
01-2884	8	53	5	SMPRL
01-1204	8	578	5	SMPRL
01-1956	1	306	5	SMPRL
01-1199	NT	577	5	SMPRL
04-2055	1	228	5	SMPRL
96-5878	2	74	6	SMPRL
02-3013	NT	448	7	SMPRL
01-2866	23F	40	8	SMPRL
00-1153	9V	156	9	SMPRL
01-3862	7F	191	10	SMPRL
02-2744	7F	191	10	SMPRL
01-4296	27	571	10	SMPRL
01-2914	20	591	11	SMPRL
01-2513	18C	818	12	SMPRL
00-2328	6A	813	13	SMPRL
S1-11	1	228	14	(Brueggemann and Spratt, 2003)

Table 3.1: Ply diversity study isolates

Isolates selected for further study of Ply diversity after initial sequencing study, with serotype, sequence type and Ply amino acid allele noted. HPA = Health Protection Agency, Colindale, London. SMPRL = Scottish Meningococcal and Pneumococcal Reference Laboratory, Stobhill Hospital, Glasgow.

					Amino acid position																
					14	136	142	150	154	167	172	224	260	265	267	270	271	273	380	402	415
Strain no.	Serotype	ST	Allele	Specific activty	N	Q		Y	T	S	T	K	E	A	I	V	K	A	D	Q	
D39	2	128	1	$4.13x10^5$																	
TIGR4	4	205	2	$4.74x10^5$															N		
01-2696	1	227	2	$4.20x10^5$															N		
2PN00495	8	404	3	$6.50x10^3$							I	R		S		DEL	DEL				
H040920498	8	944	3	$7.43x10^3$							I	R		S		DEL	DEL				
00-3645	1	227	4	--------															N		INS
01-2884	8	53	5	0				H			I	R		S		DEL	DEL				
01-1204	8	578	5	0				H			I	R		S		DEL	DEL				
01-1956	1	306	5	0				H			I	R		S		DEL	DEL				
01-1199	NT	577	5	0				H			I	R		S		DEL	DEL				
04-2055	1	228	5	0				H			I	R		S		DEL	DEL				
96-5878	2	74	6	$3.72x10^3$						F	I	R		S		DEL	DEL				
02-3013	NT	448	7	$2.95x10^5$	D							R									
01-2866	23F	40	8	$9.09x10^4$														D			
00-1153	9V	156	9	$3.14x10^5$											M						
01-3862	7F	191	10	$1.01x10^5$					M					S		DEL	DEL		N		
02-2744	7F	191	10	$9.77x10^4$					M					S		DEL	DEL		N		
01-4296	27	571	10	$1.36x10^5$					M					S		DEL	DEL		N		
01-2914	20	591	11	$5.95x10^5$		K													N		
01-2513	18C	818	12	--------															N	E	
00-2328	6A	813	13	$3.09x10^5$									D						N		
S1-11	1	228	14	--------			INS												N		

Figure 3.1: Ply amino acid alleles

Ply amino acid alleles determined by sequencing of ply genes from 22 isolates. The numbers in row 1 refer to wild-type amino acid positions, with the identity of the wild-type amino acid shown in row 2. The mutated amino acids present in specific strains are highlighted in colour below the corresponding wild-type amino acid. Strains highlighted in green are non-haemolytic and as such have a specific activity of 0. Insertion regions (INS) are highlighted in gray. Deletion of an amino acid is represented by the abbreviation DEL. Specific activity is shown in haemolytic units per milligram, as calculated using the Ply assay described. The lysate of strain 01-2513 (allele 12) was haemolytic but was not recognized by the monoclonal antibody PLY-7; therefore, no specific activity could be determined. The lysate of strain 00-3645 (allele 4) was haemolytic but at a level that did not allow calculation of specific activity. The lysate of strain S1-11 was non-haemolytic, due to the presence of an insertion sequence (IS1515) in the *ply* gene. No haemolytic activity was therefore calculated. These strains are highlighted in yellow. ST = sequence type.

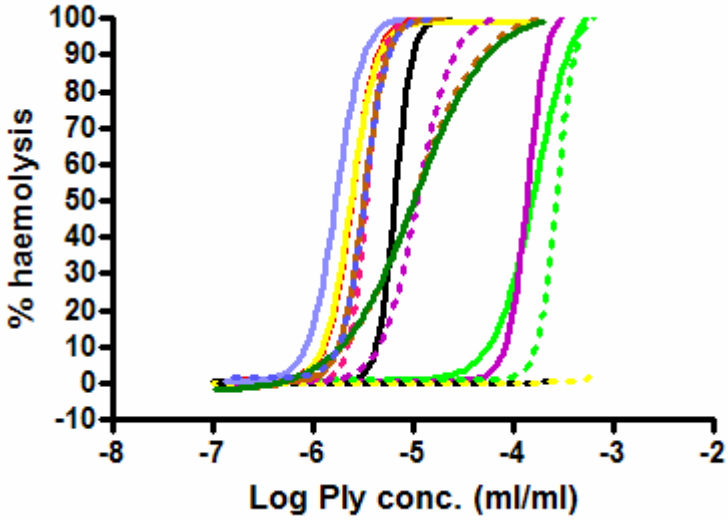


Figure 3.2: Plot of haemolytic activity of Ply amino acid alleles

Plot represents percentage haemolysis of sheep erythrocytes against log Ply concentration (mg/ml) in tested lysates. Percentage haemolysis was calculated from the haemolytic assay, whilst Ply concentration was measured by the Ply sandwich ELISA. Specific activities of Ply alleles were calculated from these curves. Purified Ply allele 1 was included as a control (black solid line). In the brackets, Ply amino acid allele and line colour are denoted. Test lysates are represented by the following lines: D39 (1 - red solid), TIGR4 (2 - blue solid), 01-2696 (2 - yellow solid), 2PN00495 (3 - green solid), H040920498 (3 - purple solid), 01-2884 (5 - pink solid), 01-1204 (5 - black dashed), 01-1956 (5 - red dashed), 01-1199 (5 - blue dashed), 04-2055 (5 - yellow dashed), 96-5878 (6 - green dashed), 02-3013 (7 - pink dashed), 01-2688 (8 - purple dashed), 00-1153 (9 - brown solid), 01-3862 (10 - brown dashed), 02-2744 (10 - dark green solid), 01-4296 (10 - dark green dashed), 01-2914 (11 - light blue solid), 00-2328 (13 - light blue dashed).

3.2.2 Comparing sequence and activity of Ply alleles

Ply amino acid alleles were aligned and numbered, with the wild-type allele (to which all others were compared) being allele 1 (Figure 3.1). This allele was expressed by the strain D39, and was encoded by the first *ply* gene to be fully sequenced (Walker *et al.*, 1987). This allele displayed a specific activity of $4.13x10^5$ HU/mg.

Allele 2, which was present in reference isolate TIGR4 (Tettelin *et al.*, 2001), possessed only one amino acid mutation in comparison to the wild-type allele – the substitution of an aspartic acid (D) residue with an asparagine (N) at position 380, represented by D380N. Further mutations were given the same nomenclature. An isolate of serotype 1, ST227 was also found to possess this allele. Allele 2 displayed similar specific activity to wild-type allele 1. These were by far the most predominant alleles uncovered in the initial screen, with 23/121 isolates possessed allele 1, and 69/121 allele 2.

Allele 3 was discovered in two serotype 8 isolates, and was identical to the Ply8 allele found previously in this serotype (Lock *et al.*, 1996). This allele had three amino acid substitutions (T172I, K224R, A265S) and a two amino acid deletion (270-271) in comparison to the wild-type (Figure 3.1), and showed a significant reduction in specific activity (~60-fold).

Allele 4, which contains an 8 amino acid insertion/duplication, has been previously reported (Kirkham *et al.*, 2006a). This allele was found in a single serotype 1, ST227 isolate, and displayed very low levels of haemolytic activity when tested, making accurate calculation of specific activity impossible.

Allele 5 was initially identified from research done in our laboratory in a number of serotype 1, ST306 isolates (Kirkham *et al.*, 2006a). This allele is similar to allele 3, with an extra mutation, Y150H. However, unlike allele 3, this allele produces a non-haemolytic variant of the Ply toxin. In this initial study, allele 5 was found in two serotype 1 isolates of ST306 and ST228, and two serotype 8 isolates of ST53 and ST578, as well as a non-typable ST577 isolate. This is the first demonstration of this non-haemolytic toxin in clones other than ST306, showing it to be more widespread than previously thought.

Allele 6, found in a serotype 2, ST74 isolate, also had similar mutations to those seen in allele 3, but had a further S167F mutation. This allele also had reduced haemolytic activity, resulting in a specific activity comparable to allele 3.

Allele 7, present in a non-typable isolate of ST448, possessed 2 mutations in comparison to wild-type allele 1, namely N14D and K224R, and had a specific activity of similar level to the wild-type allele.

Allele 8, present in a serotype 23F, ST40 isolate, had only a single mutation (A273D), and displayed slightly reduced haemolytic activity, resulting in a ~5-fold reduction in specific activity.

Allele 9, which was found in a serotype 9V, ST156 isolate, also had only a single mutation (I267M). However, this mutation had no effect on the specific activity of the Ply protein.

Allele 10 possessed 3 amino acid substitutions (T154M, A265S, D380N) and 2 amino acid deletions (270-271), and showed reduced haemolytic activity, resulting in a ~5-fold reduction in specific activity.

Allele 11 was found in an isolate of serotype 20, ST591, and although it possessed the amino acid mutations Q136K and D380N, had a specific activity comparable to the wild-type allele.

Allele 12, discovered in a serotype 18C, ST8181 isolate, had 2 mutations in comparison to the wild-type allele 1. These mutations were D380N, as in allele 2, and Q402E. Although this allele was haemolytic, it was not possible to calculate the specific activity of the Ply as the protein was not recognised by the monoclonal antibody PLY-7 (mAbPLY-7) (de los Toyos *et al.*, 1996), which was used in the sandwich ELISA, and recognises AA401-407 of the Ply protein. Production of recogniseable Ply was confirmed by Western blot using either mAbPLY-7 or polyclonal α–Ply antibody raised in rabbit as the primary antibody (Figure 3.3). This result showed that although a functional Ply protein is produced and recognised by the polyclonal antibody, it is not recognised by the PLY-7 monoclonal antibody.

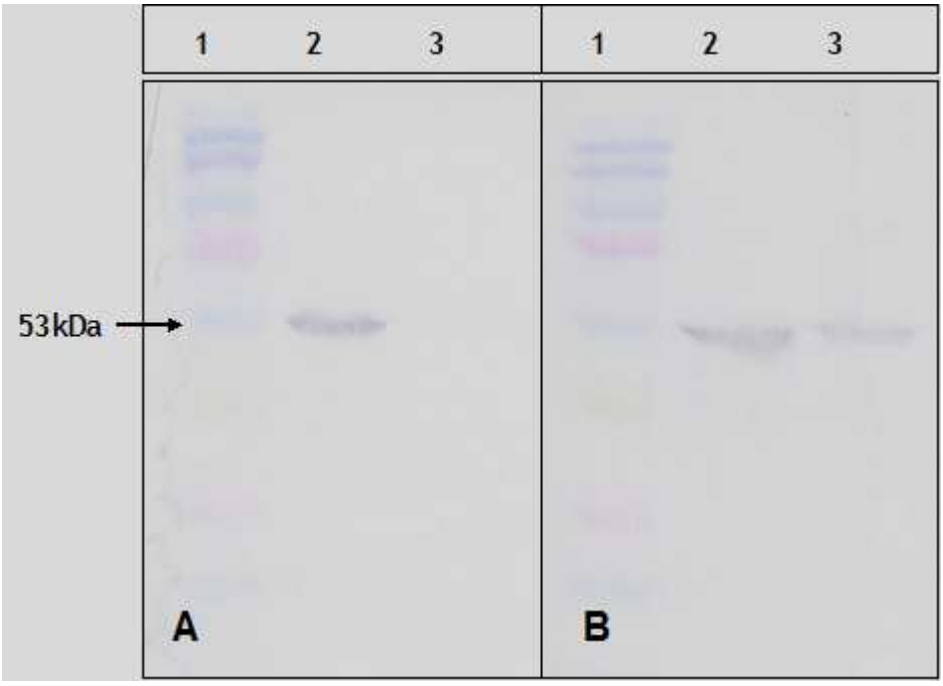


Figure 3.3: Western blot – anti-Ply – isolate 01-2513

Poly and monoclonal α-Ply Western blots of lysate from isolate 01-2513, possessing Ply allele 12. Blot A is probed with mAbPLY-7 was a primary antibody, and blot B is probed with polyclonal α–Ply antibody raised in rabbit as a primary antibody. In lane 1 of each blot is the Kaleidoscope protein marker, with the band corresponding to 53kDa highlighted. In lane 2 is the lysate of wild-type strain D39. In lane 3 is the lysate of 01-2513.

Allele 13, present in a serotype 6A, ST813 isolate, possessed the D380N mutation as in allele 2, as well as an E260D mutation. These mutations had no effect on the specific activity in comparison to wild-type.

Allele 14, present in a serotype 1 isolate of ST228, was found to be non-haemolytic. PCR amplification of the gene using primers 27R and 27S showed a PCR product of 3282bp in size, substantially larger than the wild-type gene from D39 (2411bp) (Figure 3.4). Sequencing of this PCR product revealed an insertion of 871bp at amino acid 142 in the *ply* gene, which was found to represent the mobile genetic element IS1515 (Munoz *et al.*, 1998). IS1515 was able to insert into the *ply* gene due to the presence of a recognition site within the gene, with the IS inserting in the orientation shown in Figure 3.5, and the underlying Ply being allele 2. Furthermore, the IS was found to have produced a 3bp repeat upon insertion, as expected (Munoz *et al.*, 1998). Although IS1515 is inserted in the opposite orientation to the *ply* gene, the full sequence is present, and therefore the insertion element should in theory be able to excise from the *ply* gene. Western blots with mAbPLY-7 and polyclonal α–Ply antibody raised in rabbit confirmed that no recognisable Ply protein was produced by this isolate,

due to the disruption of the *ply* gene (Figure 3.6). This is the first demonstration of a clinical isolate of *S. pneumoniae* expressing no recogniseable Ply protein.

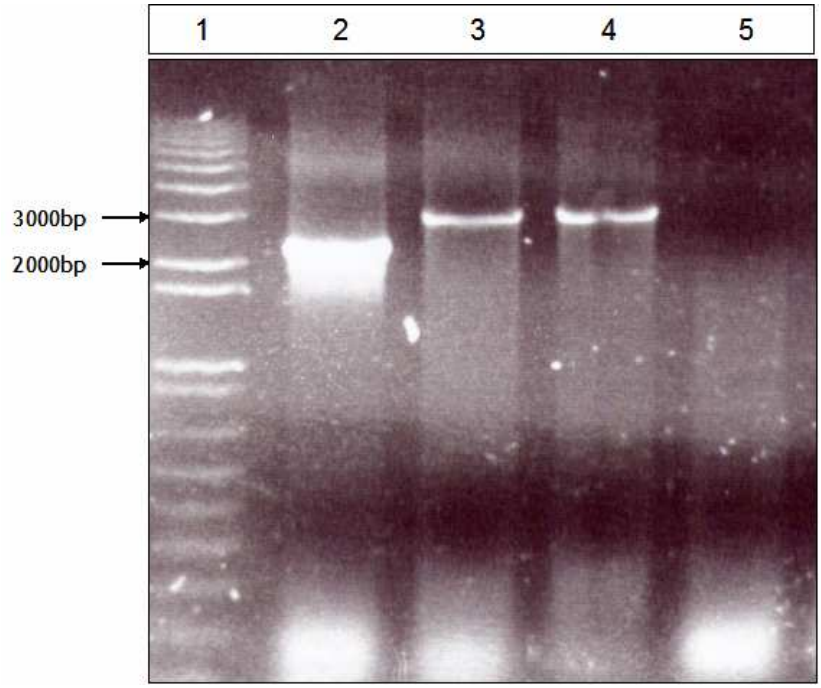


Figure 3.4: PCR of *ply* gene of isolate S1-11

Gel shows gDNA of wild-type strain D39 and serotype 1, ST228 strain S1-11. Samples are as follows: 1 – 1kb+ DNA ladder; 2 - D39 (2411bp); 3 – S1-11 (3282bp); 4 – S1-11 (3282bp); 5 – negative control (no gDNA).

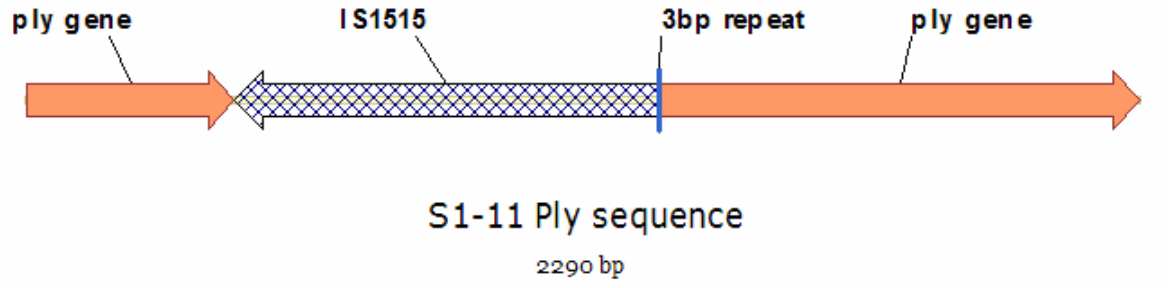


Figure 3.5: IS1515 in *ply* gene of S1-11

Schematic of IS1515 inserted into the p*ly* gene sequence from the isolate S1-11. The 871bp insertion in the *ply* gene is in the opposite orientation. Figure constructed using Vector NTI™ software (Invitrogen, Scotland).

Figure 3.6: Western blot – anti-Ply – isolate S1-11

Poly- and monoclonal α-Ply **Western blots of lysates from isolate S1-11 to test for expression of recogniseable Ply protein. The primary antibody in blot A was polyclonal α–Ply antibody raised in rabbit, whilst in blot B it was mAbPLY-7. In both blots, lane 1 was a kaleidoscope protein marker, lane 2 was the lysate from wild-type D39, lane 3 was S1-11 lysate, and lane 4 was the lysate from a knockout of Ply in a D39 background (D39 ΔPly stop – see 3.5).**

3.2.3 Western blot analysis of 14 Ply alleles

One representative from each of the 14 Ply alleles was selected for comparison using a Western blot probed with the polyclonal α-Ply antibody. The results, displayed in Figure 3.7, showed that most protein alleles were expressed and a toxin of the correct size was produced. However, a number of interesting results were observed. Firstly, alleles 3, 5 and 10 showed reduced mobility when run on the SDS-PAGE gel. This resulted in the Ply protein on the Western blot appearing slightly larger in these isolates than in the other isolates, circled in red in Figure 3.7. This result agreed with previous findings where a similar phenomenon was observed with allele 3 (Ply8 in their paper) (Lock *et al.*, 1996). Secondly, the lysate of the S1-11 isolate, which had the IS1515 insertion in it, appeared to express some recognisable protein of the correct size (circled in green). This may represent excision of the IS1515 insertion element from the *ply* gene at some point during growth of the bacteria. The weaker bands on the blot, at different sizes, were a result of over-exposure of the blot, and can be ignored.

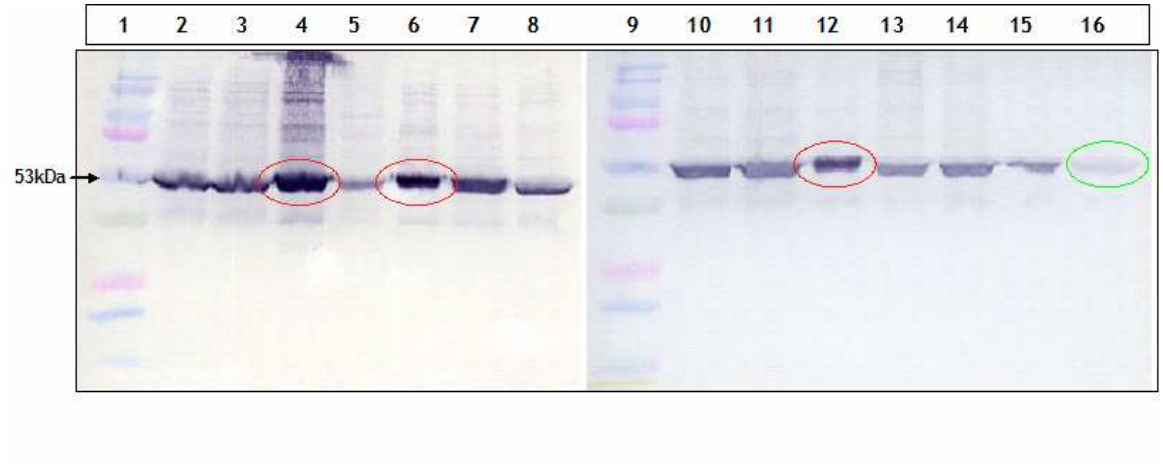


Figure 3.7: Western blot – polyclonal anti-Ply - 14 Ply alleles

Polyclonal anti-Ply Western blot of lysates possessing 14 different Ply alleles. Primary antibody used in the Western blot was polyclonal α–Ply antibody raised in rabbit. The lanes in the blots represent the following: lane 1 – Kaleidoscope protein marker; lane 2 – allele 1; lane 3 – allele 2; lane 4 – allele 3; lane 5 – allele 4; lane 6 – allele 5; lane 7 – allele 6; lane 8 – allele 7; lane 9 – kaleidoscope protein marker; lane 10 – allele 8; lane 11 – allele 9; lane 12 – allele 10; lane 13 – allele 11; lane 14 – allele 12; lane 15 – allele 13; lane 16 – allele 14.

3.2.4 Variation in CDCs at mutated positions in Ply protein

An alignment of the amino acid sequences of all known CDCs was created by Dr Graeme Cowan. The alignment included the *ply* allele 1 from D39 as a *S. pneumoniae* representative. From this alignment, values of variability were calculated for each amino acid position, using the Kabat method (Wu *et al.*, 1975), where a value of 1 showed that the amino acid was fully conserved across the CDCs, whilst a value of 2-5 indicated the amino acid was highly conserved, 5-10 equated to weakly conserved and >10 to highly variable. The variability values for each mutated amino acid position found in the study of Ply alleles were determined (Table 3.2). The results showed that a wide range of variability is present in these positions. Only one of these amino acids was fully conserved across all CDCs in the alignment, the tyrosine residue at position 150 (in red). In Ply, mutation of this amino acid to a histidine is unique to allele 5, which displays no haemolytic activity. Furthermore, the presence of this mutation is the only difference between alleles 3 and 5, implying that although allele 3 shows reduced haemolytic activity, the Y150H mutation is responsible for the full abrogation of haemolytic activity in allele 5.

Mutation	Alleles with mutation	Variability value
N14D	7	28
Q136K	11	12
Y150H	5	1
T154M	10	3.6
S167F	6	14.4
T172I	3, 5, 6	8
K224R	3, 5, 6, 7	15
E260D	3, 5, 6	5.14
A265S	3, 5, 6, 10	5.3
I267M	9	12
V270DEL	3, 5, 6, 10	28
K271DEL	3, 5, 6, 10	9.6
A273D	8	15
D380N	2, 4, 10, 11, 12, 13, 14	2.4
Q402E	12	15

Table 3.2: Variability of mutations in the Ply amino acid sequence

Mutations discovered in Ply protein, alleles that possess these mutations, and the variability present at the amino acid positions as calculated by Kabat method. A variability value of 1 indicated 100% conservation of an amino acid at a certain position. 2-5 = highly conserved, 5-10 = weakly conserved, >10 = highly variable. Highlighted in red is Y150H, a mutation at a position fully conserved across all CDCs.

3.3 Distribution of non-haemolytic Ply allele 5 in the pneumococcal population

3.3.1 eBURST analysis of *S. pneumoniae serotype 1 and 8 populations*

As the non-haemolytic Ply allele 5 was discovered in isolates of both serotypes 1 and 8, it was of interest to determine the distribution of this mutant toxin in the populations of serotype 1 and 8 isolates present in the MLST database. In order to initially assess the overall clonal diversity present in each of these serotypes, eBURST diagrams representing all serotype 1 and 8 isolates in the MLST database were constructed.

A search of the MLST database for serotype 1 pneumococci showed 51 different STs possessing this serotype. eBURST revealed the presence of 4 major clusters (Figure 3.8), representing the 3 major geographical lineages described previously (Brueggemann and Spratt, 2003). Lineage A, representing serotype 1 pneumococci from Europe and North America, is split into 2 clonal groups by eBURST. The predicted founders of these complexes were ST306 and ST305,

represented by blue dots in Figure 3.8. ST217 was the predicted founder of Lineage B, representing serotype 1 pneumococci from Africa and Israel, and was also the predicted founder of the overall serotype 1 pneumococcal population, as it possessed the highest number of single locus variants (SLVs). ST2296 was the predicted founder of Lineage C. At the time of the study by Brueggemann et al., isolates from this lineage were predominantly from Chile (Brueggemann and Spratt, 2003). However, this lineage now contains isolates from a wide range of geographical locations.

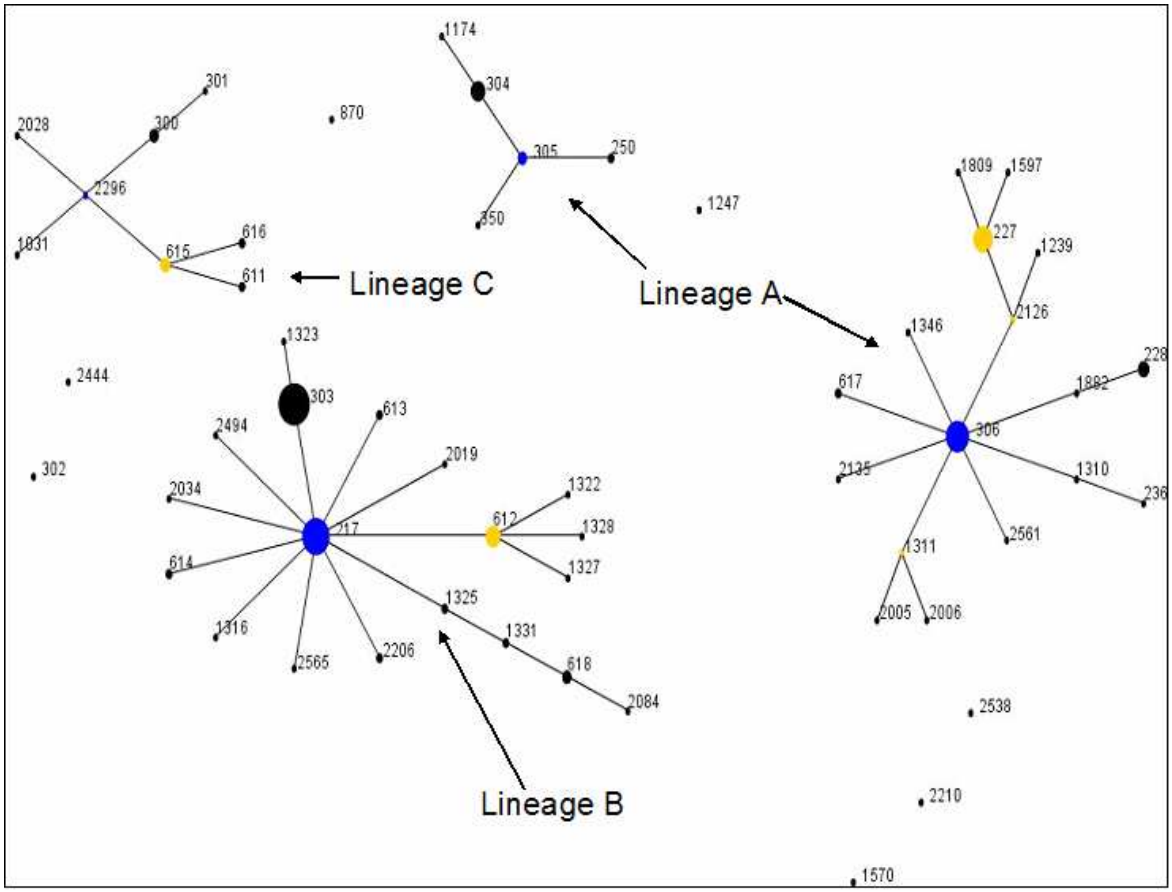


Figure 3.8: eBURST diagram of serotype 1 pneumococcal isolates

A search of the MLST database revealed 43 different STs present in the serotype 8 pneumococcal population. eBURST showed that there was one major clonal complex present with predicted founder ST53, two smaller complexes with predicted founders ST944 and ST404, and a large number of singletons (Figure 3.9).

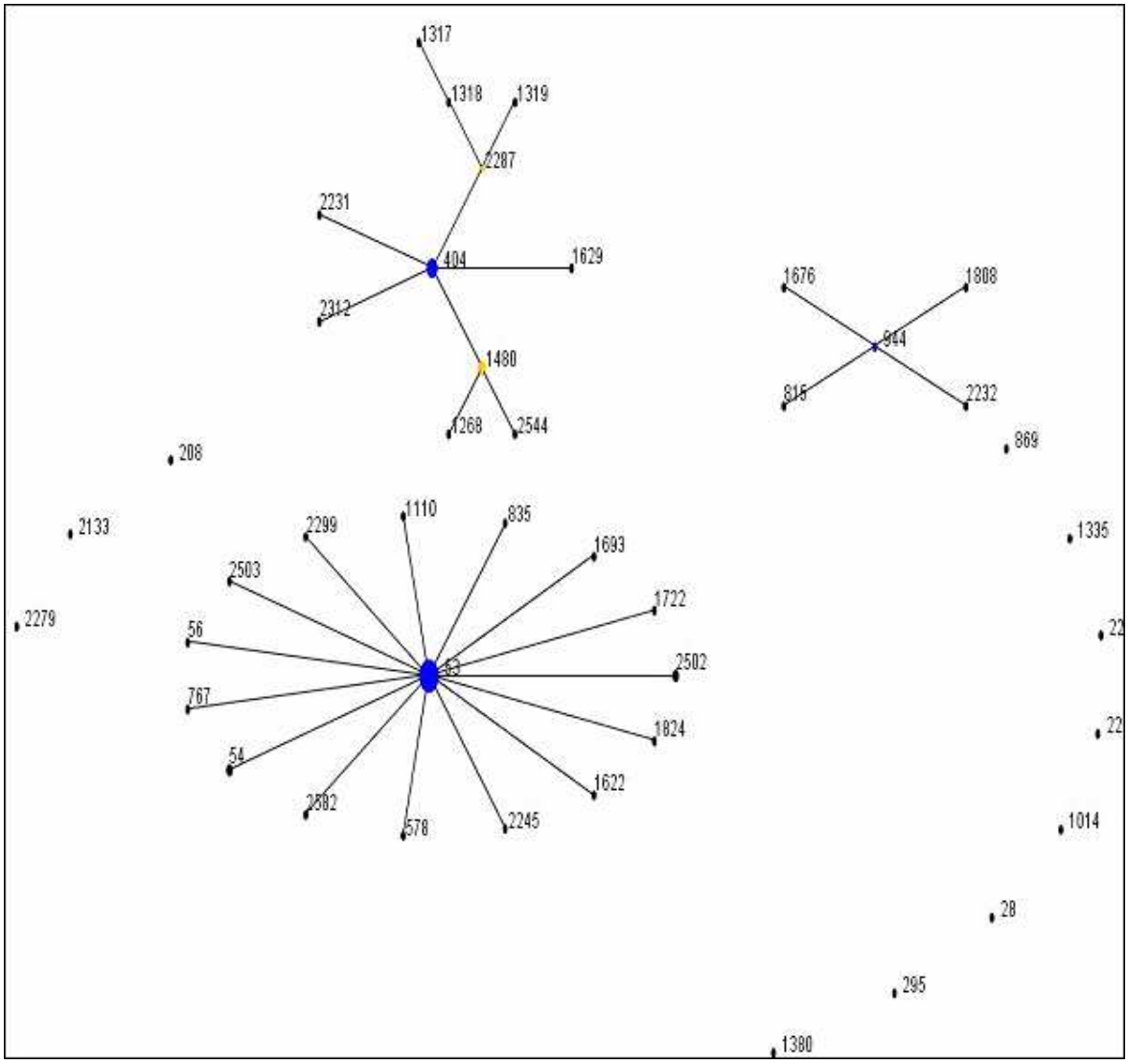


Figure 3.9: eBURST diagram of serotype 8 pneumococcal isolates

3.3.2 Distribution of Ply alleles within serotype 1 and 8 pneumococcal populations, distinguished by ST

With the diversity of pneumococci present in these populations uncovered, it was possible to determine the distribution of the non-haemolytic Ply allele 5 within these populations. 31 serotype 1 isolates were collected from a number of studies (Brueggemann and Spratt, 2003; Kirkham *et al.*, 2006a; Leimkugel *et al.*, 2005) representing 24 of the 51 serotype 1 STs in the database. Fewer serotype 8 isolates were available, but 7 isolates were collected and the Ply allele present in each isolate determined (Table 3.3).

Strain	Serotype	ST	Ply amino acid allele	Source
S1-46	1	217	1	(Brueggemann and Spratt, 2003)
P1041	1	217	2	(Leimkugel *et al.*, 2005)
01-2696	1	227	2	(Kirkham *et al.*, 2006a)
S1-2	1	227	2	(Brueggemann and Spratt, 2003)
00-3645	1	227	4	(Kirkham *et al.*, 2006a)
04-2055	1	228	5	SMPRL
S1-11	1	228	14	(Brueggemann and Spratt, 2003)
INV1871	1	300	2	(Brueggemann and Spratt, 2003)
P1039	1	303	1	(Leimkugel *et al.*, 2005)
S1-8	1	303	2	(Brueggemann and Spratt, 2003)
S1-4	1	304	2	(Brueggemann and Spratt, 2003)
S1-30	1	305	2	(Brueggemann and Spratt, 2003)
S1-3	1	306	5	(Brueggemann and Spratt, 2003)
01-1956	1	306	5	(Kirkham *et al.*, 2006a)
S1-71	1	611	2	(Brueggemann and Spratt, 2003)
S1-45	1	612	2	(Brueggemann and Spratt, 2003)
P1021	1	612	2	(Leimkugel *et al.*, 2005)
S1-125	1	613	1	(Brueggemann and Spratt, 2003)
S1-126	1	614	1	(Brueggemann and Spratt, 2003)
NCTC7465	1	615	2	(Brueggemann and Spratt, 2003)
S1-102	1	616	2	(Brueggemann and Spratt, 2003)
S1-38	1	617	5	(Brueggemann and Spratt, 2003)
S1-99	1	618	2	(Brueggemann and Spratt, 2003)
03-5343	1	1239	5[a]	SMPRL
03-5340	1	1310	5[a]	SMPRL
04-1259	1	1311	5[a]	SMPRL
04-1837	1	1346	5[a]	SMPRL
04-2889	1	1597	5[a]	SMPRL
05-1635	1	1809	5[a]	SMPRL
05-1934	1	1882	5[a]	SMPRL
01-2117	1	2126	5[a]	SMPRL
01-2884	8	53	5	SMPRL
2PN00495	8	404	3	HPA
01-1204	8	578	5	SMPRL
03-2331	8	835	5[a]	SMPRL
H040920498	8	944	3	HPA
03-2620	8	1110	5[a]	SMPRL
H043900039	8	1722	5[a]	HPA

Table 3.3: Isolates used in serotype 1 & 8 Ply study

Serotype, ST, and *ply* allele were included, as well as the source of each isolate used. SMPRL = Scottish Meningococcal and Pneumococcal Reference Laboratory, Stobhill Hospital, Glasgow, HPA = Health Protection Agency, Colindale, London. [a] *ply* gene partially sequenced to confirm presence of *ply* allele 5.

These results allowed the mapping of Ply alleles onto the eBURST diagrams for each population set, giving insights into the distribution of Ply alleles within these populations. When Ply allele identity was mapped onto the serotype 1 eBURST diagram, it was found that wild-type allele 1 was associated with 4 STs and allele 2 was associated with 12 STs (Figure 3.10). Interestingly, allele 5 was associated with 9 STs, all of which were either single- or double-locus variants (SLV/DLV) of ST306. Indeed, all tested SLVs of ST306 were found to possess this non-haemolytic toxin. This is the first demonstration of allele 5 in serotype 1 isolates other than ST306 where it was discovered. The results also showed that ST217, the predicted founder of Lineage B, was associated with alleles 1 and 2. Similarly, in Lineage A, ST228 was associated with 2 Ply alleles (5 and 14), as was ST227 (2 and 4). This is the first demonstration of intra-ST amino acid variation within Ply. Results of haemolytic assays performed on these isolates were in agreement with the haemolytic profiles previously determined for these alleles (data not shown).

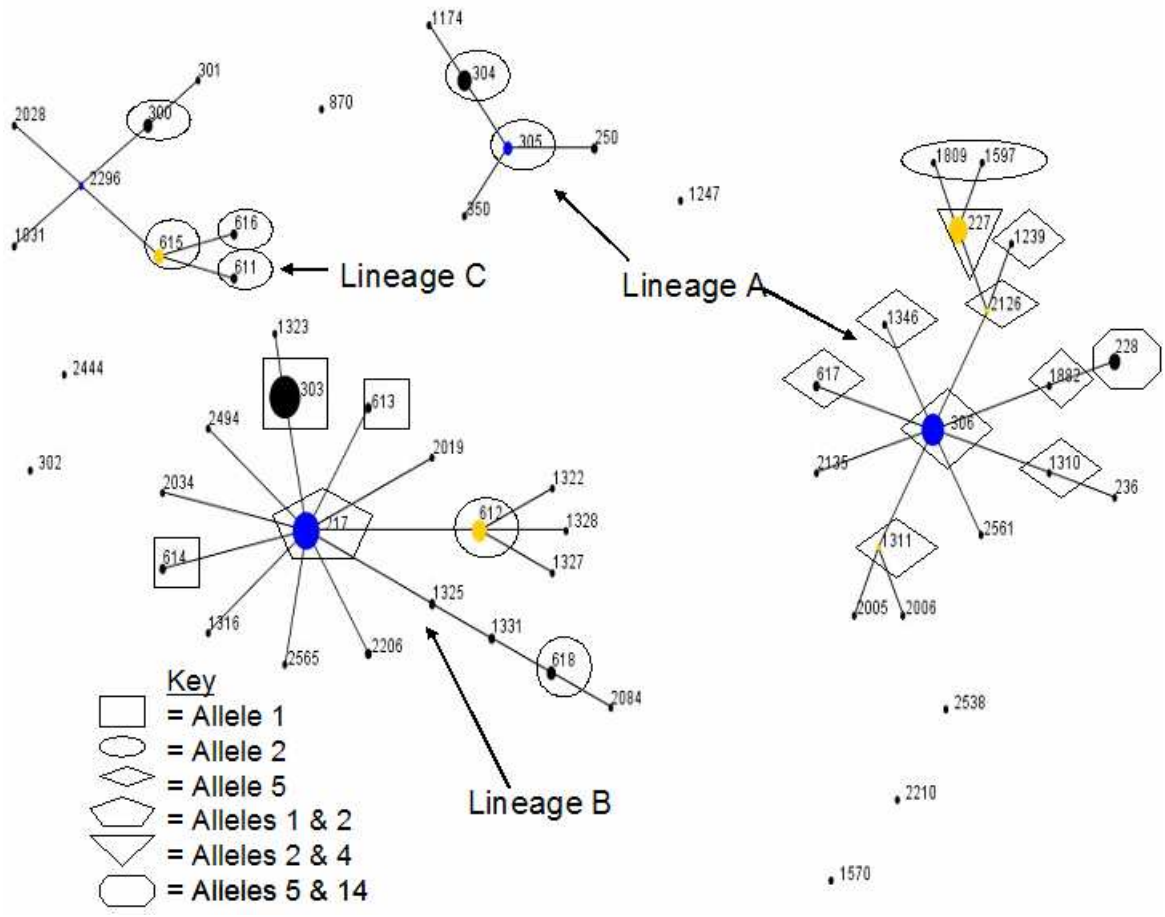


Figure 3.10: Ply amino acid allele identity of serotype 1 clones

Ply amino acid allele identity mapped onto serotype 1 eBURST diagram. Each allele or combination of alleles is represented by a different shape, as detailed in the key.

When Ply amino acid allele identity was mapped onto the serotype 8 eBURST diagram (Figure 3.11), all SLVs of ST53 tested possessed allele 5, whilst predicted founders of the two smaller clonal complexes had allele 3. This implies that allele 5 is only harboured by isolates in the ST53 clonal complex in this population.

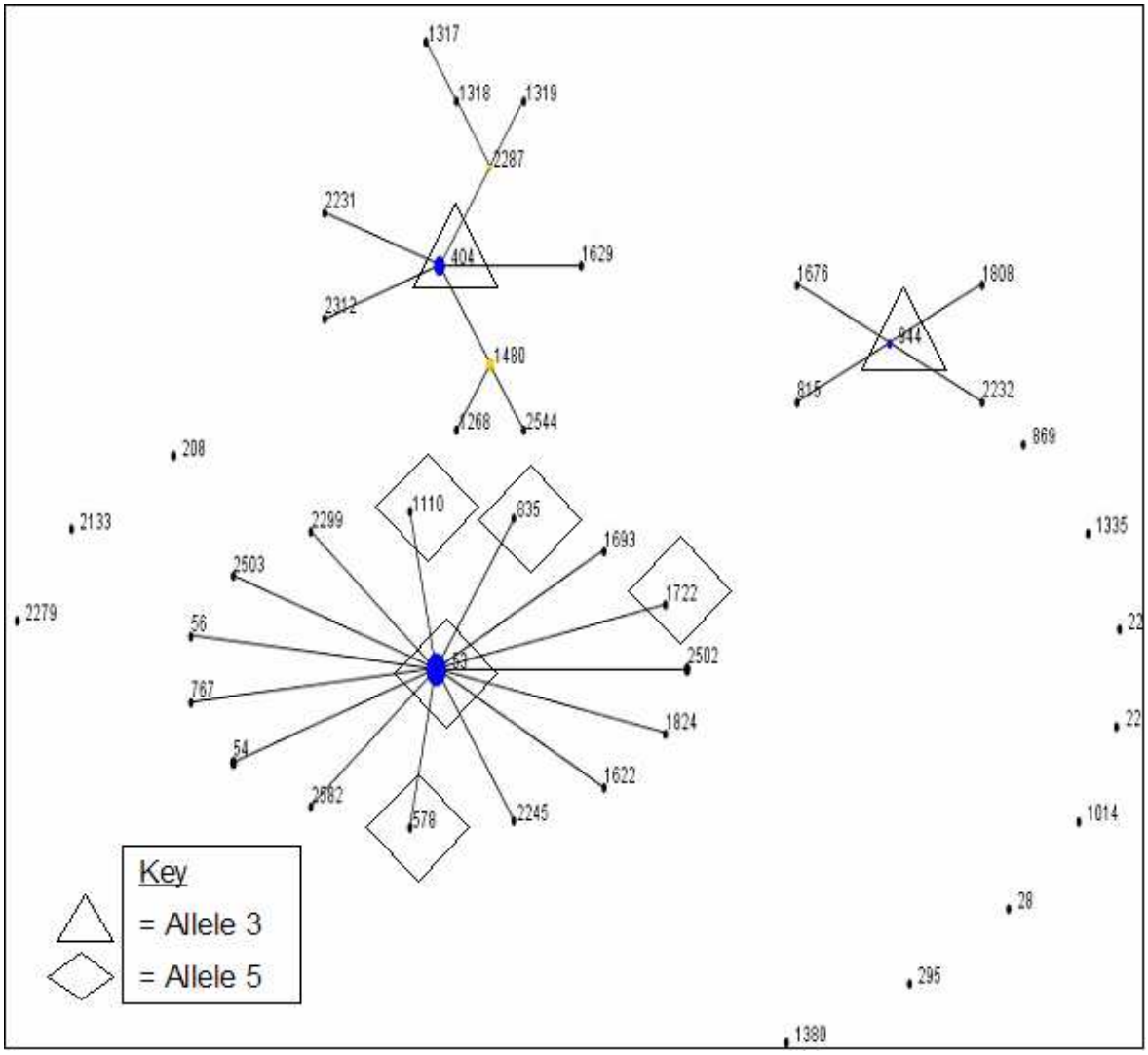


Figure 3.11: Ply amino acid allele identity of serotype 8 clones

Ply amino acid allele identity mapped onto serotype 8 eBURST diagram. Each allele or combination of alleles is represented by a different shape, as detailed in the key.

3.3.3 Prevalence of serotype 8-related IPD in Scotland

The ST306 clone possessing the non-haemolytic Ply allele 5 was shown to be undergoing clonal expansion in the Scottish IPD population (Kirkham *et al.*, 2006a). Information from SMPRL was used to determine the prevalence of serotype 8-related IPD in Scotland covering the years 2002-2005. The prevalence of the ST53 clone within serotype 8-related IPD was also investigated. Serotype 8 was not found to be an increasing cause of IPD in Scotland, having caused ~5% of the IPD in Scotland between 2002 and 2005 (Figure 3.12A). However, the prevalence data showed that ST53 was already the dominant clone within the serotype 8 IPD-causing population, representing >80% of serotype 8 isolates within this time period (Figure 3.12B). This implies that the ST53 clone has undergone previous clonal expansion within the serotype 8 IPD population.

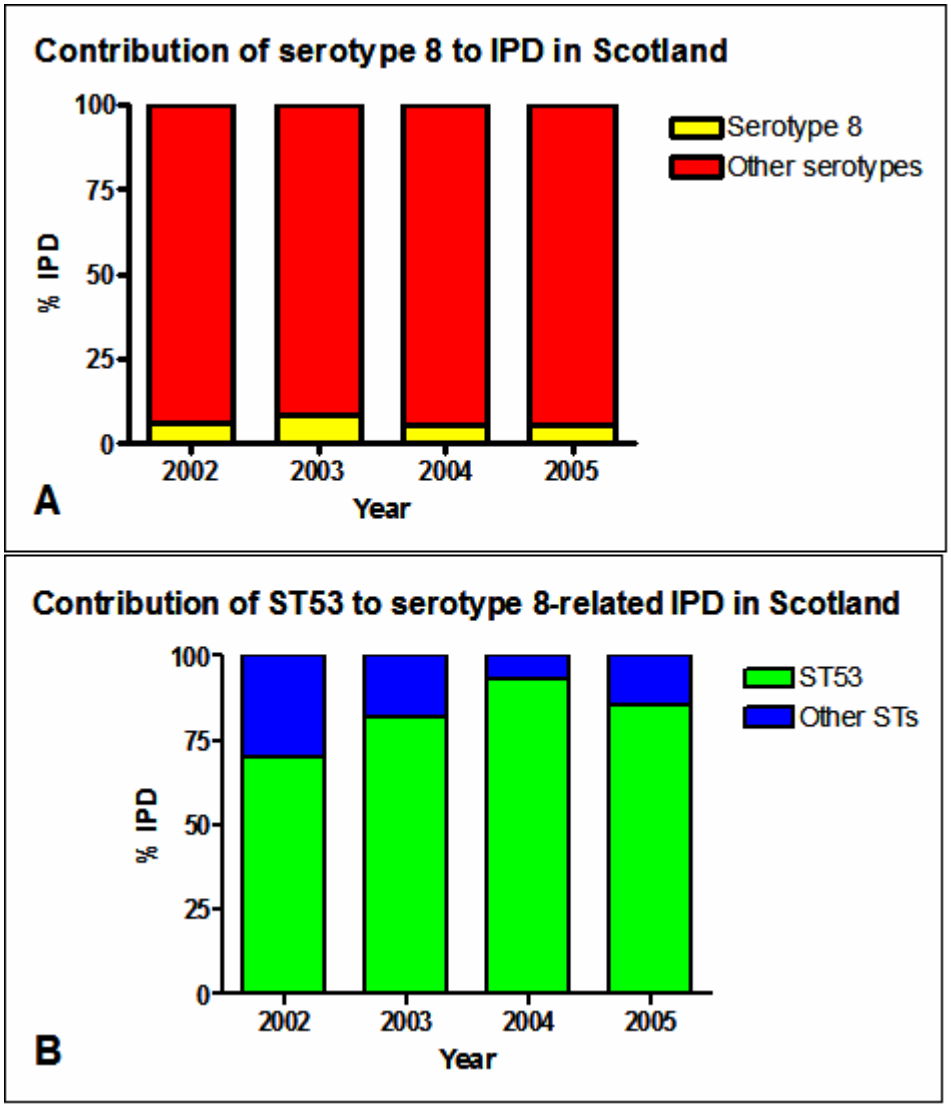


Figure 3.12: ST53 prevalence data in Scotland

A - Prevalence of serotype 8-related IPD in Scotland, in comparison to other serotypes, between 2002-2005. B - Prevalence of ST53-related IPD in Scotland, in comparison to other STs of serotype 8, between 2002-2005.

3.4 Construction and characterization of a Ply knock-out strain

3.4.1 Construction of D39 ΔPly stop mutant

It was of interest to investigate the role played by the toxin Ply in virulence of *S. pneumoniae*. One way to determine this is to disrupt the *ply* gene resulting in an isolate producing no Ply protein. This has been done previously (Berry *et al.*, 1989b; Friedland *et al.*, 1995; Wellmer *et al.*, 2002; Winter *et al.*, 1997), however mutation techniques used, involving insertion of antibiotic resistance cassettes, may have produced polar effects other than just disrupting the *ply* gene. In order to produce an unmarked mutation, a mutant expressing no Ply protein (D39 ΔPly stop) was created in a D39 background using the Janus technique described previously (Sung *et al.*, 2001). The advantage of using this technique is that it allows production of mutations without the insertion of antibiotic resistance cassette into the gene of interest, essentially producing unmarked mutations. This mutant had an insertion of 1 T base after base 6 in the D39 open reading frame, which resulted in a premature stop codon (TAA) after the second amino acid of the translated gene. This mutant was constructed as described in Chapter 2.6, by three transformation steps, and successful clones at each stage were confirmed by sequencing and antibiotic resistance profiles as described.

3.4.2 Confirmation of successful D39 ΔPly stop mutation

The success of the mutation produced in two separate clones was confirmed by haemolytic assay and Western blot using the polyclonal α–Ply antibody. The haemolytic assay result confirmed that although the wild-type D39 parent strain displayed haemolytic activity, neither D39 ΔPly stop clone had any haemolytic activity (Figure 3.13). Similarly, the results of the Western blot show that whilst the wild-type D39 parent strain and purified Ply allele 1 produced bands of the correct size on the blot, neither D39 ΔPly stop clone produced any recognisable Ply (Figure 3.14).

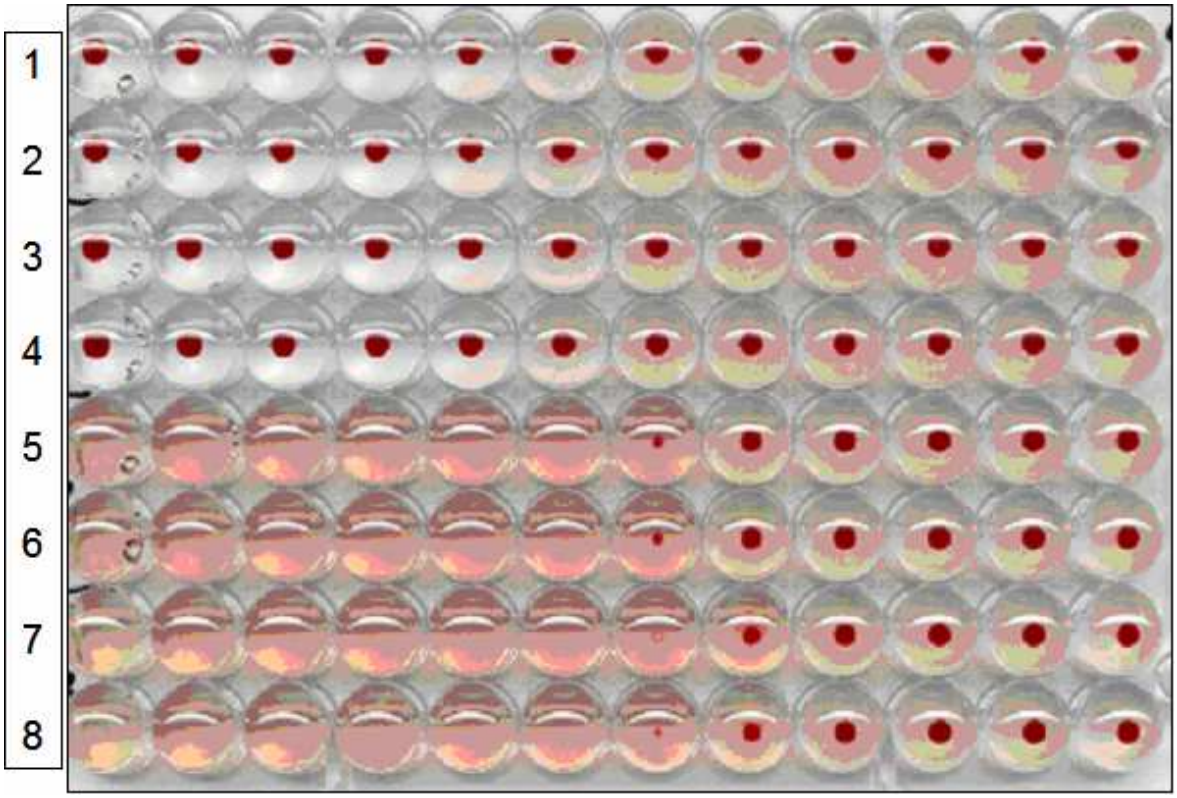


Figure 3.13: Haemolytic assay of D39 ΔPly stop

Haemolytic assay of D39 ΔPly stop clones to confirm the successful production of mutants producing no Ply. Lanes 1-2 – D39 ΔPly stop clone 1; lanes 3-4 – D39 ΔPly stop clone 2; lanes 5-6 – D39 wild-type parent strain; lanes 7-8 – purified Ply allele 1, 1/100 starting dilution. Doubling dilutions were carried out across the plate for each sample.

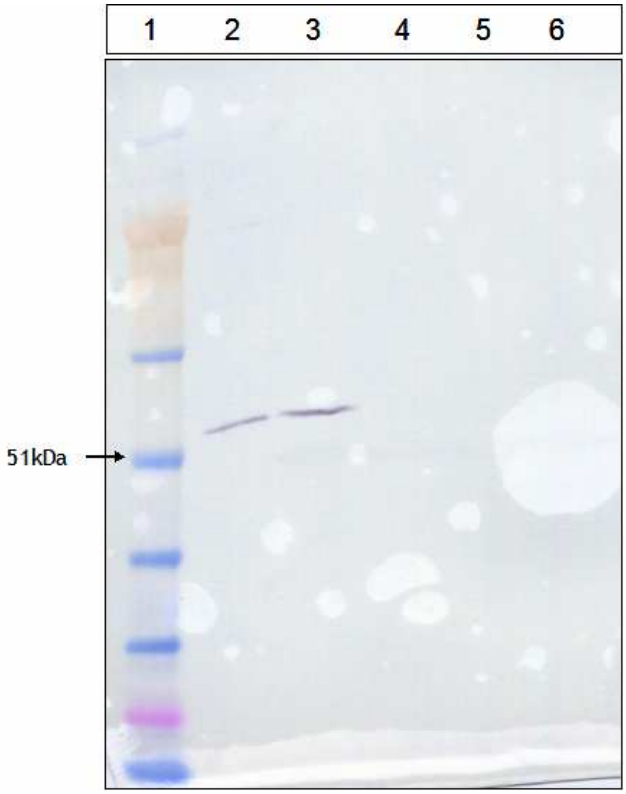


Figure 3.14: Western blot of D39 ΔPly stop

Western blot of D39 ΔPly stop clones to confirm the successful construction of mutants producing no Ply protein. Probed with polyclonal α-Ply antibodies. Lane 1 – SeeBlue +2 protein marker; lane 2 – D39 parent strain; lane 3 - purified Ply allele 1, 1/100 starting dilution; lane 4 - D39 ΔPly stop clone 1; lane 5 - D39 ΔPly stop clone 2; lane 6 – negative control – dH2O.

3.4.3 Animal studies

In order to test the effect of knocking out the Ply gene on pneumococcal virulence, studies in MF1 mice were carried out by Dr. Gill Douce, using 3 pneumococcal strains. The virulence of wild-type parent strains D39, the knock-out strain D39 ΔPly stop clone 1 (referred to from now as D39 ΔPly stop), and a third strain, D39 Δ6, (provided by Dr. Lea-Ann Kirkham, deletion of 2 amino acids (ΔA146R147) that abrogates the haemolytic activity of the toxin although a protein is still expressed) was tested. Result showed that at 24h and 48h, there was no statistical difference between viable counts of the three strains in blood (Figure 3.15A, 3.15B). However, when survival was monitored for a period of ten days following challenge, there was a significant difference between the mean survival times of D39 and D39 Δ6 ($p = 0.0399$), as well as D39 and D39 ΔPly stop ($p = 0.0144$) (Figure 3.15C, 3.15D). However, there was no significant difference in survival times between D39 Δ6 and D39 ΔPly stop ($p=0.4484$). This implies that

the haemolytic activity of the toxin is important in pneumonia, whilst other toxin activities play a lesser role.

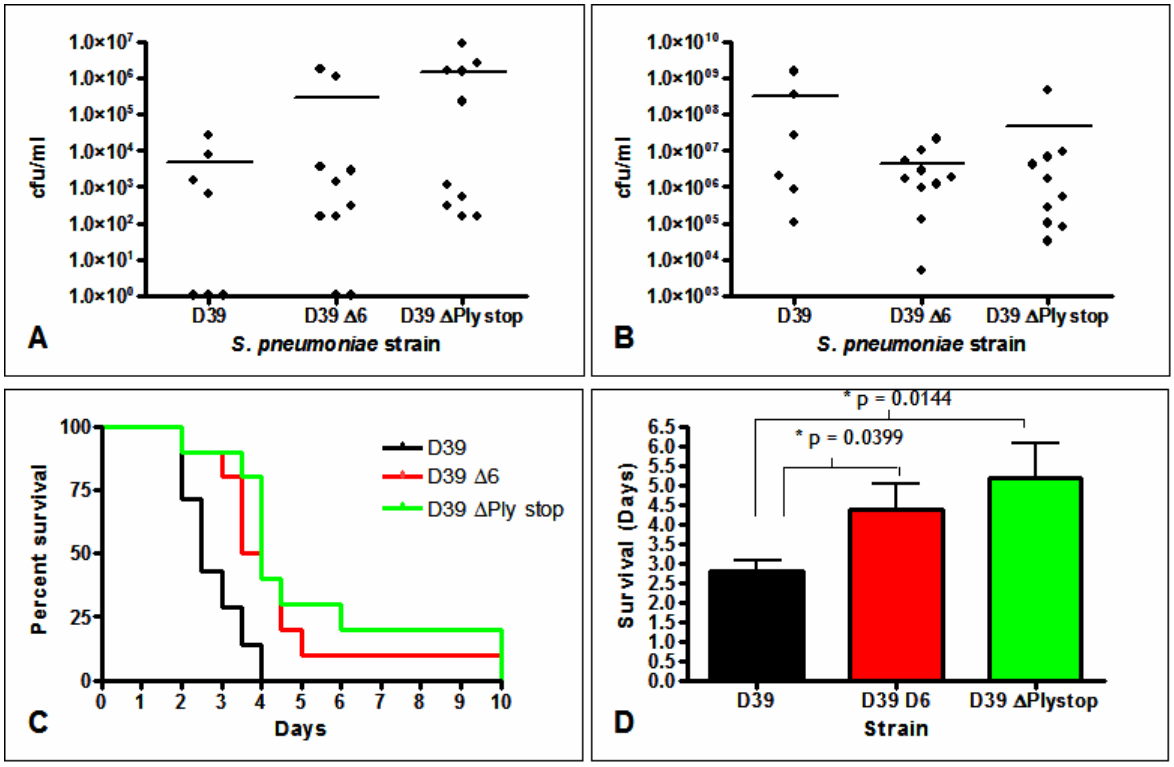


Figure 3.15: Animal studies with D39 ΔPly stop

A – viable counts of *S. pneumoniae* present in the blood 24h post-infection (p.i.). B – viable counts of *S. pneumoniae* present in the blood 48h p.i. C – plot of survival data monitored for a period of ten days for mice infected with each isolate. D – mean survival time for mice infected with each isolate, showing significant differences between the sample sets.

3.5 Discussion

Ply is a cytoplasmic toxin, and was considered to be well-conserved (Mitchell *et al.*, 1990). However, 6 Ply proteins had been identified before this study (Kirkham *et al.*, 2006a; Lock *et al.*, 1996; Tettelin *et al.*, 2001; Walker *et al.*, 1987). Nine more novel Ply amino acid alleles were uncovered here, making a total of at least 15 Ply amino acid alleles, and showing that the variation present in Ply, at the amino acid level, is greater than previously thought. Ply shows an amino acid sequence variation of 3.3%, which is slightly higher than the variation seen in the house-keeping genes (1-2%) (Enright and Spratt, 1998), but substantially lower than variation observed in pneumococcal surface proteins such as NanA (King *et al.*, 2005) and PspC (Iannelli *et al.*, 2002), or antibiotic resistance genes such as the penicillin-binding proteins (Dowson *et al.*, 1994). Variation in these proteins is thought to be due to the evolutionary pressure exerted by the host immune system or exposure to antibiotics respectively. However, the selective evolutionary pressure exerted on the Ply protein, which is cytoplasmic, is unknown. Observation of a higher level of variation in the *ply* gene than previously thought should not prevent the use of this toxin as a protein component in future pneumococcal conjugate vaccines, as polyclonal antibodies raised against wild-type Ply were able to recognise all alleles of Ply expressed, and the levels of variation are far lower than those of other surface-exposed protein vaccine candidates.

The allelic variation observed resulted in discovery of alleles with a number of different characteristics. Alleles 1 and 2 were possessed by most isolates in the initial screen, with 92/121 isolates possessing either of these alleles. This result shows that although 15 alleles were identified, the majority of isolates possess a wild-type Ply with normal haemolytic activity. However, a number of the other alleles discovered had reductions in haemolytic activity as a result of mutations present in the protein. The most drastic example of this was allele 5, originally identified in a serotype 1, ST306 isolate and lacking any haemolytic activity (Kirkham *et al.*, 2006a). In this study, allele 5 was shown to be harboured by a number of serotype 1 and 8 isolates, as well as a non-typable ST577 isolate. At the protein level, allele 5 was identical between serotype 1 and 8 isolates. However, at the DNA level, there was a single nucleotide polymorphism (SNP) which was found consistently in all alleles from serotype 8 isolates when

compared to serotype 1 isolates (C>T at nucleotide position 1113). This mutation was synonymous. The non-typable isolate possessing allele 5 shared the DNA allele present in serotype 8 isolates. Furthermore, the clone ST577 was found to be a SLV of the serotype 8 clonal complex founder ST53. Therefore, as the non-typable isolate shared Ply allele and housekeeping genes with serotype 8 pneumococci, it is possible that this isolate has serotype 8 capsular genes but does not express the capsule, or there may have been partial or full deletion of capsular genes. Further to this, it is interesting to note that this isolate was isolated from the blood of a bacteraemic patient. Although this seems to contradict data showing the pneumococcal capsule to be essential for colonisation (Magee and Yother, 2001) and survival in blood, it is not possible to confirm that this isolate was not expressing capsule whilst in the host.

Ply allele 12 was not recognised by the monoclonal antibody PLY-7. This antibody recognises an epitope GQDLTAH within the Ply amino acid sequence, relating to AA401-407 in the Ply protein (de los Toyos *et al.*, 1996). Allele 12 has a mutation in this region, Q402E, which is the likely cause of this non-recognition. mAbPLY-7 is currently used in an ELISA in many laboratories worldwide to determine presence or absence of Ply (Cima-Cabal *et al.*, 2001), and by association *S. pneumoniae*. However, the identification of an allele of Ply that was not recognised by this antibody indicated that there may be instances of false negative results when using this antibody. Therefore, care should be taken when interpreting results obtained using this antibody as a diagnostic or therapeutic tool. Allele 12 was identified in a single 18C, ST818 isolate, which was the only representative of this ST in the MLST database. When eBURST was used to analyse all strains of serogroup 18/18C, ST818 was not part of the main serotype 18C clonal complex. Furthermore, 6 other serogroup 18 isolates were present in the initial screen, and all possessed allele 2. Therefore, allele 12 appears to be confined to ST818 in the serogroup 8 population, of which there is only 1 representative in the MLST database.

Isolate S1-11 had allele 14, and was shown to express no recognisable Ply due to the presence of an 871bp insertion sequence. This finding was of interest as the isolate was isolated from a patient with pneumonia in Spain (Brueggemann and Spratt, 2003), implying this isolate possesses disease-causing ability without recogniseable Ply. S1-11 represents the first demonstration of a clinical isolate

of *S. pneumoniae* that does not produce functional Ply. The 871bp insertion sequence encodes IS1515, a characterized, functional insertion sequence found predominantly in serotype 1 pneumococci (Munoz *et al.*, 1998). Because IS1515 contains internal promoter regions, expression and therefore excision of the insertion sequence may still be possible. This hypothesis is supported by the Western blot result in Figure 3.7, which showed a weak band for Ply with this isolate at the correct size, implying that excision may have occurred during growth of the isolate. In hindsight, testing the lysate used for this Western blot would have allowed confirmation of this hypothesis, and further study would be required to allow confirmation. This could include passage of the S1-11 strain through an animal model, as the selective pressure on the strain from the host immune system may select for isolates producing Ply as a result of the excision of the insertion sequence.

S1-11 was isolated from the blood of the patient, displaying the ability to progress from the lungs of the infected patient to the blood. IS1515 is an active element capable of excision and insertion elsewhere in the pneumococcal genome, and it was therefore possible that this isolate produced a functional Ply toxin at the time when disease was established, and the insertion element inserted into the *ply* gene at some point after isolation. However, recent identification of this insertion in the *ply* gene of another serotype 1, ST228 clone isolated from a pneumonia patient in France (Garnier *et al.*, 2007) shows that this is unlikely to be the case, and that indeed a serotype 1, ST228 clone expressing no Ply is circulating and causing invasive disease in mainland Europe. The fact that this clone causes IPD without producing Ply means that any vaccine targeting Ply may select for this clone, resulting in clonal expansion. Interestingly, although the underlying Ply allele in these isolates was allele 2, the ST228 isolate from Scotland possessed non-haemolytic allele 5, demonstrating that the *ply* gene of pneumococcal clones of identical ST can vary markedly according to geographic location.

Distribution of allele 5 was not confined to ST306 in the serotype 1 pneumococcal population, although it was found only in the ST306 clonal complex. There appears to be a split in lineage A, with allele 2 being found predominantly in 1 branch and allele 5 predominating in the other. The two branches of Lineage A are linked as the predicted founders are triple-locus

variants (TLV) of each other. It is possible that a clone exists that would link these two branches to form 1 complex on the eBURST diagram. This clone may have yet to be isolated or entered into the MLST database, or it may have existed and been out-competed. This theoretical clone would be predicted to have Ply allele 2. The two *ply* DNA alleles in this lineage are unrelated, varying at 17 nucleotide positions, resulting in 7 amino acid changes, including a 2 amino acid deletion.

Allele 5 was also discovered in serotype 8 pneumococci. This allele was found only in clones of the main clonal complex, with ST53 as predicted founder. In contrast to the SLVs of ST306 found in the MLST database, which were solely Scottish isolates, SLVs of ST53 were present from a variety of geographical locations, including Spain, Norway, the Netherlands and the USA. It is therefore possible that serotype 8 clones possessing allele 5 are more internationally disseminated than serotype 1 clones. The Scottish prevalence data showed that ST53 is already the dominant clone in serotype 8-related IPD. This is in contrast to published data on ST306 prevalence in serotype 1-related IPD, which shows ongoing clonal expansion of this clone to become the dominant clone of this serotype (Kirkham *et al.*, 2006a). Taken together, these two findings imply that allele 5 may have been present in serotype 8, ST53 clones prior to acquisition by serotype 1 isolates, presumably starting with an ST306 isolate. The presence of allele 5 may then be responsible for driving the clonal expansion of these two unrelated clones. However, it is important to appreciate that the nature of the MLST database means that internationally disseminated clones of the ST306 clonal complex may exist. This is supported by published data showing that an increase in serotype 1 IPD in Sweden from 1992-1997 was as a direct result of ST306 emergence (Henriques Normark *et al.*, 2001), although none of the SLV STs from ST306 were observed, and the identity of the Ply protein in these clones was not established. Furthermore, it is possible that other genetic differences between these clones and closely related relatives account for this phenomenon. One way to test this would be to sequence and compare the genomes of these clones, for example genomic comparison of an ST306 clone and an ST227 clone would allow discovery of other genetic differences which may play a role. Although ST227 and ST306 genomes are becoming available, it

was not possible to perform this full genome analysis under the remit of this project.

The clones ST53 and ST306 are unrelated, sharing only 1 housekeeping allele (*spi*). Furthermore, the sequences of the remaining 6 housekeeping alleles were no more similar than those of 2 random clones (data not shown). This result clearly shows that there is no obvious relation between the 2 founders of the clonal complexes associated with Ply allele 5. With few exceptions, Ply allele was found to generally correlate with the ST of a pneumococcal isolate. Therefore, it was hypothesised that the pressure to maintain alternative Ply alleles may stem from the pneumococcal genome itself, suggesting that certain Ply alleles will be maintained only when present alongside a certain set of other pneumococcal proteins and that all of the proteins present in the set are required to work in concert to ensure the survival of the clone. This would lead to the hypothesis that although ST53 and ST306 isolates are unrelated by sequence type, they share some unknown factor that allows the non-haemolytic allele to be maintained whilst still allowing the clones to cause invasive pneumococcal disease.

These results suggest that Ply, and potentially further virulence factors, may have varying levels of importance in clones of different genetic backgrounds. This hypothesis was investigated during this study, by attempting to produce mutants of serotype 1 clones using the Janus technique (Sung *et al.*, 2001). The plan was to "swap" the Ply genes between a haemolytic ST227 isolate, possessing allele 2, and a non-haemolytic ST306 isolate possessing allele 5. Animal studies of these clones would allow great insight into the role and importance of the non-haemolytic toxin in the serotype 1 clones. However, despite numerous attempts, the serotype 1 clones displayed levels of transformation that did not allow the gene swap to occur, and the clones could not be created. Despite this, the clones were transformable using the replicative plasmid PVA838, showing that these isolates do maintain low levels of competence, and further work may allow these clones to be created. Continuation of this work would offer novel insight into the role of the non-haemolytic toxin, as well as the possibility of virulence factors having differing levels of importance in pathogenesis of distinct pneumococcal clones.

The mutation Y150H was unique to allele 5.The tyrosine residue at this position was fully conserved across all known CDCs, whilst other positions of mutation within Ply were shown to vary among the other CDCs. This suggests an importance of this residue in the haemolytic activity of these toxins, as only allele 5, possessing this mutation, displayed fully abrogated haemolytic activity. However, purified toxin possessing only the Y150H mutation (purified by Kirsty Ross) had a specific activity of 478 HU/mg. This represents a drastic reduction (~1000-fold) in comparison to wild-type, but also shows this mutation is not solely responsible for abrogation of haemolytic activity. This implies that the presence of the other mutations in this allele is required to fully abolish haemolytic activity. This hypothesis is supported by the fact that allele 3, which lacks Y150H but has all other mutations present in allele 5, had a ~60-fold reduction in specific activity. Indeed one of these mutations, T172I has been shown to reduce haemolytic activity of Ply previously (Lock *et al.*, 1996). However, purified allele 5 (purified by Paul Hughes, supervised project student) displayed a haemolytic activity comparable to the toxin possessing only the Y150H mutation. This result is in agreement with previous findings in Ply and PFO, which showed that a Y181A PFO mutant (equivalent to Y150A in Ply) had haemolytic activity <1% of wild type (Hotze *et al.*, 2001), and a Y150A mutant in Ply displayed similar activity (Kirkham *et al.*, 2006a). This leads to the conclusion that although these toxins display low levels of haemolytic activity at high concentration, these concentrations are not present during *in vitro* growth of the isolates, and therefore unlikely to be present *in vivo*. This hypothesis is supported by the findings of Kirkham et al. (2006) who demonstrated that crude lysates of isolates with allele 5 showed reduced binding ability to erythrocytes, and were unable to form pores and lyse the cells (Kirkham *et al.*, 2006a).

There was a relationship between the mutations carried by alleles 3, 5 and 10 and reduced mobility of these proteins in SDS-PAGE. This was in agreement with previous findings (Lock *et al.*, 1996), and results in an apparent increase in molecular weight. Lock et al. attributed the difference in electrophoretic mobility as seen in their Ply8 (allele 3) solely to the T172I mutation. However, whilst alleles 3 and 5 possess this mutation, allele 10 does not, and still has clearly reduced mobility. Interestingly, these alleles share the mutation A265S and the deletion of amino acids at positions 270 and 271. These may play a role

in the reduced mobility observed; however, allele 6 also possessed these mutations and showed no reduction in mobility, although a further S167F mutation was also present, which may have had further effects on electrophoretic mobility.

Unlike pneumococci of most other serotypes, serotype 1 pneumococci are known to be associated with outbreaks of pneumococcal disease (Dagan *et al.*, 2000; Gratten *et al.*, 1993; Leimkugel *et al.*, 2005; Mercat *et al.*, 1991; Proulx *et al.*, 2002). Serotype 8 pneumococci have also been responsible for such outbreaks (Berk *et al.*, 1985; Birtles *et al.*, 2005). In the study of 121 pneumococcal isolates, the non-haemolytic allele 5 was observed only in isolates of serotypes associated with invasive disease. There is therefore an association between this non-haemolytic variant of the toxin and serotypes with the ability to cause outbreaks of IPD. Although haemolytic activity is not required for these clones to cause IPD, other activities of Ply such as complement activation (Mitchell *et al.*, 1991; Paton *et al.*, 1984) and inflammation via interaction with TLR-4 (Malley *et al.*, 2003) may contribute to the pathogenesis of infection. However, there is no data presented in these papers on the STs of these outbreak clones, which would allow further insight. Whether there is a causal link between production of a non-haemolytic toxin variant and outbreaks of IPD would require a much larger study comparing outbreak and non-outbreak isolates. Given that outbreaks of pneumococcal disease remain rare, such a study may prove difficult.

The construction of a Ply knock-out without a selective marker has allowed initial studies into the importance of Ply to *S. pneumoniae*. Many studies have been done previously in different animal models, with differing results. A number of studies have shown that Ply is not essential for pneumococcus-induced inflammation during meningitis (Friedland *et al.*, 1995; Wellmer *et al.*, 2002; Winter *et al.*, 1997). Another study showed that pneumococci with the *ply* gene caused acute sepsis, whilst those without caused chronic bacteraemia (Benton *et al.*, 1995). The initial data reported here showed that the Ply knock-out had no reduction in bacterial counts in blood compared to wild-type and a non-haemolytic mutant. However, mice infected with either the knock-out or the non-haemolytic mutant showed a significant increase in survival compared to the wild-type. This implies that although these bacteria can cause bacteraemia at a similar level to wild-type, there is decreased virulence. Furthermore,

results show that in this model, the haemolytic property of the toxin is of more importance than other properties, as there is no significant difference between the knock-out and the non-haemolytic mutant. However, this study was preliminary and further studies are required with this mutant to fuller determine its effect on virulence. This unmarked knock-out is a valuable resource, is routinely used as a negative control in many Ply studies in our laboratory, and is currently being used by colleagues on a number of further projects.

In conclusion, Ply has a level of variation higher than previously thought, and the haemolytic activity of the toxin is not always required for IPD. Furthermore, an insertion sequence in the *ply* gene of one isolate means no functional Ply is produced. These findings imply that the roles played by Ply may vary in pneumococci of different genetic backgrounds. The non-haemolytic allele appears to be confined to unrelated clonal complexes that have undergone expansion. Furthermore, this toxin is harboured in unrelated serotypes frequently associated with disease outbreaks. Therefore, this property of Ply, in certain genetic backgrounds, may play a role in driving clonal expansion and may also have an involvement in outbreaks of pneumococcal disease.

4 Diversity studies of the virulence factor NanA

4.1 Summary

NanA is an important pneumococcal virulence factor, which has been studied as a pneumococcal vaccine candidate (Tai, 2006), and implicated as important in pathogenesis of pneumococcal diseases including p-HUS (Klein *et al.*, 1977; Seger *et al.*, 1980). Results of diversity studies can give insight into the importance of NanA to the pneumococcus as well as informing on the potential of this virulence factor as a vaccine candidate. The aim of this chapter was to assess the diversity present in the NanA protein, relate these results to the protein structure, and investigate the importance of NanA in p-HUS pathogenesis. The diversity present in the pneumococcal virulence factor NanA was investigated and results showed that the protein had levels of diversity significantly higher than Ply, with mutations present at 13.7% of the amino acids of the wild-type protein. In total, 18 protein alleles were identified from 33 isolates (full DNA alignment in Appendix I). Three regions of diversity, two insertion regions and one region of 20AA repeats were identified, and had been previously reported (King *et al.*, 2005). Furthermore, a large number of point mutations were identified along the length of the protein. Despite this, all residues previously identified as essential to the neuraminidase activity of the protein were fully conserved. The diverse nature of this protein was expected due to its surface-exposed nature (Camara *et al.*, 1994), however, the diversity present should not reduce the potential of NanA as a vaccine candidate, as polyclonal antibodies display ability to recognise significantly divergent NanA alleles.

The structure of the enzymatic site of the NanA protein was solved by collaborators, and allowed mapping of mutations present in each allele onto the structural model of the active site. From this result, three alleles were chosen for purification, with activities varying by up to 7-fold from wild-type. Furthermore, a region of diversity termed mosaic block C mapped to a small region inserted into the active site domain, with two distinct variants of this region identified, sharing only 57% sequence identity. Although the function of this region was unknown, the surface shape and electrostatic potential of the two variants was shown to differ markedly, implying they may play different roles in different clones.

Pneumococcal NanA has been implicated as important for the progression of HUS, a complication of IPD. The incidence of HUS following on from invasive pneumococcal disease (p-HUS) has been increasing. This form of the disease has higher levels of mortality than normal, verotoxin-associated HUS (Fitzpatrick *et al.*, 1993). Very little is known about the pathogenesis of p-HUS, although a model for the mechanism of pathogenesis has been suggested, implicating the action of pneumococcal neuraminidase (Klein *et al.*, 1977; Seger *et al.*, 1980). This study attempted to correlate NanA activity to p-HUS by determining sequence and activity of NanA in these clones. In this study of 9 p-HUS isolates and 6 matched controls, no correlation was found between p-HUS and NanA protein allele, or overall neuraminidase activity. These results show that although an active neuraminidase enzyme may be required to cause p-HUS, no particular NanA allele is associated with p-HUS isolates, and other genetic factors may dictate which pneumococcal clones are able to cause p-HUS.

4.2 Diversity of NanA in *S. pneumoniae* clinical isolates

NanA is an important virulence factor of the human pathogen *S. pneumoniae*. Like Ply, it is being considered as a vaccine candidate for next-generation pneumococcal protein vaccines (Tai, 2006). It is therefore of importance to determine the diversity present in this virulence factor in the pneumococcal population, as variation uncovered may give insight into suitability of NanA as a pneumococcal vaccine candidate. The *nanA* genes of 18 clinical isolates were sequenced and analysed. These isolates were either isolated from p-HUS, or were matched control isolates (see section 4.6). Furthermore, the *nanA* genes from 15 sequenced pneumococcal genomes were analysed (Dopazo *et al.*, 2001; Hiller *et al.*, 2007; Lanie *et al.*, 2007; Tettelin *et al.*, 2001). All isolates tested can be found in Table 4.1. These genes were aligned to wild-type D39 using Vector NTI™ software, resulting in mutations at 457 positions in the DNA sequence, representing 34.9% of the wild-type sequence, with 3 regions of insertion or repeat also present. Translation and alignment of these resulted in mutations at 142 amino acid positions along the protein. This represented 13.7% of the amino acids in the wild-type protein, which is significantly higher than the diversity observed in Ply. The mutations were clustered into several groups, with the presence of three regions of diversity, two insertions of 16 and 5 amino acids, a frameshift mutation resulting in a stop codon, and a region of 20 amino acid repeats occurring from 0-3 times (Figure 4.1). The presence of these insertion and repeat regions increased the diversity present in the NanA protein compared to Ply. The frameshift mutation resulted in a premature stop codon at AA801, and was only present in TIGR4. The 16AA insertion (EEGENITLPAEHVESV) was present after amino acid 84 in the protein, whilst the 5AA insertion observed (EMGKG) was found to be an insertion-duplication of AA822-826. Region of diversity I (RDI) ran from AA163-337, whilst RDII ran from AA393-568, and RDIII from AA837-891. Interestingly, when comparing the diversity present in each region of diversity, two distinct variants appear present in each region (Figure 4.2). All of these features of NanA diversity have been previously reported in a partial study of the *nanA* gene (King *et al.*, 2005). The regions of diversity correspond to mosaic blocks A, C and D from this study respectively, and will be described as such from now on. Mosaic block B was not observed in this study. Further to these mutations, a large number of point mutations were present within the alleles. The aligned NanA proteins were assigned an allele number

based on alignment to the wild-type D39 protein, termed allele 1. 18 protein alleles were discovered, with one representative of each shown in Figure 4.2. The alleles were further separated into 7 allele groups, determined by the presence of diversity features as shown in Table 4.2. A number of amino acids essential to the activity of the NanA protein have been identified, including four Aspartic boxes conserved across all sialidase enzymes (SXDXGXTW motifs) (Camara *et al.*, 1994) 3 arginine-triad residues (R347, R663, R721) and other essential residues Y752, E647 (Yesilkaya *et al.*, 2006) and D372 (Chong *et al.*, 1992). Despite the high levels of variation present, these amino acids are fully conserved across all protein alleles identified. Indeed, although one aspartic box is present within a mosaic block region, it remains conserved in both variants of the block. Neuraminidase activity was compared for available isolates, and there was no significant difference in activity associated with particular protein alleles (See section 4.6).

Strain	Serotype	ST	NanA allele	NanA allele group	Source
D39	2	125	1	1	(Walker *et al.*, 1987)
OXC141	3	180	1	1	(Brueggemann *et al.*, 2003)
02-1198	3	180	1	1	SMPRL
H050600025	3	180	1	1	HPA, (Waters *et al.*, 2007)
SP3-BS71	3	180	1	1	(Hiller *et al.*, 2007)
05_1308	19A	199	2	2	SMPRL
05_2084	19A	199	2	2	SMPRL
H051740086	19A	199	2	2	HPA, (Waters *et al.*, 2007)
H052300328	19A	199	2	2	HPA, (Waters *et al.*, 2007)
TIGR4	4	205	3	3	(Tettelin *et al.*, 2001)
SP6-BS73	6	460	4	4	(Hiller *et al.*, 2007)
SP11-BS70	11	62	5	4	(Hiller *et al.*, 2007)
SP14-BS69	14	124	6	4	(Hiller *et al.*, 2007)
G54	19F	n/a	7	4	(Dopazo *et al.*, 2001)
INV104B	1	227	8	4	Sanger
SP18-BS74	6	n/a	9	4	(Hiller *et al.*, 2007)
Spain $23F^{-1}$	23F	81	10	4	Sanger
SP9-BS68	9	1269	11	4	(Hiller *et al.*, 2007)
SP19-BS75	19	485	12	5	(Hiller *et al.*, 2007)
06_1011	14	9	12	5	SMPRL
H053940080	14	9	12	5	HPA, (Waters *et al.*, 2007)
H053940083	14	9	12	5	HPA, (Waters *et al.*, 2007)
INV200	14	9	12	5	Sanger
SP23-BS72	23	37	13	5	(Hiller *et al.*, 2007)
05_2426	19A	1201	14	5	SMPRL
H050940049	19A	1201	14	5	HPA, (Waters *et al.*, 2007)
A66.1 Xen	3	180	15	5	Xenogen
H060160064	19A	199	16	6	HPA, (Waters *et al.*, 2007)
H040400333	19A	199	16	6	HPA, (Waters *et al.*, 2007)
1427	15B/C	199	16	6	SMPRL
2028	15B/C	199	16	6	SMPRL
06_2715	6A	65	17	7	SMPRL
H043640049	6A	65	18	7	HPA, (Waters *et al.*, 2007)

Table 4.1: Strains from NanA diversity study

Strains used in the study of NanA diversity, with serotype and ST included. Also included are NanA amino acid allele, allele group and source of each isolate. SMPRL = Scottish Meningococcal and Pneumococcal Reference Laboratory, HPA = Health Protection Agency. STs noted as n/a were undefined.

Amino acid allele group	Regions of diversity present	Amino acid alleles in group
1	Wild-type – Three 20AA repeats	1
2	Two 20AA repeats	2
3	Frameshift mutation	3
4	5AA insertion, mosaic block D, no 20AA repeats	4, 5, 6, 7, 8, 9, 10, 11
5	5AA insertion, mosaic blocks C & D, no 20AA repeats	12, 13, 14, 15
6	16AA and 5AA insertions, mosaic blocks C & D, no 20AA repeats	16
7	16AA and 5AA insertions, mosaic blocks A & D, no 20AA repeats	17, 18

Table 4.2: NanA amino acid allele groups

NanA allele groups determined by presence or absence of 7 diversity markers present in NanA protein. Alleles present in each group also included.

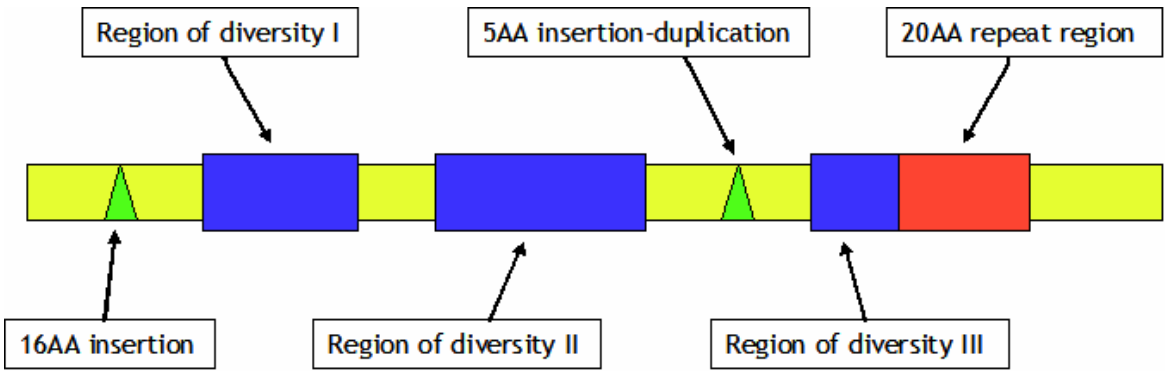


Figure 4.1: Diagram of NanA diversity

Schematic showing regions of diversity present in NanA protein alleles. Three main regions of diversity were identified, in blue, along with two insertion regions, in green, and one repeat region in red. Regions of diversity I-III correspond to previously identified mosaic blocks A, C and D (King *et al*., 2005).

															Amino acid position								
					7	14	71	85	93	134	163	166	176	179	180	186	188	189	193	194	199	200
Strain	**Serotype**	**ST**	**Allele**	**Group**	**R**	**N**	**E**	**----**	**K**	**G**	**K**	**A**	**A**	**K**	**D**	**A**	**Y**	**N**	**T**	**L**	**S**	**D**
D39	2	125	1	1																		
05_1308	19A	199	2	2		I	G															
TIGR4	4	205	3	3		I																
SP6-BS73	6	460	4	4		I																
SP11-BS70	11	62	5	4		I																
SP14-BS69	14	124	6	4		I				D												
G54	19F	63	7	4		I				D												
INV104B	1	227	8	4		I																
SP18-BS74	6	n/a	9	4		I																
Spain 23F^{-1}	23F	81	10	4		I																
SP9-BS68	9	1269	11	4		I			N		S	R	D	E								
06_1011	14	9	12	5		I																
SP23-BS72	23	37	13	5		I																
05_2426	19A	1201	14	5		I																
A66.1 Xen	3	180	15	5		I	G															
H060160064	19A	199	16	6		I	G	INS														
06_2715	6A	65	17	7	Q	I	D	INS		A	S	R	D	E	N	S	L	D	L	I	A	N
H043640049	6A	65	18	7	Q	I	D	INS			S	R	D	E	N	S	L	D	L	I	A	N

					Amino acid position																	
					202	206	207	209	216	229	232	236	253	265	266	267	269	275	277	286	293	297
Strain	**Serotype**	**ST**	**Allele**	**Group**	**K**	**N**	**N**	**N**	**K**	**K**	**A**	**K**	**R**	**T**	**H**	**V**	**I**	**A**	**N**	**I**	**N**	**T**
D39	2	125	1	1																		
05_1308	19A	199	2	2																		
TIGR4	4	205	3	3																		
SP6-BS73	6	460	4	4																		
SP11-BS70	11	62	5	4								N	K									
SP14-BS69	14	124	6	4								N	K									
G54	19F	63	7	4								N	K									
INV104B	1	227	8	4								N	K									
SP18-BS74	6	n/a	9	4		G					P		K									
Spain 23F^{-1}	23F	81	10	4		G					P		K									
SP9-BS68	9	1269	11	4																		
06_1011	14	9	12	5	Q	G							K									
SP23-BS72	23	37	13	5								N	K									
05_2426	19A	1201	14	5								N	K									
A66.1 Xen	3	180	15	5																		
H060160064	19A	199	16	6	Q	G							K									
06_2715	6A	65	17	7	E	D	K	T	R	Q		H	K	N	Q	A	L	G	K	V	D	S
H043640049	6A	65	18	7	E	D	K	T	R	Q		H	K	N	Q	A	L	G	K	V	D	S

											Amino acid position										
					303	309	311	322	323	324	327	329	332	334	336	337	392	393	395	397	404	406
Strain	**Serotype**	**ST**	**Allele**	**Group**	**K**	**K**	**S**	**A**	**L**	**T**	**T**	**I**	**S**	**R**	**G**	**K**	**D**	**R**	**T**	**T**	**K**	**S**
D39	2	125	1	1																		
05_1308	19A	199	2	2				**T**														
TIGR4	4	205	3	3												**N**						
SP6-BS73	6	460	4	4																		
SP11-BS70	11	62	5	4																		
SP14-BS69	14	124	6	4																		
G54	19F	63	7	4																		
INV104B	1	227	8	4													**N**					
SP18-BS74	6	n/a	9	4				**V**														
Spain 23F^{-1}	23F	81	10	4				**V**														**F**
SP9-BS68	9	1269	11	4												**N**						
06_1011	14	9	12	5				**V**										**K**	**V**	**S**	**E**	**K**
SP23-BS72	23	37	13	5														**K**	**V**	**S**	**E**	**K**
05_2426	19A	1201	14	5				**V**										**K**	**V**	**S**	**E**	**K**
A66.1 Xen	3	180	15	5												**N**		**K**	**V**	**S**	**E**	**K**
H060160064	19A	199	16	6				**V**										**K**	**V**	**S**	**E**	**K**
06_2715	6A	65	17	7	**T**	**E**	**G**	**K**	**I**	**S**	**E**	**V**	**G**	**M**	**N**							
H043640049	6A	65	18	7	**T**	**E**	**G**	**K**	**I**	**S**	**E**	**V**	**G**	**M**	**N**							

									Amino acid position													
					409	410	411	414	425	440	441	442	446	447	448	449	451	454	455	457	458	460
Strain	**Serotype**	**ST**	**Allele**	**Group**	**S**	**I**	**G**	**V**	**E**	**K**	**G**	**I**	**S**	**S**	**Q**	**K**	**E**	**K**	**K**	**D**	**G**	**T**
D39	2	125	1	1																		
05_1308	19A	199	2	2															**E**			
TIGR4	4	205	3	3																		
SP6-BS73	6	460	4	4															**E**			
SP11-BS70	11	62	5	4															**E**			
SP14-BS69	14	124	6	4															**E**			
G54	19F	63	7	4															**E**			**I**
INV104B	1	227	8	4																		
SP18-BS74	6	n/a	9	4																		
Spain 23F^{-1}	23F	81	10	4																		
SP9-BS68	9	1269	11	4																		
06_1011	14	9	12	5	**A**	**A**	**P**	**L**	**T**	**R**	**A**	**V**	**P**	**K**	**T**	**P**	**K**	**E**		**G**	**D**	
SP23-BS72	23	37	13	5	**A**	**A**	**P**	**L**	**T**	**R**	**A**	**V**	**P**	**K**	**T**	**P**	**K**	**E**		**G**	**D**	
05_2426	19A	1201	14	5	**A**	**A**	**P**	**L**	**T**	**R**	**A**	**V**	**P**	**K**	**T**	**P**	**K**	**E**		**G**	**D**	
A66.1 Xen	3	180	15	5	**A**	**A**	**P**	**L**	**T**	**R**	**A**	**V**	**P**	**K**	**T**	**P**	**K**	**E**		**G**	**D**	
H060160064	19A	199	16	6	**A**	**A**	**P**	**L**	**T**	**R**	**A**	**V**	**P**	**K**	**T**	**P**	**K**	**E**		**G**	**D**	
06_2715	6A	65	17	7															**E**			
H043640049	6A	65	18	7															**E**			

													Amino acid position								
					466	467	470	471	472	475	480	483	484	485	486	487	488	496	498	499	501	503
Strain	**Serotype**	**ST**	**Allele**	**Group**	**R**	**E**	**K**	**G**	**A**	**I**	**T**	**T**	**P**	**D**	**G**	**K**	**A**	**D**	**V**	**K**	**A**	**S**
D39	2	125	1	1																		
05_1308	19A	199	2	2				**E**														
TIGR4	4	205	3	3																		
SP6-BS73	6	460	4	4				**E**														
SP11-BS70	11	62	5	4				**E**														
SP14-BS69	14	124	6	4				**E**														
G54	19F	63	7	4				**E**														
INV104B	1	227	8	4																		
SP18-BS74	6	n/a	9	4																		
Spain 23F^{-1}	23F	81	10	4																		
SP9-BS68	9	1269	11	4																		
06_1011	14	9	12	5	**K**	**Q**	**S**		**H**	**V**	**E**	**N**	**A**	**Q**	**N**	**Q**	**K**	**N**	**T**	**E**	**G**	**R**
SP23-BS72	23	37	13	5	**K**	**Q**	**S**		**H**	**V**	**E**	**N**	**A**	**Q**	**N**	**Q**	**K**	**N**	**T**	**E**	**G**	**R**
05_2426	19A	1201	14	5	**K**	**Q**	**S**		**H**	**V**	**E**	**N**	**A**	**Q**	**N**	**Q**	**K**	**N**	**T**	**E**	**G**	**R**
A66.1 Xen	3	180	15	5	**K**	**Q**	**S**		**H**	**V**	**E**	**N**	**A**	**Q**	**N**	**Q**	**K**	**N**	**T**	**E**	**G**	**R**
H060160064	19A	199	16	6	**K**	**Q**	**S**		**H**	**V**	**E**	**N**	**A**	**Q**	**N**	**Q**	**K**	**N**	**T**	**E**	**G**	**R**
06_2715	6A	65	17	7																		
H043640049	6A	65	18	7																		

					Amino acid position																	
					507	512	513	515	521	522	523	524	525	526	530	532	533	552	557	558	559	560
Strain	**Serotype**	**ST**	**Allele**	**Group**	**D**	**N**	**Q**	**L**	**T**	**T**	**N**	**K**	**T**	**S**	**I**	**K**	**D**	**Q**	**M**	**V**	**K**	**A**
D39	2	125	1	1																		
05_1308	19A	199	2	2		**D**																
TIGR4	4	205	3	3		**D**																
SP6-BS73	6	460	4	4		**D**																
SP11-BS70	11	62	5	4		**D**																
SP14-BS69	14	124	6	4		**D**																
G54	19F	63	7	4		**D**																
INV104B	1	227	8	4																		
SP18-BS74	6	n/a	9	4		**D**																
Spain 23F^{-1}	23F	81	10	4																		
SP9-BS68	9	1269	11	4		**D**																
06_1011	14	9	12	5	**N**	**Q**	**E**	**I**	**A**	**H**	**S**	**T**	**K**	**N**	**V**	**N**	**T**	**R**	**G**	**L**	**R**	**K**
SP23-BS72	23	37	13	5	**N**	**Q**	**E**	**I**	**A**	**H**	**S**	**T**	**K**	**N**	**V**	**N**	**T**	**R**	**G**	**L**	**R**	**K**
05_2426	19A	1201	14	5	**N**	**Q**	**E**	**I**	**A**	**H**	**S**	**T**	**K**	**N**	**V**	**N**	**T**	**R**	**G**	**L**	**R**	**K**
A66.1 Xen	3	180	15	5	**N**	**Q**	**E**	**I**	**A**	**H**	**S**	**T**	**K**	**N**	**V**	**N**	**T**	**R**	**G**	**L**	**R**	**K**
H060160064	19A	199	16	6	**N**	**Q**	**E**	**I**	**A**	**H**	**S**	**T**	**K**	**N**	**V**	**N**	**T**	**R**	**G**	**L**	**R**	**K**
06_2715	6A	65	17	7		**D**																
H043640049	6A	65	18	7		**D**																

												Amino acid position										
					568	599	606	641	706	720	787	792	795	800	801	811	837	839	844	850	853	856
Strain	**Serotype**	**ST**	**Allele**	**Group**	**V**	**N**	**I**	**R**	**E**	**K**	**D**	**D**	**S**	**K**	**V**	**----**	**R**	**M**	**T**	**T**	**S**	**M**
D39	2	125	1	1																		
05_1308	19A	199	2	2		D	V															
TIGR4	4	205	3	3		D	V				E	N		N	*							
SP6-BS73	6	460	4	4		D	V				E	N				INS	T	L	S	A	K	I
SP11-BS70	11	62	5	4			V	K				N				INS	T	L	S	A	K	I
SP14-BS69	14	124	6	4			V		N	N	E		F			INS	T	L	S	A	K	I
G54	19F	63	7	4			V		N							INS	T	L	S	A	K	I
INV104B	1	227	8	4			V		N		E					INS	T	L	S	A	K	I
SP18-BS74	6	n/a	9	4		D	V	K			E	N				INS	T	L	S	A	K	I
Spain 23F^{-1}	23F	81	10	4			V	K	N		E					INS	T	L	S	A	K	I
SP9-BS68	9	1269	11	4			V	K			E					INS	T	L	S	A	K	I
06_1011	14	9	12	5	T		V	E								INS	T	L	S	A	K	I
SP23-BS72	23	37	13	5	T		V	E								INS	T	L	S	A	K	I
05_2426	19A	1201	14	5	T		V	E								INS	T	L	S	A	K	I
A66.1 Xen	3	180	15	5	T		V	K								INS	T	L	S	A	K	I
H060160064	19A	199	16	6	T	D	V					N				INS	T	L	S	A	K	I
06_2715	6A	65	17	7			V				E	N				INS	T	L	S	A	K	I
H043640049	6A	65	18	7			V				E	N				INS	T	L	S	A	K	I

										Amino acid position												
					859	860	861	863	865	867	877	878	881	882	883	884	885	887	891	898	912	967	1020
Strain	**Serotype**	**ST**	**Allele**	**Group**	**K**	**V**	**T**	**L**	**E**	**A**	**S**	**V**	**T**	**K**	**L**	**S**	**N**	**M**	**E**	**3**	**G**	**T**	**L**
D39	2	125	1	1																**3**			
05_1308	19A	199	2	2																**2**			
TIGR4	4	205	3	3																**----**			
SP6-BS73	6	460	4	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
SP11-BS70	11	62	5	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**		**I**	
SP14-BS69	14	124	6	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
G54	19F	63	7	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
INV104B	1	227	8	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
SP18-BS74	6	n/a	9	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
Spain 23F^{-1}	23F	81	10	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
SP9-BS68	9	1269	11	4	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
06_1011	14	9	12	5	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
SP23-BS72	23	37	13	5	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
05_2426	19A	1201	14	5	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			
A66.1 Xen	3	180	15	5		**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**	**R**		
H060160064	19A	199	16	6	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**		**I**	
06_2715	6A	65	17	7	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			**P**
H043640049	6A	65	18	7	**E**	**I**	**I**	**I**	**K**	**S**	**N**	**L**	**A**	**R**	**V**	**P**	**G**	**V**	**K**	**0**			**P**

Figure 4.2: Mutations in NanA amino acid alleles

Mutations present in 18 amino acid alleles of NanA discovered in diversity screen. Numbers in row 1 represent amino acid numbers in wild-type allele 1, with row 2 showing amino acids present at these positions in the wild-type genes. Mutated amino acids are shown below these in their respective alleles. STs noted as n/a were undefined. Abbreviation INS indicated region of insertion. At amino acid position 898, the number represents the number of 20AA repeats present in the allele.

4.3 NanA alleles mapped onto structural model

The structure of a NanA fragment (AA318-792) from TIGR4 (allele 3) was solved by collaborators (Drs. Heinz Gut and Martin Walsh, MRC France, Grenoble, France). This represented the neuraminidase enzymatic 6-bladed ß-propeller, circled in red, as well as an insertion region of unknown origin, circled in green (Figure 4.3). This fragment contained the enzymatically active part of the protein, and allowed the mapping of mutations present in discovered NanA protein alleles onto a structural model. Results gave insight into the positions of mutations on the NanA structure, allowing predictions regarding mutations which may affect the activity of the enzyme. The mutations present in alleles 1-18 were mapped onto the structural model using PyMol software (DeLano Scienfitic, USA), and results can be found in Figure 4.4. Mutations were in comparison to D39 and not TIGR4, due to the frameshift mutation present later in the gene. As the solved region only represents amino acids 318-792, the mutations outside this region were not mapped.

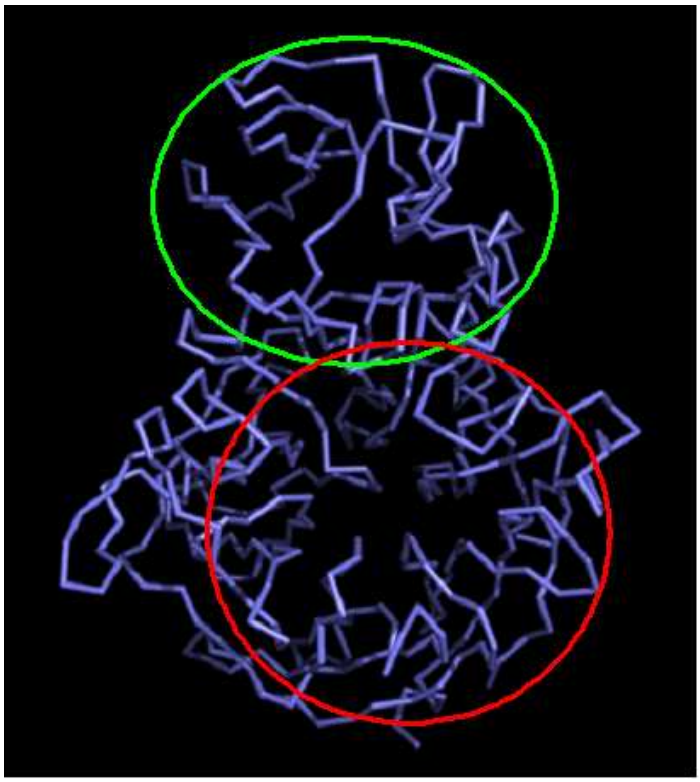
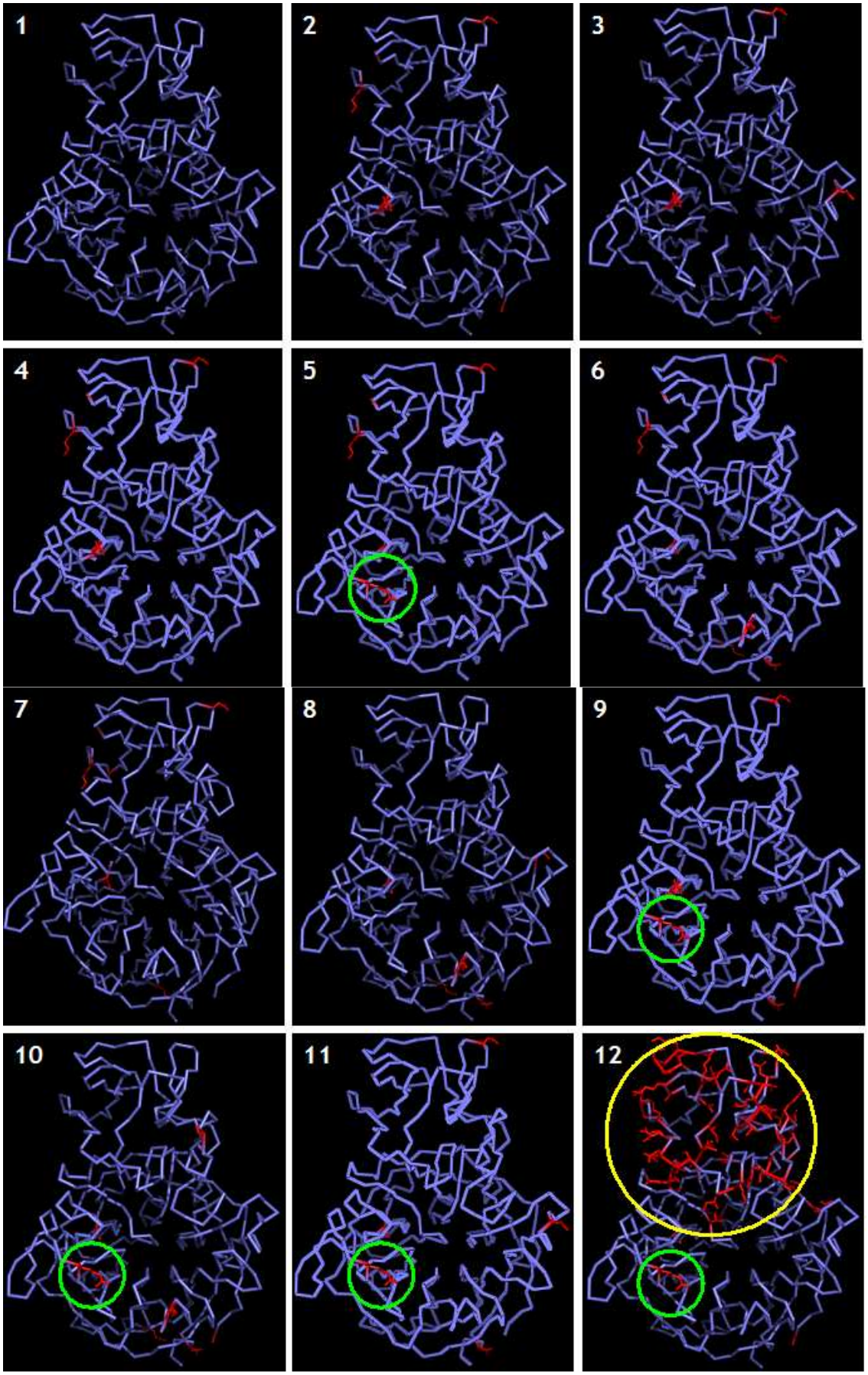


Figure 4.3: Model of NanA crystal structure

NanA protein fragment structure model, with enzymatic 6-bladed β-propeller circled in red and insertion region circled in green.


1
2
3
4
5
6
7
8
9
10
11
12

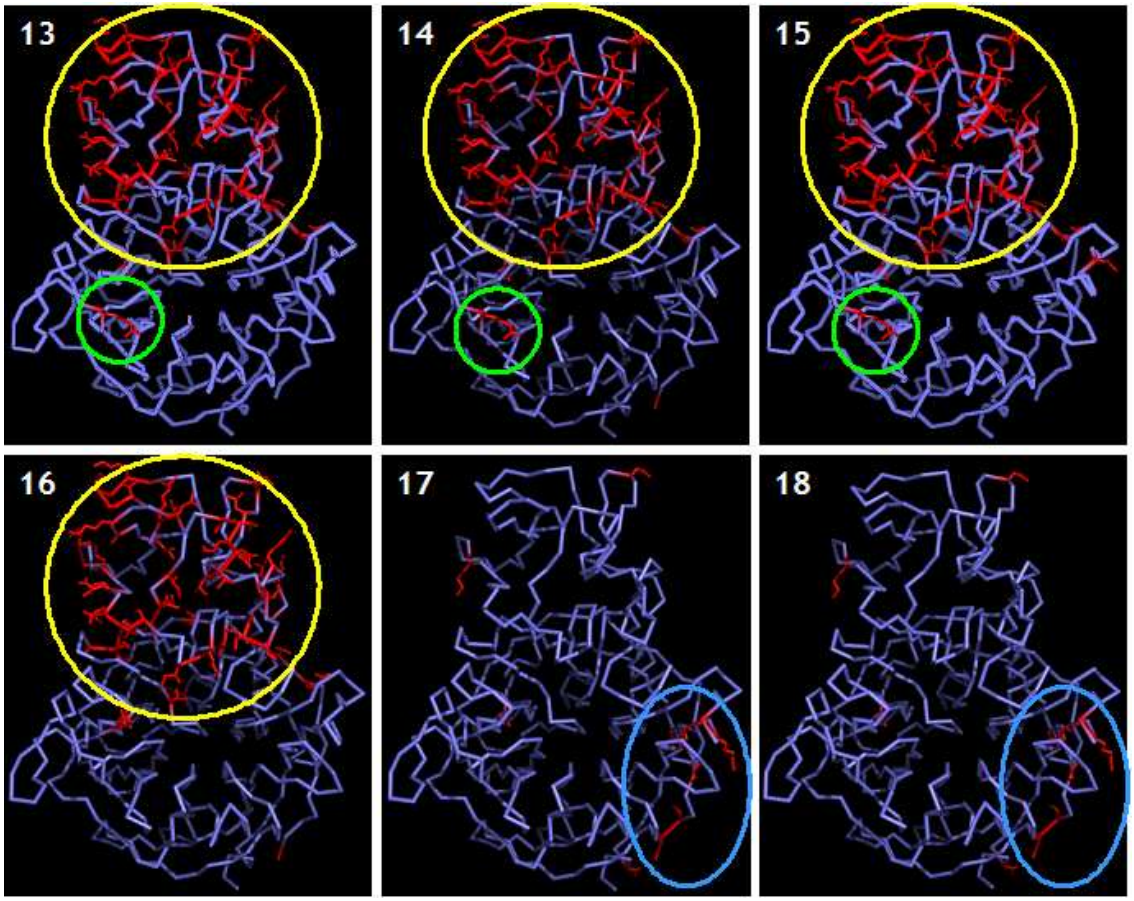


Figure 4.4: NanA mutations mapped onto crystal structure

NanA mutations mapped onto structural model of enzymatic region of the protein (318-792). Regions highlighted in red represent amino acid positions that are altered in a particular allele compared to D39. Circled in green is amino acid position 641, circled in blue is the end of mosaic block A, circled in yellow is the distinct variant of mosaic block C, the inserted region.

Results showed that a variety of point mutations were present within this region (Figure 4.4). In particular, a mutation discovered at amino acid 641 was hypothesised to affect the activity of the enzyme due to proximity to the entrance of the active ß-propeller potentially affecting substrate recognition (Figure 4.4, circled in green). Mutation at this amino acid position occurred in 8/18 NanA alleles, where an arginine residue was substituted with either a lysine or glutamic acid residue. A small section of mosaic block A was present at the start of the fragment in 2/18 alleles, circled in blue in Figure 4.4. Interestingly, mosaic block C mapped to the region inserted into the ß-propeller, and is circled in yellow in 5/18 alleles in Figure 4.4. Although point mutations existed in this region in some alleles, there were two distinct variants of this region present. These two variants varied markedly from each other, sharing only 57% sequence identity. The mosaic C block variant region was BLASTed against the nucleotide

database (http://blast.ncbi.nlm.nih.gov/Blast) to discover sequences related to this variant of unknown origin, as previous work had failed to determine homologues to this sequence (King *et al.*, 2005). The version of the region not present in the wild-type allele was found to share most sequence identity (66%) with a sialidase gene from *S. agalactiae*.

In order to determine the effect of this stark difference on electrostatic potential of this region, a homology model was constructed. The model, constructed using PyMol software (DeLano Scientific, USA), compared alleles 1 and 16, as each possessed a different variant of this region. The colours in the diagram represented electrostatic potentials, so regions sharing red colour had similar negative charges, and regions sharing blue colours shared positive charge. Comparison of electrostatic potentials of insert regions of alleles 1 and 16 can be seen in Figure 4.5A and B, with 90º right rotations of the respective regions in Figure 4.5C and D. Results showed that major differences in electrostatic potential, viewed as colour differences, existed between the two variants of mosaic block C, the inserted region (circled in green). Furthermore, Figure 4.5 showed that the structural shape of the region was significantly altered by the amino acid differences observed between these variants.

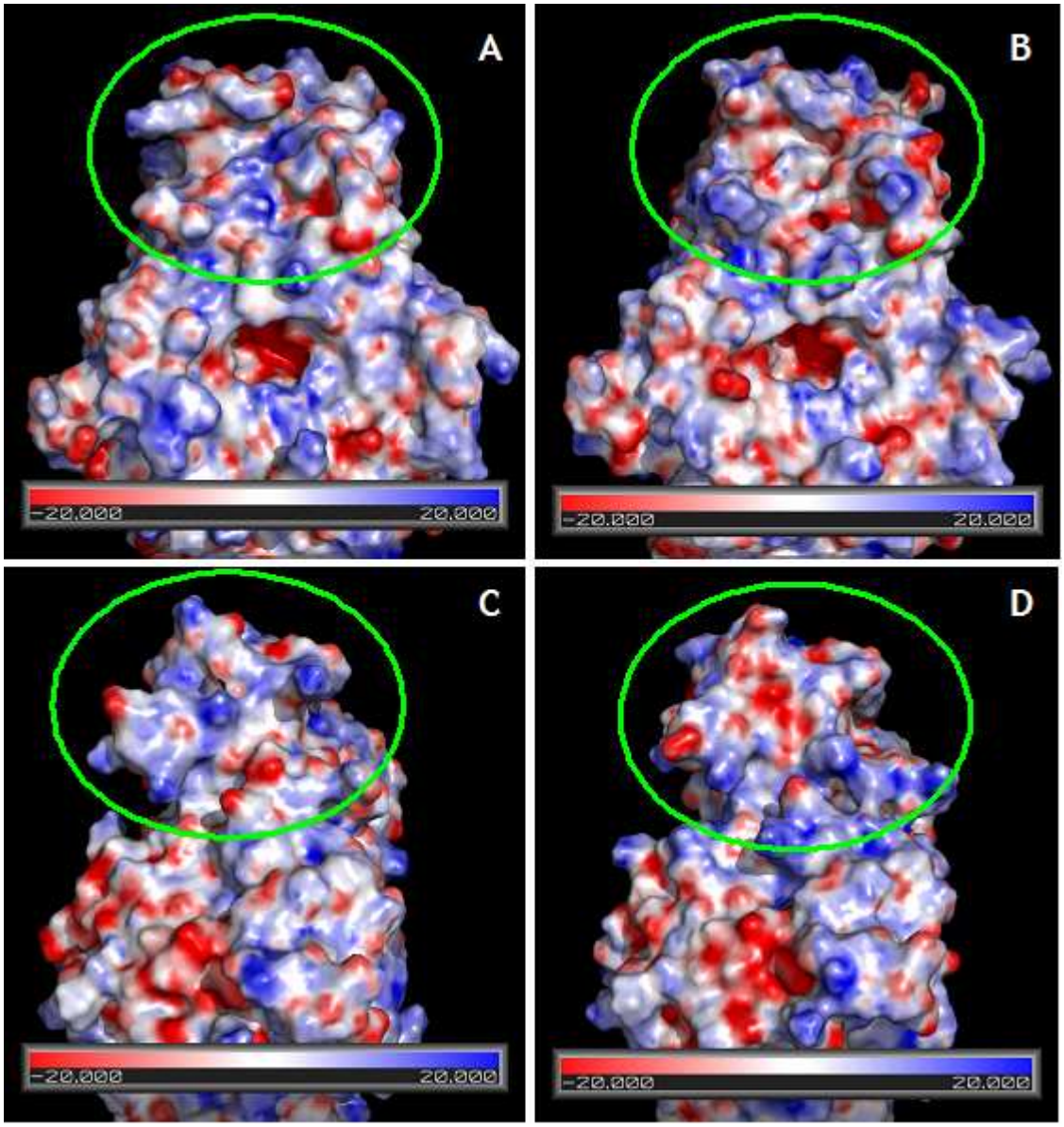


Figure 4.5: Homology model of NanA insertion region

Homology model of region of unknown function inserted in enzymatic β–propeller of NanA, showing shape and electrostatic potential of two distinct variants of the region inserted into the NanA β-propeller. Figure 4.5A represents the region in allele 1, with the same model at 90º rotation to the right in Figure 4.5C. Similarly, Figure 4.5B represents the region in allele 16, with the same model at 90º rotation to the right in Figure 4.5D.

4.4 NanA activity of purified NanA active fragments of different sizes

The NanA protein possesses a lectin-binding domain located before the enzyme region in the amino acid sequence. In order to assess the effect of the lectin-binding domain at the start of the NanA protein on overall neuraminidase activity, the activity of the NanA allele 3 purified fragment (AA318-792) was compared to a larger fragment of the same allele (AA65-792, created by Dr. Heinz Gut), which included the lectin-binding domain of the NanA protein. The plot from the neuraminidase assay (Figure 4.6) was used to calculate specific activity for the two enzyme fragments. To adjust for the different molecular weights, concentrations were calculated in mM. NanA 318-792 had specific activity of $3.43x10^2$ enzymatic units/mM (EU/mM), whilst NanA 65-792 had specific activity of $2.47x10^2$ EU/mM. The similarity of these results implied that the lectin-binding domain of the NanA protein had no effect on neuraminidase activity. Furthermore, this result showed that identical fragments of region 318-792 purified from further NanA alleles should have activities comparable to their full-length proteins and allow comparison between activity of NanA alleles.

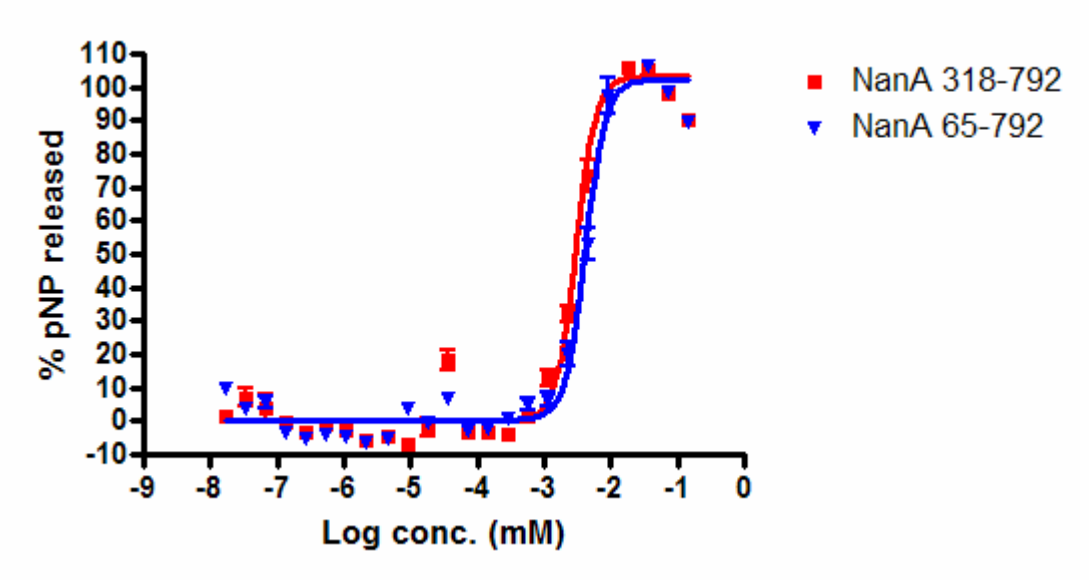


Figure 4.6: Effect of lectin domain on NanA activity

NanA activity curves of 2 purified fragments of different sizes, one with the lectin domain, one without. The two fragments represent NanA AA318-792 (red) and NanA AA65-792 (blue). Specific enzymatic activity was calculated from plots, in enzymatic units/mM.

4.5 Studies of purified fragments of NanA alleles

4.5.1 Cloning, expression and purification

NanA alleles 12, 16 and 18 were chosen for cloning, expression and purification of the enzymatic fragment (AA318-792) in order to compare activity directly to wild-type TIGR4 allele 3 activity. The alleles purified were from isolates H053940080, 1427 and H043640049 respectively. These alleles were selected due to major amino acid differences present in comparison to the wild-type protein. Allele 12 possessed the highly divergent variant of the mosaic block C, as well as a mutation of residue R641E, and A322V and I606V mutations. Allele 16 had the variant of the mosaic block C, as well as A322V, N599D, I606V and D792N mutations, but lacked the R641E mutation. Allele 18 had the last 8 mutations from mosaic block A as well as K455E, N512D I606V, D787E and D792N mutations. Comparison of the activities of these alleles should allow determination of the effect of the variant mosaic block C and block A regions and the R641E mutation on NanA activity.

Protein fragments were cloned and expressed in both B834 and Rosetta expressing cell lines. Crude lysates of clones were run on SDS-PAGE gel after

small-scale overnight expression to confirm successful production of NanA protein fragment. Results showed that whilst Rosetta cells produced a large band at the correct size (55kDa, circled in red), no band was observed with B834 cells (Figure 4.7). Rosetta cells were therefore used for purification of NanA fragments.

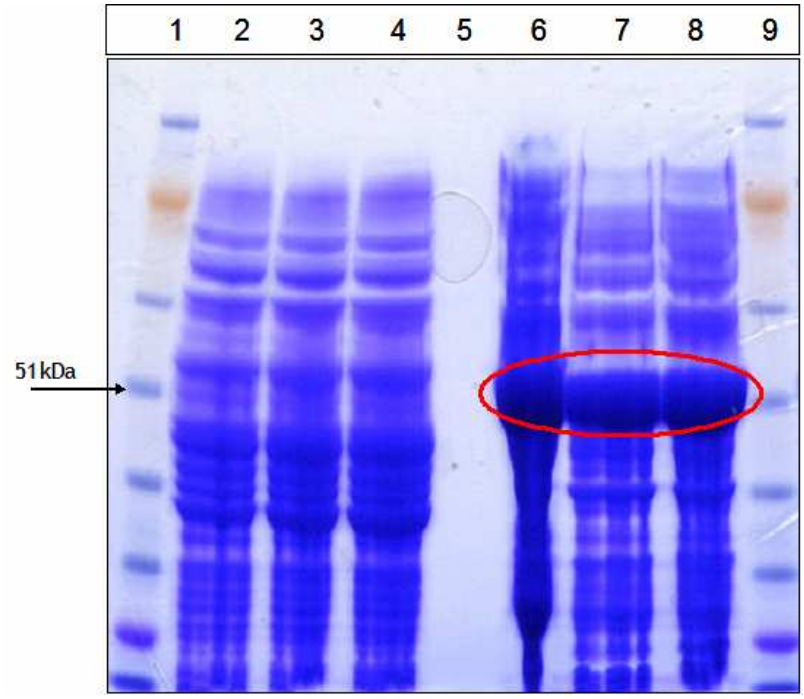


Figure 4.7: Expression of NanA fragments

SDS-PAGE gel of crude lysates from NanA-expressing clones. Gel layout: lane 1 – SeeBlue +2 prestained marker; lane 2 – NanA allele 12 in B834; lane 3 – NanA allele 16 in B834; lane 4 – NanA allele 18 in B834; Lane 5 – blank; lane 6 – NanA allele 12 in Rosetta; lane 7 – NanA allele 16 in Rosetta; lane 8 – NanA allele 18 in Rosetta; lane 9 – SeeBlue +2 prestained marker. Expressed NanA fragment circled in red.

Samples from all stages of purification were collected and tested for presence and purity of NanA fragment by SDS-PAGE. Results of purification of allele 12 show that C3 cleavage and gel filtration steps were required to produce purified NanA at a concentration of 3.06mg/ml (Figure 4.8). In comparison, only C3 cleavage was required to purify NanA allele 16 at a concentration of 16.42mg/ml (Figure 4.9). C3 cleavage and gel filtration steps were required to produce purified NanA allele 18 at a concentration of 2.95mg/ml (Figure 4.10).

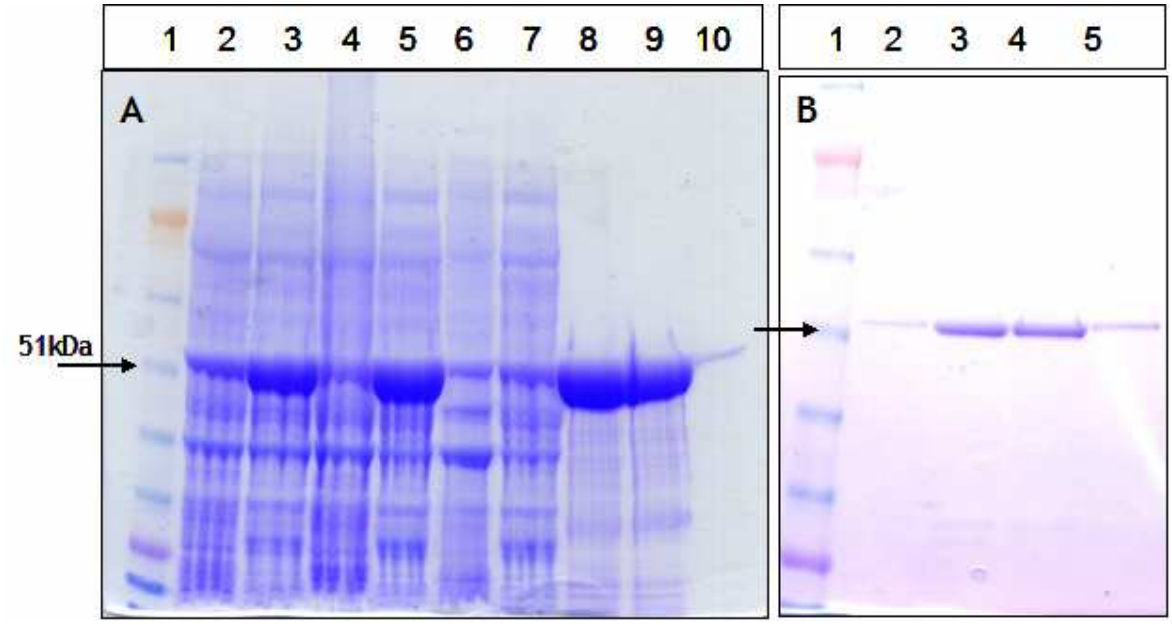


Figure 4.8: NanA allele 12 fragment

Purification of NanA allele 12 fragment AA318-792. SDS-PAGE gels of samples collected during purification. Gel A shows samples from initial column purification and C3 cleavage. Gel B shows fractions 34-37 collected from gel filtration. Gel A layout: Lane 1 – SeeBlue +2 prestained marker; lane 2 – crude lysate; lane 3 – spin 1 supernatant; lane 4 – spin 1 pellet; lane 5 – spin 2 supernatant; lane 6 – spin 2 pellet; lane 7 – column flow-through; lane 8 – column elution fraction; lane 9 – C3 cleaved flow-through; lane 10 – C3-cleaved elution fraction. Gel B Layout: lane 1 - SeeBlue +2 prestained marker; lane 2 – fraction 34; lane 3 – fraction 35; lane 4 – fraction 36; lane 5 – fraction 37.

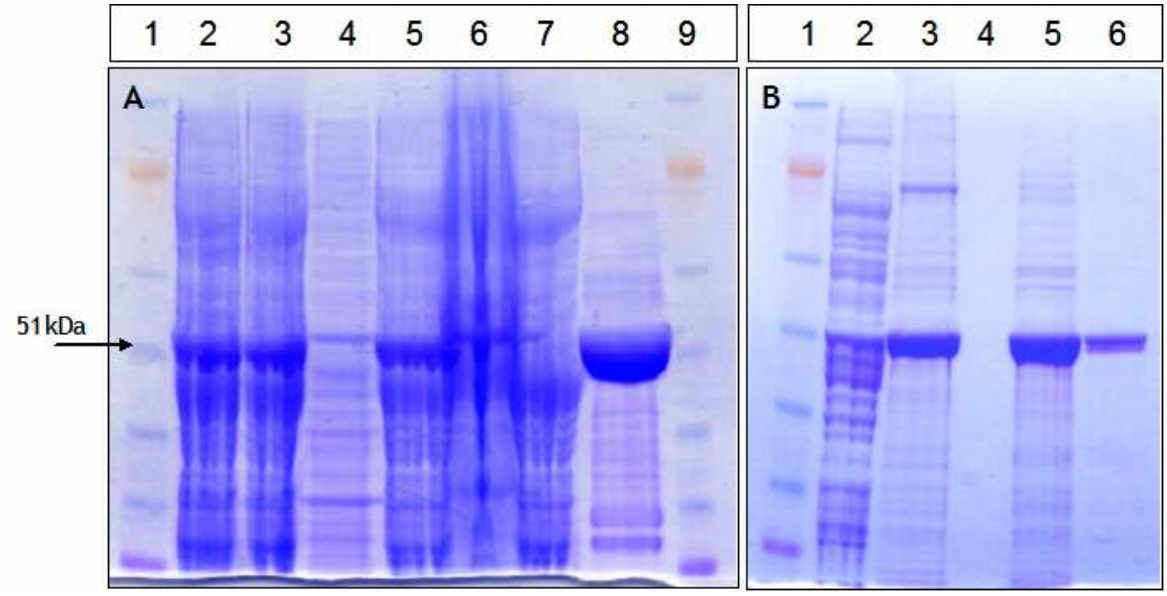


Figure 4.9: NanA allele 16 fragment

Purification of NanA allele 16 fragment AA318-792. SDS-PAGE gels of samples collected during purification. Gel A shows samples from initial column purification. Gel B shows samples from second column purification after C3 protease cleavage. Gel A layout: Lane 1 – SeeBlue +2 prestained marker; lane 2 – crude lysate; lane 3 – spin 1 supernatant; lane 4 – spin 1 pellet; lane 5 – spin 2 supernatant; lane 6 – spin 2 pellet; lane 7 – column flow-through; lane 8 – column elution fraction; lane 9 – SeeBlue +2 prestained marker. Gel B layout: lane 1 – SeeBlue +2 prestained marker; lane 2 – as gel A, lane 7; lane 3 – as gel A, lane 8; lane 4 – blank; lane 5 – column elution fraction; lane 6 – column flow-through – shows purified NanA fragment.

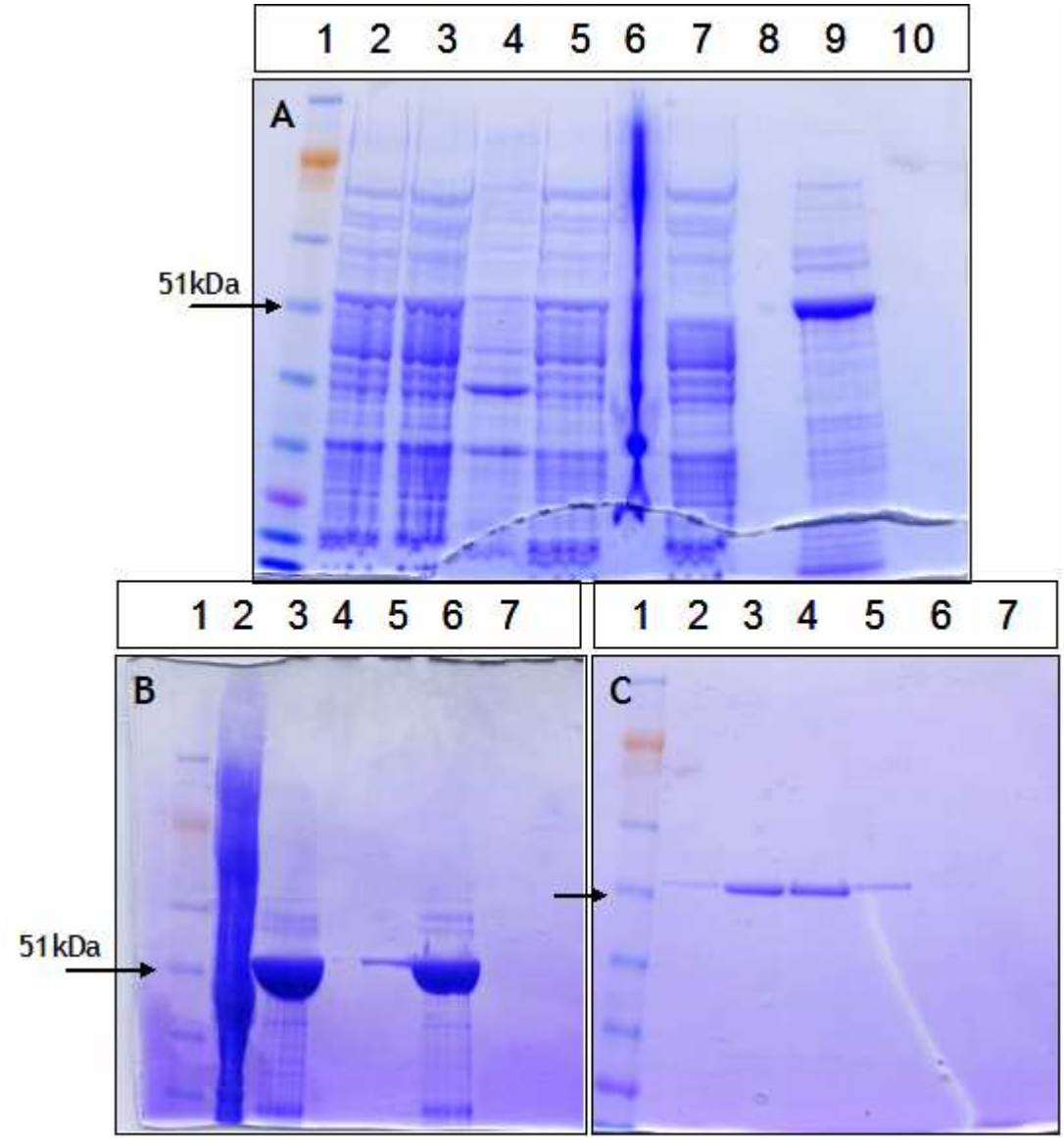


Figure 4.10: NanA allele 18 fragment

Purification of NanA allele 18 fragment AA318-792. SDS-PAGE gels of samples collected during purification. Gel A shows samples from initial column purification. Gel B shows samples from second column purification after C3 protease cleavage. Gel C shows fractions 36-40 collected after gel filtration. Gel A layout: Lane 1 – SeeBlue +2 prestained marker; lane 2 – crude lysate; lane 3 – spin 1 supernatant; lane 4 – spin 1 pellet; lane 5 – spin 2 supernatant; lane 6 – spin 2 pellet; lane 7 – column flow-through; lane 8 – blank; lane 9 – column elution fraction; lane 10 – blank. Gel B layout: lane 1 – SeeBlue +2 prestained marker; lane 2 – as gel A, lane 7; lane 3 – as gel A, lane 9; lane 4 – blank; lane 5 – column elution fraction; lane 6 – column flow-through – shows NanA fragment. Gel C layout: lane 1 – SeeBlue +2 prestained marker; lane 2 – fraction 36; lane 3 – fraction 37; lane 4 – fraction 38; lane 5 – fraction 39; lane 6 – fraction 40; lane 7 – blank.

4.5.2 Neuraminidase activity of purified NanA fragments

The neuraminidase activity of the purified NanA allele fragments was measured using the neuraminidase assay, and compared to the wild-type TIGR4 purified fragment by creating plots of activity (Figure 4.11). Specific activities were calculated from the plots, and results showed that alleles had activities varying from $5.29x10^3$ to $4.03x10^4$ enzymatic units/mg (Table 4.3). The most active NanA allele was allele 18, which showed activity 7.62-fold higher than TIGR4. Fragments of alleles 12, 16 and 18 all had activity higher than the TIGR4 allele 3 fragment, and therefore differences were observed in NanA activity between alleles.

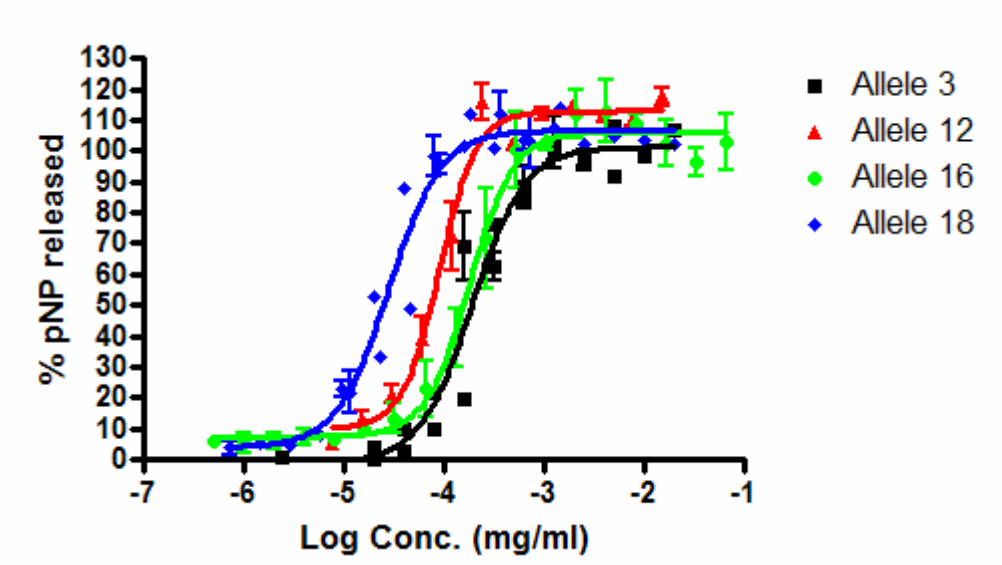


Figure 4.11: NanA activity plots of purified fragments

Plots represent concentration of purified NanA allele fragments against percentage of pNP released during assay at t=30 min. Specific activity of enzyme fragments were calculated by taking the reciprocal of the concentration required to release 50% of pNP from the assay.

Strain	NanA allele	Allele group	Specific activity	Ratio
TIGR4	3	3	5.29E+03	1
H053940080	12	5	1.35E+04	2.55
1427	16	6	6.40E+03	1.21
H043640049	18	7	4.03E+04	7.62

Table 4.3: Specific activity of NanA alleles

Specific activities of NanA allele fragments calculated in pNP neuraminidase assay, and measured in enzymatic units/mg purified protein. Ratios show the fold-increase in activity compared to the TIGR4 purified fragment.

4.5.3 Recognition of NanA alleles by polyclonal antibodies

One method of assessing the potential cross-serotype protection afforded by a protein vaccine targeting NanA is to assess the recognition of different NanA alleles by polyclonal antibodies raised against wild-type NanA allele 1. In order to do this, a Western blot was carried out with the 4 purified NanA fragments, representing alleles 3, 13, 16 and 18, and probed with α-NanA polyclonal antibodies raised in rabbit against NanA allele 1. Results show that all NanA alleles are recognised by these polyclonal antibodies (Figure 4.12). This result has important implications as it shows that the high levels of diversity observed in the NanA protein may not act as a barrier for cross-serotype protection of a vaccine targeting this virulence factor. An interesting observation is that alleles 12 and 16 appear to have increase elctrophoretic mobility compared to alleles 3 and 18. Interestingly, both of these alleles possess the variant version of mosaic block C, and this divergent sequence may result in a difference in electrophoretic mobility similar to that seen in Ply alleles 3, 5 and 10.

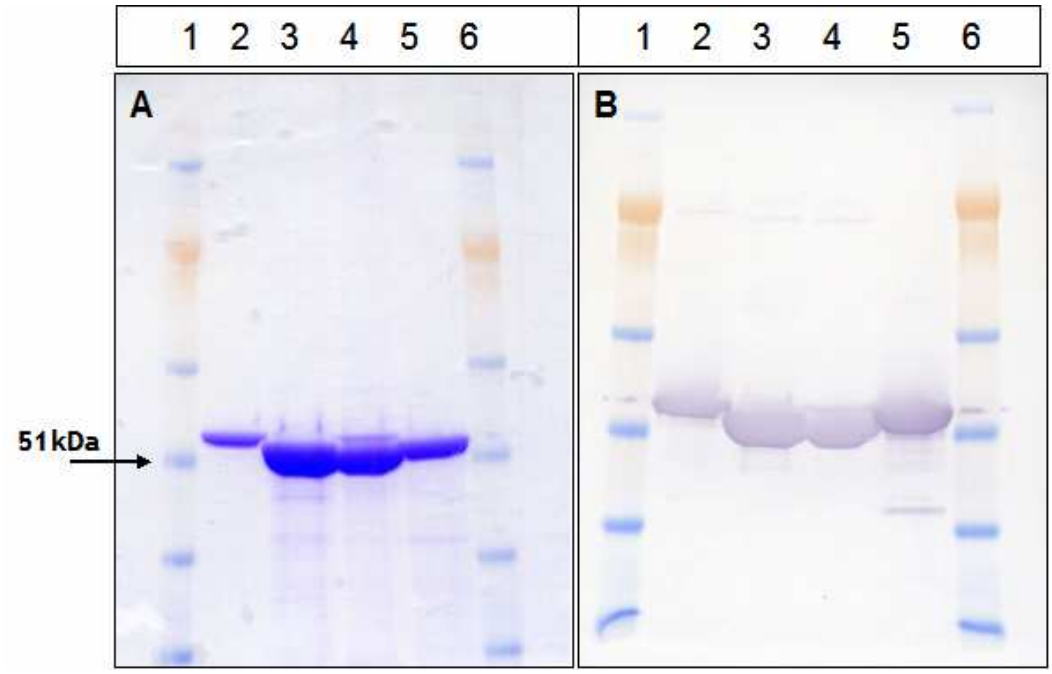


Figure 4.12: Recognition of NanA proteins by polyclonal antibodies

SDS-PAGE (A) and Western blot (B) of 4 purified NanA allele fragments. In both the gel and the membrane, identity of lanes are as follows: lane 1 – SeeBlue +2 protein marker; lane 2 – NanA allele 3; lane 3 – NanA allele 12; lane 4 – NanA allele 16; lane 5 – NanA allele 18; lane 6 - SeeBlue +2 protein marker.

4.6 Exploring the importance of NanA in progression of p-HUS

4.6.1 Comparison of NanA protein alleles in p-HUS isolates and matched controls

The pneumococcus is increasing as a cause of HUS, associated with high levels of mortality (Fitzpatrick *et al.*, 1993). Neuraminidase activity has been implicated as important in progression of p-HUS (Klein *et al.*, 1977; Seger *et al.*, 1980). The NanA protein alleles from 9 p-HUS isolates (Waters *et al.*, 2007) and 6 controls, sharing serotype and ST but not causing p-HUS were aligned and compared. This was to determine whether a particular NanA allele was associated with isolates able to cause p-HUS. Results show that NanA alleleswere generally associated with STs, and therefore no particular NanA allele was shared by the isolates causing p-HUS (Table 4.5).

Strain	Serotype	ST	Disease	NanA allele
02_1198	3	180	Neonatal death	1
H050600025	3	180	p-HUS	1
05_1308	19A	199	Pneumonia	2
05_2084	19A	199	Pneumonia	2
H051740086	19A	199	p-HUS	2
H052300328	19A	199	p-HUS	2
06_1011	14	9	Pneumonia	12
H053940080	14	9	p-HUS	12
H053940083	14	9	p-HUS	12
05_2426	19A	1201	Pneumonia	14
H050940049	19A	1201	p-HUS	14
H060160064	19A	199	p-HUS	16
H040400333	19A	199	p-HUS	16
06_2715	6A	65	Pneumonia	17
H043640049	6A	65	p-HUS	18

Table 4.4: NanA identitiy of p-HUS isolates

NanA alleles assigned to p-HUS isolates and matched controls. For full alignment of NanA alleles, see Chapter 4. Clinical records showed that control isolates did not cause p-HUS.

4.6.2 Comparison of NanA activity in p-HUS isolates and matched controls

Since no particular NanA allele was associated with isolates causing p-HUS, it was hypothesised that p-HUS isolates may display higher levels of neuraminidase activity as a result of over-expression of the protein. Overall neuraminidase activity was measured in p-HUS isolates and matched controls. Isolates were either grown un-induced or induced by the substrate N-acetlyneuraminic acid (Neu5ac). Results were separated into secreted and cell-wall associated neuraminidase activity, and showed that neuraminidase activity was consistently higher in the secreted fraction, implying that more NanA is secreted by the pneumococcus than is attached to the cell wall. A knock-out of NanA was included in the assay to determine the neuraminidase activity of related enzymes NanB and NanC in the assay, and results showed that knocking out NanA had a significant decrease on overall neuraminidase activity ($p=0.0022$), and therefore these enzymes have minimal activity in comparison to NanA in this assay. When looking at individual isolates, although differences were present in neuraminidase activity, these did not correlate with p-HUS causing ability, whether induced or un-induced (Figure 4.13). These results also showed that after induction of neuraminidase activity during growth, the amount of NanA

secreted increased significantly ($p<0.0001$), whilst the neuraminidase attached to the cell wall showed a smaller increase, which was still significant ($p=0.0013$). When comparing these expression differences, the increase in secreted NanA was over 12-fold larger than the increase in cell wall-associated NanA, representing a statistically significant difference ($p=0.0019$). However, when overall neuraminidase activity was calculated, there was no significant difference in overall activity between p-HUS isolates and controls when induced ($p=0.7756$) or un-induced ($p=0.9546$). Taken together, these results show that there is no over-expression of NanA enzyme in p-HUS isolates, and that these isolates do not have an increased ability to be induced by presence of a substrate.

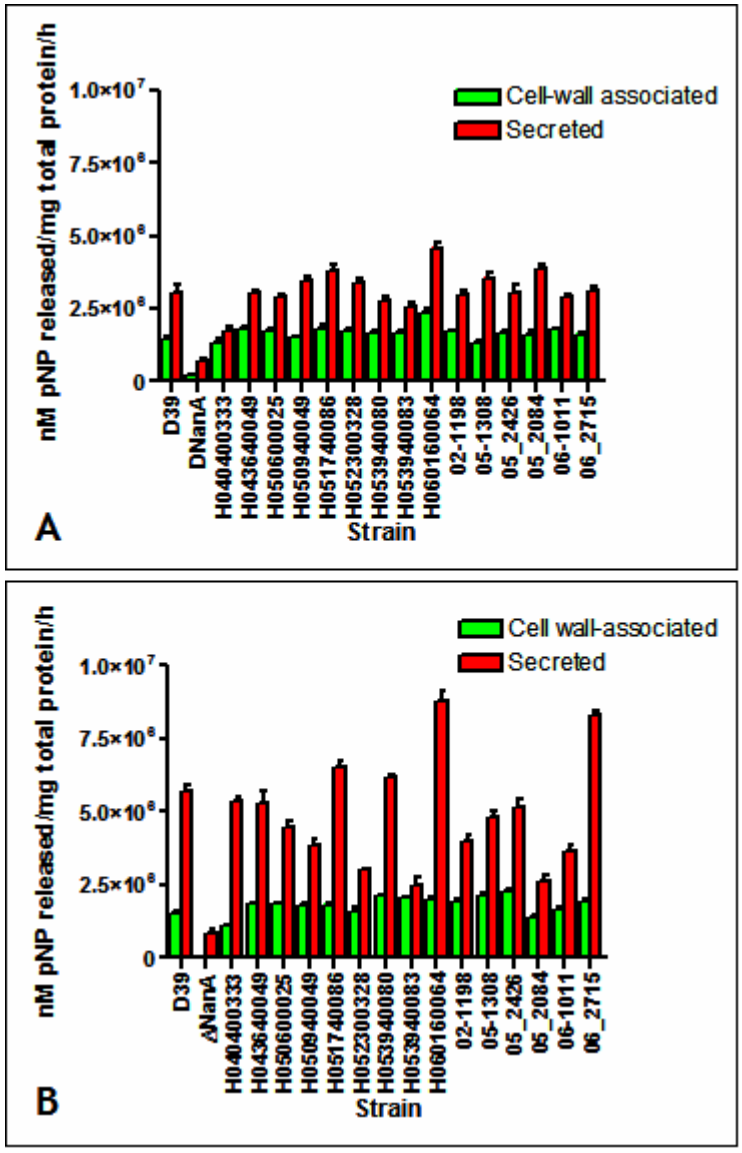


Figure 4.13: Neuraminidase activity of p-HUS isolates

Results of neuraminidase assay of isolates from p-HUS study, either un-induced (A) or induced with the substrate Neu5ac (B). 15 Test isolates were included in the assay as well as positive control D39, and NanA knock-out strain D39 ΔNanA. Results were separated into neuraminidase activity from cell-wall associated enzyme (green) or secreted enzyme (red). Results are calculated from triplicate repeats in each isolate.

4.6.3 The effect of NanA on the cytotoxicity of Ply

NanA has been shown to cleave sialic acid from a variety of host cells, including erythrocytes. It was hypothesised that the action of NanA may increase the susceptibility of erythrocytes to the pneumococcal toxin Ply, possibly resulting in the haemolytic anaemia observed in p-HUS. However, results showed that there was no difference in haemolytic activity of wild-type allele 1 Ply when erythrocytes were incubated with PBS, 6.38µg/ml or 63.8µg/ml NanA allele 3 (Figure 4.14). Therefore, the cleaving of sialic acid from glycolipids on the surface of erythrocytes has no effect on the haemolytic activity of the toxin Ply, and a synergism between these two virulence factors is not involved in the haemolytic anaemia observed in p-HUS patients.

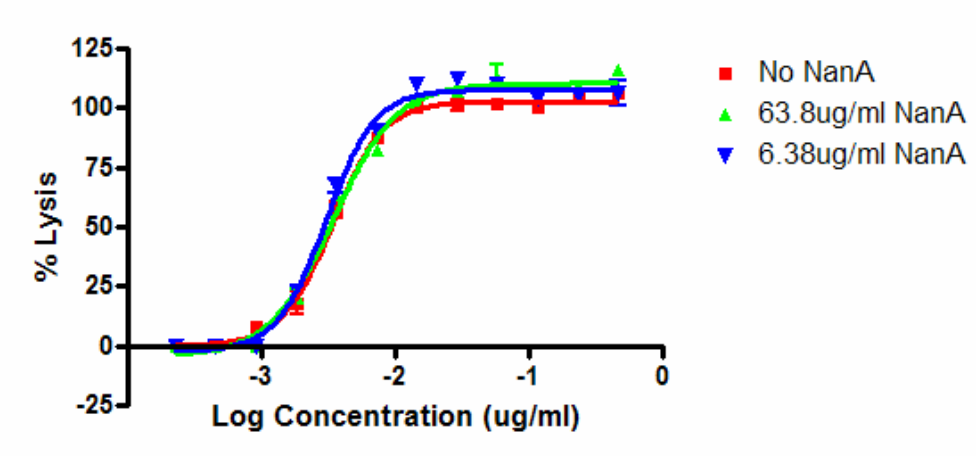


Figure 4.14: Effect of NanA on cytotoxic activity of Ply

Plot shows curves of % lysis of erythrocytes by Ply after pre-incubation of erythrocytes with differing concentrations of NanA.

4.7 Discussion

An analysis of diversity present in the NanA gene of *S. pneumoniae* revealed high levels of sequence diversity. Mutations are present at 142 amino acid positions, representing 13.7% of the amino acids in the wild type protein. This was a similar level of diversity as found in a previous study of the *nanA* gene (King *et al.*, 2005). Furthermore, diversity was increased by the presence of regions of insertion and repeat across the protein. In comparison to Ply, where 3.3% sequence diversity was present at the amino acid level, this represents a significantly higher level of variation. This was expected due to the surface-exposed nature of the NanA protein (Camara *et al.*, 1994), presumably resulting in selective pressure exerted by exposure to the host immune system. Despite the high levels of diversity present, all residues previously identified as essential (Camara *et al.*, 1994; Chong *et al.*, 1992; Yesilkaya *et al.*, 2006) were fully conserved across all protein alleles. However, the high levels of variation present within the NanA protein raise questions regarding use of this protein in a potential vaccine, as protection elicited by a vaccine targeting one NanA allele may not kill pneumococci possessing other alleles. Despite this, previous studies have shown that immunization with NanA protects against nasopharyngeal colonisation (Tong *et al.*, 2005) and otitis media (Long *et al.*, 2004) in a chinchilla model. Crucially, however, animals in the study were challenged with the same strains of *S. pneumoniae* from which the NanA was purified for immunization. Therefore, these results do not take into account the effect the diversity present in the NanA protein would have on immunogenicity and coverage of a NanA-based protein vaccine. One method for evaluating the potential protection elicited to NanA alleles is to determine whether polyclonal antibodies rasied against one NanA allele are capable of recognising other NanA alleles possesing significant diversity. This would determine whether antibodies raised by a vaccine targeting one particular NanA allele would recognise other NanA alleles, and thus potentially afford cross-serotype protection. The results of this study show for the first time that polyclonal antibodies raised against NanA allele 1 are capable of recognising the purified fragments of 4 significantly divergent alleles (Figure 4.12). This result shows that the high levels of diversity observed within the NanA protein should not act as a barrier to the goal of cross-serotype protection when targeting this virulence factor with a protein-based pneumococcal vaccine.

The structure of the enzymatic 6-barrel ß-propeller of NanA (Figure 4.3), solved by collaborators, allowed further study into the positioning of mosaic blocks and point mutations on the protein, and effects this may have on the activity of the NanA enzyme. Mutations were compared to D39 NanA despite the purified fragment being from TIGR4. This was to be concurrent with the full allele mutation analysis, as the NanA from TIGR4 possesses a frameshift inducing a stop mutation at AA801. Mapping of the mutations present in each protein allele onto the structural model showed that a number of mutations were positioned close to the active site (Figure 4.4). Of most interest was a mutation at amino acid position 641, which mapped to the entrance of the active site of the enzyme. In comparison to wild-type, alleles possessed substitutions of an arginine residue with either lysine or glutamic acid. Whilst arginine and lysine were both positively charged, glutamic acid displayed negative charge, and may therefore be more likely to have an effect on substrate binding. However, alleles with this mutation showed no reduction in neuraminidase activity, implying that the mutations at this position do not affect the ability of the enzyme to bind to the substrate (Table 4.3).

A novel region inserted into the enzymatic ß-propeller was identified (Figure 4.3). Interestingly, mosaic block C mapped to this region, and two distinct variants were present in this region. This inserted region has an unknown function, and was found to be absent from Influenza neuraminidase, showing it is not a region conserved across all neuraminidase proteins. Therefore, it is unlikely that this region affects the neuraminidase activity of the protein, a hypothesis supported by the finding that alleles possessing either variant in this region show little difference in activity (Table 4.3). However, it is likely that this region plays some role in pneumococcal pathogenesis, as it appears present in all alleles of this virulence factor (Figure 4.4). Preliminary studies by collaborators (Prof. Nizet, La Jolla, California) using knock-outs of this region have identified a potential role for this region in allowing the pneumococcus to cross the blood-brain barrier during the progression of meningitis. Intriguingly, the shape and electrostatic potential of this region were found to differ markedly between variants (Figure 4.5). These observed physical differences, along with the low levels of sequence identity (57%) between the variants, imply that they may play different roles. The observation that the divergent variant in

this region shared most sequence identity (66%) with a sialidase gene from *S. agalactiae* implies that the exchange of genetic information between these species may occur at some level. Furthermore, a related domain, attached to the β-propeller in a similar way, is found in the 3D-structure of *Macrobdella decora* NanL and is also predicted to be present in *S. pneumoniae* NanB, *Clostridium perfringens* NanI (Newstead *et al.*, 2008) and *Clostridium tertium* NanH proteins (Luo *et al.*, 1998).

The NanA protein possesses a candidate lectin-binding domain at its N-terminus. The role of this domain is thought to be to bind to host lectins, allowing cleavage of terminal sialic acids to occur via neuraminidase activity. There was no difference in neuraminidase activity when comparing TIGR4 NanA protein fragment possessing the lectin-binding and neuraminidase domains (AA65-792) and one possessing only the neuraminidase domain (AA318-792, Figure 4.6). This showed that the lectin-binding domain does not affect the neuraminidase activity of the NanA protein in this background.

The TIGR4 purified NanA fragment (AA318-792) differed from wild-type D39 NanA at 6 amino acid positions (K337N, N512D, N599D, I606V, D787E, D792N). None of these mutations were found to be near the neuraminidase active site (Figure 4.4-3), and it was therefore likely that this fragment purified from TIGR4 produced neuraminidase activity comparable to wild-type D39. When comparing activity of the three purified allele fragments (alleles 12, 16 and 18), it was found that alleles 12 and 16 did not differ significantly in activity from TIGR4. However, allele 18 was found to have neuraminidase activity >7-fold greater. Compared to wild-type, this allele had the last 8 amino acids of mosaic block A, as well as K455E, N512D I606V, D787E and D792N mutations. Since 4 of these mutations are also present in TIGR4, only the region of mosaic block A or amino acid mutation K455E, absent from the other 3 tested alleles, could have an affect on the neuraminidase activity of this allele. However, as none of these amino acids are positioned near the active site (Figure 4.4-18), there is no obvious structural reason for this increased activity.

Neuraminidase activity has been suggested as important in pathogenesis of p-HUS (Klein *et al.*, 1977; Seger *et al.*, 1980). This complication of IPD is increasing, and has higher mortality levels in comparison to typical, shiga-toxin

related HUS (Fitzpatrick *et al.*, 1993). However, very little is known about the pathogenesis of this disease and the virulence factors involved. Establishment of a concrete role for NanA may improve understanding of this complication, and assist in treatment and prevention of this fatal disease. It has been hypothesised that neuraminidase activity may play a role by cleaving neuraminic acid from host cells and exposing T Antigen (Klein *et al.*, 1977), leading to the symptoms of p-HUS. The results presented here showed that no particular NanA allele was harboured in isolates causing HUS. Furthermore, there was no significant difference in NanA activity between p-HUS isolates and matched controls. Furthermore, a knock-out of NanA showed that neither of the related enzymes NanB or NanC display significant levels of activity in comparison to NanA. This implies that although an active neuraminidase enzyme may be required for progression to p-HUS, other genetic factors may play a more significant role in allowing particular clones to cause this disease.

The results from the neuraminidase assay gave a number of further insights into pneumococcal NanA. Firstly, in all isolates tested, the majority of neuraminidase activity was detected in the supernatant, showing that the pneumococcus secretes more NanA into the environment than it attaches to the cell wall (Figure 5.13A). Furthermore, growth in the presence of the substrate can induce an increase in NanA expression by the bacteria (Figure 5.13B). Upon the induction of expression, there is a greater increase in secreted NanA than in cell surface-associated enzyme. These results imply that the pneumococcus may sense an increase in substrate in the environment and over-express and over-secrete NanA to compensate for this. This hypothesis makes biological sense when relating to the role of NanA, cleaving sialic acid from glycolipids of host cells, exposing surface receptors to allow pneumococcal adherence (Tong *et al.*, 1999). The pneumococcus would therefore benefit from increasing expression and secretion of NanA in response to increased substrate concentrations, as it would have an increased possibility of exposing host receptors, allowing attachment, leading to colonisation and possibly IPD. Finally, the results of the assay on the NanA knock-out in a D39 background showed that NanB and NanC display very small levels of activity in comparison to NanA, due to a significant decrease in activity compared to wild-type.

In conclusion, the diversity present in the NanA protein is significantly larger than that seen in Ply, with a number of mosaic blocks, insertions and point mutations present in 18 protein alleles. However, this level of variation should not reduce the potential of NanA as a cross-serotype protein-based vaccine candidate, since divergent alleles were recognised by polyclonal antibodies. Mapping of mutations uncovered onto the structural model of the NanA enzymatic fragment allowed identification of an insertion region, with two variants present, relating to mosaic block C. These two variants were found to differ significantly in surface shape and electrostatic potential, although the role of this insertion remains unknown. Finally, attempting to relate differences in NanA to disease showed that there was no difference in sequence or activity of NanA between p-HUS isolates and matched controls. This implies that although neuraminidase activity may be involved in p-HUS pathogenesis, other bacterial factors may play a role, whilst host genetics and environmental factors may also be important.

5 Homologues of Ply and NanA in related VGS species

5.1 Summary

The presence of homologues of pneumococcal virulence factors in related VGS species has been reported previously. Most notably, homologues of Ply, LytA and NanA have been observed in species of *S. mitis, S. oralis* and *S. pseudopneumoniae* (Balsalobre *et al.,* 2006; King *et al.,* 2005; Neeleman *et al.,* 2004; Obregon *et al.,* 2002; Whatmore *et al.,* 2000). However, these homologues remain poorly characterised, and study into this area will further our understanding of the relationships between these species, as well as offering insight as to the suitability of pneumococcal virulence factors as vaccine candidates. The aim of this study was therefore to further characterise homologues of Ply and NanA present in related VGS isolates.

A new CDC was identified and characterised in a subset of isolates of *S. mitis*, a commensal organism closely related to *S. pneumoniae* (Facklam, 2002). This work was carried out in collaboration with Dr. Johanna Jefferies and Leena Nieminen. The toxin was termed mitilysin (Mly). This toxin was shown to be genetically distinct from Ply, and actively secreted by a number of *S. mitis* isolates by an unknown mechanism. Furthermore, another homologue of Ply was identified in isolates of recently identified species *S. pseudopneumoniae*, and called pseudopneumolysin (pPly). This toxin was much more closely related to Mly than Ply.

For the first time, a NanA homologue was identified in *S. mitis* and *S. pseudopneumoniae* isolates. This enzyme was very closely related between all three species, in contrast to homologues of Ply. This may imply that it plays a more significant role in the commensal VGS species, which is supported by the defined role of NanA and not Ply in colonization. The presence of these homologues in related species, paired with the pneumococcus' ability to acquire altered genes by horizontal gene transfer (Chi *et al.*, 2007; King *et al.*, 2005; Poulsen *et al.*, 1998; Sibold *et al.*, 1994) gives rise to the possibility of evasion of a pneumococcal vaccine targeting these virulence factors. This may occur as selective pressure induced by the pneumococcal vaccine gives an evolutionary advantage to pneumococci with altered genes and thus evading the scope of a vaccine. Therefore, these findings should be considered when selecting vaccine candidates for future protein-based pneumococcal vaccines.

5.2 Ply homologues in related VGS species

5.2.1 Identification and characterization of a new CDC, Mly

Homologues of a number of pneumococcal genes in related VGS species have been discovered previously (King *et al.*, 2005; Neeleman *et al.*, 2004; Poulsen *et al.*, 1998; Sibold *et al.*, 1994; Whatmore *et al.*, 2000), although few are well characterized. Identification and characterization of homologues is important to develop understanding of the relationships between these species, and also in the context of pneumococcal vaccine candidates, such as Ply, with homologues present in the commensal streptococcal species. A novel CDC, given the name Mly was discovered, in collaboration with colleagues, in a subset of *S. mitis* isolates. This toxin was shown to be fully haemolytic, and secreted extracellularly by some isolates, by an unknown mechanism. It was also demonstrated that a number of alleles of the toxin were not recognised by the monoclonal antibody PLY-7, and based on ELISA results, the presence of a further haemolytic toxin in a number of isolates was suggested.

The sequence of 7 *mly* genes was determined by PCR and sequencing as described for *ply*. Results show that across the 7 genes, mutations in the DNA sequence occur at 98 positions when compared to wild-type *ply* DNA allele 1. This number is much higher than in the Ply screen, where 37 mutated base positions were discovered from a much larger sample set, and gives nucleotide sequence diversity of 6.9% compared to D39 (See Appendix I). Alignment of translated sequences resulted in mutations at 13 amino acid positions, giving 7 Mly alleles, with each isolate possessing a unique allele (Figure 5.1). Of the 13 amino acid mutations present, two were also present in the Ply alleles, namely N14D and K224R, with a resulting protein sequence diversity of 2.7%. This is remarkably low, due to the low numbers in the sample set, and the large number of synonymous mutations present in the protein. If a larger sample set of *S. mitis* isolates were studied, it is likely that Mly would display greater sequence diversity than Ply.

						Amino acid position											
					2	14	116	222	224	231	246	378	386	396	398	423	443
Strain	**Identification**	**Origin**	**Isolation site**	**Mly allele**	**A**	**N**	**E**	**D**	**K**	**E**	**V**	**T**	**H**	**A**	**D**	**V**	**T**
R75 I	*S. mitis*	UK	NPA	1		D			R				Y				
R77	*S. mitis*	UK	NPA	2	T	D			R	D							
COL15	*S. mitis*	UK	Sputum	3	T	D			R		I		Y			I	
990123	*S. mitis*	UK	Dental abscess	4	T	D		G	R	D		A	Y	S	N		
QH17	*S. mitis*	UK	Blood	5		D	K		R			A	Y	S	N	I	
R75 II	*S. mitis*	UK	NPA	6		D			R	D			Y		N		N
R76	*S. mitis*	UK	NPA	7	T	D			R	D			Y		N	I	

Figure 5.1: Mly amino acid alleles

Amino acid sequences of Mly alleles aligned to Ply allele 1 from D39. Identification of species was done by *sodA* gene sequencing and alignment as described. Row 1 represents the amino acid position, whilst row 2 represents the amino acid present in wild-type Ply allele 1.

5.2.2 A Ply homologue in S. pseudopneumoniae isolates

A sample set of 13 isolates of the newly classified species *S. pseudopneumoniae* (Arbique *et al.*, 2004) were supplied by Dr. Ralf Reinert (National Reference Centre for *streptococci*, Aachen, Germany). These isolates have been shown to possess genotypic and phenotypic traits of both *S. pneumoniae* and *S. mitis*, without fitting into the brackets of either species (Arbique *et al.*, 2004). These isolates also possessed *ply* gene homologues (Franken *et al.* poster presentation, Europneumo 7). It was of interest to determine the sequence of the Ply homologues present in these isolates, as well as their haemolytic abilities. The *ply* gene was amplified from gDNA using primers 27R and 27S as described. Results showed that a *ply* gene homologue was amplified from all isolates except PS866 (Figure 5.2). Interestingly the PCR products differed in size in a number of isolates compared to the amplified product from D39 gDNA. Alignment of 3 sequenced *ply* homologues (circled in white) from *S. pseudopneumoniae* isolates to Ply allele 1 uncovered mutations at 74 base positions (see Appendix 1). Alignment of translated sequences showed these equating to 7 amino acid mutations, resulting in 3 unique alleles in these isolates (Figure 5.3). Of the 7 amino acid mutations present, 2 were shared with both Ply and Mly alleles (N14D and K224R), whilst a further 4 were shared only with Mly (A2T, H386Y, D398N, V423I). The toxin was termed pPly. Although the *pply* genes sequenced were from isolates with PCR products differing in size from *ply*, no size difference was observed in the gene. The observed size differences were likely to be due to differences in flanking DNA, as primers used for PCR amplification were external to the *ply* gene. The pPly proteins in these isolates possessed haemolytic activity similar to wild-type Ply (data not shown).

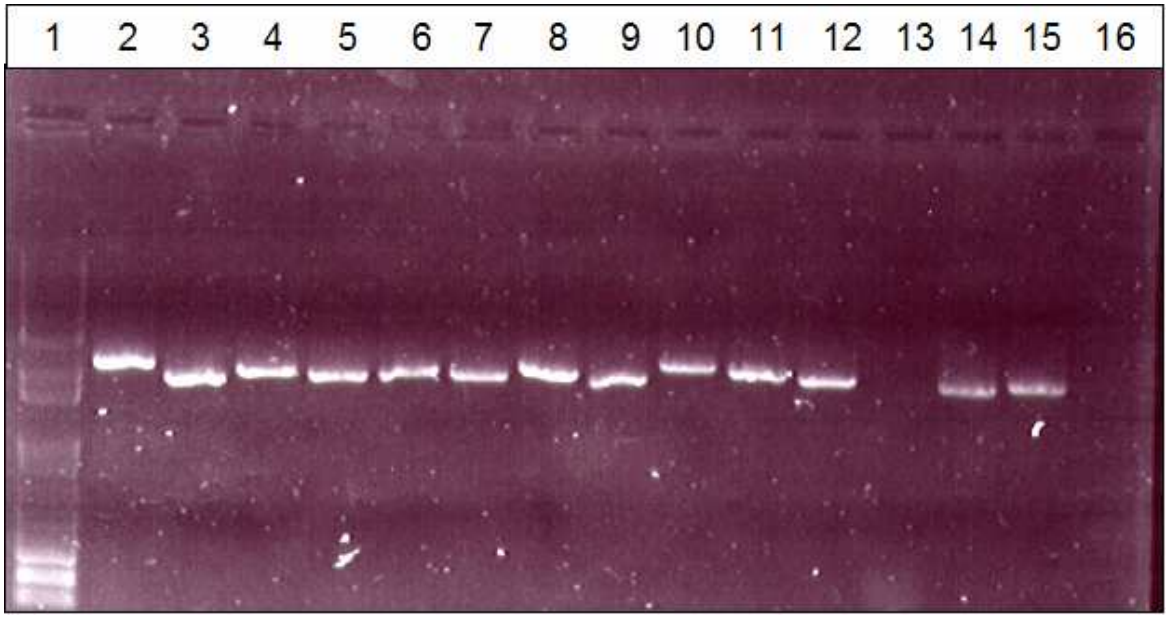


Figure 5.2: PCR of *ply* homologues in *S. pseudopneumoniae*

Agarose gel with products of PCR with 27R and S primers on *S. pseudopneumoniae* gDNA, viewed under UV light. Circled in white are the isolates for which the pply gene was sequenced. A number of the PCR products differ markedly in size from the *S. pneumoniae* wild-type control *ply* from D39. The identity of samples in the lanes was as follows: 1 – 1Kb+ DNA ladder; 2 – D39; 3 - RRR468; 4 – RRR472; 5 – CCUG49455; 6 – RRR474; 7 – RRR473; 8 – PW2108; 9 – RRR471; 10 – PW2113; 11 – CCUG48465; 12 – RRR475; 13 – PS866; 14 – A76; 15 – RKI1158; 16 – negative control (no gDNA).

		Amino acid position						
		2	14	224	323	386	398	423
Strain	pPly allele	A	N	K	D	H	D	V
RRR468	1	T	D	R		Y		I
RRR471	2			R	N			
RRR472	3	T	D	R		Y	N	I

Figure 5.3: pPly amino acid alleles

Amino acid sequences of pPly alleles aligned to Ply allele 1 from D39. Row 1 represents the amino acid position, whilst row 2 represents the amino acid present in wild-type Ply allele 1.

5.2.3 Genetic relationship between ply, mly and pply

The genetic relatedness of the three CDCs *ply, mly* and *pply* was explored using MEGA4 software to construct a minimum evolution tree to show the relationships between the toxin genes at the DNA level (Figure 5.4). Results showed that whilst the *ply* alleles all cluster together, the *mly* and *pply* alleles are part of the same branch of the tree. This implied that in terms of genetic relationships, the *mly* and *pply* genes are more similar to each other than either is to *ply*. When looking solely at the Ply branch of the tree, the isolates were shown to cluster by Ply allele, although there was mixed clustering between isolates possessing alleles 1 and 2.

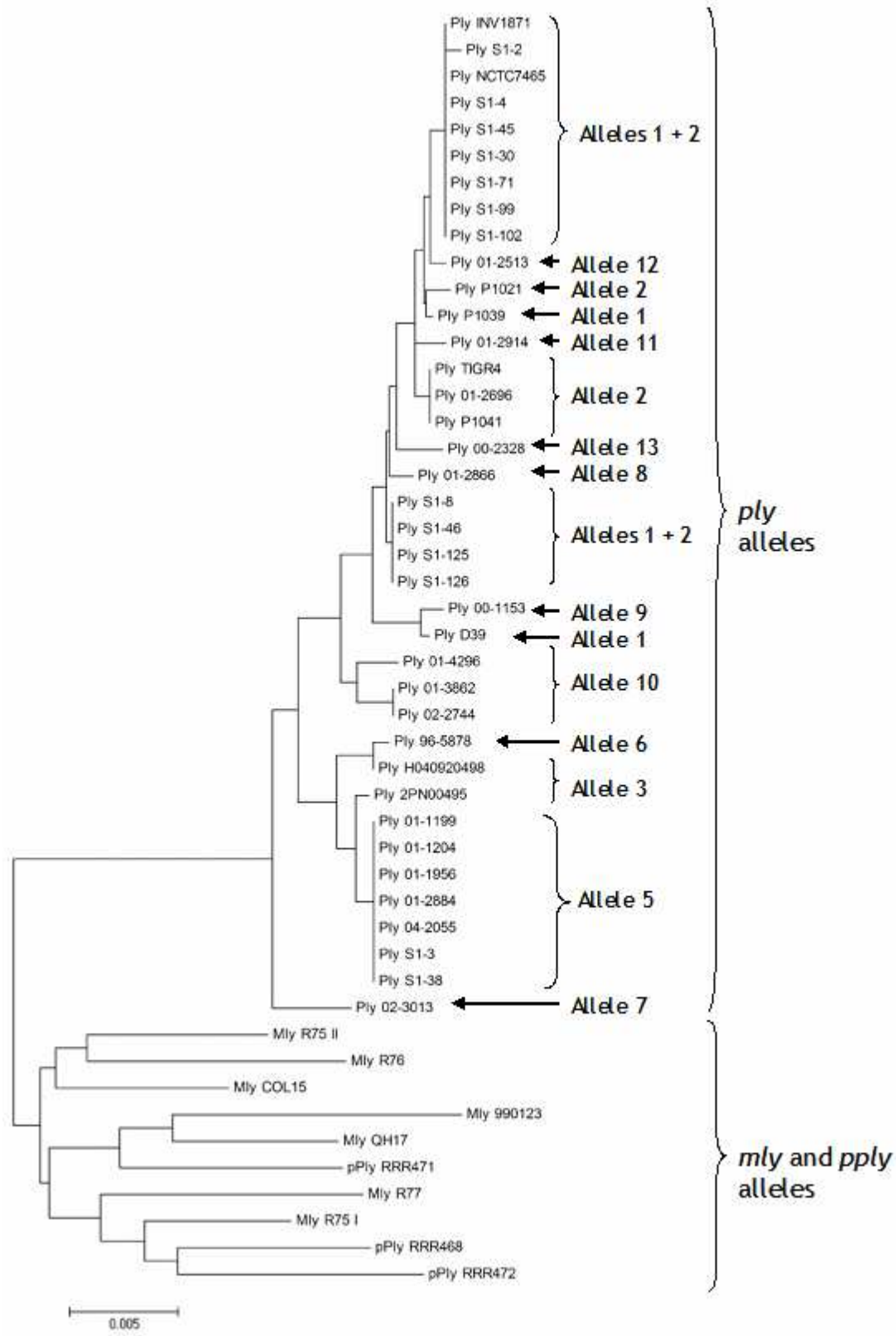


Figure 5.4: Relatedness of *ply* gene homologues

Minimum evolution tree of *ply* genes and homologues in related streptococcal species. Tree was drawn using MEGA4 software, and demonstrates the relatedness of these toxins at the DNA level. The units at the bottom of the tree indicate distance between sequence pairs.

5.3 NanA homologues in related VGS species

5.3.1 NanA alleles present in related VGS isolates

A screen of 29 VGS isolates resulted in discovery of a *nanA* homologue gene in a number of isolates. A homologue of NanA was discovered for the first time in all five *S. pseudopneumoniae* isolates tested, as well as in five *S. mitis* isolates. The sequence of the *nanA* gene from four *S. pseudopneumoniae* and one *S. mitis* isolates allowed translation into protein, and alignment and comparison with the NanA alleles discovered in *S. pneumoniae*. Results showed that whilst high levels of variation were present within these alleles, they displayed striking similarities with the NanA alleles of *S. pneumoniae*. From the five sequences, there were 4 VGS NanA alleles identified. The information on these alleles can be found in Figure 4.5. VGS NanA alleles 1-3, present in *S. pseudopneumoniae* and *S. mitis* isolates, shared 99.7%, 99.6% and 99.4% sequence identity with pneumococcal NanA allele 14 respectively, and were aligned to this allele in Figure 4.5. However, VGS NanA allele 4, present in 2 *S. pseudopneumoniae* isolates, shared 94.9% sequence identity with pneumococcal NanA allele 7, and was aligned to this allele. Results showed that all alleles possessed the same diversity features as the pneumococcal NanA alleles, with only 5 amino acid mutations present in the VGS alleles but absent from the pneumococcal alleles. Furthermore, both variants of mosaic block C were present in the related VGS population, whilst only one variant of mosaic blocks A and D was present in these isolates. The 5AA insertion-duplication was present in all isolates, whilst none possessed the 16AA insertion found in several pneumococcal alleles. Therefore, in contrast to Ply, there was very little difference between the NanA proteins found in the pneumococcus and those found in related species *S. mitis* and *S. pseudopneumoniae*.

									Amino acid position									
VGS NanA allele	Species	Present in	Compared to pneumococcal NanA allele	44	134	253	322	442	460	471	488	599	641	641	721	787	792	990
				V	D	K	V	D	I	E	E	N	E	R	K	D	D	E
1	*S. pseudopneumoniae*	R82 E6	14				A											
2	*S. mitis*	R83 E2	14				A	V			K							
3	*S. pseudopneumoniae*	R98 A5	14				A					D	K			E	N	K
4	*S. pseudopneumoniae*	R89 G3, R89 G5	7	M	G	R			T	G				K	N			

Figure 5.5: Amino acid sequence comparison of NanA homologues

Amino acid alleles of NanA in species of VGS related to *S. pneumoniae*. Alleles were similar to different pneumococcal NanA alleles, and were compared to their most closely related counterparts. The amino acids in the 3rd row represent those present in either pneumococcal NanA allele 7 or 14 depending on which was used for comparison.

5.3.2 NanA activity of related VGS species

The NanA activity of the four *S. pseudopneumoniae* and one *S. mitis* isolates was measured using the neuraminidase assay as described previously. The results showed that whilst all five isolates displayed activity, there were significant differences between them (Figure 5.6). Whilst the *S. pseudopneumoniae* isolates R89 G3 and R89 G5 had activities very similar to the *S. pneumoniae* isolate D39, the *S. mitis* isolate R83 E2 and *S. pseudopneumoniae* R98 A5 showed reduced activity in comparison. Interestingly, almost all of the activity detected in *S. pseudopneumoniae* isolate R82 E6 was in the secreted fraction, a similar phenomenon to that observed in TIGR4. However, unlike TIGR4, this isolate did not possess the frameshift mutation which resulted in a premature stop codon in the protein, and loss of the LPXTG motif.

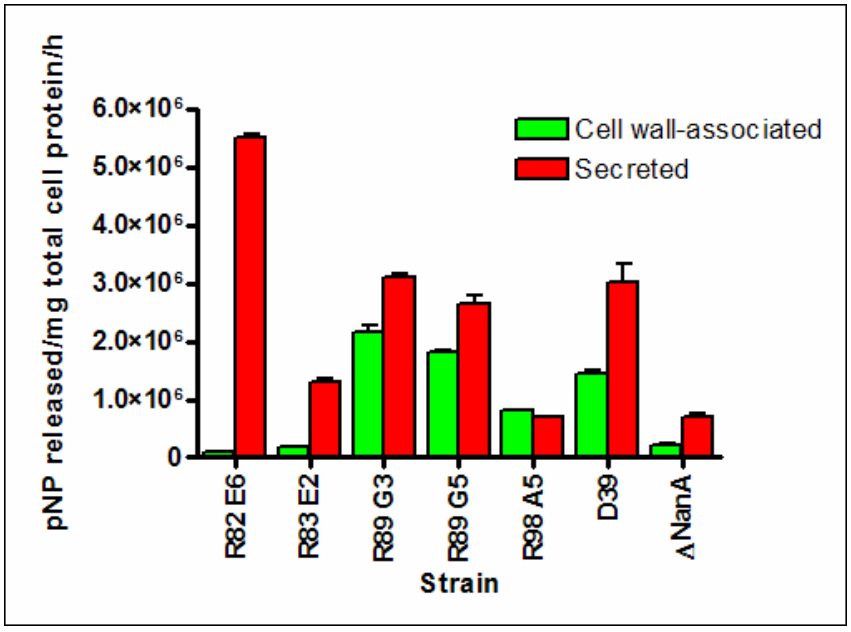


Figure 5.6: Neuraminidase activity of NanA homologues

Neuraminidase activity of cell wall-associated and secreted NanA from related VGS species. The activity of the cell wall associated NanA is in green, whilst the secreted NanA is in red. S. pneumoniae isolates D39 and NanA knock-out strain ΔNanA were included as +ve and –ve controls respectively. NanA activity was measured in nM pNP released/mg total cell protein/hour.

5.3.3 Genetic relationship between pneumococcal and VGS nanA alleles

The genetic relatedness of the *nanA* alleles from the pneumococcus and related VGS species was explored using MEGA4 software to construct a minimum evolution tree to show the relationships between the *nanA* genes at the DNA level (Figure 5.7). Results showed that the pneumococcal isolates cluster by allele in the tree. However, the *nanA* sequences from *S. mitis* and *S. pseudopneumoniae* are not separated from the pneumococcal sequences, and are closely linked to their nearest pneumococcal relatives. Therefore, unlike *ply*, no evolutionary distinction between these species was observed when looking at the NanA enzyme. However, the greater diversity observed in the *nanA* gene in comparison to *ply* is reflected in the large separations between pneumococcal *nanA* alleles in the tree.

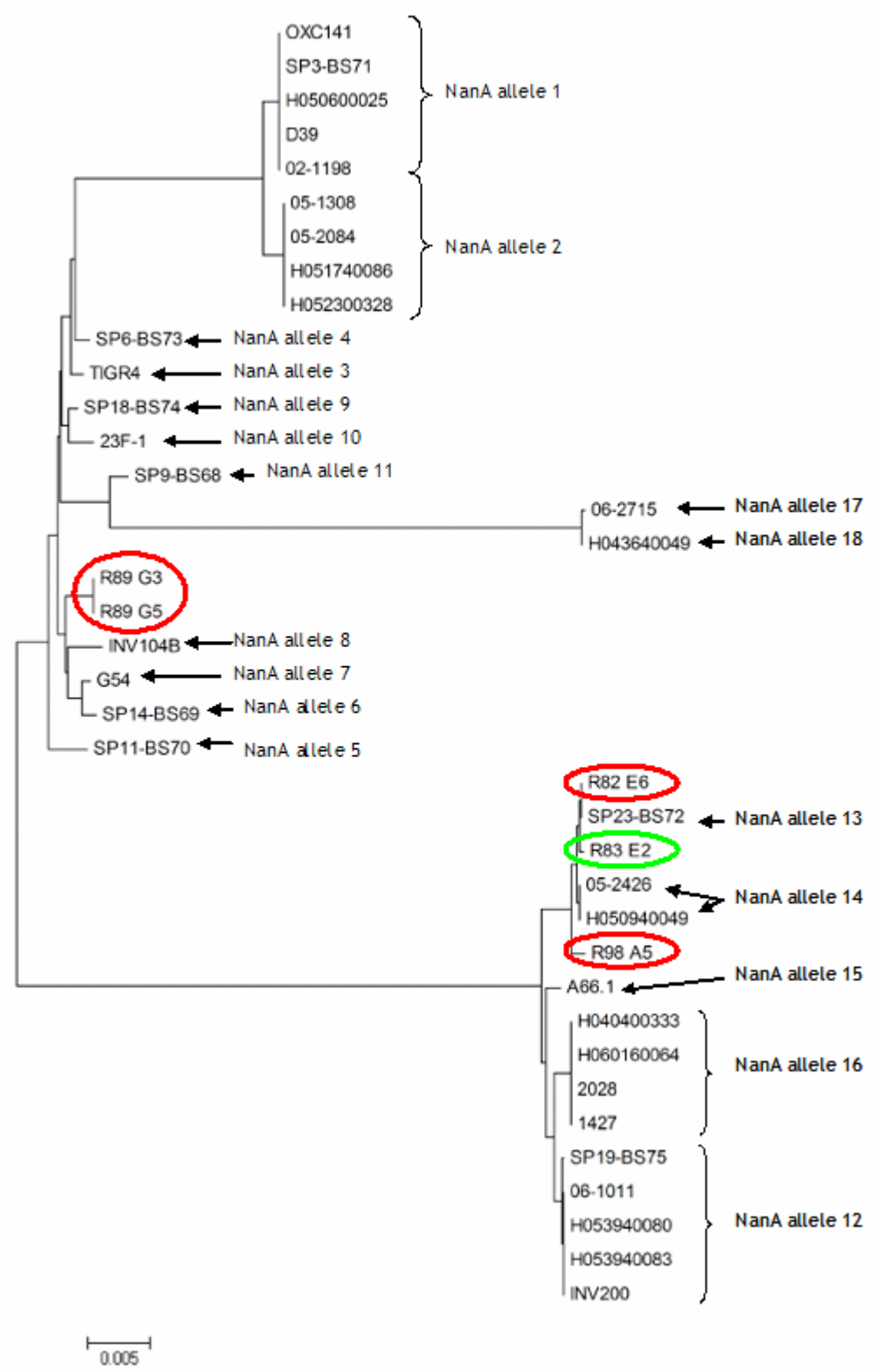


Figure 5.7: Relatedness of pneumococcal *nanA* gene homologues

Minimum evolution tree of *nanA* genes and homologues in related streptococcal species. Tree was drawn using MEGA4 software, and demonstrates the relatedness of these enzymes at the DNA level. The units at the bottom of the tree indicate distance between sequence pairs.

5.4 Discussion

Although the presence of a *ply* homologue in *S. mitis* isolates has been documented previously (Neeleman *et al.*, 2004; Whatmore *et al.*, 2000), this work has for the first time characterized the toxin and coined the name Mly. *mly* genes display higher levels of sequence variation than *ply*. This is in agreement with published findings which analysed the sequence variation in a *ply* gene fragment from pneumococcal and non-typable, non-pneumococcal isolates, and found that non-pneumococcal isolates were much more diverse when the mean genetic difference between the two groups was compared (Hanage *et al.*, 2005). This result implies that there is less selective pressure for the *mly* gene to be conserved by *S. mitis* species, possibly due to the mostly commensal nature of this species, meaning the toxin is not as essential as in the pathogen *S. pneumoniae*, where it is a virulence factor important in a plethora of diseases as discussed in Section 1.6.

A number of the *S. mitis* isolates were found to actively secrete Mly, which is in contrast to the proposed mechanism of Ply release from *S. pneumoniae* by autolysis, and in agreement with previous findings (Whatmore *et al.*, 2000). A previous study found that release of Ply in *S. pneumoniae* serotype 3 isolate WU2 was not dependent on autolysin or choline, thus excluding all pneumococcal lytic enzymes as mechanisms for release of Ply. The authors conclude that Ply release in WU2 is due to a non-type II, non-type III mechanism (Balachandran *et al.*, 2001), although there is some debate as to whether this isolate may possess a bacteriophage, with the lytic activity of this phage accounting for Ply release. Therefore, it is possible a similar mechanism to that observed in WU2, be it phage-related or not, accounts for secretion of Mly. Furthermore, a number of Mly alleles were recognised by the monoclonal antibody PLY-7, which is used diagnostically in laboratories (Cima-Cabal *et al.*, 2001). This raises the possibility of false positive results when using this antibody. Therefore, when coupled with the non-recognition of Ply allele 12 by this antibody, both false positive and false negative results are possible. This highlights the need for care to be taken when using antibodies either diagnostically or therapeutically to combat pneumococcal infection.

The discovery of Mly in *S. mitis* isolates was not unexpected as *S. mitis*, along with *S. oralis*, was the closest relation to *S. pneumoniae*, determined by 16S rRNA analysis (Facklam, 2002). More recently, a species termed *S. pseudopneumoniae* was identified as distinct from both *S. pneumoniae* and *S. mitis*, and was shown to be even more closely related to *S. pneumoniae (Arbique et al., 2004)*. This work has identified and begun to characterize a further Ply homologue in *S. pseudopneumoniae* isolates, pPly, which was shown to be much more closely related to *mly* than *ply* in a minimum evolution tree. Indeed whilst the variations in the *ply* gene separated most *ply* alleles on the tree, the *mly* and *pply* sequences were mixed together as part of the same branch of the tree. This result implies that *S. pseudopneumoniae* isolates are more closely related to *S. mitis* than to *S. pneumoniae,* at least in terms of this toxin. This was due to the fact that in comparison to *ply,* both *mly* and *pply* share many synonymous mutations. This segregation is in agreement with phenotypic traits such as bile solubility, lack of capsule and optochin resistance, which all link *S. pseudopneumoniae* to *S. mitis (Arbique et al., 2004)*. However, genotypic tests such as *sodA* sequencing, ply PCR or commercial kits, which generally distinguish *S. mitis* and *S. pneumoniae,* do not distinguish between *S. pseudopneumoniae* and *S. pneumoniae* (Arbique *et al.*, 2004). Furthermore, when comparing the sequences of *nanA* gene homologues identified in *S. mitis* and *S. pseudopneumoniae* isolates to pneumococcal *nanA* alleles, there was no segregation of the isolates into different species in a minimum evolution tree. These two results, showing the differences between *ply, mly* and *pply* genes, and the lack of difference between *nanA* genes, gives further evidence to the argument that a smooth continuum may exist between these species, with pneumococcal genes and homologues randomly dispersed in isolates identified as *S. mitis* or *S. pseudopneumoniae* (Hakenbeck *et al.*, 2001). These results also seem to agree with the recent hypothesis that species such as *S. mitis* and *S. pseudopneumoniae* are in the process evolving towards commensality from a pneumococcus-like virulent ancestor, a theory which seems to sit together with the idea of a continuum between species, and also allows for the seeming random spread of homologues to important pneumococcal virulence factors (Kilian *et al.*, 2008). However, this theory is not fully accepted, as it is also possible that evolution is occurring in the other direction, with acquisition of virulence genes allowing the pneumococcus to evolve away from a commensal

state. The results found here could in theory also support this hypothesis, and therefore further study is required to untangle the evolutionary relationships between these species.

This study also represents the first demonstration of a NanA homologue in a subset of *S. mitis* and *S. pseudopneumoniae* isolates. Sialidase genes have previously been found in other related VGS species however, in particular in *S. oralis*, a close relation of the pneumococcus (Byers *et al.*, 2000; King *et al.*, 2005). The sequences of the protein alleles discovered here were very similar to pneumococcal NanA, which is in contrast to the finding for Ply homologues, where many synonymous and a number of non-synonymous mutations were present. This may indicate that NanA is of more importance to these commensal species than Mly. The presence of this pneumococcal virulence factor in commensal relatives is less of a surprise than Ply, as NanA has a highly defined role in pneumococcal colonization (Manco *et al.*, 2006; Tong *et al.*, 1999; Tong *et al.*, 2000), and although it has been implicated as important in a number of pneumococcal diseases, particularly otitis media (Long *et al.*, 2004; Tong *et al.*, 2000), the enzyme is most important for pneumococcal colonization. In contrast, Ply has a much less defined role in pneumococcal colonization, but well-defined roles in fatal diseases such as pneumonia and meningitis (Hirst *et al.*, 2008; Jounblat *et al.*, 2003; Rubins *et al.*, 1995; Wellmer *et al.*, 2002). Therefore, it may be expected that NanA would also be important for the colonization of commensal organisms, whereas the benefit of possessing a cytolytic toxin, and a possible role for the toxin in colonization and pathogenesis of this subset of *S. mitis* isolates remains unclear. This potential difference in importance to the commensal organisms may be reflected in the apparent conservation of NanA when compared to the pneumococcal NanA protein. In comparison, Mly and pPly appear less conserved compared to Ply, possibly due to the fact that they are less important to the organisms.

The presence of homologues of Ply and NanA in these closely related species raises issues when discussing these virulence factors as protein vaccine candidate. Results show that a number of the Mly alleles were not recognised by monoclonal antibody PLY-7, presumably due to variation within the alleles, showing that small alterations in the protein can have significant effects. Furthermore, these streptococcal species are able to co-inhabit the same niche,

in the nasopharynx of the host, and have been shown to exchange genetic material (Chi *et al.*, 2007; King *et al.*, 2005; Poulsen *et al.*, 1998; Sibold *et al.*, 1994). Indeed, homologous recombination in the *nanA* gene has already been shown to occur between *S. pneumoniae* and *S. oralis*, whilst the increase of penicillin resistance in pneumococci was shown to be as a result of horizontal gene transfer of the *pbp2x* gene from *S. mitis* (King *et al.*, 2004; Sibold *et al.*, 1994). These results together give rise to the possibility that a vaccine targeting Ply or NanA will give a selective advantage to pneumococci acquiring altered variants of these genes. When these homologues are present in related species, it is possible that pneumococci could acquire such altered genes by horizontal gene transfer, and thus evade the scope of a vaccine targeting them. This is one reason why a multi-factorial protein vaccine is likely to be more successful than a vaccine targeting a single pneumococcal protein (Ogunniyi *et al.*, 2000; Ogunniyi *et al.*, 2007a).

In conclusion, this study has characterized and named a new CDC, Mly, present in a subset of *S. mitis* isolates. Further homologues of Ply and NanA have been identified in *S. mitis* and *S. pseudopneumoniae* species, although further characterization of these is required. Presence of gene homologues in closely related commensal species not only informs further on the relationships between these *Streptococci*, but further blurs the lines between these species, one a debilitating pathogen, the others generally commensal organisms, and brings into question the importance of Ply and NanA in pneumococcal virulence. The presence of these homologues also gives rise to the possibility of pneumococci evading a Ply-targeted vaccine by acquisition of gene variants through horizontal gene transfer. Further study and characterization of these homologues is required to fully understand their roles in these species, and the potential importance they may play pneumococcal vaccine strategies.

6 Inhibition studies of NanA with viral neuraminidase inhibitors

6.1 Summary

The aim of this chapter was to investigate the effects of inhibitors of influenza NA on pneumococcal NanA, as well as determining the effects of this inhibition *in vivo*. This work demonstrated for the first time that specific inhibitors of influenza NA were able to inhibit purified pneumococcal NanA *in vitro*. In particular, oseltamivir carboxylate (OC, the active form of Tamiflu, Roche) showed high potency when inhibiting wild-type NanA (IC_{50}=1.608μM), although studies of inhibition potency with other NanA alleles require further work. This inhibitor was also shown to inhibit NanB, but with a potency ~350–fold lower than NanA. Studies of NanA inhibition with competitive inhibitor DANA and antiviral Zanamivir resulted in IC_{50}s ~19- and ~9400-fold higher than with OC, respectively. These results suggest that OC may have potential as a treatment of secondary pneumococcal pneumonia following influenza infection. Furthermore, results suggest the influenza NA may essentially play the same role as NanA, uncovering pneumococcal adherence receptors *in vivo*, and priming the host for pneumococcal colonisation. This finding may help explain the high levels of secondary pneumococcal pneumonia following influenza infection, and agrees with a previously presented hypothesis (McCullers and Bartmess, 2003).

Animal studies in a pneumonia model showed that pre-treatment of mice with purified NanA before intranasal challenge with *S. pneumoniae* increased the severity of symptoms compared to PBS controls, as well as producing significantly increased bacterial counts in the lungs and blood at 72h post-infection. Conversely, pre-treatment of mice with OC before intranasal challenge with *S. pneumoniae* decreased the severity of symptoms, and significantly reduced the bacterial counts in the lungs at 72h post-infection. These two results, taken together, show that the neuraminidase activity of NanA plays an important role in pneumococcal pneumonia. Furthermore, they confirm that OC, or derivatives, have potential for treatment of secondary pneumococcal infection, either replacing or in conjunction with antibiotics, although further studies are required. As well as acting therapeutically, this could have the benefit of helping tackle the over-reliance on antibiotics for treating pneumococcal diseases.

6.2 Inhibition of purified NanA by viral inhibitors of Influenza

6.2.1 Comparison of inhibition of wild-type NanA fragment by viral neuraminidase inhibitors

The solved structure of the NanA active site has allowed new insight into the activity of the protein as well as similarity with other sialidase enzymes. For example, whilst the protein sequences of pneumococcal NanA and neuraminidase (NA) from Influenza shared <10% sequence identity, the enzymatic 6-bladed ß–propeller was highly conserved between these proteins, (Figure 6.1, circled in red) along with six catalytic residues. However, the insertion region present in pneumococcal NanA, circled in green, is not present in the Influenza NA (Figure 6.1).

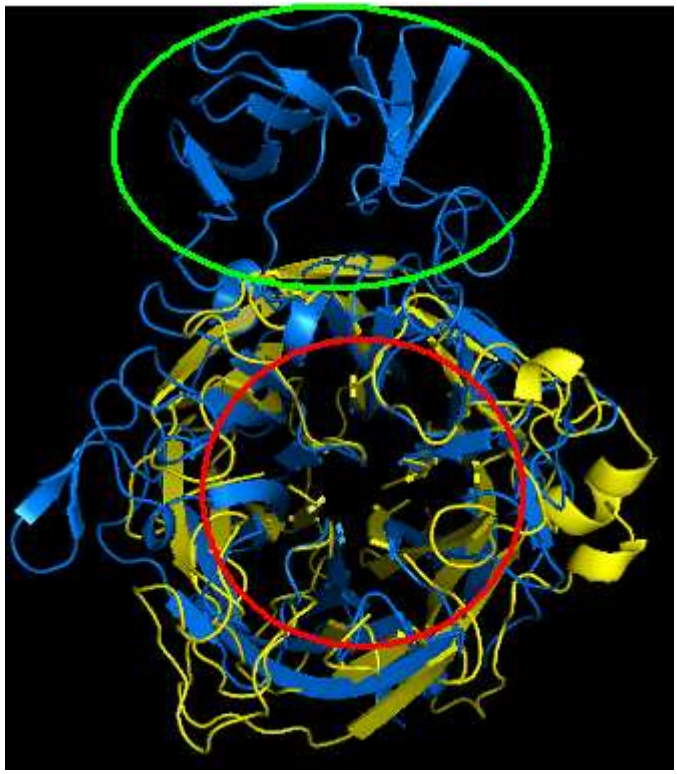


Figure 6.1: Structural comparison of pneumococcal NanA and Influenza NA

Overlay of crystal structures of pneumococcal NanA and influenzal NA enzymes. NanA is in blue whilst NA is in yellow. The 6-bladed β-propeller, active site circled in red, is highly conserved in both of these enzymes. The mosaic block C insertion region in pneumococcal NanA, absent from Influenza NA, is circled in green.

The inhibitory effect of inhibitors of influenza NA on pneumococcal NanA was investigated. The inhibitors tested were the competitive inhibitor DANA, Oseltamivir carboxylate (OC, active form of Tamiflu, Roche) and Zanamivir (marketed as Relenza, GSK). The inhibition of purified NanB (created by Dr. Heinz Gut) with OC was also investigated. Results showed that all three viral inhibitors also had an inhibitory effect on pneumococcal NanA, and whilst OC had an inhibitory effect on NanB, this was significantly lower than that seen with NanA (Figure 6.2). From these plots of inhibition, the concentration of inhibitors required to inhibit 50% of the enzyme activity (IC_{50} values) were calculated for each inhibitor with NanA, and with OC for NanB (Table 6.1). This is the first demonstration of viral neuraminidase inhibitors having a direct effect on the pneumococcal enzymes NanA and NanB. Comparison of the IC_{50} values for each inhibitor showed that OC inhibited NanA with the greatest affinity, whilst DANA inhibited NanA with 18.8-fold less affinity, and Zanamivir had a very low inhibition, with an IC_{50} value 9392-fold greater than OC. Comparing OC inhibition of NanA and NanB showed that the inhibitor had a much greater affinity for NanA, inhibiting with an IC_{50} 345.5-fold lower than with NanB.

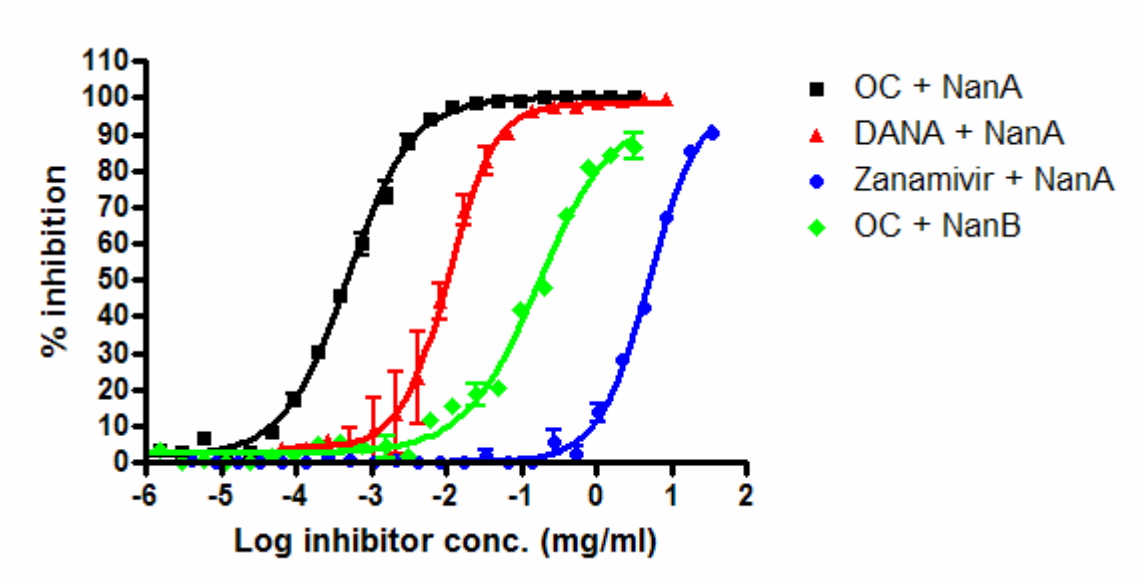


Figure 6.2: Inhibition of pneumococcal neuraminidases

Inhibition curves for viral neuraminidase inhibitors with pneumococcal neuraminidase enzymes. Curves were drawn from triplicate repeats for each inhibitor. IC_{50} values were calculated for each inhibitor from these curves.

Neuraminidase	Inhibitor	IC_{50} (μM)	Ratio
NanA	Oseltamivir carboxylate	1.608	1
NanB	Oseltamivir carboxylate	555.60	345.5
NanA	DANA	30.27	18.8
NanA	Zanamivir	15102	9392

Table 6.1: IC_{50} values of selected inhibitors

IC_{50} values of viral NA inhibitors inhibiting pneumococcal neuraminidases, along with ratios of inhibition for comparison.

6.2.2 Comparison of Oseltamivir carboxylate inhibition of purified NanA protein fragments of varying amino acid sequence

OC was found to inhibit NanA with the greatest affinity. It was therefore of interest to determine whether the mutations present in the purified fragments of NanA alleles 12, 16 and 18 had an effect of inhibition by OC. The inhibition assays were repeated with these allele fragments, and results showed that the alleles were inhibited with differing levels of affinity by OC (Figure 6.3). Calculation of IC_{50} values showed that the mutations present in alleles 12, 16 and 18 appeared to lower the affinity of OC towards the NanA active site (Table 6.2). Indeed, allele 18 showed a 100-fold reduction in IC_{50} compared to allele 3. However, there was a correlation between the activity of the alleles and their inhibition with OC; the inhibitor had lower affinity to alleles with higher activity. This may have skewed the results (see Section 6.4, Discussion).

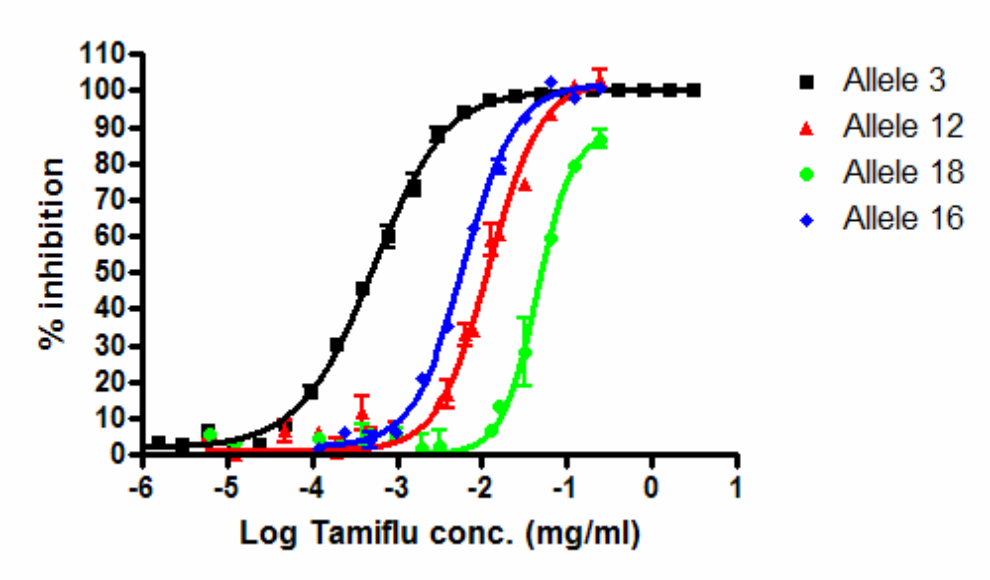


Figure 6.3: Inhibition of NanA amino acid alleles by OC

Inhibition curves of OC with purified NanA amino acid allele fragments 3, 12, 16 and 18. Curves were drawn from triplicate repeats for each inhibitor. IC_{50} values for OC with each allele were calculated from these curves.

NanA amino acid allele	OC IC_{50} (μM)	Ratio
3	1.608	1
12	37.16	23.1
16	18.53	11.5
18	160.44	99.8

Table 6.2: IC_{50} values of OC with NanA amino acid alleles

IC_{50} values of viral NA inhibitor OC with NanA amino acid alleles, along with ratios of inhibition for comparison.

6.3 Studies of effect of Oseltamivir carboxylate inhibition of NanA and NanB on *S. pneumoniae in vivo*

NanA and NanB have been shown to be involved in the virulence of the pneumococcus (Manco *et al.*, 2006; Tong *et al.*, 2000). The novel discovery that inhibitors of influenza NA were able to inhibit pneumococcal neuraminidases *in vitro* led to a study of OC as an inhibitor of pneumococcal virulence *in vivo*. The *S. pneumoniae* isolate used in *in vivo* studies was the bioluminescent strain A66.1 (Xenogen, USA), allowing imaging of bacterial infection in real-time using IVIS CCD Camera (Xenogen, USA). This isolate possessed allele 15, which had the divergent version of mosaic block C. *In vivo* studies were carried out by Kirsty Ross, and will be presented fully in her thesis. However, *in vitro* and *in vivo*

studies were carried out in parallel, and *in vivo* results strengthen the hypothesis presented, and offer further insight into the relationship between influenza and the pneumococcus. These results will therefore be briefly discussed.

6.3.1 Effect of pre-treatment with neuraminidase A on pneumococcal pneumonia

In order to explore the importance of NanA in progression of pneumococcal pneumonia, mice were pre-treated with the purified NanA allele 3 enzymatic fragment before challenge with *S. pneumoniae* in a pneumonia model. Results showed that pre-treatment with NanA caused significant weight loss compared to controls (14%, $p<0.001$), an indicator of pneumococcal infection. Furthermore, imaging at 72h post-infection revealed infection in the ventral thoracic cavity of 7/10 mice pre-treated with NanA, and only 4/10 in the control mice, with one death in each group. The results assume a positive signal was present in mice that died before imaging. Surviving mice were culled at 72h post-infection, and bacterial counts from selected organs measured. Pre-treatment with NanA caused a significant increase in bacterial counts in lungs ($p=0.0188$) and blood ($p=0.0142$), whilst a trend towards increased bacterial counts was present in the spleen and liver. A similar, although less significant result was observed upon pre-treatment of mice with influenza NA, thus mimicking the real-life interaction between pathogens.

6.3.2 Effect of pre-treatment with Oseltamivir carboxylate on pneumococcal pneumonia

To determine the potential protective effect of OC against *S. pneumoniae* infection, OC-treated mice were challenged with *S. pneumoniae*. There was no significant difference in survival time between OC-treated and control mice, with median survival times of 65h and 94h respectively, although a trend towards significance is observed. However, it appeared that mice pre-treated with OC developed disease symptoms, determined by pain score, at a slower rate than control mice, implying a protective effect of OC. Imaging of mice at 48h and 72h allowed the tracking of pneumococcal pneumonia progression. A signal from the ventral thoracic cavity was detected in 3/19 OC-treated mice at 48h post-infection, compared to 8/20 control mice. Similarly, a signal was detected from 5/19 OC-treated mice at 72h post-infection, compared to 11/20

controls. There was a significant decrease ($p=0.009$) in bacterial lung counts in the OC-treated group after 72h, supporting the imaging data. Furthermore, trends towards lower bacterial counts were observed in the spleen, liver and blood of the OC-treated group. These results imply that OC, presumably by directly inhibiting the activity of NanA and possibly NanB, decreases the virulence of *S. pneumoniae* in this murine pneumonia model.

6.4 Discussion

This study has demonstrated for the first time the inhibition of pneumococcal NanA activity by inhibitors specifically targeting the NA enzyme of influenza. Results showed that OC inhibited NanA with a 20-fold greater potency than the competitive inhibitor DANA (Table 4.4), whilst in influenza, studies have shown OC to inhibit with 1000-fold greater potency than DANA (Klumpp and Graves, 2006). Furthermore, OC was shown to inhibit NanA with ~350-fold greater potency than NanB, implying that the active site of NanA shares more structural homology with influenza NA than that of NanB does. These results offer insight not only into potential treatments of pneumococcal disease, but into the relationships between the pathogens *S. pneumoniae* and influenza. A large amount of work has explored the relationship between these organisms, as the pneumococcus has been shown as the major cause of secondary bacterial pneumonia following influenza infection, greatly increasing mortality, particularly during Influenza pandemics (Abrahams *et al.*, 1919; McCullers, 2006; Morens and Fauci, 2007; Stone and Swift, 1919). A number of recent studies have explored the possible mechanisms of this interaction. Studies in a mouse model have shown that pneumococcal adherence in the lungs is increased by presence of influenza NA (McCullers and Bartmess, 2003) and that influenza NA activity correlated with ability of a viral infection to predispose to secondary bacterial pneumonia (Peltola *et al.*, 2005). The findings presented here are concurrent with the hypothesis suggested by Peltola and colleagues (2005) that infection with influenza predisposes to subsequent pneumococcal infection due to the action of the viral NA. The authors suggested that this enzyme exposes receptors for pneumococcal adherence, increasing the chance of pneumococcal colonization and pneumonia (Peltola *et al.*, 2005). This implies that viral NA and pneumococcal NanA have similar roles and substrates, a hypothesis supported by our data showing the similarity of the active site of these enzymes (Figure 6.1), as well as the inhibition of NanA with viral NA inhibitors (Table 1.). Furthermore, in a mouse model of secondary pneumococcal pneumonia, treatment with a selection of classes of antivirals showed that the NA inhibitor, OC, increased survival, from 0% to 75%, whilst other antivirals had no effect. This was reported to be due to effective treatment of the predisposing viral infection (McCullers, 2004). However, the data presented here suggest that this result may in part be due to the inhibitory effect of OC directly on *S. pneumoniae*. Indeed, treatment

of mice with OC was shown to significantly reduce pneumococcal counts in the lungs and blood, showing that in the previous study, the OC would have had an inhibitory effect on both primary and secondary pathogens in the model.

When comparing the inhibition of different NanA alleles by OC, large differences in inhibition were observed in the assay used. The most active allele was inhibited with ~100–fold less potency in comparison to TIGR4 allele 3 (Table 6.2). However, the nature of the assay means the results may be misleading. In the inhibition assay, a concentration of NanA sufficient to cause 100% release of pNP from the substrate by the least active allele was used. This produces problems when attempting to compare inhibition of alleles with different activities, as since one allele has ~7-fold increased activity, if these alleles are inhibited to the same degree by OC, a 7-fold increase in OC would be required to inhibit the more active allele to the same degree, resulting in an apparent difference in IC_{50} where there may be no difference. This is due to the difference in activity between alleles, where the same molar concentration, and therefore the same number of NanA active sites is 7 times more active in one allele than the other. Conversely, if a concentration of NanA sufficient to cause 100% release of pNP from the substrate had been calculated individually for each allele, the molar concentrations, and therefore the number of active sites, would have been reduced ~7-fold in the most active allele, resulting in further inaccuracies in the comparison of IC_{50}s for OC with each allele. Therefore, although a large difference in potency of OC towards different NanA alleles was shown, these results should be regarded as preliminary. The reason this method was selected was due to the restricted availability of the active OC compound. Further studies of Michealis-Menten enzyme kinetics and interactions between enzyme, substrate and inhibitor would allow accurate comparison of IC_{50}s between alleles of differing NanA activity.

Taken together, the *in vivo* studies with NanA and OC showed that the enzymatic activity of NanA plays an important role in progression from colonization to pneumococcal pneumonia in a mouse model, agreeing with previous findings (Berry and Paton, 2000; Orihuela *et al.*, 2004). Pre-treatment of the respiratory tract with purified NanA was shown to enhance infection, and may partly reflect the mechanism by which a preceding influenza infection enhances susceptibility to the pneumococcus. This hypothesis was supported by the observation that a

similar trend is observed when mice are treated with purified viral NA prior to pneumococcal challenge, mimicking the interaction between pathogens *in vivo*. Furthermore, treatment of mice with OC was shown to decrease the severity of pneumococcal infection, demonstrating that when neuraminidase activity is inhibited, progression of pneumococcal pneumonia is impaired. This agrees with previous results showing that NanA is important in colonization and virulence of *S. pneumoniae (Manco et al., 2006)*. These findings suggest that inhibition of NanA, potentially by OC or other derivatives, could play a role in decreasing the burden of secondary bacterial infection complications after viral infection. Therefore this treatment may have benefit in replacement or complementation of current antibiotics used to treat such complications. This could be important not only in reducing the burden of secondary bacterial infection caused by *S. pneumoniae,* but also potentially in reducing the global spread of antibiotic resistance in pathogenic strains of *S. pneumoniae* by tackling the misuse and over-reliance on antibiotics for treatment of pneumococcal infections.

A further potential use of these compounds is to treat or prevent pneumococcal diseases in which neuraminidase activity plays an important role, such as p-HUS. Although in the case of p-HUS, OC would have no therapeutic potential once the disease is established, it may be of use in preempting the onset orf p-HUS. For example, if a paediatric patient presented with IPD caused by a pneumococcal clone known to be associated with p-HUS, such as serotype 19A, ST199, then co-treatment of this infection with both antibiotics and OC would not only treat the infection but also inhibit the neuraminidase activity of the pathogen, reducing the likelihood of p-HUS occurring.

In conclusion, this study has demonstrated the *in vitro* inhibition of NanA by inhibitors of influenza NA for the first time. This novel discovery, coupled with *in vivo* studies, demonstrated the role of neuraminidase activity in pneumonia, as well as giving further insight into the well-studied relationship between influenza and subsequent pneumococcal infections. Furthermore, this study has established the potential benefit of using OC as a replacement for antibiotics to treat secondary pneumococcal infection. Further study in this area may lead to development of inhibitors that are able to prevent both primary and secondary infections, thus lowering the burden of disease and mortality caused by the synergism between these two debilitating pathogens. Finally, OC may be of use

to treat or prevent pneumococcal diseases where neuraminidase activity it actively involved in pathogenesis, such as p-HUS.

7 Genomic studies using microarray technology

7.1 Summary

The research presented in this project has so far focussed on investigation of diversity within individual genes and proteins. However, recent advances in technology have led to usability and affordability of genome-wide genetic comparison tools in the research lab. One of these tools is microarray analysis, which has recently been used by our lab to show genetic diversity between isolates identical by serotype and ST (Silva *et al.*, 2006). This chapter reports on results of using microarray analysis to explore differences in genetic content and gene expression between isolates. Specifically, this chapter will attempt to link genetic differences to p-HUS pathogenesis, as well as exploring differences between *S. pneumoniae* and related species.

The genetic differences between p-HUS isolates and matched controls were investigated, by comparative genomic hybridization, in an attempt to identify genes important in p-HUS progression. Results show that although differences exist in genetic content between these isolates, none correlate with ability to cause p-HUS. Furthermore, the isolates tested were relatively clonal, sharing large regions of diversity when compared to the reference isolate TIGR4.

Differential gene expression between p-HUS isolates and matched controls was investigated, using RNA microarrays, as differences in gene expression may be important in p-HUS. A total of 12 genes were identified as being differentially regulated in p-HUS isolates compared to controls. These may be of importance in p-HUS pathogenesis. Two of the genes, *ghdA* and *guaA*, were involved in the glutamate metabolism pathway, and regulation of these resulted in increased glutamate metabolism. It was hypothesised that this may be in response to high circulating ammonia or ammonium concentrations in p-HUS patients, and preliminary results implied that p-HUS isolates have higher tolerance to ammonium in growth medium than controls.

Microarray technology is limited to informing on genes present in the reference isolate. This may ignore genetic information carried by bacteriophages or mobile genetic elements. The presence of lysogenic bacteriophages in the p-HUS isolates was determined. Two isolates were found to harbour phages, although their presence did not correlate with p-HUS-causing ability. However, although

further study is required, use of this technology has identified several candidate genes with potential roles in p-HUS pathogenesis. Understanding the pathogenic methods employed by the pneumococcus will allow more effective diagnosis and treatment of p-HUS, and reduce the levels of mortality observed.

The pneumococcus is known to exchange genetic information with related species such as *S. mitis* and *S. oralis*. Indeed, diversity in pneumococcal virulence factors has been shown to result from homologous recombination with these species, with homologues of a number of pneumococcal virulence factors identified (King *et al.*, 2005; Poulsen *et al.*, 1998; Sibold *et al.*, 1994; Whatmore *et al.*, 2000). These homologues can be important in diagnosis and treatment of pneumococcal disease, as homologues in related species may allow acquisition of altered virulence factors by the pneumococcus upon selective pressure from a vaccine targeting a particular virulence factor. This may allow pneumococci to escape the scope of vaccines targeting virulence factors with homologues in these related species. In this study, gDNA from related VGS species was hybridised to a microarray of genes from a pneumococcal reference to determine the presence of homologues to pneumococcal genes in these related species.

Results showed that when comparing the genes present in *S. mitis* and *S. oralis* isolates, homologues to 72-84% of genes were present. 25 genes essential to pneumococcal virulence (Hava and Camilli, 2002) were found to be absent from all of these isolates. These included genes from islets shown to be involved in pneumococcal virulence, such as the pilus islet RlrA, or the PsrP islet. Essential genes in these regions may make interesting pneumococcal vaccine candidates. 312 genes deemed essential to pneumococcal virulence (Hava and Camilli, 2002) were found to have homologues in the related VGS population. These included virulence factors with defined roles in pneumococcal virulence, many of which were also vaccine candidates.

Microarray results from a *S. pseudopneumoniae* isolate showed a higher level of homology with *S. pneumoniae*, implying that this species is more closely related to *S. pneumoniae* than either *S. mitis* or *S. oralis*, agreeing with previous findings (Arbique *et al.*, 2004). Furthermore, the finding that pneumococcal virulence genes are distributed within the *S. mitis* population, present in some

isolates but not others, agrees with the hypothesis that this isolate may be evolving from a pneumococcus-like ancestor into a commensal state by loss of virulence factors (Kilian *et al.*, 2008). However, it is also possible that evolution is occurring in the opposite direction, with the pneumococcus evolving away from a commensal ancestor, and these results would also support this hypothesis. Further studies are therefore required to unravel the evolutionary relationships between these species.

7.2 Microarray analysis of p-HUS isolates and matched controls

7.2.1 Rationale for microarray analysis and strain selection

There appears no difference, in terms of sequence or expression, between the NanA proteins of p-HUS isolates and matched controls. Therefore, DNA and RNA microarrays were employed to determine genetic differences between p-HUS isolates and controls. Comparative genomic hybridization using DNA microarrays was used to detect presence or absence of genes, whilst the analysis of gene expression using RNA microarrays was used to detect differential expression of genes, both compared to the reference isolate TIGR4. These analyses allowed identification of genes encoding factors that may play a role in progression to p-HUS. Six serotype 19A, ST199 isolates (4-p-HUS isolates (Waters *et al.*, 2007) and 2 matched controls) were chosen for the microarray study, as differences identified between these clones were more likely to be related to p-HUS causing ability. These isolates are found in Table 1. Had isolates with varying serotypes or STs been selected, it would have been impossible to determine whether observed differences were linked to p-HUS or simply differences related to serotype or ST.

Strain	Serotype	ST	Disease
05-1308	19A	199	Pneumonia
05-2084	19A	199	Pneumonia
H040400333	19A	199	p-HUS
H051740086	19A	199	p-HUS
H052300328	19A	199	p-HUS
H060160064	19A	199	p-HUS

Table 7.1: Isolates used for microarray study of p-HUS isolates

All isolates were serotype 19A, ST199. 4 isolates caused p-HUS whilst 2 matched controls caused pneumonia without p-HUS.

7.2.2 Analysis of genetic content of p-HUS isolates and matched controls

The results of the DNA microarray study showed that the 6 test isolates shared large regions of diversity in comparison to the reference isolate TIGR4. 12 of these regions, representing genes missing or too diverse for hybridization in the test isolates when compared to the reference, were present and varied in size from 2-22 genes (Full alignment found in Appendix II). These regions included, as expected, the capsule locus, as reference isolate TIGR4 was a serotype 4 isolate. Other regions included the pilus locus, ABC transporters and regions of hypothetical proteins. However, although the 6 isolates tested shared the same serotype and ST, there were also differences in the genetic content of these isolates. These were represented by regions of diversity present in some test isolates but not others, or by single genes not hybridized to the array in certain test isolates. When the association with p-HUS was explored, only two genes were identified as present in all p-HUS isolates and absent from controls. These were the putative transcriptional regulator SpTIGR4-0306 and hypothetical protein SpTIGR4-0309. However, these genes were present in a region of varying results in different test isolates (Figure 7.1), and PCR validation of the microarray results showed these two genes to be present in all isolates. Conversely, no genes were found to be absent from p-HUS isolates and present in control isolates. Therefore, none of the differences in genetic content identified between these isolates correlated to p-HUS causing ability.

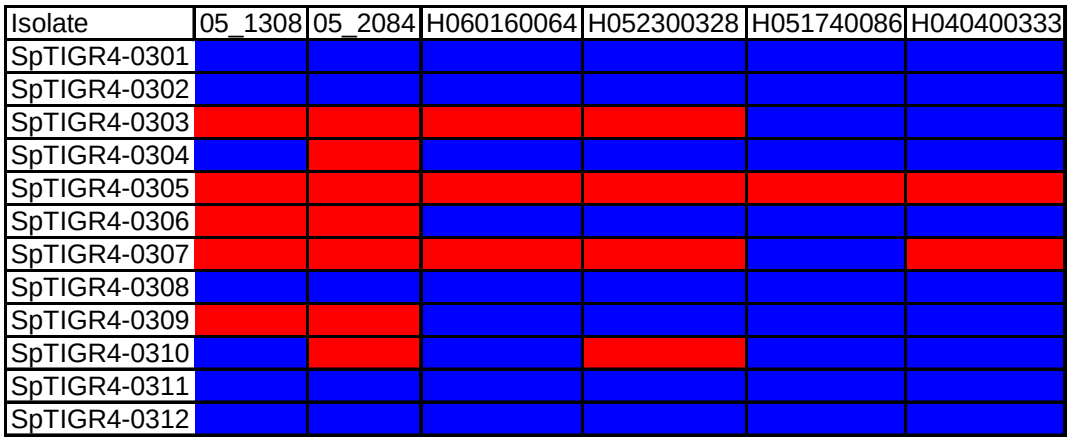


Figure 7.1: Diversity present between p-HUS isolates

Region of diversity after comparative genomic hybridization, where two genes were found to be present in all p-HUS isolates and absent from controls (SpTIGR4-0306 and SpTIGR4-0309). Blue represents presence of gene in both test and reference isolates. Red represents presence of gene in reference isolate and absence from test isolate.

7.2.3 Analysis of gene expression of p-HUS isolates and matched controls

7.2.3.1 RNAvsRNA Microarrays

Comparison of gene expression between the HUS isolates and matched controls identified twelve genes differentially expressed in isolates causing p-HUS. These can be found in Table 7.2. Of these, nine were down-regulated in p-HUS isolates and three up-regulated. Interestingly, two of the genes, NADP-specific glutamate dehydrogenase (*gdhA)* and glutamine amidotransferase (*guaA)*, were found to be involved in the glutamate metabolism pathway, and while *gdhA* is up-regulated and *guaA* is down-regulated, both contribute to increased cellular levels of glutamine and glutamate in isolates causing HUS. Other genes down-regulated encode the phosphomevalonate kinase *mvaK2*, which is involved in the biosynthesis of steroids, a gene encoding a NifU family protein and another encoding competence-inducable protein CinA. Other up-regulated genes encoded for a RecF protein and ABC-N/P, the ATP-binding permease protein of an ABC transporter.

Gene	Gene name	Abbrev.	Fold change
SpTIGR4-0098	Hypothetical protein		1.73
SpTIGR4-0383	Phosphomevalonate kinase	*mvaK2*	1.33
SpTIGR4-0409	Conserved hypothetical protein	*mip*	4.1
SpTIGR4-0716	Transcriptional regulator, putative		8.6
SpTIGR4-0738	Conserved domain protein		1.57
SpTIGR4-0870	NifU family protein	*nifU*	3.19
SpTIGR4-1306	NADP-specific glutamate dehydrogenase	*gdhA*	1.53
SpTIGR4-1434	ABC transporter, ATP-binding/premease protein	ABC-N/P	2.5
SpTIGR4-1732	Serine/threonine protein kinase	*pkn2*	1.31
SpTIGR4-1941	Competence/damage inducible protein CinA	*cinA*	1.27
SpTIGR4-2072	Glutamine amidotransferase	*guaA*	5.47
SpTIGR4-2227	RecF protein	*recF*	2.11

Table 7.2: Differential gene expression in p-HUS isolates

Genes differentially expressed in HUS isolates compared to matched controls. Fold change represents the change in gene expression in HUS isolates compared to controls. Changes highlighted in red were up-regulated in HUS isolates while those highlighted in blue were down-regulated. Abbrev = abbreviation of gene.

When comparing the expression of the *nanA* gene between HUS isolates and matched controls, there was no significant difference between expression levels (p=0.2496), which agreed with the findings of NanA activity in these isolates in Chapter 4. Furthermore, when comparing these levels of *nanA* gene expression to the reference isolate TIGR4, there was a significant decrease in expression in the tested isolates (p=0.0007). There was also a significant decrease in *nanB* expression when the test isolates and controls were compared to TIGR4 (p=0.0002). Interestingly, the *nanC* gene was only found in 1 isolate, the p-HUS isolate H040400333, and in this isolate, expression was increased 21.8-fold compared to TIGR4, implying that it may be important in this particular clone.

7.2.3.2 Validation of microarray results by RT-PCR

Two genes, SpTIGR4-1306 and SpTIGR4-2072, were selected for analysis by RT-PCR. These gene numbers represented *gdhA* and *guaA* respectively, and were found to be differentially regulated in p-HUS isolates compared to controls by RNA microarray. The expression of these two genes was explored in one p-HUS isolate (H060160064) and one control (05-1308). The results of RT-PCR show that the *guaA* gene was down-regulated 28.3-fold in the p-HUS isolate compared to the control (Figure 7.2A). Conversely, the *gdhA* gene was up-regulated 5.38-fold in the p-HUS isolate (Figure 7.2B). These ratios, although slightly higher, were in agreement with the findings of RNA microarray analysis for these genes, and validated the RNA microarray results obtained for other genes.

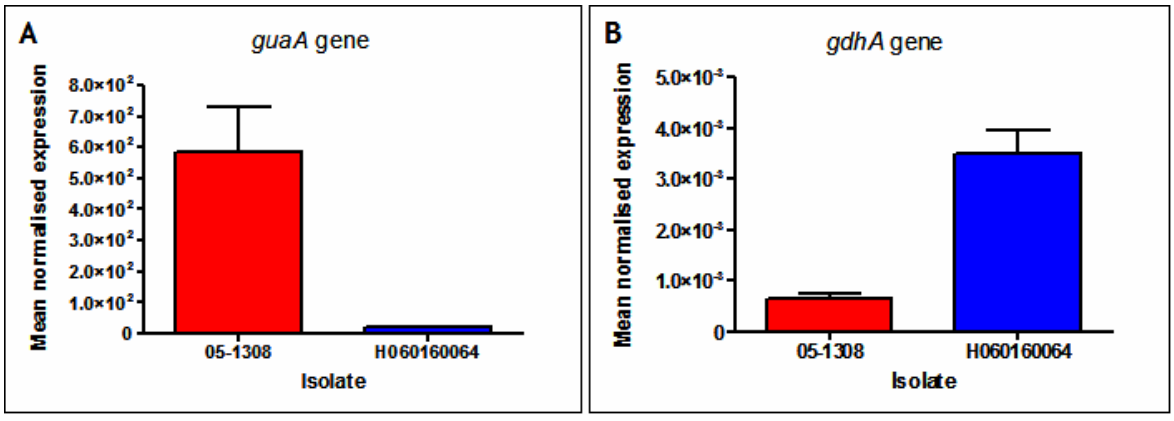


Figure 7.2: RT-PCR validation of RNA microarray

RT-PCR results. Mean normalised expression levels of guaA and gdhA genes in p-HUS isolate (H060160064) and control (05-1308), as measured by RT-PCR.

7.2.3.3 Growth of p-HUS isolates in differing ammonium concentrations

The effect of differing ammonium concentration on growth of p-HUS isolates and controls was investigated. Two isolates were compared in the study, namely the p-HUS isolate H060160064 and control 05-1308, as in the RT-PCR study. Preliminary results showed that the control isolate was highly sensitive to concentrations of ammonium, as low as 1µg/ml, whereas growth of the p-HUS isolate was only affected at concentrations as high as 500µg/ml (Figure 7.3). However, the same phenomenon was not observed upon repeat of the experiment, and further study is therefore required to confirm the observed difference. Despite this, this is an example of using the results generated from microarray studies to begin to elucidate the involvement of certain genes in pathogenesis.

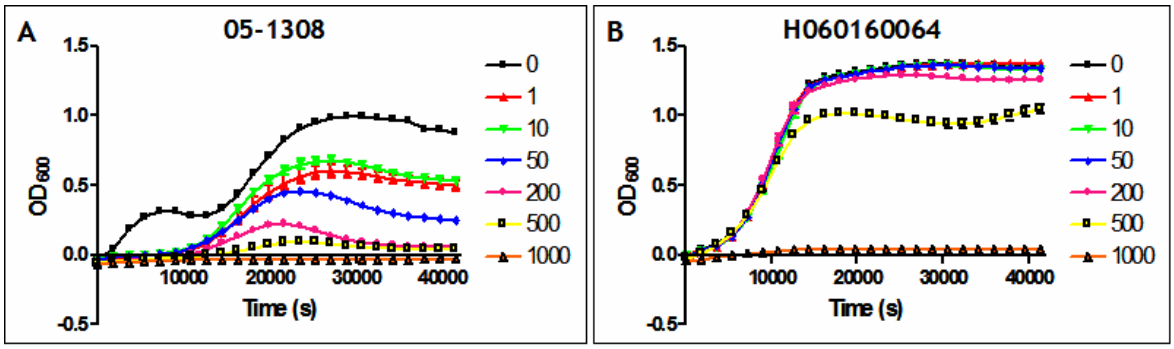


Figure 7.3: Ammonium growth curves

Growth curves of control isolate (05-1308, plot A) and p-HUS isolate (H060160064, plot B), in differing concentrations of ammonium. Concentrations of ammonium are shown on right had side, in µg/ml.

7.3 Analysis of lysogenic bacteriophage presence in p-HUS isolates and matched controls

The six serotype 19A, ST199 isolates were tested for the presence of lysogenic bacteriophages. Phages were identified in two of the six isolates tested, namely H040400333 and H051740086 (Figure 7.4). The bacteriophage possessed by H04040033 was found to be a type 1 pneumococcal phage by phage-specific PCR, whilst PCR failed to identify a phage type in H051740086 (Bacteriophage identity determined by Dr. Patricia Romero, Romero et al. 2008, manuscript in preparation). However, the two other p-HUS isolates did not possess lysogenic bacteriophages, showing that the presence of these phages was not required for the isolates to cause p-HUS. Therefore, genes important for progression of p-HUS were not carried on lysogenic bacteriophages.

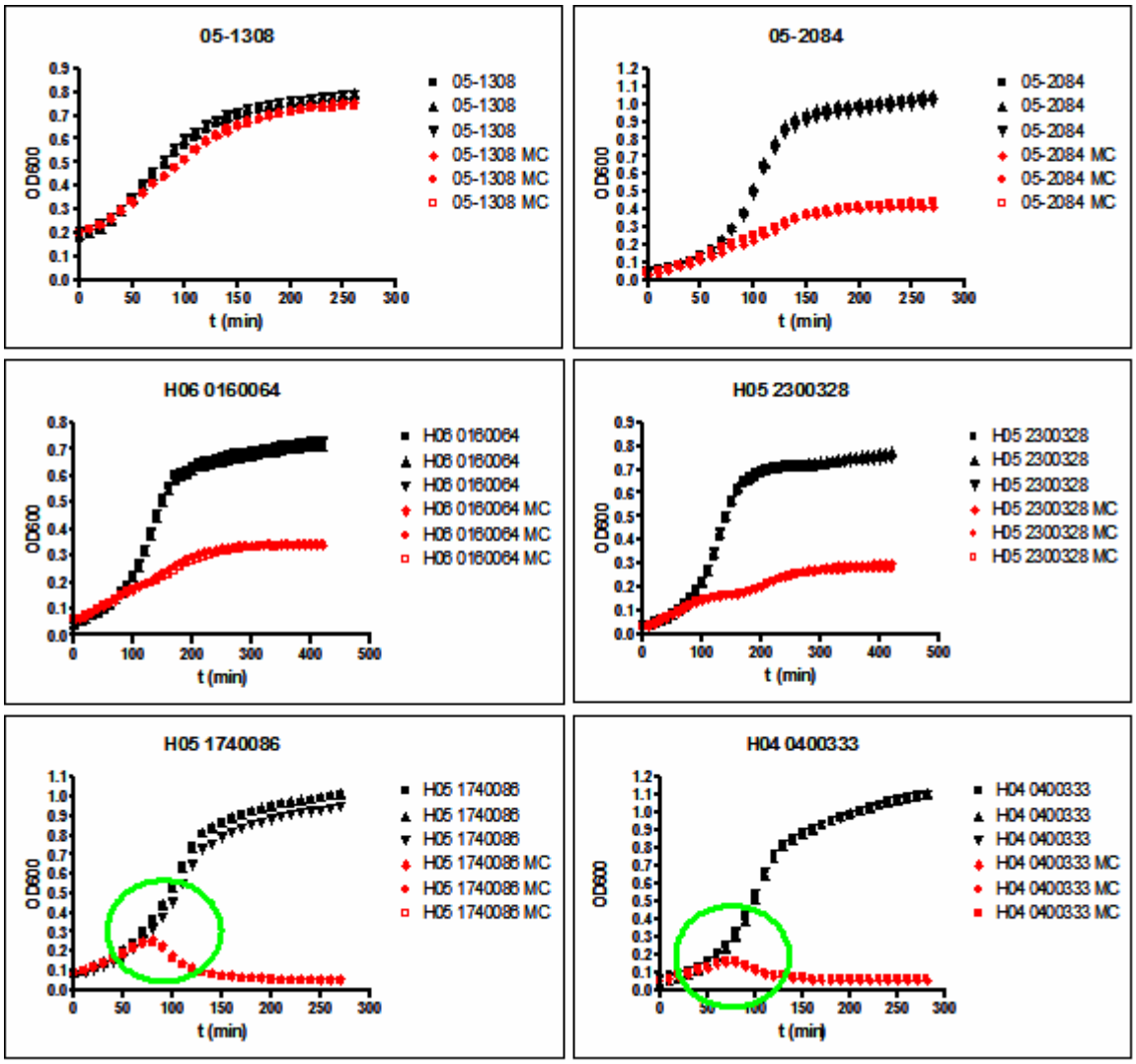


Figure 7.4: Presence of bacteriophage in p-HUS isolates

Results of mitomycin C test for presence of temperate bacteriophage in p-HUS isolates and controls. Test and negative control growth curves were measured in triplicate. Possession of a phage is shown by a clear decrease in OD600 after introduction of mitomycin C. This can be seen for isolates H040400333 and H051740086, circled in green.

7.4 Microarray study of pneumococcal gene homologues in related VGS species

7.4.1 Study plan and selection of isolates

The presence of homologues of pneumococcal virulence genes in the related VGS population has been previously documented (King *et al.*, 2005; Poulsen *et al.*, 1998; Sibold *et al.*, 1994; Whatmore *et al.*, 2000). These homologues may allow alteration of their pneumococcal counterparts by homologous recombination, which has implications in treatment of pneumococcal diseases and vaccine design. This study aimed to take a global approach to homologue identification by hybridizing gDNA from related VGS species to a pneumococcal DNA microarray. This allowed genome-wide identification of homologues of pneumococcal genes, and in particular virulence factors, present in the VGS population. To study the presence of homologues in *S. pseudopneumoniae*, *S. mitis* and *S. oralis* isolates, 7 isolates were selected from a variety of isolation sites (Table 7.3). As well as looking at overall presence of gene homologues, this allowed stratification of results into commensal or opportunistically virulent strains, and comparison of virulence genes present. The species of each isolate was determined by *sodA* sequencing and alignment to reference isolates as previously described (Hoshino *et al.*, 2005), and minimum evolution tree showing the results can be found in Figure 7.5. *S. pneumoniae* and *S. pseudopneumoniae* are indistinguishable by this method, and optochin resistance was used to confirm these species where necessary.

Strain	Species	Isolation site	Disease
NCTC12261	*S. mitis*	Reference isolate	NCTC
R75 A1	*S. mitis*	Nasopharynx	Commensal
R77 C4	*S. mitis*	Nasopharynx	Commensal
Col	*S. mitis*	Blood	Endocarditis
990123	*S. mitis*	Pus	Dental abscess
Sv29	*S. oralis*	Nasopharynx	Commensal
R98 A5	*S. pseudopneumoniae*	Nasopharynx	Commensal

Table 7.3: VGS isolates in microarray study

Isolates used for study of homologues of pneumococcal genes in related VGS species. Isolation site and disease caused are also listed. Species identification was by sodA sequencing and alignment.

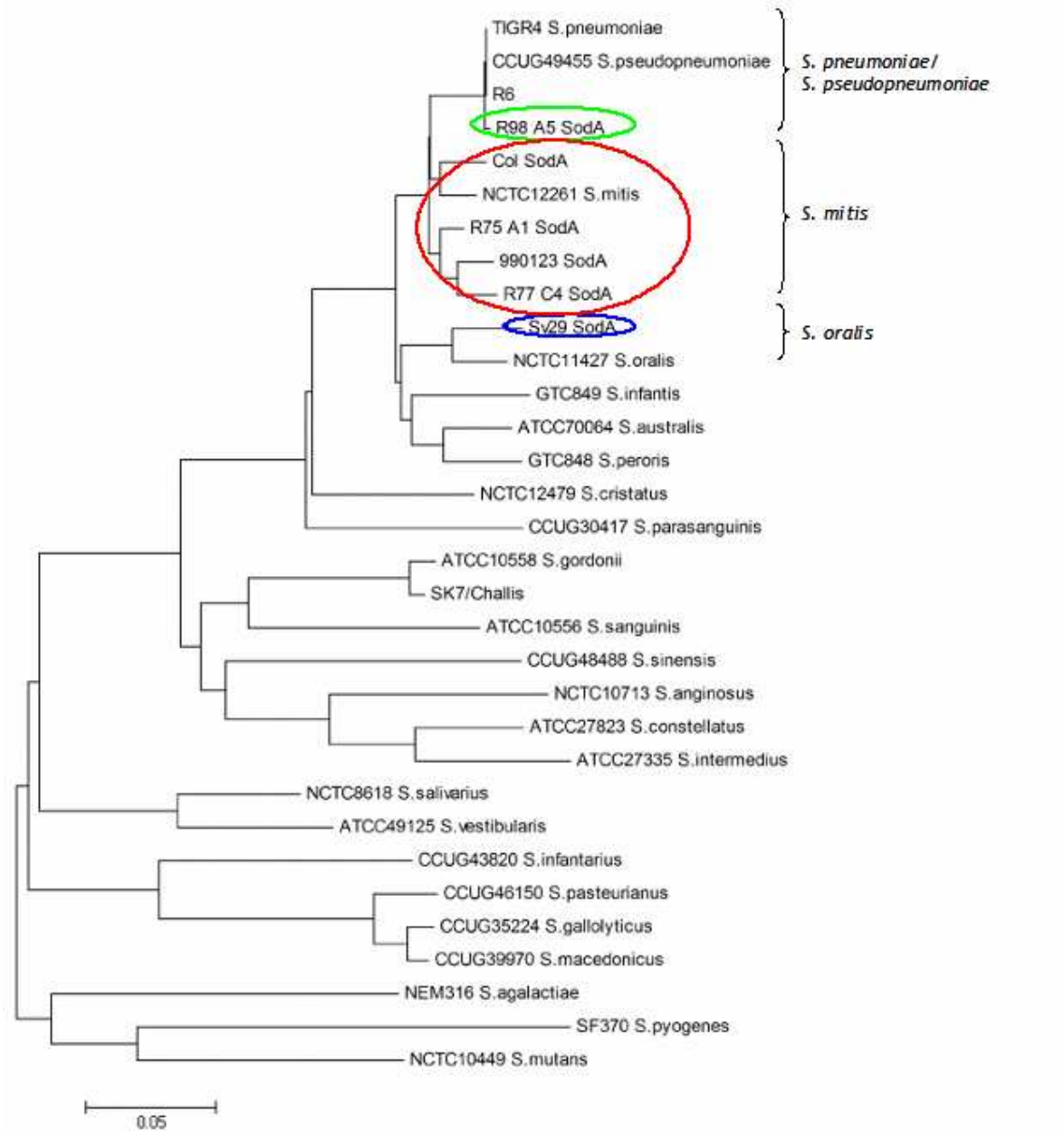


Figure 7.5: Species identification by *sodA* comparison

Minimum evolution tree drawn from alignment of *sodA* sequences from isolates for VGS microarray study to *sodA* sequences from reference isolates. Tree drawn using MEGA4 software. Highlighted with open brackets are the species *S. pneumoniae/S. pseudopneumoniae, S. mitis* and *S. oralis*. Test isolates are highlighted in green (*S. pseudopneumoniae*), red (*S. mitis*) and blue (*S. oralis*). The units at the bottom of the tree indicate distance between sequence pairs.

7.4.2 Microarray study of S. mitis and S. oralis isolates

7.4.2.1 Microarray results

The results of the DNA microarrays for each *S. mitis* or *S. oralis* isolate were aligned (full alignment in Appendix II). A high number of homologues of pneumococcal genes were present, with homologues to between 72 and 84% of pneumococcal gene identified. This high level of hybridization was expected due to the close relationship between the species. The alignment allowed the identification of 143 pneumococcal genes absent or non-hybridized from all of these isolates. Of these genes, 25 have been reported as essential for pneumococcal pneumonia in a mouse model (Table 6.2) (Hava and Camilli, 2002), and were therefore thought of as pneumococcal virulence factors.

Gene number	Gene details
	Cell envelope
SpTIGR4-0136	"glycosyl transferase, family 2"
SpTIGR4-1770	"glycosyl transferase, family 8"
SpTIGR4-1771	"glycosyl transferase, family 2/glycosyl transferase family 8"
SpTIGR4-2017	membrane protein
	Cellular processes
SpTIGR4-0071	immunoglobulin A1 protease
SpTIGR4-0314	hyaluronidase
SpTIGR4-1154	immunoglobulin A1 protease
	DNA metabolism
SpTIGR4-1040	site-specific recombinase, resolvase family
SpTIGR4-1431	type II DNA modification methyltransferase, putative
	Energy metabolism
SpTIGR4-2167	L-fuculose kinase fucK, putative
	Hypothetical proteins
SpTIGR4-1143	conserved hypothetical protein
SpTIGR4-1344	conserved hypothetical protein
SpTIGR4-1760	conserved domain protein
SpTIGR4-1952	hypothetical protein
SpTIGR4-2159	fucolectin-related protein
	Protein fate
SpTIGR4-1343	prolyl oligopeptidase family protein
	Regulatory functions
SpTIGR4-0246	transcriptional regulator, DeoR family
SpTIGR4-0461	transcriptional regulator, putative
SpTIGR4-1433	transcriptional regulator, araC family
	Signal transduction
SpTIGR4-0396	PTS system, mannitol-specific IIA component
SpTIGR4-0474	PTS system, cellobiose-specific IIC component
SpTIGR4-2162	PTS system, IIC component
SpTIGR4-2164	PTS system, IIA component
	Transport and binding proteins
SpTIGR4-1321	v-type sodium ATP synthase, subunit K
SpTIGR4-1328	sodium:solute symporter family protein
SpTIGR4-1434	ABC transporter, ATP-binding/permease protein
SpTIGR4-2086	phosphate ABC transporter, permease protein
	Unknown function
SpTIGR4-0320	oxidoreductase, short chain dehydrogenase/reductase family

Table 7.4: Essential pneumococcal genes absent from related species

Genes essential for pneumococcal pneumonia in a mouse model (Hava and Camilli, 2002) and absent from all *S. mitis* and *S. oralis* isolates tested by microarray hybridization. Bold titles in second column represent the functional groups of each gene, as reported previously (Hava and Camilli, 2002).

These genes included a trio of genes, two glycosyl transferases (SpTIGR4-1770 and 1771) and the cell wall surface anchor protein PsrP (SpTIGR4-1772), present in a pathogenicity island shown to be important for pneumococcal virulence (Obert *et al.*, 2006). Indeed most of the remaining genes in this 19 gene islet were also absent from the VGS species tested, indicating that the majority of this pathogenicity island may be absent from VGS species. Also absent was the transcriptional activator *rlrA*, shown to regulate expression of the pneumococcal pilus (Barocchi *et al.*, 2006). The RlrA islet, encoding the pneumococcal pilus, was fully absent from the VGS isolates tested, with the exception of the *S. oralis* isolate, which hybridized to 3 of the 8 genes in the islet. Other regions absent from all isolates included a PTS system encoded by the genes SpTIGR4-1615-1621, a phosphate ABC transporter encoded by genes SpTIGR4-2084-87, of which one gene, SpTIGR4-2086, was essential for pneumococcal virulence, and a PTS system and fucose operon encoded by genes SpTIGR4-2158-68, of which 4 were essential. However, when separating the isolates into commensal and pathogenic/ opportunistic isolates, there was no obvious difference in genetic content relating to ability to cause disease.

Further to this, many regions of diversity were observed in some isolates and not others, representing a number of ABC transporters, PTS systems, and the pneumococcal capsular genes. Interestingly, a number of pneumococcal capsular genes were found in two of the isolates, namely Sv29 and R77 C4. This is in agreement with previous findings that the isolate Sv29 possessed the pneumococcal capsular gene *cpsA,* although the isolate was negative for the presence of capsular polysaccharide by the Quellung reaction.

Since 25 genes identified as essential to pneumococcal pneumonia were absent from all isolates tested, a remaining 312 essential genes were found to have homologues in these VGS isolates. These included a large number of pneumococcal virulence genes, with well-defined roles in pneumococcal pathogenesis, many of which are also protein pneumococcal vaccine candidates (Table 7.5) (Tai, 2006).

Virulence factor	encoding gene	Presence in VGS	Role in virulence
Ply	*ply*	3/6	Cytotoxicity and complement activation.
NanA	*nanA*	2/6	Exposure of host receptors for adhesion.
Pneumococcal surface protein A	*pspA*	5/6	Choline-binding protein – evasion of host immune system (Ren et al. 2004) and lactoferrin binding (Shaper et al. 2004).
Pneumococcal surface antigen A	*psaA*	6/6	Mn^{2+} transport (Novak et al. 1998) and adhesion (Briles et al. 2000).
Pneumococcal surface protein C	*pspC (cbpA)*	3/6	Choline-binding protein – adhesion to host cells (Rosenow et al. 1997).
Pneumococcal histidine triad proteins	*phtA, B, D, E*	3/6, 4/6, 6/6, 1/6	Unknown, possibly metal and nucleose binding (Adamou et al. 2001).
Autolysin	*lytA*	3/6	Peptidoglycan degradation/cell lysis (Berry et al. 1989).

Table 7.5: Homologues of pneumococcal vaccine candidates

Pneumococcal virulence factors with homologues present in the related VGS population, genes encoding these factors, and roles played in pneumococcal virulence. These virulence factors are well-established pneumococcal vaccine candidates.

7.4.2.2 PCR validation

The microarray results were validated by PCR for 15 genes (Figure 7.6). The results showed that genes that had no positive hits for homologues by microarray had the same result by PCR validation (Table 7.6). However, where genes had homologues in many of the isolates tested by microarray, consistently fewer homologues were identified by PCR validation (Table 7.6). This implied that the PCR method, used to validate the microarray results of pneumococcal isolates, was less sensitive when used to validate results from related species. This was most likely due to the increased level of variation present in the genome of these species compared to the primers that were designed against the TIGR4 pneumococcal genome. As a result, the microarray results were found to be more informative on presence of gene homologues than those of the PCR validation method in this case.

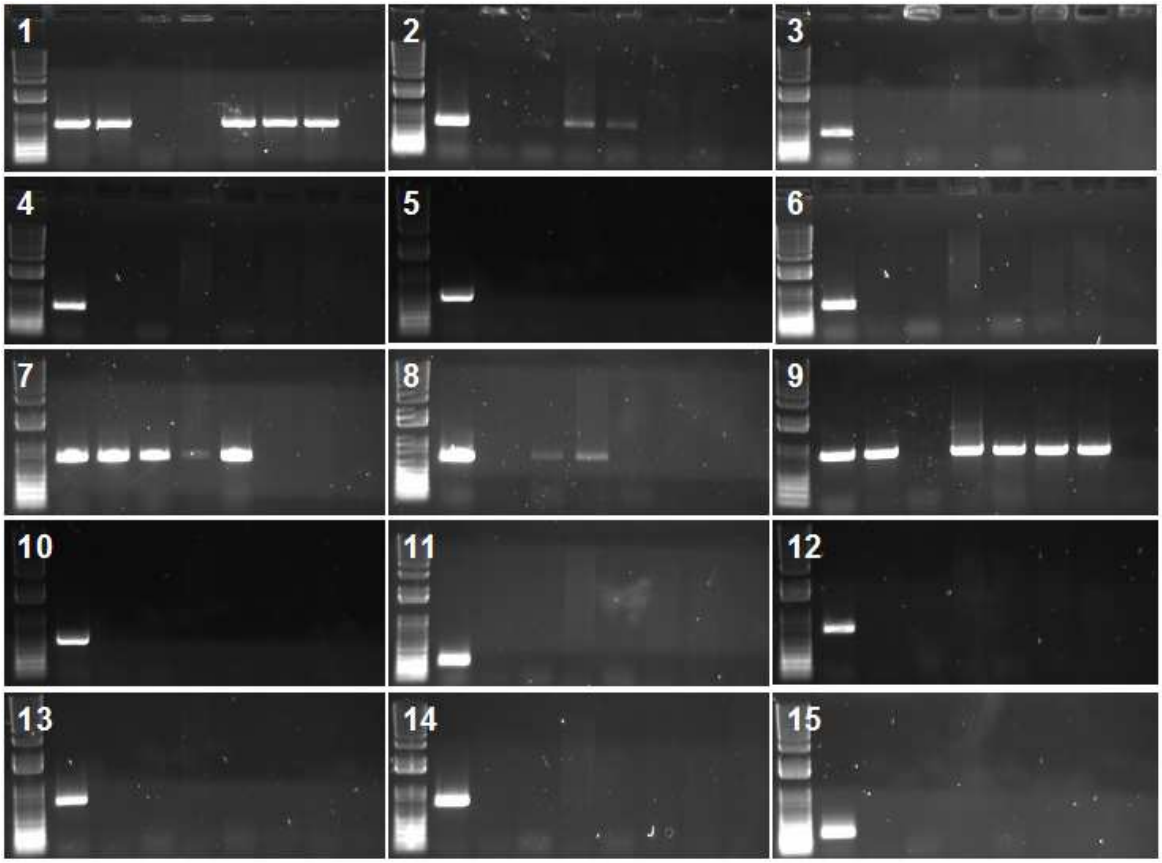


Figure 7.6: PCR validation of DNA microarray

PCR validation results for 15 genes listed in Table 6.4. Layout of gels: lane A – 1Kb+ DNA ladder; lane 2 – TIGR4; lane 3 – Col; lane 4 – Sv29; lane 5 – R75 A1; lane 6 – R77 C4; lane 7 – NCTC12261; lane 8 – 990123; lane 9 – negative control – no gDNA. The numbers on the gel represent the gene numbers assigned in Table 7.6.

No.	Gene number	gene details	microarray result	PCR result
1	SpTIGR4-0268	alkaline amylopullulanase, putative	6/6	4/6
2	SpTIGR4-0314	hyaluronidase*	0/6	2/6
3	SpTIGR4-0377	choline binding protein C	5/6	0/6
4	SpTIGR4-0461	transcriptional regulator, putative - RlrA*	0/6	0/6
5	SpTIGR4-0463	cell wall surface anchor protein - Pilus protein	0/6	0/6
6	SpTIGR4-0468	sortase, putative - Pilus sortase*	1/6	0/6
7	SpTIGR4-0530	BlpC ABC transporter, ATP-binding protein	4/6	4/6
8	SpTIGR4-0966	adherence and virulence protein A	6/6	2/6
9	SpTIGR4-0978	competence protein CoiA*	6/6	5/6
10	SpTIGR4-1154	immunoglobulin A1 protease*	0/6	0/6
11	SpTIGR4-1396	phosphate ABC transporter, ATP-binding protein	6/6	0/6
12	SpTIGR4-1770	glycosyl transferase, family 8*	0/6	0/6
13	SpTIGR4-1771	glycosyl transferase/glycosyl transferase*	0/6	0/6
14	SpTIGR4-1772	cell wall surface anchor family protein	2/6	0/6
15	SpTIGR4-2190	choline binding protein A*	4/6	0/6

Table 7.6: Comparison of microarray and PCR results

List of genes selected for PCR validation of DNA microarray results with VGS isolates. Both results from microarray and PCR are included to allow comparison. Genes with an * have been shown to be essential for pneumococcal lung infection (Hava and Camilli, 2002). No. represents number of gel in Figure 7.6.

7.4.3 Microarray study of S. pseudopneumoniae isolate

One *S. pseudopneumoniae* isolate, R98 A5, was hybridized to the pneumococcal DNA microarray to explore the presence of pneumococcal gene homologues in this species. Results showed that 91.4% of the 2240 pneumococcal TIGR4 genes present on the microarray were present in this isolate by hybridization. This represented a larger percentage than observed with the *S. mitis* and *S. oralis* isolates, where 84% was the highest level observed. This implied that there was less divergence between this species and *S. pneumoniae* than between *S. pneumoniae* and *S. mitis* or *S. oralis*. This result was supported when the percentage gene hybridization to the probes for the TIGR4 reference isolate in these species was compared by species. Results showed that, unsurprisingly, *S. pneumoniae* isolates had the highest levels of hybridization (results from Section 7.2.2). After this, *S. pseudopneumoniae* had significantly higher levels of hybridization that either *S. mitis* or *S. oralis* (Figure 7.7).

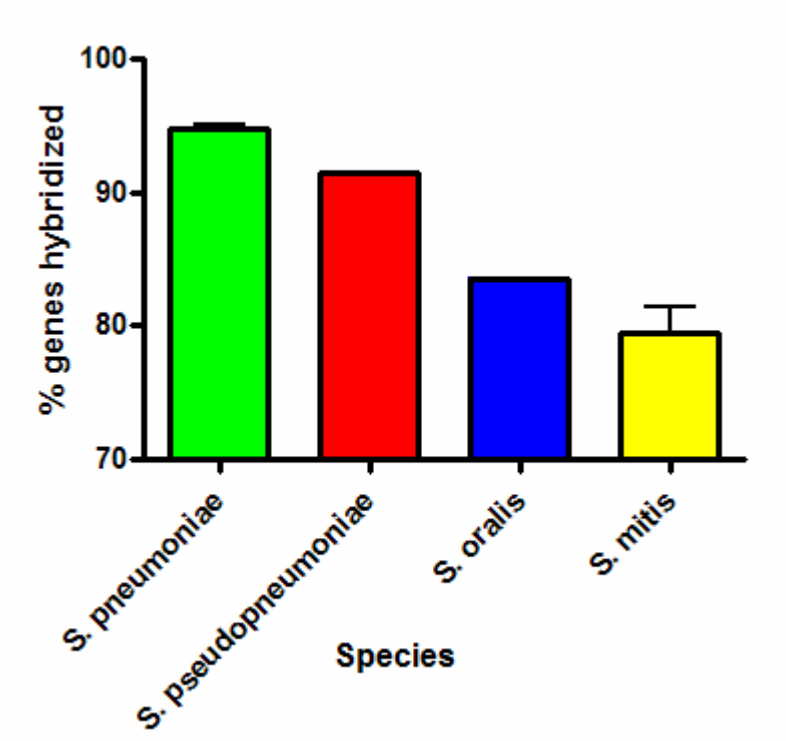


Figure 7.7: Comparison of hybridizations in microarray studies

Mean percentage of genes on DNA microarray which hybridised for each separate species. ***S. pneumoniae*** **n=6,** ***S. pseudopneumoniae*** **n=1,** ***S. oralis*** **n=1,** ***S. mitis*** **n=5.**

However, when looking at the presence or absence of virulence factors, a similar picture to that seen with *S. mitis* and *S. oralis* isolates was observed. The largest regions that appeared absent from the *S. pseudopneumoniae* isolate were predominantly the same as those absent from the other VGS isolates (Table 7.7). The full alignment can be found in Appendix II.

Region	Virulence factors	Presence in other VGS
SpTIGR4-0067-0074	IgA1 protease	6/6 isolates
SpTIGR4-0163-0169	PlcR transcriptional regulator	2/6 isolates
SpTIGR4-0348-0360	Capsule	2/6 isoaltes
SpTIGR4-0460-0468	RlrA islet - pilus	Partially, in *S. oralis*
SpTIGR4-1054-1064	ABC transporter	0/6 isolates
SpTIGR4-1129-1146	Hypothetical proteins	Partially, in 1 *S. mitis* isolate
SpTIGR4-1314-1336	Sodium ATP synthase, NanC,	Partially, in 2 isolates
SpTIGR4-1338-1351	ABC transporter	Partially, in 3 isolates
SpTIGR4-1432-1438	AraC regulator, ABC transporter	0/6 isolates
SpTIGR4-1615-1622	PTS system	0/6 isolates
SpTIGR4-1755-1773	PsrP islet	Partially, in 6/6 isolates
SpTIGR4-1948-1955	Bacteriocin	Partially, in *S. oralis*

Table 7.7: *S. pseudopneumoniae* microarray results

Main regions of diversity absent or variable in *S. pseudopneumoniae*, with associated encoded factors, and presence of these in 6 VGS isolates previously tested.

7.5 Discussion

These genomic studies have allowed further insight into pneumococcal virulence and pathogenesis, exploring genetic differences at a genome-wide level rather than focussing on individual virulence genes. The studies have allowed identification of factors with possible roles in pathogenesis of pneumococcal p-HUS as well as further exploration into the relationship between the pneumococcus and related pathogenic species. Two genes important in the glutamate metabolism pathway were differentially regulated in p-HUS isolates compared to controls. The NADP-specific glutamate dehyrdogenase gene (*gdhA)*, involved in metabolism, was up-regulated, whilst a glutamine amidotransferase gene (*guaA)*, involved in catabolism, was down-regulated. These expression changes resulted in increased cellular levels of both glutamate and glutamine, which are the main pneumococcal nitrogen donors (Kloosterman *et al.*, 2006), in p-HUS isolates. *gdhA,* which is also involved in nitrogen metabolism, is independently repressed by both the nitrogen regulatory protein GlnR and the pleiotropic regulator CodY, which seems to play a more important role (Kloosterman *et al.*, 2006). In *B. subtilus*, CodY is responsible for regulation of cellular nutritional and energy status (Kim *et al.*, 2003; Shivers and Sonenshein, 2005), and therefore *gdhA* has been suggested as an important control point for nutritional and energy regulation in the pneumococcal cell (Kloosterman *et al.*, 2006). The up-regulation of this gene in p-HUS isolates may represent an increased energy requirement in these isolates in the environment generated by p-HUS symptoms. Studies have shown that a deletion mutant of *gdhA* had significantly reduced adherence to Detroit cells (20% of WT). GdhA may therefore have an indirect role in pneumococcal adhesion. An interesting study would be to knock this gene out in p-HUS isolates and controls and observe the difference in adhesion, as the product of this gene appears to be important in penumococcal adhesion, which may be linked to p-HUS pathogenesis. However, the *gdhA* knock-out showed no significant difference in virulence when compared to the wild-type strain D39 in colonisation, pneumonia or bacteraemia models *in vivo* (Hendriksen *et al.,* 2008). The repressor GlnR is up-regulated by the presence of high levels of glutamate and glutamine, as well as high ammonia and ammonium concentrations (Hendriksen *et al.*, 2008). Due to the symptom of kidney dysfunction, p-HUS patients are likely to have high levels of circulating ammonia and ammonium (Nath *et al.*, 1991). These high levels should in theory

result in repression of *gdhA* by GlnR, however, *gdhA* was found to be significantly up-regulated in p-HUS isolates and there was no difference in expression of the *glnR* gene between p-HUS isolates and controls. This was also the case for the *codY* gene, which also represses *gdhA*. This implies that another factor or environmental stimulus may be involved in regulation of glutamate metabolism and nitrogen metabolism via up-regulation of *gdhA*. It is possible that this unknown factor also acts on *guaA*, but has the opposite effect, resulting in a net increase in glutamate metabolism. This altered gene expression may be involved in ability to cause p-HUS or survive in the altered environment of a host with p-HUS symptoms.

Preliminary results indicated that p-HUS isolates may survive better in high ammonium concentrations, which are present in p-HUS patients. However, the differences observed between the p-HUS isolate and the control, when grown in varying concentrations of ammonium, require further study. Whilst a large difference was observed between the isolates, this result was not observed upon repeat. However, it is likely that some phenotypic differences exist between the isolates, possibly as a result of altered gene expression, as such a stark difference in susceptibility to ammonium was observed. Further study is required to validate this hypothesis. This, however, would not represent a mechanism of pathogenesis employed by these isolates to cause p-HUS, but rather a reaction to the symptoms caused in the host. However, it may be that due to unknown genetic factors, not all pneumococcal isolates are able to adapt and survive in this environment.

Of the other genes differentially regulated in p-HUS isolates compared to controls, none play specific roles with obvious links to p-HUS-causing ability. As a result, it is difficult to draw conclusions on the importance of their up- or down-regulation in the p-HUS isolates. Two of the twelve differentially regulated genes, namely CinA and ABC-N/P, have previously been identified as essential for pneumococcal virulence in a mouse model (Hava and Camilli, 2002). Whilst expression of CinA is down-regulated in p-HUS isolates, ABC-N/P expression is up-regulated. However, since these genes are present in both p-HUS isolates and controls, and simply differentially regulated, it is unlikely that their essentiality has any effect on the isolates. One drawback of the expression microarray study was the inability to recreate the in vivo growth conditions to which the

pneumococci are exposed when causing p-HUS, since so little is known about disease progression. If this had been possible, it may have resulted in clearer patterns of gene expression in the p-HUS isolates compared to the controls. However, these results show that despite this, differences in gene expression levels correlating to p-HUS can be observed in these isolates.

One of the main criticisms of microarray studies is that only genetic differences in genes present in the reference strain, in this case TIGR4, are detectable. It is therefore possible that genes important in p-HUS are present in these p-HUS isolates and absent from the controls, but also absent from the reference strain. In this case, these genetic differences would be overlooked by the microarray analyses. Similarly, differential expression of these genes may give insight into their involvement in p-HUS, but is not detectable using this technology. Furthermore, mutations in genes may go undetected as the genes may be detectable by DNA microarray, and have regular levels of expression, so mutations would also go unnoticed by the RNA microarray. However, as described with Ply allele 5 (see Chapter 3), mutations can have drastic effects on the activities of virulence factors. This is a further drawback of microarray technology, as mutations may be undetected, whilst clearly having a major impact on virulence factor activity and consequently pneumococcal virulence. A method that would overcome all of these problems is full-genome sequencing of the pneumococcal genome. Full pneumococcal genome sequencing projects have resulted in large advances in understanding of pneumococcal biology, as well as identification of vaccine candidates (Hiller *et al.*, 2007; Hoskins *et al.*, 2001; Lanie *et al.*, 2007; Tettelin *et al.*, 2001). Furthermore, this is becoming increasingly affordable, and our laboratory will shortly be able to sequence pneumococcal genomes at a relatively high-throughput level. It would be of obvious interest to sequence the genomes of these isolates in order to identify true differences in genetic content, and mutations, as this should offer more definitive insight into genes involved in p-HUS. When this technique becomes routine, it is likely that microarray technology will become obsolete the purpose of full genome analysis, although it may still have uses as a diagnostic tool, where probes for a number of pathogenic bacteria are present. Furthermore, a microarray for use in determining pneumococcal serotype is in development by

the Bacterial Microarray Group at St. Georges Hospital, London, and may replace the techniques currently in use, although cost is currently prohibitive.

Taking into consideration the fact that the isolates may differ in genetic content not present on the array, it was of interest to determine if virulence factors present on mobile genetic elements were associated with isolates causing p-HUS. Bacteriophages have been shown to harbour important virulence factors in a number of pathogenic species including cholera toxin in *Vibrio cholerae* (Waldor and Mekalanos, 1996) and Shiga-like toxin in enterohaemorrhagic *E. coli* (EHEC) (O'Brien *et al.*, 1984). Indeed the Shiga-like toxin that allows EHEC to cause typical HUS is encoded on a phage inserted in the *E. coli* genome (O'Brien *et al.*, 1984). It was hypothesised that the genetic information possessed on mobile lysogenic bacteriophages in the pneumococcus may be associated with isolates causing p-HUS. However, although two isolates tested possessed lysogenic phages, there was no correlation between possession of lysogenic bacteriophages and p-HUS. This result shows that unlike EHEC, where the bacteriophage-encoded Shiga-like toxin mediates HUS (Banatvala *et al.*, 2001), virulence factors carried on lysogenic bacteriophages are not required to cause p-HUS.

Studies have shown that particular serotypes and sequence types (STs) are more commonly isolated from p-HUS patients (Vanderkooi *et al.*, 2003; Waters *et al.*, 2007). This implies that there are genetic differences between these and other serotypes and STs that increase their ability to cause p-HUS. Although there were differences in genetic content and gene expression between the p-HUS isolates tested and matched controls, the isolates, which were all serotype 19A, ST199, were relatively clonal, and shared large regions of diversity in comparison to the reference isolate TIGR4. The 6 isolates (4 HUS and 2 controls) shared 12 regions of gene absence or diversity in comparison to TIGR4. These varied in size from 2 genes to 22 genes. In contrast to these major differences from the reference isolate TIGR4, relatively few differences were observed between the isolates. This shows that the isolates are fairly clonal. It is therefore possible that although the controls used in this study did not cause p-HUS, they possess the genetic factors that allow them to do so, in common with the p-HUS isolates tested. If this is the case, it suggests that although bacterial genetics may play a role, other factors such as host genetics or environmental factors may have a more significant influence on which cases of invasive

pneumococcal disease progress to cause p-HUS. Although this study is merely a starting point, further understanding of the bacterial methods of pathogenesis employed may improve diagnosis and treatment of p-HUS, and reduce the mortality levels observed.

The presence of homologues of pneumococcal virulence genes in related VGS species has been demonstrated previously, including NanA (King *et al.*, 2005), penicillin binding protein 2x (Sibold *et al.*, 1994), IgA proteases (Poulsen *et al.*, 1998), autolysin A (Obregon *et al.*, 2002; Whatmore *et al.*, 2000) and Ply (Neeleman *et al.*, 2004; Whatmore *et al.*, 2000). The results presented here show that homologues of a large proportion of pneumococcal genes are present in species of *S. mitis*, *S. oralis* or *S. pseudopneumoniae*. This was not entirely unexpected, as these species are known to be the closest relations to *S. pneumoniae (Arbique et al., 2004; Facklam, 2002)*. Generally, however, *S. mitis* and *S. oralis* are commensal organisms, residing asymptomatically in the host nasopharynx and oral cavity. This contrasts strikingly with the role of *S. pneumoniae* as a major human pathogen causing a plethora of lethal diseases. Therefore, although these species are relatively closely related, and share many homologues, it was hypothesised that the differences observed between the species would lie in the presence of virulence factors with known roles in pneumococcal pathogenesis. However, the results presented here show that this is not the case, with many important pneumococcal virulence factors having homologues in the related VGS species (Table 7.5). Many of the virulence factors present in the *S. mitis* population are currently being considered as pneumococcal protein vaccine candidates, and the presence of homologues in related species is of detriment to their potential as vaccine candidates, since selective pressure from a vaccine may allow acquisition of altered variants of these factors by pneumococci, and evasion of vaccines targeting them. A similar phenomenon has been observed in the transfer of penicillin resistance gene *pbp2x* from *S. mitis* to *S. pneumoniae,* upon exposure to penicillin, resulting in penicillin resistant pneumococci (Chi *et al.*, 2007).

This work discovered a number of pneumococcal islets, with factors shown to be essential to pneumococcal virulence, which were absent from the related VGS isolates tested. Targeting the essential factors in these regions may be of benefit as the possibility of vaccine evasion by acquisition of altered genes

would be negated. However, although two islets, the RlrA islet and the PsrP islet, have been shown to be involved in pneumococcal virulence (Barocchi *et al.*, 2006; Obert *et al.*, 2006), they are not universally present in the pneumococcal population. Indeed, the RlrA islet was shown to be present in only 27-30.6% of pneumococci (Aguiar *et al.*, 2008; Moschioni *et al.*, 2008), whilst the PsrP islet was also shown to be variably present in *S. pneumoniae* isolates (Dr. Donald Inverarity, personal communication). It is therefore unsurprising that these regions were also absent from related VGS species. As a result of this, potential targeting of these regions would require to be as part of a multifactorial protein vaccine targeting various pneumococcal virulence factors to ensure full coverage. This possibility has been described previously as having the most potential for creating a successful protein vaccine (Ogunniyi *et al.*, 2000; Ogunniyi *et al.*, 2007a). A number of other regions were identified to be fully absent from tested isolates, and although less is known about the roles of these regions, a number of essential virulence genes were located within them (Hava and Camilli, 2002). However, many of the proteins in these regions remain hypothetical, and further study may uncover genes in these regions that may be of interest as pneumococcal vaccine candidates if they are found to be absent from all related VGS isolates.

A recent study showed that whilst the house-keeping genes in the pneumococcus are fairly conserved, shared alleles of 4 house keeping genes (*ddl, gdh, rpoB* and *sodA*) in *S. mitis, S. oralis* or *S. infantis* species were almost non-existent (Kilian *et al.*, 2008). This implied that a much larger level of variation was present within these species that in the pneumococcus. Furthermore, by normal taxonomic methods, the authors demonstrated that many *S. mitis* isolates did not share enough homology to belong to the same species. These results, coupled with the observation that pneumococcal genomes were on average 0.3Mb larger than *S. mitis* genomes, allowed them to hypothesise that these species shared a common, pneumococcus-like ancestor, and that *S. mitis*, as a species, has been evolving towards a fully commensal lifestyle by loss of virulence genes (Kilian *et al.*, 2008). The results of the microarray study, with large levels of variation present between *S. mitis* isolates, and pneumococcal virulence factors variably distributed within the *S. mitis* population, appear to support with this hypothesis (See appendix II). However, it also seems possible

that evolution is occurring in the opposite direction, with the pneumococcus evolving away from a commensal ancestor, with *S. mitis* and *S. pseudopneumoniae* isolates in the process of the same evolution. Although results may be used to support either hypothesis, the authors do not address the possibility. As a result, further studies are required to determine the evolutionary relationships between these species.

The picture of *S. pneumoniae* as a human pathogen and related species as commensals is blurred by their potential to cause opportunistic diseases including endocarditis (Hsu and Lin, 2006), dental abscesses (Wickremesinghe and Russell, 1976) and bacteraemia (Han *et al.*, 2006; Hoshino *et al.*, 2005; Smith *et al.*, 2004). Furthermore, a number of *S. mitis* isolates have been isolated from the CSF of patients with meningitis, implying that this species may also cause opportunistic meningitis (Echeverria *et al.*, 1998; Kutlu *et al.*, 2008). This shows that although these species are generally commensal, they possess virulence factors that allow them to cause opportunistic disease in certain environments, or in hosts with compromising factors. This is supported by results which showed no clear difference in virulence factor homologues in VGS isolated from disease compared to commensal VGS. This implied that although the commensal isolates were not causing opportunistic disease when isolated, they may possess the virulence factors to do so under the correct conditions. However, recent findings showed that the *S. mitis* species consisted of many different, loosely related lineages (Kilian *et al.*, 2008). It is therefore possible that whilst some of these lineages may be able to cause opportunistic disease, other lineages have evolved to such an extent as to be fully commensal. By the taxonomic methods used, these would all fall under the bracket of *S. mitis*.

The PCR validation of 15 genes in *S. mitis* and *S. oralis* showed that the hybridization reaction was more sensitive with these species. In almost all cases, when positive results were obtained, there were fewer positives in the PCR validation than the initial hybridization (Table 7.6). This was the opposite to what was seen when validating results of pneumococcal hybridizations (Section 7.2.2, Dr. Donald Inverarity, personal communication). A possible reason for this was the level of diversity, which was obviously increased in these isolates when compared to the pneumococcus (Figure 7.7). Furthermore, the probes were designed against the sequenced genome of the reference pneumococcal isolate

TIGR4 (Tettelin *et al.*, 2001), and therefore were not optimised for hybridization with related VGS species. Therefore, since the hybridization relied only on one probe for a positive result, whilst the PCR required recognition of two primers for successful amplification, it was likely that the hybridization reaction would be more successful than PCR, as a result of increased diversity in these isolates. This, however, does not rule out the use of PCR as a screen for penumococcal homologoues, it simply shows that care must be taken to design PCR primers in regions with as little variation as possible, to prevent failure of the PCR due to high diversity levels.

As described previously, microarray technology, when used for genome-wide analysis, has a number of drawbacks. In this case, as in the HUS study, full genome sequencing of a number of isolates from each species would allow a much more robust and informative comparison between species. For example, whilst homologues of many pneumococcal virulence factors have been identified in this study, the level of diversity present in these is not known. Furthermore, non-hybridization to the microarray may be as a result of increased diversity in the probed region, rather than absence of a homologue. It is known that the level of diversity can act as a barrier to homologous recombination in pneumococci (Mortier-Barriere *et al.*, 1997). Therefore, although homologues are present to pneumococcal genes, the genetic divergence between the homologues may be such that homologous recombination between the species is not possible. Full genome sequencing would not only allow confirmation of homologue presence in these related species, but also determination of the diversity present, and likelihood of homologous recombination allowing vaccine escape.

The comparison between hybridized genes in *S. pneumoniae, S. pseudopneumoniae, S. mitis* and *S. oralis* implied that there was less diversity between *S. pneumoniae* and *S. pseudopneumoniae* than *S. pneumoniae* and *S. mitis* or *S. oralis* (Figure 6.2). These results give support to the hypothesis that *S. pseudopneumoniae* is an independent species, distinct from both *S. pneumoniae* and *S. mitis* (Arbique et al. 2004, Kilian et al. 2008). However, only one *S. pseudopneumoniae* isolate was tested in this study, and study of further isolates would confirm this hypothesis.

In conclusion, this study has shown that despite a number of limitations, the use of microarray technology can give important insight into the genetic differences between isolates of *S. pneumoniae*. This study identified a number of genes differentially regulated in p-HUS isolates, which may play a role in pathogenesis. However, the relative clonality of the p-HUS and control isolates, implies that although bacterial genetics may be important, other factors such as host genetics or environmental factors may play a more significant role in pathogenesis of p-HUS. Furthermore, this study shows for the first time that pneumococcal microarrays may be used to assess presence of virulence gene homologues in related species. A large number of homologues were identified, allowing novel insight into the evolutionary relationship between these species and the genetic diversity that results in stark differences in pathogenic potential. Furthermore, these results should be taken into consideration when studying virulence factors as candidates for future protein-based pneumococcal vaccines.

8 Final Discussion

8.1 Final Discussion

The initial aim of this project was to explore the diversity present in pneumococcal virulence factors, and relate differences to pneumococcal disease. The diversity present in the pneumococcal virulence factor Ply was found to be greater than expected for a cytoplasmic protein, with the 14 protein alleles discovered in this study showing a diversity of 3.3%. This is higher than pneumococcal house-keeping genes, where 1-2% variation is observed (Enright and Spratt, 1998). However, this level of diversity is significantly lower than in surface-exposed pneumococcal virulence factors such as PspA, NanA and PspC (Hollingshead *et al.*, 2000; Iannelli *et al.*, 2002; King *et al.*, 2005). The diversity present resulted in Ply alleles with varying haemolytic activity, with allele 5 completely non-haemolytic. We are the first to report a pneumococcal clinical isolate producing no recogniseable Ply, due to interruption of the *ply* gene by the insertion sequence IS1515. Furthermore, in Ply allele 12, a point mutation resulted in non-recognition by a diagnostic monoclonal antibody, giving rise to the possibility of false-negative diagnosis of pneumococcal disease. However, the diversity present in the Ply protein should not prevent its use in future protein-based pneumococcal vaccines, as all alleles identified were recognised by polyclonal antibodies and diversity remains lower than in many other surface-exposed vaccine candidates.

The non-haemolytic allele 5 has been documented in serotype 1 isolates by our lab previously (Kirkham *et al.*, 2006a). The toxin mutant was shown to be harboured by the ST306 clone, which had undergone recent expansion in the Scottish serotype 1 IPD population. This study has reported the presence of this non-haemolytic variant in a further 8 serotype 1 clones, all closely related to ST306, demonstrating that this mutant is wide-spread in the serotype 1 IPD population. Furthermore, allele 5 was identified for the first time in the dominant clone of serotype 8 pneumococci, as well as related clones. This mutant may therefore play a role in driving expansion of these unrelated clones within the pneumococcal IPD population. Furthermore, these two serotypes have both been associated with rare outbreaks of pneumococcal disease, and the non-haemolytic toxin harboured by these clones may play a role in this association.

The reasons for these associations remain unclear, although results of other studies have shown that serotype 1 clones generally cause mild disease (Sjostrom *et al.*, 2006), which may be as a result of loss of haemolytic activity in certain clones. This loss of severity, and associated patient survival, may increase the potential of transmission of these clones, resulting in an apparent expansion within the population, as well as tendency to cause outbreaks of IPD. It is unclear if a similar phenomenon occurs in serotype 8 clones, but this observation may link the non-haemolytic allele to the clonal expansion observed, although further study would be required to confirm this. The full sequenced genomes of ST306 and related clone ST227 are now available, and comparison of these would allow further insight into the genomic differences between them, and possibly identify other genetic differences involved in driving the observed clonal expansion.

Construction of an unmarked Ply knock-out in a D39 background allowed preliminary study into the importance of Ply in pneumococcal pneumonia. Results showed that this mutant was less virulent that wild-type, implicating Ply as important in pneumonia, and agreeing with previous studies (Wellmer *et al.*, 2002; Winter *et al.*, 1997). Furthermore, comparison with a non-haemolytic isolate showed no difference in virulence, implying that the haemolytic property of the toxin, in this background and disease model, is the most important property of Ply. However, further study including in vivo imaging studies, would allow greater insight into the differences in pathogencity of these mutants, and elucidate further the roles of Ply in virulence. In addition, this unmarked Ply mutant is now used routinely in our laboratory for a variety of Ply studies.

A previous study reported high levels of diversity within the NanA protein (King *et al.*, 2005). This study confirms these high levels of diversity, with diversity at 13.7% of amino acid positions, mosaic blocks and insertions identified. From 33 isolates, 18 protein alleles identified. It was thought that this high level of diversity may restrict the use of NanA in pneumococcal vaccines, as it may not afford full cross-serotype protection. However, antibodies raised against wild-type NanA recognised 4 highly divergent alleles, implying that the diversity present in the NanA protein should not restrict the use of this virulence factor as a vaccine candidate with the potential of cross-serotype protection. Mapping of mutations onto the solved structure of the enzymatic domain of NanA allowed

identification of mutations in this region, and purification of alleles showed up to 7-fold differences in activity between different alleles. Furthermore, a novel insertion region within the NanA active site was identified, and mosaic block C was found to map to this region. Two distinct variants of the region were discovered, with striking differences in surface shape and electrostatic potential, although the role of this region in pneumococcal pathogenesis remains unknown.

The activity of pneumococcal neuraminidases has been implicated as important in pathogenesis of p-HUS, which displays high mortality rates compared to typical HUS (Klein *et al.*, 1977; Vanderkooi *et al.*, 2003). In order to explore this hypothesis further, and attempt to link specific NanA amino acid alleles to pneumococcal disease, the importance of NanA in the progression of p-HUS was investigated. However, results showed there was no correlation between NanA amino acid allele or activity and ability to cause p-HUS. Therefore, although neuraminidase activity may be required for pneumococci to cause p-HUS, these results imply that other bacterial factors, host genetics or environmental factors may play a more important role.

Homologues of Ply and NanA have been reported in related VGS species previously, although these are not well characterised (Byers *et al.*, 2000; King *et al.*, 2005; Neeleman *et al.*, 2004; Whatmore *et al.*, 2000). This study has named and characterised a new CDC in a subset of *S. mitis* isolates, called Mly. This novel toxin was shown to be genetically distinct from Ply, whilst actively secreted in a number of isolates. Furthermore, this study began characterization of a toxin in *S. pseudopneumoniae*, termed pPly. This toxin was more closely linked to Mly than Ply. A NanA homologue was also identified for the first time in isolates of *S. mitis* and *S. pseudopneumoniae,* although the NanA proteins in these species were very closely related to pneumococcal NanA. The roles of these homologues in these commensal species remains unknown, however, the relative conservation of NanA homologues compared to Ply homologues implies it may play a more significant role in these commensal species. This may be due to the importance of NanA in adherence and colonisation of the bacteria, whilst Ply is more important in disease. Furthermore, the presence of these homologues in commensal *Streptococci*, coupled with the ability of the pneumococcus to alter its genetic information by horizontal gene transfer, gives rise to the possibility of

evasion of pneumococcal vaccines targeting these virulence factors. This is a concrete threat, with evasion of PCV7 by capsular switch between pneumococci already a major problem in the USA (Brueggemann *et al.*, 2007; Pai *et al.*, 2005b), and an increase in penicillin resistance in pneumococci due to interspecies recombination (Chi *et al.*, 2007). These findings demonstrate that the evasion of a vaccine targeting a virulence factor with high levels of diversity, or homologues in related species is possibly due to the remarkable plasticity of the pneumococcal genome.

Structural studies of the pneumococcal enzyme NanA showed that despite <10% sequence similarity, the enzymatic β-propeller of pneumococcal NanA shared striking structural homology with influenza NA. This study presents the novel discovery that pneumococcal NanA is inhibited in vitro by specific viral NA inhibitors. In particular, OC showed high potency in inhibiting wild type NanA, with an IC_{50} of 1.608μM. However, further studies are required to determine accurate levels of inhibition of other pneumococcal NanA alleles with this inhibitor. This novel discovery was further investigated in an *in vivo* model of pneumococcal pneumonia, with pre-treatment with OC delaying the onset of pneumonia and reducing severity of symptoms. Conversely, pre-treatment with purified NanA increased severity of symptoms. These findings offer further insight into the synergistic relationship between influenza and the pneumococcus, as the similarity between these enzymes adds weight to the hypothesis that influenza NA may expose host receptors, priming the host for pneumococcal infection (McCullers and Bartmess, 2003). Furthermore, although further studies are required, these results confirm that OC has potential for treatment of secondary pneumococcal pneumonia, and for the first time demonstrate that this inhibitor has inhibitory effects on the secondary as well as primary pathogen in this lethal synergism. Finally, optimization of this drug to pneumococcal NanA may increase its potential as a pneumocidal drug, and derivatives may also have a role in preventing other pneumococcal diseases, such as p-HUS, where NanA is thought to play an important role.

Further to these studies exploring diversity within individual virulence factors, microarray technology was used to explore genetic diversity at a full genome level. Although no differences in genetic content were observed between p-HUS

isolates and matched controls, a number of differences in gene expression were observed. These may play a role in allowing certain clones to cause p-HUS. Two of these genes, *ghdA* and *guaA* play roles in glutamate and glutamine metabolism, and differential regulation resulted in increased metabolism in p-HUS isolates. Expression of *gdhA* is repressed by high concentrations of ammonium, circulating in p-HUS patients. Nevertheless, this gene is up-regulated in p-HUS isolates, implying control by another factor. Although further study is required to unravel the pathogenesis of p-HUS, this study has highlighted several candidate genes with potential roles in p-HUS pathogenesis and provided a baseline for further study to elucidate the mechanisms of p-HUS pathogenesis. Understanding the pathogenic methods employed by the pneumococcus will allow more effective diagnosis and treatment of p-HUS.

Microarray studies of related VGS species with pneumococcal microarrays gave novel insight into the relationship between these species. Species of *S. mitis* and *S. oralis* were shown to possess homologues to 72-84% of pneumococcal genes, whilst *S. pseudopneumoniae* possessed 91.4%. This result supported previous findings that *S. pseudopneumoniae* is a species distinct from both *S. pneumoniae* and *S. mitis* (Arbique *et al.*, 2004). Furthermore, this study reports the presence of many pneumococcal virulence factors in these commensal relatives. The scattered distribution of these virulence factors in commensal species supports the hypothesis that these species are evolving towards commensality from a pathogenic ancestor (Kilian *et al.*, 2008), although evolution in the other direction is also a possibility. Furthermore, these results can inform on the potential of vaccine escape by horizontal gene transfer, and should be considered when evaluating the potential of specific pneumococcal virulence factors as vaccine candidates.

To conclude, this study has offered numerous novel insights into the biology and pathogenesis of the pneumococcus. Distribution and diversity studies of Ply and NanA presented here contribute significantly to the body of knowledge on these important virulence factors, as well as offering insight for pneumococcal vaccine design. Notably, the diversity present in both Ply and NanA should not restrict their use as pneumococcal vaccine candidates. Identification of pneumococcal homologues in related commensal species has furthered understanding of the relationships between these species, whilst also adding to our understanding of

S. mitis biology, and furthering knowledge for pneumococcal vaccine design. Furthermore, this work has begun to unravel the pathogenesis of p-HUS, furthered our understanding of the lethal synergism between the pneumococcus and influenza, and identified potential inhibitors to co-treat these infections.

Appendices

- Appendices can be found on CD at the back of this thesis. The contents of the Appendices are as follows:

- Appendix I – Sequence alignments and sequence information.
    1. Full spreadsheet of Ply protein alleles discovered in large-scale screen (Chapter 3).
    2. DNA alignment of *ply* alleles of all isolates fully sequenced during study, with the exception of S1-11, possessing the insertion sequence IS1515 (47 isolates) (Chapter 3).
    3. Full DNA alignment of *nanA* sequences (Chapter 4).
    4. *mly* alleles from *S. mitis* aligned to *ply* allele 1 from *S. pneumoniae* D39 (Chapter 5).
    5. *pply* alleles from *S. pseudopneumoniae* aligned to *ply* allele 1 from *S. pneumoniae* D39 (Chapter 5).
- Appendix II – Microarray alignments.
    1. General layout of spreadsheets - Genes aligned in order of presence in reference TIGR4 genome, with those highlighted in blue present in the test isolate, and those in red absent. Genes highlighted in colour are essential for pneumococcal pneumonia in a mouse model (Hava and Camilli, 2002).
    2. Full alignment of comparative genomic hybridization results from 6 serotype 19A, ST199 isolates in p-HUS study (4 p-HUS isolates and 2 controls, Chapter 7).
    3. Full alignment of comparative genomic hybridization results from 5 *S. mitis* and 1 *S. oralis* isolates (Chapter 7)
    4. Full alignment of comparative genomic hybridization results from 1 *S. pseudopneumoniae* isolate (Chapter 7)
- PDF copies of full thesis and associated publications (see next page) are also included.

Publications

Jefferies, J., Nieminen, L., Kirkham, L., Johnston, C., Smith, A., Mitchell, T.J. (2007) Identification of a secreted cholesterol-dependent cytolysin (mitilysin) from *Streptococcus mitis*. *J Bacteriol* **189 (2)**: 627-32.

Jefferies, J.[¥], Johnston, C.[¥], Kirkham, L., Cowan, G., Ross, K., Smith, A., Clarke, S., Brueggemann, A., George, R., Pichon, B., Pluschke, G., Pfluger, V., Mitchell, T.J. (2007) Presence of non-haemolytic pneumolysin in serotypes of *Streptococcus pneumoniae* associated with disease outbreaks. *J Infect Dis* **196 (6)**: 936-44. [¥] **Contributed equally to the article.**

Gut, H., Ross, K., Johnston, C., Mitchell, T.J., Walsh, M. (2008) Inhibition of *Streptococcus pneumoniae* NanA with Oseltamivir reduces pneumonia. (Atricle in press, PLOS Pathogens)

Johnston, C., Romero, P., Inverarity, D., Gut, H., Paterson, G., Hinds, J., Walsh, M., George, R., Leonard, A., Smith, A., Mitchell, T.J. (2008) Investigating the impact of pneumococcal diversity on haemolytic uraemic syndrome. (Article in press, Journal of Clinical Microbiology).

Johnston, C., Hinds, J., Smith, A., Mitchell, T.J. (2008) Homologues of pneumococcal virulence genes in related *streptococci* has implications for pneumococcal vaccine design. (Manuscript in preparation).

Conference contributions (presenting author underlined)

Johnston, C., Romero, P., Gut, H., Walsh, M., Pichon, B., Slack, M., George, R., Smith, A., Mitchell, T.J. Genetic analysis of *Streptococcus pneumoniae* isolates associated with haemolytic uraemic syndrome. (2008, poster) 6th International Symposium on Pneumococci and Pneumococcal Diseases (ISPPD-6), Reykjavik, Iceland.

Gut, H., Ross, K., Johnston, C., Mitchell, T. J., Walsh, M. Inhibition of *Streptococcus pneumoniae* NanA with Oseltamivir reduces pneumonia. (2008, poster) 6th International Symposium on Pneumococci and Pneumococcal Diseases (ISPPD-6), Reykjavik, Iceland.

Johnston, C., Jefferies, J., Smith, A., Mitchell, T.J. Presence of homologues of *Streptococcus pneumoniae* virulence genes in related viridans group *streptococci* has implications for pneumococcal vaccine design. (2008, poster) 9th European Oral Microbiology Workshop (EOWM), Helsinki, Finland.

Johnston, C., Jefferies, J., Kirkham, L., Cowan, G., Ross, K., Smith, A., Clarke, S., Brueggemann, A., George, R., Pichon, B., Pluschke, G., Pfluger, V., Mitchell, T.J. Presence of non-haemolytic pneumolysin in serotypes of *Streptococcus pneumoniae* associated with disease outbreaks. (2007, poster) Europneumo, Lisbon, Portugal.

Johnston, C., Jefferies, J., Kirkham, L., Cowan, G., Ross, K., Smith, A., Clarke, S., Brueggemann, A., George, R., Pichon, B., Pluschke, G., Pfluger, V., Mitchell, T.J. Presence of non-haemolytic pneumolysin in serotypes of *Streptococcus pneumoniae* associated with disease outbreaks. (2007, poster) Society for General Microbiology AGM, Manchester, UK.

List of References

Abeyta, M., Hardy, G.G., and Yother, J. (2003) Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. *Infect Immun* **71**: 218-225.

Abrahams, A., Hallows, N., and French, H. (1919) A further investigation into influenzo-pneumococcal and influenzo-streptococcal septicaemia. *Lancet*: 1-11.

Abramson, J.S., Giebink, G.S., and Quie, P.G. (1982) Influenza A virus-induced polymorphonuclear leukocyte dysfunction in the pathogenesis of experimental pneumococcal otitis media. *Infect Immun* **36**: 289-296.

Adamou, J.E., Heinrichs, J.H., Erwin, A.L., Walsh, W., Gayle, T., Dormitzer, M., Dagan, R., Brewah, Y.A., Barren, P., Lathigra, R., Langermann, S., Koenig, S., and Johnson, S. (2001) Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis. *Infect Immun* **69**: 949-958.

Aguiar, S.I., Serrano, I., Pinto, F.R., Melo-Cristino, J., and Ramirez, M. (2008) The presence of the pilus locus is a clonal property among pneumococcal invasive isolates. *BMC Microbiol* **8**: 41.

Alcantara, R.B., Preheim, L.C., and Gentry, M.J. (1999) Role of Pneumolysin's complement-activating activity during pneumococcal bacteremia in cirrhotic rats. *Infect Immun* **67**: 2862-2866.

Alcantara, R.B., Preheim, L.C., and Gentry-Nielsen, M.J. (2001) Pneumolysin-induced complement depletion during experimental pneumococcal bacteremia. *Infect Immun* **69**: 3569-3575.

Alexander, J.E., Lock, R.A., Peeters, C.C., Poolman, J.T., Andrew, P.W., Mitchell, T.J., Hansman, D., and Paton, J.C. (1994) Immunization of mice with pneumolysin toxoid confers a significant degree of protection against at least nine serotypes of *Streptococcus pneumoniae*. *Infect Immun* **62**: 5683-5688.

Alon, U., Adler, S.P., and Chan, J.C. (1984) Hemolytic-uremic syndrome associated with *Streptococcus pneumoniae*. Report of a case and review of the literature. *Am J Dis Child* **138**: 496-499.

Alouf, J.E. (2000) Cholesterol-binding cytolytic protein toxins. *Int J Med Microbiol* **290**: 351-356.

Amdahl, B.M., Rubins, J.B., Daley, C.L., Gilks, C.F., Hopewell, P.C., and Janoff, E.N. (1995) Impaired natural immunity to pneumolysin during human immunodeficiency virus infection in the United States and Africa. *Am J Respir Crit Care Med* **152**: 2000-2004.

Arbique, J.C., Poyart, C., Trieu-Cuot, P., Quesne, G., Carvalho Mda, G., Steigerwalt, A.G., Morey, R.E., Jackson, D., Davidson, R.J., and Facklam, R.R. (2004) Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae sp.* nov. *J Clin Microbiol* **42**: 4686-4696.

Arditi, M., Mason, E.O., Jr., Bradley, J.S., Tan, T.Q., Barson, W.J., Schutze, G.E., Wald, E.R., Givner, L.B., Kim, K.S., Yogev, R., and Kaplan, S.L. (1998) Three-year multicenter surveillance of pneumococcal meningitis in children: clinical characteristics, and outcome related to penicillin susceptibility and dexamethasone use. *Pediatrics* **102**: 1087-1097.

Austrian, R., Douglas, R.M., Schiffman, G., Coetzee, A.M., Koornhof, H.J., Hayden-Smith, S., and Reid, R.D. (1976) Prevention of pneumococcal pneumonia by vaccination. *Trans Assoc Am Physicians* **89**: 184-194.

Austrian, R. (1981) Some observations on the pneumococcus and on the current status of pneumococcal disease and its prevention. *Rev Infect Dis* **3 Suppl**: S1-17.

Baba, H., Kawamura, I., Kohda, C., Nomura, T., Ito, Y., Kimoto, T., Watanabe, I., Ichiyama, S., and Mitsuyama, M. (2001) Essential role of domain 4 of pneumolysin from *Streptococcus pneumoniae* in cytolytic activity as determined by truncated proteins. *Biochem Biophys Res Commun* **281**: 37-44.

Bagnoli, F., Moschioni, M., Donati, C., Dimitrovska, V., Ferlenghi, I., Facciotti, C., Muzzi, A., Giusti, F., Emolo, C., Sinisi, A., Hilleringmann, M., Pansegrau, W., Censini, S., Rappuoli, R., Covacci, A., Masignani, V., and Barocchi, M.A. (2008) A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. *J Bacteriol* **190**: 5480-5492.

Balachandran, P., Hollingshead, S.K., Paton, J.C., and Briles, D.E. (2001) The autolytic enzyme LytA of *Streptococcus pneumoniae* is not responsible for releasing pneumolysin. *J Bacteriol* **183**: 3108-3116.

Balsalobre, L., Ferrandiz, M.J., Linares, J., Tubau, F., and de la Campa, A.G. (2003) Viridans group *streptococci* are donors in horizontal transfer of topoisomerase IV genes to *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* **47**: 2072-2081.

Balsalobre, L., Hernandez-Madrid, A., Llull, D., Martin-Galiano, A.J., Garcia, E., Fenoll, A., and de la Campa, A.G. (2006) Molecular characterization of disease-associated *streptococci* of the mitis group that are optochin susceptible. *J Clin Microbiol* **44**: 4163-4171.

Banatvala, N., Griffin, P.M., Greene, K.D., Barrett, T.J., Bibb, W.F., Green, J.H., and Wells, J.G. (2001) The United States National Prospective Hemolytic Uremic Syndrome Study: microbiologic, serologic, clinical, and epidemiologic findings. *J Infect Dis* **183**: 1063-1070.

Barocchi, M.A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Masignani, V., Hultenby, K., Taddei, A.R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D.W., Normark, S., Rappuoli, R., and Henriques-Normark, B. (2006) A pneumococcal pilus influences virulence and host inflammatory responses. *Proc Natl Acad Sci U S A* **103**: 2857-2862.

Basset, A., Trzcinski, K., Hermos, C., O'Brien, K.L., Reid, R., Santosham, M., McAdam, A.J., Lipsitch, M., and Malley, R. (2007) Association of the pneumococcal pilus with certain capsular serotypes but not with increased virulence. *J Clin Microbiol* **45**: 1684-1689.

Batt, S.L., Charalambous, B.M., McHugh, T.D., Martin, S., and Gillespie, S.H. (2005) Novel PCR-restriction fragment length polymorphism method for determining serotypes or serogroups of *Streptococcus pneumoniae* isolates. *J Clin Microbiol* **43**: 2656-2661.

Beghetto, E., Gargano, N., Ricci, S., Garufi, G., Peppoloni, S., Montagnani, F., Oggioni, M., Pozzi, G., and Felici, F. (2006) Discovery of novel Streptococcus pneumoniae antigens by screening a whole-genome lambda-display library. *FEMS Microbiol Lett* **262**: 14-21.

Begue, R., Dennehy, P.H., and Peter, G. (1991) Hemolytic uremic syndrome associated with *Streptococcus pneumoniae*. *N Engl J Med* **325**: 133-134.

Beighton, D., Carr, A.D., and Oppenheim, B.A. (1994) Identification of viridans *streptococci* associated with bacteraemia in neutropenic cancer patients. *J Med Microbiol* **40**: 202-204.

Bentley, S.D., Aanensen, D.M., Mavroidi, A., Saunders, D., Rabbinowitsch, E., Collins, M., Donohoe, K., Harris, D., Murphy, L., Quail, M.A., Samuel, G., Skovsted, I.C., Kaltoft, M.S., Barrell, B., Reeves, P.R., Parkhill, J., and Spratt, B.G. (2006) Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet* **2**: e31.

Benton, K.A., Everson, M.P., and Briles, D.E. (1995) A pneumolysin-negative mutant of *Streptococcus pneumoniae* causes chronic bacteremia rather than acute sepsis in mice. *Infect Immun* **63**: 448-455.

Berk, S., Gage, K., Holtsclaw-Berk, S., and Smith, J. (1985) Type 8 pneumococcal pneumonia: an outbreak on an oncology ward. *South Med J.* **78**: 159-161.

Berrow, N.S., Alderton, D., Sainsbury, S., Nettleship, J., Assenberg, R., Rahman, N., Stuart, D.I., and Owens, R.J. (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res* **35**: e45.

Berry, A.M., Paton, J.C., Glare, E.M., Hansman, D., and Catcheside, D.E. (1988) Cloning and expression of the pneumococcal neuraminidase gene in *Escherichia coli*. *Gene* **71**: 299-305.

Berry, A.M., Lock, R.A., Hansman, D., and Paton, J.C. (1989a) Contribution of autolysin to virulence of *Streptococcus pneumoniae*. *Infect Immun* **57**: 2324-2330.

Berry, A.M., Yother, J., Briles, D.E., Hansman, D., and Paton, J.C. (1989b) Reduced virulence of a defined pneumolysin-negative mutant of *Streptococcus pneumoniae*. *Infect Immun* **57**: 2037-2042.

Berry, A.M., Lock, R.A., Thomas, S.M., Rajan, D.P., Hansman, D., and Paton, J.C. (1994) Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*. *Infect Immun* **62**: 1101-1108.

Berry, A.M., Alexander, J.E., Mitchell, T.J., Andrew, P.W., Hansman, D., and Paton, J.C. (1995) Effect of defined point mutations in the pneumolysin gene on the virulence of *Streptococcus pneumoniae*. *Infect Immun* **63**: 1969-1974.

Berry, A.M., and Paton, J.C. (1996) Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*. *Infect Immun* **64**: 5255-5262.

Berry, A.M., Ogunniyi, A.D., Miller, D.C., and Paton, J.C. (1999) Comparative virulence of *Streptococcus pneumoniae* strains with insertion-duplication, point, and deletion mutations in the pneumolysin gene. *Infect Immun* **67**: 981-985.

Berry, A.M., and Paton, J.C. (2000) Additive Attenuation of Virulence of *Streptococcus pneumoniae* by Mutation of the Genes Encoding Pneumolysin and Other Putative Pneumococcal Virulence Proteins. *Infect. Immun.* **68**: 133-140.

Billington, S.J., Jost, B.H., and Songer, J.G. (2000) Thiol-activated cytolysins: structure, function and role in pathogenesis. *FEMS Microbiol Lett* **182**: 197-205.

Birtles, A., McCarthy, N., Sheppard, C.L., Rutter, H., Guiver, M., Haworth, E., and George, R.C. (2005) Multilocus sequence typing directly on DNA from clinical samples and a cultured isolate to investigate linked fatal

pneumococcal disease in residents of a shelter for homeless men. *J Clin Microbiol* **43**: 2004-2008.

Black, R.E., Morris, S.S., and Bryce, J. (2003) Where and why are 10 million children dying every year? *Lancet* **361**: 2226-2234.

Black, S., Shinefield, H., Fireman, B., Lewis, E., Ray, P., Hansen, J.R., Elvin, L., Ensor, K.M., Hackell, J., Siber, G., Malinoski, F., Madore, D., Chang, I., Kohberger, R., Watson, W., Austrian, R., and Edwards, K. (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**: 187-195.

Black, S., and Shinefield, H. (2002) Safety and efficacy of the seven-valent pneumococcal conjugate vaccine: evidence from Northern California. *Eur J Pediatr* **161 Suppl 2**: S127-131.

Blue, C.E., Paterson, G.K., Kerr, A.R., Berge, M., Claverys, J.P., and Mitchell, T.J. (2003) ZmpB, a novel virulence factor of *Streptococcus pneumoniae* that induces tumor necrosis factor alpha production in the respiratory tract. *Infect Immun* **71**: 4925-4935.

Bogaert, D., De Groot, R., and Hermans, P.W. (2004) *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* **4**: 144-154.

Bohr, V., Rasmussen, N., Hansen, B., Gade, A., Kjersem, H., Johnsen, N., and Paulson, O. (1985) Pneumococcal meningitis: an evaluation of prognostic factors in 164 cases based on mortality and on a study of lasting sequelae. *J Infect* **10**: 143-157.

Bondy, J., Berman, S., Glazner, J., and Lezotte, D. (2000) Direct expenditures related to otitis media diagnoses: extrapolations from a pediatric medicaid cohort. *Pediatrics* **105**: E72.

Boulnois, G.J., Paton, J.C., Mitchell, T.J., and Andrew, P.W. (1991) Structure and function of pneumolysin, the multifunctional, thiol-activated toxin of Streptococcus pneumoniae. *Mol Microbiol* **5**: 2611-2616.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Brandt, C.T., Holm, D., Liptrot, M., Ostergaard, C., Lundgren, J.D., Frimodt-Moller, N., Skovsted, I.C., and Rowland, I.J. (2008) Impact of bacteremia on the pathogenesis of experimental pneumococcal meningitis. *J Infect Dis* **197**: 235-244.

Brandt, J., Wong, C., Mihm, S., Roberts, J., Smith, J., Brewer, E., Thiagarajan, R., and Warady, B. (2002) Invasive pneumococcal disease and hemolytic uremic syndrome. *Pediatrics* **110**: 371-376.

Braun, J.S., Sublett, J.E., Freyer, D., Mitchell, T.J., Cleveland, J.L., Tuomanen, E.I., and Weber, J.R. (2002) Pneumococcal pneumolysin and H(2)O(2) mediate brain cell apoptosis during meningitis. *J Clin Invest* **109**: 19-27.

Briles, D.E., Ades, E., Paton, J.C., Sampson, J.S., Carlone, G.M., Huebner, R.C., Virolainen, A., Swiatlo, E., and Hollingshead, S.K. (2000a) Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of *Streptococcus pneumoniae*. *Infect Immun* **68**: 796-800.

Briles, D.E., Hollingshead, S., Brooks-Walter, A., Nabors, G.S., Ferguson, L., Schilling, M., Gravenstein, S., Braun, P., King, J., and Swift, A. (2000b) The potential to use PspA and other pneumococcal proteins to elicit protection against pneumococcal infection. *Vaccine* **18**: 1707-1711.

Briles, D.E., Hollingshead, S.K., Paton, J.C., Ades, E.W., Novak, L., van Ginkel, F.W., and Benjamin, W.H., Jr. (2003) Immunizations with pneumococcal surface protein A and pneumolysin are protective against pneumonia in a murine model of pulmonary infection with *Streptococcus pneumoniae*. *J Infect Dis* **188**: 339-348.

Briles, D.E. (2004) Protection of the elderly from pneumococcal pneumonia with a protein-based vaccine? *Mech Ageing Dev* **125**: 129-131.

Brooks-Walter, A., Briles, D.E., and Hollingshead, S.K. (1999) The pspC gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia. *Infect Immun* **67**: 6533-6542.

Brown, J.S., Ogunniyi, A.D., Woodrow, M.C., Holden, D.W., and Paton, J.C. (2001) Immunization with components of two iron uptake ABC transporters protects mice against systemic *Streptococcus pneumoniae* infection. *Infect Immun* **69**: 6702-6706.

Bruckner, D.A., and Colonna, P. (1997) Nomenclature for aerobic and facultative bacteria. *Clin Infect Dis* **25**: 1-10.

Brueggemann, A.B., Griffiths, D.T., Meats, E., Peto, T., Crook, D.W., and Spratt, B.G. (2003) Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. *J Infect Dis* **187**: 1424-1432.

Brueggemann, A.B., and Spratt, B.G. (2003) Geographic distribution and clonal diversity of *Streptococcus pneumoniae* serotype 1 isolates. *J Clin Microbiol* **41**: 4966-4970.

Brueggemann, A.B., Pai, R., Crook, D.W., and Beall, B. (2007) Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. *PLoS Pathog* **3**: e168.

Brundage, J.F., and Shanks, G.D. (2008) Deaths from bacterial pneumonia during 1918-19 influenza pandemic. *Emerg Infect Dis* **14**: 1193-1199.

Buck, J.M., Lexau, C., Shapiro, M., Glennen, A., Boxrud, D.J., Koziol, B., Whitney, C.G., Beall, B., Danila, R., and Lynfield, R. (2006) A community outbreak of conjunctivitis caused by nontypeable *Streptococcus pneumoniae* in Minnesota. *Pediatr Infect Dis J* **25**: 906-911.

Butler, J.C., Breiman, R.F., Lipman, H.B., Hofmann, J., and Facklam, R.R. (1995) Serotype distribution of *Streptococcus pneumoniae* infections among preschool children in the United States, 1978-1994: implications for development of a conjugate vaccine. *J Infect Dis* **171**: 885-889.

Byers, H.L., Tarelli, E., Homer, K.A., and Beighton, D. (2000) Isolation and characterisation of sialidase from a strain of *Streptococcus oralis*. *J Med Microbiol* **49**: 235-244.

Byington, C.L., Spencer, L.Y., Johnson, T.A., Pavia, A.T., Allen, D., Mason, E.O., Kaplan, S., Carroll, K.C., Daly, J.A., Christenson, J.C., and Samore, M.H. (2002) An epidemiological investigation of a sustained high rate of pediatric parapneumonic empyema: risk factors and microbiological associations. *Clin Infect Dis* **34**: 434-440.

Byington, C.L., Korgenski, K., Daly, J., Ampofo, K., Pavia, A., and Mason, E.O. (2006) Impact of the pneumococcal conjugate vaccine on pneumococcal parapneumonic empyema. *Pediatr Infect Dis J* **25**: 250-254.

Cabrera, G.R., Fortenberry, J.D., Warshaw, B.L., Chambliss, C.R., Butler, J.C., and Cooperstone, B.G. (1998) Hemolytic uremic syndrome associated with invasive *Streptococcus pneumoniae* infection. *Pediatrics* **101**: 699-703.

Camara, M., Mitchell, T.J., Andrew, P.W., and Boulnois, G.J. (1991) Streptococcus pneumoniae produces at least two distinct enzymes with

neuraminidase activity: cloning and expression of a second neuraminidase gene in *Escherichia coli*. *Infect Immun* **59**: 2856-2858.
Camara, M., Boulnois, G.J., Andrew, P.W., and Mitchell, T.J. (1994) A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. *Infect Immun* **62**: 3688-3695.
Canvin, J.R., Marvin, A.P., Sivakumaran, M., Paton, J.C., Boulnois, G.J., Andrew, P.W., and Mitchell, T.J. (1995) The role of pneumolysin and autolysin in the pathology of pneumonia and septicemia in mice infected with a type 2 pneumococcus. *J Infect Dis* **172**: 119-123.
Cao, J., Chen, D., Xu, W., Chen, T., Xu, S., Luo, J., Zhao, Q., Liu, B., Wang, D., Zhang, X., Shan, Y., and Yin, Y. (2007) Enhanced protection against pneumococcal infection elicited by immunization with the combination of PspA, PspC, and ClpP. *Vaccine* **25**: 4996-5005.
Cardozo, D.M., Nascimento-Carvalho, C.M., Andrade, A.L., Silvany-Neto, A.M., Daltro, C.H., Brandao, M.A., Brandao, A.P., and Brandileone, M.C. (2008) Prevalence and risk factors for nasopharyngeal carriage of *Streptococcus pneumoniae* among adolescents. *J Med Microbiol* **57**: 185-189.
Carrol, E.D., Mankhambo, L.A., Corless, C., and Guiver, M. (2008) Bacteremia Is Associated with a Worse Outcome in Pneumococcal Meningitis. *J Infect Dis* **198**: 626-627.
Carruthers, M.M., and Kanokvechayant, R. (1973) Effect of pneumococcal neuraminidase on canine cerebral cortex in experimental meningitis. *Infect Immun* **7**: 370-372.
Chen, S., Paterson, G.K., Tong, H.H., Mitchell, T.J., and DeMaria, T.F. (2005) Sortase A contributes to pneumococcal nasopharyngeal colonization in the chinchilla model. *FEMS Microbiol Lett* **253**: 151-154.
Chi, F., Nolte, O., Bergmann, C., Ip, M., and Hakenbeck, R. (2007) Crossing the barrier: evolution and spread of a major class of mosaic pbp2x in *Streptococcus pneumoniae, S. mitis* and *S. oralis*. *Int J Med Microbiol* **297**: 503-512.
Chong, A.K., Pegg, M.S., Taylor, N.R., and von Itzstein, M. (1992) Evidence for a sialosyl cation transition-state complex in the reaction of sialidase from influenza virus. *Eur J Biochem* **207**: 335-343.
Cima-Cabal, M.D., Mendez, F.J., Vazquez, F., del Mar Garcia-Suarez, M., and de los Toyos, J.R. (2001) A specific and ultrasensitive chemiluminescent sandwich ELISA test for the detection and quantitation of pneumolysin. *J Immunoassay Immunochem* **22**: 99-112.
Cima-Cabal, M.D., Mendez, F.J., Vazquez, F., Aranaz, C., Rodriguez-Alvarez, J., Garcia-Garcia, J.M., Fleites, A., Martinez Gonzalez-Rio, J., Molinos, L., de Miguel, D., and de los Toyos, J.R. (2003) Immunodetection of pneumolysin in human urine by ELISA. *J Microbiol Methods* **54**: 47-55.
Coffey, T., Daniels, M., Enright, C., and Spratt, B. (1999) Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of *Streptococcus pneumoniae* arose by large recombinational replacements of the cpsA--pbp1a region. *Microbiology* **145**: 2023-2031.
Coffey, T.J., Dowson, C.G., Daniels, M., and Spratt, B.G. (1993) Horizontal spread of an altered penicillin-binding protein 2B gene between *Streptococcus pneumoniae and Streptococcus oralis*. *FEMS Microbiol Lett* **110**: 335-339.
Coffey, T.J., Enright, M.C., Daniels, M., Morona, J.K., Morona, R., Hryniewicz, W., Paton, J.C., and Spratt, B.G. (1998a) Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype

changes among natural isolates of *Streptococcus pneumoniae*. *Mol Microbiol* **27**: 73-83.

Coffey, T.J., Enright, M.C., Daniels, M., Wilkinson, P., Berron, S., Fenoll, A., and Spratt, B.G. (1998b) Serotype 19A variants of the Spanish serotype 23F multiresistant clone of *Streptococcus pneumoniae*. *Microb Drug Resist* **4**: 51-55.

Colino, J., and Snapper, C.M. (2003) Two distinct mechanisms for induction of dendritic cell apoptosis in response to intact *Streptococcus pneumoniae*. *J Immunol* **171**: 2354-2365.

Comis, S.D., Osborne, M.P., Stephen, J., Tarlow, M.J., Hayward, T.L., Mitchell, T.J., Andrew, P.W., and Boulnois, G.J. (1993) Cytotoxic effects on hair cells of guinea pig cochlea produced by pneumolysin, the thiol activated toxin of *Streptococcus pneumoniae*. *Acta Otolaryngol* **113**: 152-159.

Constantinescu, A.R., Bitzan, M., Weiss, L.S., Christen, E., Kaplan, B.S., Cnaan, A., and Trachtman, H. (2004) Non-enteropathic hemolytic uremic syndrome: causes and short-term course. *Am J Kidney Dis* **43**: 976-982.

Copelovitch, L., and Kaplan, B.S. (2007) *Streptococcus pneumoniae*-associated hemolytic uremic syndrome. *Pediatr Nephrol*.

Czajkowsky, D.M., Hotze, E.M., Shao, Z., and Tweten, R.K. (2004) Vertical collapse of a cytolysin prepore moves its transmembrane beta-hairpins to the membrane. *Embo J* **23**: 3206-3215.

Dagan, R., Gradstein, S., Belmaker, I., Porat, N., Siton, Y., Weber, G., Janco, J., and Yagupsky, P. (2000) An outbreak of *Streptococcus pneumoniae* serotype 1 in a closed community in southern Israel. *Clin Infect Dis* **30**: 319-321.

Dang, T.X., Hotze, E.M., Rouiller, I., Tweten, R.K., and Wilson-Kubalek, E.M. (2005) Prepore to pore transition of a cholesterol-dependent cytolysin visualized by electron microscopy. *J Struct Biol* **150**: 100-108.

Dave, S., Carmicle, S., Hammerschmidt, S., Pangburn, M.K., and McDaniel, L.S. (2004) Dual roles of PspC, a surface protein of *Streptococcus pneumoniae*, in binding human secretory IgA and factor H. *J Immunol* **173**: 471-477.

de los Toyos, J.R., Mendez, F.J., Aparicio, J.F., Vazquez, F., Del Mar Garcia Suarez, M., Fleites, A., Hardisson, C., Morgan, P.J., Andrew, P.W., and Mitchell, T.J. (1996) Functional analysis of pneumolysin by use of monoclonal antibodies. *Infect Immun* **64**: 480-484.

DeMaria, A., Jr., Browne, K., Berk, S.L., Sherwood, E.J., and McCabe, W.R. (1980) An outbreak of type 1 pneumococcal pneumonia in a men's shelter. *Jama* **244**: 1446-1449.

Diaz, E., Lopez, R., and Garcia, J.L. (1992) Role of the major pneumococcal autolysin in the atypical response of a clinical isolate of *Streptococcus pneumoniae*. *J Bacteriol* **174**: 5508-5515.

Dintilhac, A., Alloing, G., Granadel, C., and Claverys, J.P. (1997) Competence and virulence of *Streptococcus pneumoniae*: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. *Mol Microbiol* **25**: 727-739.

Dochez, A.R., and Gillespie, L.J. (1913) A biologic classification of pneumococci by means of immunity reactions. *A.M.A.*: 727-732.

Dopazo, J., Mendoza, A., Herrero, J., Caldara, F., Humbert, Y., Friedli, L., Guerrier, M., Grand-Schenk, E., Gandin, C., de Francesco, M., Polissi, A., Buell, G., Feger, G., Garcia, E., Peitsch, M., and Garcia-Bustos, J.F. (2001) Annotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate. *Microb Drug Resist* **7**: 99-125.

Douglas, C.W., Heath, J., Hampton, K.K., and Preston, F.E. (1993) Identity of viridans *streptococci* isolated from cases of infective endocarditis. *J Med Microbiol* **39**: 179-182.

Dowson, C.G., Hutchison, A., Brannigan, J.A., George, R.C., Hansman, D., Linares, J., Tomasz, A., Smith, J.M., and Spratt, B.G. (1989) Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A* **86**: 8842-8846.

Dowson, C.G., Coffey, T.J., Kell, C., and Whiley, R.A. (1993) Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of Streptococcus mitis in the formation of a low affinity PBP2B in S. pneumoniae. *Mol Microbiol* **9**: 635-643.

Dowson, C.G., Coffey, T.J., and Spratt, B.G. (1994) Origin and molecular epidemiology of penicillin-binding-protein-mediated resistance to beta-lactam antibiotics. *Trends Microbiol* **2**: 361-366.

Eastham, K.M., Freeman, R., Kearns, A.M., Eltringham, G., Clark, J., Leeming, J., and Spencer, D.A. (2004) Clinical features, aetiology and outcome of empyema in children in the north east of England. *Thorax* **59**: 522-525.

Echeverria, M.J., Lopez de Goicoechea, M.J., Ayarza, R., Berdonces, P., Audicana, J., Fernandez, R., and Soriano, F. (1998) [Meningitis caused by *Streptococcus mitis* in a cocaine addict]. *Enferm Infecc Microbiol Clin* **16**: 441-442.

Engel, F., Blatz, R., Kellner, J., Palmer, M., Weller, U., and Bhadki, S. (1995) Breakdown of the round window membrane permeability barrier evoked by streptolysin O: possible etiologic role in development of sensorineural hearing loss in acute otitis media. *Infect Immun* **63**: 1305-1310.

Enright, M., and Spratt, B. (1998) A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology* **144**: 3049-3060.

Facklam, R. (2002) What happened to the *streptococci*: overview of taxonomic and nomenclature changes. *Clin Microbiol Rev* **15**: 613-630.

Feld, L.G., Springate, J.E., Darragh, R., and Fildes, R.D. (1987) Pneumococcal pneumonia and hemolytic uremic syndrome. *Pediatr Infect Dis J* **6**: 693-695.

Feldman, C., Mitchell, T.J., Andrew, P.W., Boulnois, G.J., Read, R.C., Todd, H.C., Cole, P.J., and Wilson, R. (1990) The effect of *Streptococcus pneumoniae* pneumolysin on human respiratory epithelium in vitro. *Microb Pathog* **9**: 275-284.

Feldman, C., Munro, N.C., Jeffery, P.K., Mitchell, T.J., Andrew, P.W., Boulnois, G.J., Guerreiro, D., Rohde, J.A., Todd, H.C., and Cole, P.J. (1991) Pneumolysin induces the salient histologic features of pneumococcal infection in the rat lung in vivo. *Am J Respir Cell Mol Biol* **5**: 416-423.

Fenoll, A., Martinez-Suarez, J.V., Munoz, R., Casal, J., and Garcia, J.L. (1990) Identification of atypical strains of *Streptococcus pneumoniae* by a specific DNA probe. *Eur J Clin Microbiol Infect Dis* **9**: 396-401.

Ferrante, A., Rowan-Kelly, B., and Paton, J.C. (1984) Inhibition of in vitro human lymphocyte response by the pneumococcal toxin pneumolysin. *Infect Immun* **46**: 585-589.

Finland, M., and Sutliff, W.D. (1932) Specific antibody response of human subjectsto intracutaneous injection of pneumococcus products. *J Exp Med*: 853-865.

Fitzpatrick, M.M., Walters, M.D., Trompeter, R.S., Dillon, M.J., and Barratt, T.M. (1993) Atypical (non-diarrhea-associated) hemolytic-uremic syndrome in childhood. *J Pediatr* **122**: 532-537.

Flannery, B., Heffernan, R.T., Harrison, L.H., Ray, S.M., Reingold, A.L., Hadler, J., Schaffner, W., Lynfield, R., Thomas, A.R., Li, J., Campsmith, M., Whitney, C.G., and Schuchat, A. (2006) Changes in invasive Pneumococcal disease among HIV-infected adults living in the era of childhood pneumococcal immunization. *Ann Intern Med* **144**: 1-9.

Francis, T., and Tillett, W.S. (1930) Cutaneous reactions in pneumonia. The development of antibodies following the intradermal injection of type-specific polysaccharide. *J Exp Med*: 573-589.

Frazao, N., Brito-Avo, A., Simas, C., Saldanha, J., Mato, R., Nunes, S., Sousa, N.G., Carrico, J.A., Almeida, J.S., Santos-Sanches, I., and de Lencastre, H. (2005) Effect of the seven-valent conjugate pneumococcal vaccine on carriage and drug resistance of *Streptococcus pneumoniae* in healthy children attending day-care centers in Lisbon. *Pediatr Infect Dis J* **24**: 243-252.

French, N., Nakiyingi, J., Carpenter, L.M., Lugada, E., Watera, C., Moi, K., Moore, M., Antvelink, D., Mulder, D., Janoff, E.N., Whitworth, J., and Gilks, C.F. (2000) 23-valent pneumococcal polysaccharide vaccine in HIV-1-infected Ugandan adults: double-blind, randomised and placebo controlled trial. *Lancet* **355**: 2106-2111.

French, N. (2003) Use of pneumococcal polysaccharide vaccines: no simple answers. *J Infect* **46**: 78-86.

Friedland, I.R., Paris, M.M., Hickey, S., Shelton, S., Olsen, K., Paton, J.C., and McCracken, G.H. (1995) The limited role of pneumolysin in the pathogenesis of pneumococcal meningitis. *J Infect Dis* **172**: 805-809.

Garcia-Suarez Mdel, M., Cima-Cabal, M.D., Florez, N., Garcia, P., Cernuda-Cernuda, R., Astudillo, A., Vazquez, F., De los Toyos, J.R., and Mendez, F.J. (2004) Protection against pneumococcal pneumonia in mice by monoclonal antibodies to pneumolysin. *Infect Immun* **72**: 4534-4540.

Garcia-Suarez Mdel, M., Florez, N., Astudillo, A., Vazquez, F., Villaverde, R., Fabrizio, K., Pirofski, L.A., and Mendez, F.J. (2007) The role of pneumolysin in mediating lung damage in a lethal pneumococcal pneumonia murine model. *Respir Res* **8**: 3.

Garnier, F., Janapatla, R.P., Charpentier, E., Masson, G., Grelaud, C., Stach, J.F., Denis, F., and Ploy, M.C. (2007) Insertion sequence 1515 in the ply gene of a type 1 clinical isolate of *Streptococcus pneumoniae* abolishes pneumolysin expression. *J Clin Microbiol* **45**: 2296-2297.

Gebo, K.A., Moore, R.D., Keruly, J.C., and Chaisson, R.E. (1996) Risk factors for pneumococcal disease in human immunodeficiency virus-infected patients. *J Infect Dis* **173**: 857-862.

Gibney, K.B., Morris, P.S., Carapetis, J.R., Skull, S.A., Smith-Vaughan, H.C., Stubbs, E., and Leach, A.J. (2005) The clinical course of acute otitis media in high-risk Australian Aboriginal children: a longitudinal study. *BMC Pediatr* **5**: 16.

Giddings, K.S., Zhao, J., Sims, P.J., and Tweten, R.K. (2004) Human CD59 is a receptor for the cholesterol-dependent cytolysin intermedilysin. *Nat Struct Mol Biol* **11**: 1173-1178.

Giebink, G.S., Berzins, I.K., Marker, S.C., and Schiffman, G. (1980) Experimental otitis media after nasal inoculation of *Streptococcus pneumoniae* and influenza A virus in chinchillas. *Infect Immun* **30**: 445-450.

Giebink, G.S. (1999) Otitis media: the chinchilla model. *Microb Drug Resist* **5**: 57-72.
Gilbert, R.J., Heenan, R.K., Timmins, P.A., Gingles, N.A., Mitchell, T.J., Rowe, A.J., Rossjohn, J., Parker, M.W., Andrew, P.W., and Byron, O. (1999a) Studies on the structure and mechanism of a bacterial protein toxin by analytical ultracentrifugation and small-angle neutron scattering. *J Mol Biol* **293**: 1145-1160.
Gilbert, R.J., Jimenez, J.L., Chen, S., Tickle, I.J., Rossjohn, J., Parker, M., Andrew, P.W., and Saibil, H.R. (1999b) Two structural transitions in membrane pore formation by pneumolysin, the pore-forming toxin of *Streptococcus pneumoniae*. *Cell* **97**: 647-655.
Glezen, W.P. (1982) Serious morbidity and mortality associated with influenza epidemics. *Epidemiol Rev* **4**: 25-44.
Gosink, K.K., Mann, E.R., Guglielmo, C., Tuomanen, E.I., and Masure, H.R. (2000) Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infect Immun* **68**: 5690-5695.
Gratten, M., Morey, F., Dixon, J., Manning, K., Torzillo, P., Matters, R., Erlich, J., Hanna, J., Asche, V., and Riley, I. (1993) An outbreak of serotype 1 *Streptococcus pneumoniae* infection in central Australia. *Med J Aust* **159**: 215.
Green, B.A., Zhang, Y., Masi, A.W., Barniak, V., Wetherell, M., Smith, R.P., Reddy, M.S., and Zhu, D. (2005) PppA, a surface-exposed protein of *Streptococcus pneumoniae*, elicits cross-reactive antibodies that reduce colonization in a murine intranasal immunization and challenge model. *Infect Immun* **73**: 981-989.
Grewal, P.K., Uchiyama, S., Ditto, D., Varki, N., Le, D.T., Nizet, V., and Marth, J.D. (2008) The Ashwell receptor mitigates the lethal coagulopathy of sepsis. *Nat Med* **14**: 648-655.
Hakenbeck, R., Balmelle, N., Weber, B., Gardes, C., Keck, W., and de Saizieu, A. (2001) Mosaic Genes and Mosaic Chromosomes: Intra- and Interspecies Genomic Variation of *Streptococcus pneumoniae*. *Infect. Immun.* **69**: 2477-2486.
Hamel, J., Charland, N., Pineau, I., Ouellet, C., Rioux, S., Martin, D., and Brodeur, B.R. (2004) Prevention of pneumococcal disease in mice immunized with conserved surface-accessible proteins. *Infect Immun* **72**: 2659-2670.
Hammerschmidt, S., Wolff, S., Hocke, A., Rosseau, S., Muller, E., and Rohde, M. (2005) Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infect Immun* **73**: 4653-4667.
Han, X.Y., Kamana, M., and Rolston, K.V. (2006) Viridans *streptococci* isolated by culture from blood of cancer patients: clinical and microbiologic analysis of 50 cases. *J Clin Microbiol* **44**: 160-165.
Hanage, W.P., Auranen, K., Syrjanen, R., Herva, E., Makela, P.H., Kilpi, T., and Spratt, B.G. (2004) Ability of pneumococcal serotypes and clones to cause acute otitis media: implications for the prevention of otitis media by conjugate vaccines. *Infect Immun* **72**: 76-81.
Hanage, W.P., Kaijalainen, T., Herva, E., Saukkoriipi, A., Syrjanen, R., and Spratt, B.G. (2005) Using multilocus sequence data to define the pneumococcus. *J Bacteriol* **187**: 6223-6230.
Handley, P., Coykendall, A., Beighton, D., Hardie, J.M., and Whiley, R.A. (1991) *Streptococcus crista* sp. nov., a viridans streptococcus with tufted fibrils, isolated from the human oral cavity and throat. *Int J Syst Bacteriol* **41**: 543-547.

Hausdorff, W.P., Bryant, J., Kloek, C., Paradiso, P.R., and Siber, G.R. (2000a) The contribution of specific pneumococcal serogroups to different disease manifestations: implications for conjugate vaccine formulation and use, part II. *Clin Infect Dis* **30**: 122-140.

Hausdorff, W.P., Bryant, J., Paradiso, P.R., and Siber, G.R. (2000b) Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. *Clin Infect Dis* **30**: 100-121.

Hava, D.L., and Camilli, A. (2002) Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol* **45**: 1389-1406.

Havarstein, L.S., Gaustad, P., Nes, I.F., and Morrison, D.A. (1996) Identification of the streptococcal competence-pheromone receptor. *Mol Microbiol* **21**: 863-869.

Havarstein, L.S., Hakenbeck, R., and Gaustad, P. (1997) Natural competence in the genus Streptococcus: evidence that *streptococci* can change pherotype by interspecies recombinational exchanges. *J Bacteriol* **179**: 6589-6594.

Hayden, F.G., Atmar, R.L., Schilling, M., Johnson, C., Poretz, D., Paar, D., Huson, L., Ward, P., and Mills, R.G. (1999a) Use of the selective oral neuraminidase inhibitor oseltamivir to prevent influenza. *N Engl J Med* **341**: 1336-1343.

Hayden, F.G., Treanor, J.J., Fritz, R.S., Lobo, M., Betts, R.F., Miller, M., Kinnersley, N., Mills, R.G., Ward, P., and Straus, S.E. (1999b) Use of the oral neuraminidase inhibitor oseltamivir in experimental human influenza: randomized controlled trials for prevention and treatment. *Jama* **282**: 1240-1246.

Hendriksen, W.T., Bootsma, H.J., Estevao, S., Hoogenboezem, T., de Jong, A., de Groot, R., Kuipers, O.P., and Hermans, P.W. (2008) CodY of *Streptococcus pneumoniae*: link between nutritional gene regulation and colonization. *J Bacteriol* **190**: 590-601.

Henrichsen, J. (1995) Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* **33**: 2759-2762.

Henriques Normark, B., Kalin, M., Ortqvist, A., Akerlund, T., Liljequist, B.O., Hedlund, J., Svenson, S.B., Zhou, J., Spratt, B.G., Normark, S., and Kallenius, G. (2001) Dynamics of penicillin-susceptible clones in invasive pneumococcal disease. *J Infect Dis* **184**: 861-869.

Heuck, A.P., Tweten, R.K., and Johnson, A.E. (2003) Assembly and topography of the prepore complex in cholesterol-dependent cytolysins. *J Biol Chem* **278**: 31218-31225.

Hill, J., Andrew, P.W., and Mitchell, T.J. (1994) Amino acids in pneumolysin important for hemolytic activity identified by random mutagenesis. *Infect Immun* **62**: 757-758.

Hiller, N.L., Janto, B., Hogg, J.S., Boissy, R., Yu, S., Powell, E., Keefe, R., Ehrlich, N.E., Shen, K., Hayes, J., Barbadora, K., Klimke, W., Dernovoy, D., Tatusova, T., Parkhill, J., Bentley, S.D., Post, J.C., Ehrlich, G.D., and Hu, F.Z. (2007) Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome. *J Bacteriol* **189**: 8186-8195.

Hirst, R.A., Sikand, K.S., Rutman, A., Mitchell, T.J., Andrew, P.W., and O'Callaghan, C. (2000) Relative roles of pneumolysin and hydrogen peroxide from *Streptococcus pneumoniae* in inhibition of ependymal ciliary beat frequency. *Infect Immun* **68**: 1557-1562.

Hirst, R.A., Kadioglu, A., O'Callaghan, C., and Andrew, P.W. (2004a) The role of pneumolysin in pneumococcal pneumonia and meningitis. *Clin Exp Immunol* **138**: 195-201.

Hirst, R.A., Mohammed, B.J., Mitchell, T.J., Andrew, P.W., and O'Callaghan, C. (2004b) *Streptococcus pneumoniae*-Induced Inhibition of Rat Ependymal Cilia Is Attenuated by Antipneumolysin Antibody. *Infect Immun* **72**: 6694-6698.

Hirst, R.A., Gosai, B., Rutman, A., Guerin, C.J., Nicotera, P., Andrew, P.W., and O'Callaghan, C. (2008) *Streptococcus pneumoniae* deficient in pneumolysin or autolysin has reduced virulence in meningitis. *J Infect Dis* **197**: 744-751.

Hoffmann, O.M., Becker, D., and Weber, J.R. (2007) Bacterial hydrogen peroxide contributes to cerebral hyperemia during early stages of experimental pneumococcal meningitis. *J Cereb Blood Flow Metab* **27**: 1792-1797.

Hollingshead, S.K., Becker, R., and Briles, D.E. (2000) Diversity of PspA: mosaic genes and evidence for past recombination in *Streptococcus pneumoniae*. *Infect Immun* **68**: 5889-5900.

Hoshino, T., Fujiwara, T., and Kilian, M. (2005) Use of phylogenetic and phenotypic analyses to identify nonhemolytic *streptococci* isolated from bacteremic patients. *J Clin Microbiol* **43**: 6073-6085.

Hoskins, J., Alborn, W.E., Jr., Arnold, J., Blaszczak, L.C., Burgett, S., DeHoff, B.S., Estrem, S.T., Fritz, L., Fu, D.J., Fuller, W., Geringer, C., Gilmour, R., Glass, J.S., Khoja, H., Kraft, A.R., Lagace, R.E., LeBlanc, D.J., Lee, L.N., Lefkowitz, E.J., Lu, J., Matsushima, P., McAhren, S.M., McHenney, M., McLeaster, K., Mundy, C.W., Nicas, T.I., Norris, F.H., O'Gara, M., Peery, R.B., Robertson, G.T., Rockey, P., Sun, P.M., Winkler, M.E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C.A., Baltz, R.H., Jaskunas, S.R., Rosteck, P.R., Jr., Skatrud, P.L., and Glass, J.I. (2001) Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J Bacteriol* **183**: 5709-5717.

Hotze, E.M., Wilson-Kubalek, E.M., Rossjohn, J., Parker, M.W., Johnson, A.E., and Tweten, R.K. (2001) Arresting pore formation of a cholesterol-dependent cytolysin by disulfide trapping synchronizes the insertion of the transmembrane beta-sheet from a prepore intermediate. *J Biol Chem* **276**: 8261-8268.

Hotze, E.M., Heuck, A.P., Czajkowsky, D.M., Shao, Z., Johnson, A.E., and Tweten, R.K. (2002) Monomer-Monomer Interactions Drive the Prepore to Pore Conversion of a beta -Barrel-forming Cholesterol-dependent Cytolysin. *J. Biol. Chem.* **277**: 11597-11605.

Houldsworth, S., Andrew, P.W., and Mitchell, T.J. (1994) Pneumolysin stimulates production of tumor necrosis factor alpha and interleukin-1 beta by human mononuclear phagocytes. *Infect Immun* **62**: 1501-1503.

Howard, L.V., and Gooder, H. (1974) Specificity of the autolysin of *Streptococcus (Diplococcus) pneumoniae*. *J Bacteriol* **117**: 796-804.

Hsu, R.B., and Lin, F.Y. (2006) Effect of penicillin resistance on presentation and outcome of nonenterococcal streptococcal infective endocarditis. *Cardiology* **105**: 234-239.

Huang, F.Y., and Lin, D.S. (1998) Pneumococcal meningitis complicated with hemolytic uremic syndrome: report of two cases. *Zhonghua Min Guo Xiao Er Ke Yi Xue Hui Za Zhi* **39**: 58-61.

Huang, Y.H., Lin, T.Y., Wong, K.S., Huang, Y.C., Chiu, C.H., Lai, S.H., and Hsia, S.H. (2006) Hemolytic uremic syndrome associated with pneumococcal pneumonia in Taiwan. *Eur J Pediatr* **165**: 332-335.

Iannelli, F., Oggioni, M.R., and Pozzi, G. (2002) Allelic variation in the highly polymorphic locus pspC of *Streptococcus pneumoniae*. *Gene* **284**: 63-71.
Iannelli, F., Chiavolini, D., Ricci, S., Oggioni, M.R., and Pozzi, G. (2004) Pneumococcal surface protein C contributes to sepsis caused by *Streptococcus pneumoniae* in mice. *Infect Immun* **72**: 3077-3080.
Ibrahim, Y.M., Kerr, A.R., Silva, N.A., and Mitchell, T.J. (2005) Contribution of the ATP-dependent protease ClpCP to the autolysis and virulence of *Streptococcus pneumoniae*. *Infect Immun* **73**: 730-740.
Ip, M., Chau, S.S., Chi, F., Tang, J., and Chan, P.K. (2007) Fluoroquinolone resistance in atypical pneumococci and oral *streptococci*: evidence of horizontal gene transfer of fluoroquinolone resistance determinants from *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* **51**: 2690-2700.
Jacobs, T., Cima-Cabal, M.D., Darji, A., Mendez, F.J., Vazquez, F., Jacobs, A.A., Shimada, Y., Ohno-Iwashita, Y., Weiss, S., and de los Toyos, J.R. (1999) The conserved undecapeptide shared by thiol-activated cytolysins is involved in membrane binding. *FEBS Lett* **459**: 463-466.
Janoir, C., Podglajen, I., Kitzis, M.D., Poyart, C., and Gutmann, L. (1999) In vitro exchange of fluoroquinolone resistance determinants between *Streptococcus pneumoniae* and viridans *streptococci* and genomic organization of the parE-parC region in *S. mitis*. *J Infect Dis* **180**: 555-558.
Jedrzejas, M.J. (2001) Pneumococcal virulence factors: structure and function. *Microbiol Mol Biol Rev* **65**: 187-207; first page, table of contents.
Jefferies, J., Clarke, S.C., Diggle, M.A., Smith, A., Dowson, C., and Mitchell, T. (2003) Automated pneumococcal MLST using liquid-handling robotics and a capillary DNA sequencer. *Mol Biotechnol* **24**: 303-308.
Johnson, M.K., and Allen, J.H. (1975) The role of cytolysin in pneumococcal ocular infection. *Am J Ophthalmol* **80**: 518-521.
Johnson, M.K., Boese-Marrazzo, D., and Pierce, W.A., Jr. (1981) Effects of pneumolysin on human polymorphonuclear leukocytes and platelets. *Infect Immun* **34**: 171-176.
Johnson, S.E., Dykes, J.K., Jue, D.L., Sampson, J.S., Carlone, G.M., and Ades, E.W. (2002) Inhibition of pneumococcal carriage in mice by subcutaneous immunization with peptides from the common surface protein pneumococcal surface adhesin a. *J Infect Dis* **185**: 489-496.
Johnston, R.B., Jr. (1991) Pathogenesis of pneumococcal pneumonia. *Rev Infect Dis* **13 Suppl 6**: S509-517.
Jomaa, M., Kyd, J.M., and Cripps, A.W. (2005a) Mucosal immunisation with novel *Streptococcus pneumoniae* protein antigens enhances bacterial clearance in an acute mouse lung infection model. *FEMS Immunol Med Microbiol* **44**: 59-67.
Jomaa, M., Yuste, J., Paton, J.C., Jones, C., Dougan, G., and Brown, J.S. (2005b) Antibodies to the iron uptake ABC transporter lipoproteins PiaA and PiuA promote opsonophagocytosis of *Streptococcus pneumoniae*. *Infect Immun* **73**: 6852-6859.
Jones, N., Huebner, R., Khoosal, M., Crewe-Brown, H., and Klugman, K. (1998) The impact of HIV on *Streptococcus pneumoniae* bacteraemia in a South African population. *Aids* **12**: 2177-2184.
Jonsson, S., Musher, D.M., Chapman, A., Goree, A., and Lawrence, E.C. (1985) Phagocytosis and killing of common bacterial pathogens of the lung by human alveolar macrophages. *J Infect Dis* **152**: 4-13.
Jounblat, R., Kadioglu, A., Mitchell, T.J., and Andrew, P.W. (2003) Pneumococcal Behavior and Host Responses during Bronchopneumonia Are

Affected Differently by the Cytolytic and Complement-Activating Activities of Pneumolysin. *Infect. Immun.* **71**: 1813-1819.
Kadioglu, A., Taylor, S., Iannelli, F., Pozzi, G., Mitchell, T.J., and Andrew, P.W. (2002) Upper and lower respiratory tract infection by *Streptococcus pneumoniae* is affected by pneumolysin deficiency and differences in capsule type. *Infect Immun* **70**: 2886-2890.
Kalin, M., Kanclerski, K., Granstrom, M., and Mollby, R. (1987) Diagnosis of pneumococcal pneumonia by enzyme-linked immunosorbent assay of antibodies to pneumococcal hemolysin (pneumolysin). *J Clin Microbiol* **25**: 226-229.
Kaplan, B.S., Cleary, T.G., and Obrig, T.G. (1990) Recent advances in understanding the pathogenesis of the hemolytic uremic syndromes. *Pediatr Nephrol* **4**: 276-283.
Kaplan, S.L., Mason, E.O., Jr., Barson, W.J., Wald, E.R., Arditi, M., Tan, T.Q., Schutze, G.E., Bradley, J.S., Givner, L.B., Kim, K.S., and Yogev, R. (1998) Three-year multicenter surveillance of systemic pneumococcal infections in children. *Pediatrics* **102**: 538-545.
Kastenbauer, S., and Pfister, H.W. (2003) Pneumococcal meningitis in adults: spectrum of complications and prognostic factors in a series of 87 cases. *Brain* **126**: 1015-1025.
Kawamura, Y., Hou, X.G., Sultana, F., Miura, H., and Ezaki, T. (1995) Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus Streptococcus. *Int J Syst Bacteriol* **45**: 406-408.
Kawamura, Y., Hou, X.G., Todome, Y., Sultana, F., Hirose, K., Shu, S.E., Ezaki, T., and Ohkuni, H. (1998) *Streptococcus peroris sp. nov.* and *Streptococcus infantis sp. nov.*, new members of the *Streptococcus mitis* group, isolated from human clinical specimens. *Int J Syst Bacteriol* **48 Pt 3**: 921-927.
Keane, W.M., Potsic, W.P., Rowe, L.D., and Konkle, D.F. (1979) Meningitis and hearing loss in children. *Arch Otolaryngol* **105**: 39-44.
Kelly, D.F., Moxon, E.R., and Pollard, A.J. (2004) Haemophilus influenzae type b conjugate vaccines. *Immunology* **113**: 163-174.
Kerr, A.R., Paterson, G.K., McCluskey, J., Iannelli, F., Oggioni, M.R., Pozzi, G., and Mitchell, T.J. (2006) The contribution of PspC to pneumococcal virulence varies between strains and is accomplished by both complement evasion and complement-independent mechanisms. *Infect Immun* **74**: 5319-5324.
Kharat, A.S., and Tomasz, A. (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**: 2758-2765.
Kilian, M., Poulsen, K., Blomqvist, T., Havarstein, L.S., Bek-Thomsen, M., Tettelin, H., and Sorensen, U.B. (2008) Evolution of *Streptococcus pneumoniae* and its close commensal relatives. *PLoS ONE* **3**: e2683.
Kilpper-Balz, R., Wenzig, P., and Schleifer, K.H. (1985) Molecular relationships and classification of some viridans *Streptococci* as *Streptococcus oralis* and emended description of *S. oralis*. *Int J Syst Bacteriol*: 482-488.
Kim, H.J., Kim, S.I., Ratnayake-Lecamwasam, M., Tachikawa, K., Sonenshein, A.L., and Strauch, M. (2003) Complex regulation of the *Bacillus subtilis* aconitase gene. *J Bacteriol* **185**: 1672-1680.
Kim, Y.D., Prakash, U., Weber, G.F., and Hargie, M. (1979) Nature of human serum blood group T antibodies. *Immunol Commun* **8**: 397-406.

King, S.J., Hippe, K.R., Gould, J.M., Bae, D., Peterson, S., Cline, R.T., Fasching, C., Janoff, E.N., and Weiser, J.N. (2004) Phase variable desialylation of host proteins that bind to *Streptococcus pneumoniae* in vivo and protect the airway. *Mol Microbiol* **54**: 159-171.

King, S.J., Whatmore, A.M., and Dowson, C.G. (2005) NanA, a neuraminidase from *Streptococcus pneumoniae*, shows high levels of sequence diversity, at least in part through recombination with *Streptococcus oralis*. *J Bacteriol* **187**: 5376-5386.

King, S.J., Hippe, K.R., and Weiser, J.N. (2006) Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Mol Microbiol* **59**: 961-974.

Kirkham, L.A., Jefferies, J.M., Kerr, A.R., Jing, Y., Clarke, S.C., Smith, A., and Mitchell, T.J. (2006a) Identification of invasive serotype 1 pneumococcal isolates that express nonhemolytic pneumolysin. *J Clin Microbiol* **44**: 151-159.

Kirkham, L.A., Kerr, A.R., Douce, G.R., Paterson, G.K., Dilts, D.A., Liu, D.F., and Mitchell, T.J. (2006b) Construction and immunological characterization of a novel nontoxic protective pneumolysin mutant for use in future pneumococcal vaccines. *Infect Immun* **74**: 586-593.

Klein, P.J., Bulla, M., Newman, R.A., Muller, P., Uhlenbruck, G., Schaefer, H.E., Kruger, G., and Fisher, R. (1977) Thomsen-Friedenreich antigen in haemolytic-uraemic syndrome. *Lancet* **2**: 1024-1025.

Kloosterman, T.G., Hendriksen, W.T., Bijlsma, J.J., Bootsma, H.J., van Hijum, S.A., Kok, J., Hermans, P.W., and Kuipers, O.P. (2006) Regulation of glutamine and glutamate metabolism by GlnR and GlnA in *Streptococcus pneumoniae*. *J Biol Chem* **281**: 25097-25109.

Klumpp, K., and Graves, B.J. (2006) Optimization of small molecule drugs binding to highly polar target sites: lessons from the discovery and development of neuraminidase inhibitors. *Curr Top Med Chem* **6**: 423-434.

Kohda, C., Kawamura, I., Baba, H., Nomura, T., Ito, Y., Kimoto, T., Watanabe, I., and Mitsuyama, M. (2002) Dissociated linkage of cytokine-inducing activity and cytotoxicity to different domains of listeriolysin O from *Listeria monocytogenes*. *Infect Immun* **70**: 1334-1341.

Kontiainen, S., and Sivonen, A. (1987) Optochin resistance in *Streptococcus pneumoniae* strains isolated from blood and middle ear fluid. *Eur J Clin Microbiol* **6**: 422-424.

Korchev, Y.E., Bashford, C.L., Pederzolli, C., Pasternak, C.A., Morgan, P.J., Andrew, P.W., and Mitchell, T.J. (1998) A conserved tryptophan in pneumolysin is a determinant of the characteristics of channels formed by pneumolysin in cells and planar lipid bilayers. *Biochem J* **329 (Pt 3)**: 571-577.

Kostyukova, N.N., Volkova, M.O., Ivanova, V.V., and Kvetnaya, A.S. (1995) A study of pathogenic factors of *Streptococcus pneumoniae* strains causing meningitis. *FEMS Immunol Med Microbiol* **10**: 133-137.

Kutlu, S.S., Sacar, S., Cevahir, N., and Turgut, H. (2008) Community-acquired *Streptococcus mitis* meningitis: a case report. *Int J Infect Dis*.

Kwon, H.Y., Ogunniyi, A.D., Choi, M.H., Pyo, S.N., Rhee, D.K., and Paton, J.C. (2004) The ClpP protease of *Streptococcus pneumoniae* modulates virulence gene expression and protects against fatal pneumococcal challenge. *Infect Immun* **72**: 5646-5653.

Laible, G., Spratt, B.G., and Hakenbeck, R. (1991) Interspecies recombinational events during the evolution of altered PBP 2x genes in penicillin-resistant

clinical isolates of *Streptococcus pneumoniae*. *Mol Microbiol* **5**: 1993-2002.
LaMarco, K.L., Diven, W.F., and Glew, R.H. (1986) Experimental alteration of chinchilla middle ear mucosae by bacterial neuraminidase. *Ann Otol Rhinol Laryngol* **95**: 304-308.
Lancefield, R. (1933) A serological differenitation of human and other groups of *streptococci*. *J Exp Med*: 441-458.
Lanie, J.A., Ng, W.L., Kazmierczak, K.M., Andrzejewski, T.M., Davidsen, T.M., Wayne, K.J., Tettelin, H., Glass, J.I., and Winkler, M.E. (2007) Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. *J Bacteriol* **189**: 38-51.
Lau, G.W., Haataja, S., Lonetto, M., Kensit, S.E., Marra, A., Bryant, A.P., McDevitt, D., Morrison, D.A., and Holden, D.W. (2001) A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. *Mol Microbiol* **40**: 555-571.
Laurell, A.B. (1959) Neuraminidase-like factors in cultures of pneumococci, alpha-*streptococci* and Pasteurella pseudo-tuberculosis. *Acta Pathol Microbiol Scand* **47**: 182-190.
Lee, C.F., Liu, S.C., Lue, K.H., Chen, J.P., and Sheu, J.N. (2006) Pneumococcal pneumonia with empyema and hemolytic uremic syndrome in children: report of three cases. *J Microbiol Immunol Infect* **39**: 348-352.
Lee, C.J., and Li, Z.M. (2001) Protective immunity and gene expression related to pneumococcal glycoconjugate. *Adv Exp Med Biol* **491**: 505-514.
Lee, C.J., Wang, T.R., and Frasch, C.E. (2001) Immunogenicity in mice of pneumococcal glycoconjugate vaccines using pneumococcal protein carriers. *Vaccine* **19**: 3216-3225.
Lee, L.T., and Howe, C. (1966) Pneumococcal neuraminidase. *J Bacteriol* **91**: 1418-1426.
Leimkugel, J., Adams Forgor, A., Gagneux, S., Pfluger, V., Flierl, C., Awine, E., Naegeli, M., Dangy, J.P., Smith, T., Hodgson, A., and Pluschke, G. (2005) An outbreak of serotype 1 *Streptococcus pneumoniae* meningitis in northern Ghana with features that are characteristic of Neisseria meningitidis meningitis epidemics. *J Infect Dis* **192**: 192-199.
LeMessurier, K.S., Ogunniyi, A.D., and Paton, J.C. (2006) Differential expression of key pneumococcal virulence genes in vivo. *Microbiology* **152**: 305-311.
Lexau, C.A., Lynfield, R., Danila, R., Pilishvili, T., Facklam, R., Farley, M.M., Harrison, L.H., Schaffner, W., Reingold, A., Bennett, N.M., Hadler, J., Cieslak, P.R., and Whitney, C.G. (2005) Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine. *Jama* **294**: 2043-2051.
Lexau, C.A. (2008) The changing epidemiology of pneumococcal pulmonary disease in the era of the heptavalent vaccine. *Curr Infect Dis Rep* **10**: 229-235.
Light, R.W. (1995) A new classification of parapneumonic effusions and empyema. *Chest* **108**: 299-301.
Lloyd-Evans, N., O'Dempsey, T.J., Baldeh, I., Secka, O., Demba, E., Todd, J.E., McArdle, T.F., Banya, W.S., and Greenwood, B.M. (1996) Nasopharyngeal carriage of pneumococci in Gambian children and in their families. *Pediatr Infect Dis J* **15**: 866-871.
Lock, R.A., Paton, J.C., and Hansman, D. (1988a) Comparative efficacy of pneumococcal neuraminidase and pneumolysin as immunogens protective against *Streptococcus pneumoniae*. *Microb Pathog* **5**: 461-467.

Lock, R.A., Paton, J.C., and Hansman, D. (1988b) Purification and immunological characterization of neuraminidase produced by *Streptococcus pneumoniae*. *Microb Pathog* **4**: 33-43.
Lock, R.A., Hansman, D., and Paton, J.C. (1992) Comparative efficacy of autolysin and pneumolysin as immunogens protecting mice against infection by *Streptococcus pneumoniae*. *Microb Pathog* **12**: 137-143.
Lock, R.A., Zhang, Q.Y., Berry, A.M., and Paton, J.C. (1996) Sequence variation in the *Streptococcus pneumoniae* pneumolysin gene affecting haemolytic activity and electrophoretic mobility of the toxin. *Microb Pathog* **21**: 71-83.
Long, J.P., Tong, H.H., and DeMaria, T.F. (2004) Immunization with native or recombinant *Streptococcus pneumoniae* neuraminidase affords protection in the chinchilla otitis media model. *Infect Immun* **72**: 4309-4313.
Lucas, V.S., Beighton, D., Roberts, G.J., and Challacombe, S.J. (1997) Changes in the oral streptococcal flora of children undergoing allogeneic bone marrow transplantation. *J Infect* **35**: 135-141.
Luo, Y., Li, S.C., Chou, M.Y., Li, Y.T., and Luo, M. (1998) The crystal structure of an intramolecular trans-sialidase with a NeuAc alpha2-->3Gal specificity. *Structure* **6**: 521-530.
Luotonen, J., Herva, E., Karma, P., Timonen, M., Leinonen, M., and Makela, P.H. (1981) The bacteriology of acute otitis media in children with special reference to *Streptococcus pneumoniae* as studied by bacteriological and antigen detection methods. *Scand J Infect Dis* **13**: 177-183.
Magee, A.D., and Yother, J. (2001) Requirement for Capsule in Colonization by *Streptococcus pneumoniae*. *Infect. Immun.* **69**: 3755-3761.
Malley, R., Henneke, P., Morse, S.C., Cieslewicz, M.J., Lipsitch, M., Thompson, C.M., Kurt-Jones, E., Paton, J.C., Wessels, M.R., and Golenbock, D.T. (2003) Recognition of pneumolysin by Toll-like receptor 4 confers resistance to pneumococcal infection. *Proc Natl Acad Sci U S A* **100**: 1966-1971.
Manco, S., Hernon, F., Yesilkaya, H., Paton, J.C., Andrew, P.W., and Kadioglu, A. (2006) Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infect Immun* **74**: 4014-4020.
Marra, A., Lawson, S., Asundi, J.S., Brigham, D., and Hromockyj, A.E. (2002) In vivo characterization of the psa genes from *Streptococcus pneumoniae* in multiple models of infection. *Microbiology* **148**: 1483-1491.
Martinot, A., Hue, V., Leclerc, F., and Chenaud, M. (1989) Haemolytic-uraemic syndrome associated with *Streptococcus pneumoniae* meningitis. *Eur J Pediatr* **148**: 648-649.
McAuley, J.L., Hornung, F., Boyd, K.L., Smith, A.M., McKeon, R., Bennink, J., Yewdell, J.W., and McCullers, J.A. (2007) Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe* **2**: 240-249.
McCullers, J.A., and Rehg, J.E. (2002) Lethal synergism between influenza virus and *Streptococcus pneumoniae*: characterization of a mouse model and the role of platelet-activating factor receptor. *J Infect Dis* **186**: 341-350.
McCullers, J.A., and Bartmess, K.C. (2003) Role of neuraminidase in lethal synergism between influenza virus and *Streptococcus pneumoniae*. *J Infect Dis* **187**: 1000-1009.
McCullers, J.A. (2004) Effect of antiviral treatment on the outcome of secondary bacterial pneumonia after influenza. *J Infect Dis* **190**: 519-526.

McCullers, J.A. (2006) Insights into the interaction between influenza virus and pneumococcus. *Clin Microbiol Rev* **19**: 571-582.
McCullers, J.A., Iverson, A.R., McKeon, R., and Murray, P.J. (2008) The platelet activating factor receptor is not required for exacerbation of bacterial pneumonia following influenza. *Scand J Infect Dis* **40**: 11-17.
McEllistrem, M.C., Adams, J., Mason, E.O., and Wald, E.R. (2003) Epidemiology of acute otitis media caused by *Streptococcus pneumoniae* before and after licensure of the 7-valent pneumococcal protein conjugate vaccine. *J Infect Dis* **188**: 1679-1684.
McLeod, C.M., Hodges, R.D., Heidelberger, M., and Bernhard, W.G. (1947) Prevention of pneumococcal pneumonia by immunization with specific capsular serotypes. *J Exp Med*: 445-465.
Medina, M., Villena, J., Vintini, E., Hebert, E.M., Raya, R., and Alvarez, S. (2008) Nasal immunization with *Lactococcus lactis* expressing the pneumococcal protective protein A induces protective immunity in mice. *Infect Immun* **76**: 2696-2705.
Melegaro, A., and Edmunds, W.J. (2004) The 23-valent pneumococcal polysaccharide vaccine. Part I. Efficacy of PPV in the elderly: a comparison of meta-analyses. *Eur J Epidemiol* **19**: 353-363.
Mercat, A., Nguyen, J., and Dautzenberg, B. (1991) An outbreak of pneumococcal pneumonia in two men's shelters. *Chest* **99**: 2-3.
Michon, F., Fusco, P.C., Minetti, C.A., Laude-Sharp, M., Uitz, C., Huang, C.H., D'Ambra, A.J., Moore, S., Remeta, D.P., Heron, I., and Blake, M.S. (1998) Multivalent pneumococcal capsular polysaccharide conjugate vaccines employing genetically detoxified pneumolysin as a carrier protein. *Vaccine* **16**: 1732-1741.
Miles, A. A. and Misra, S. S. (1938) The estimation of the bactericidal power of blood. *J Hyg* **38**: 732
Miller, E., Waight, P., Efstratiou, A., Brisson, M., Johnson, A., and George, R. (2000) Epidemiology of invasive and other pneumococcal disease in children in England and Wales 1996-1998. *Acta Paediatr Suppl* **89**: 11-16.
Mitchell, T.J., Walker, J.A., Saunders, F.K., Andrew, P.W., and Boulnois, G.J. (1989) Expression of the pneumolysin gene in *Escherichia coli*: rapid purification and biological properties. *Biochim Biophys Acta* **1007**: 67-72.
Mitchell, T.J., Mendez, F., Paton, J.C., Andrew, P.W., and Boulnois, G.J. (1990) Comparison of pneumolysin genes and proteins from *Streptococcus pneumoniae* types 1 and 2. *Nucleic Acids Res* **18**: 4010.
Mitchell, T.J., Andrew, P.W., Saunders, F.K., Smith, A.N., and Boulnois, G.J. (1991) Complement activation and antibody binding by pneumolysin via a region of the toxin homologous to a human acute-phase protein. *Mol Microbiol* **5**: 1883-1888.
Mitchell, T.J., and Andrew, P.W. (1997) Biological properties of pneumolysin. *Microb Drug Resist* **3**: 19-26.
Mohammed, B.J., Mitchell, T.J., Andrew, P.W., Hirst, R.A., and O'Callaghan, C. (1999) The effect of the pneumococcal toxin, pneumolysin on brain ependymal cilia. *Microb Pathog* **27**: 303-309.
Moore, M.R., Gertz, R.E., Jr., Woodbury, R.L., Barkocy-Gallagher, G.A., Schaffner, W., Lexau, C., Gershman, K., Reingold, A., Farley, M., Harrison, L.H., Hadler, J.L., Bennett, N.M., Thomas, A.R., McGee, L., Pilishvili, T., Brueggemann, A.B., Whitney, C.G., Jorgensen, J.H., and Beall, B. (2008) Population snapshot of emergent *Streptococcus pneumoniae* serotype 19A in the United States, 2005. *J Infect Dis* **197**: 1016-1027.

Morens, D.M., and Fauci, A.S. (2007) The 1918 influenza pandemic: insights for the 21st century. *J Infect Dis* **195**: 1018-1028.
Morgan, P.J., Hyman, S.C., Rowe, A.J., Mitchell, T.J., Andrew, P.W., and Saibil, H.R. (1995) Subunit organisation and symmetry of pore-forming, oligomeric pneumolysin. *FEBS Lett* **371**: 77-80.
Morris, P.S., Leach, A.J., Silberberg, P., Mellon, G., Wilson, C., Hamilton, E., and Beissbarth, J. (2005) Otitis media in young Aboriginal children from remote communities in Northern and Central Australia: a cross-sectional survey. *BMC Pediatr* **5**: 27.
Morrison, D.A. (1997) Streptococcal competence for genetic transformation: regulation by peptide pheromones. *Microb Drug Resist* **3**: 27-37.
Mortier-Barriere, I., Humbert, O., Martin, B., Prudhomme, M., and Claverys, J.P. (1997) Control of recombination rate during transformation of *Streptococcus pneumoniae*: an overview. *Microb Drug Resist* **3**: 233-242.
Moschioni, M., Donati, C., Muzzi, A., Masignani, V., Censini, S., Hanage, W.P., Bishop, C.J., Reis, J.N., Normark, S., Henriques-Normark, B., Covacci, A., Rappuoli, R., and Barocchi, M.A. (2008) *Streptococcus pneumoniae* contains 3 rlrA pilus variants that are clonally related. *J Infect Dis* **197**: 888-896.
Mulholland, K. (1999) Magnitude of the problem of childhood pneumonia. *Lancet* **354**: 590-592.
Muller-Graf, C.D., Whatmore, A.M., King, S.J., Trzcinski, K., Pickerill, A.P., Doherty, N., Paul, J., Griffiths, D., Crook, D., and Dowson, C.G. (1999) Population biology of *Streptococcus pneumoniae* isolated from oropharyngeal carriage and invasive disease. *Microbiology* **145 (Pt 11)**: 3283-3293.
Muller, P.Y., Janovjak, H., Miserez, A.R., and Dobbie, Z. (2002) Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* **32**: 1372-1374, 1376, 1378-1379.
Munoz, R., Fenoll, A., Vicioso, D., and Casal, J. (1990) Optochin-resistant variants of *Streptococcus pneumoniae*. *Diagn Microbiol Infect Dis* **13**: 63-66.
Munoz, R., Lopez, R., and Garcia, E. (1998) Characterization of IS1515, a functional insertion sequence in *Streptococcus pneumoniae*. *J Bacteriol* **180**: 1381-1388.
Musher, D.M. (1992) Infections caused by *Streptococcus pneumoniae*: clinical spectrum, pathogenesis, immunity, and treatment. *Clin Infect Dis* **14**: 801-807.
Musher, D.M., Mediwala, R., Phan, H.M., Chen, G., and Baughn, R.E. (2001) Nonspecificity of assaying for IgG antibody to pneumolysin in circulating immune complexes as a means to diagnose pneumococcal pneumonia. *Clin Infect Dis* **32**: 534-538.
Myers, C., and Gervaix, A. (2007) *Streptococcus pneumoniae* bacteraemia in children. *Int J Antimicrob Agents* **30 Suppl 1**: S24-28.
Nagamune, H., Ohkura, K., Sukeno, A., Cowan, G., Mitchell, T.J., Ito, W., Ohnishi, O., Hattori, K., Yamato, M., Hirota, K., Miyake, Y., Maeda, T., and Kourai, H. (2004) The human-specific action of intermedilysin, a homolog of streptolysin o, is dictated by domain 4 of the protein. *Microbiol Immunol* **48**: 677-692.
Nandoskar, M., Ferrante, A., Bates, E.J., Hurst, N., and Paton, J.C. (1986) Inhibition of human monocyte respiratory burst, degranulation, phospholipid methylation and bactericidal activity by pneumolysin. *Immunology* **59**: 515-520.

Nath, K.A., Hostetter, M.K., and Hostetter, T.H. (1991) Increased ammoniagenesis as a determinant of progressive renal injury. *Am J Kidney Dis* **17**: 654-657.
Navarre, W.W., and Schneewind, O. (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* **63**: 174-229.
Neeleman, C., Klaassen, C.H., Klomberg, D.M., de Valk, H.A., and Mouton, J.W. (2004) Pneumolysin is a key factor in misidentification of macrolide-resistant *Streptococcus pneumoniae* and is a putative virulence factor of *S. mitis* and other *streptococci*. *J Clin Microbiol* **42**: 4355-4357.
Nelson, A.L., Ries, J., Bagnoli, F., Dahlberg, S., Falker, S., Rounioja, S., Tschop, J., Morfeldt, E., Ferlenghi, I., Hilleringmann, M., Holden, D.W., Rappuoli, R., Normark, S., Barocchi, M.A., and Henriques-Normark, B. (2007) RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol Microbiol* **66**: 329-340.
Neuhaus, T.J., Calonder, S., and Leumann, E.P. (1997) Heterogeneity of atypical haemolytic uraemic syndromes. *Arch Dis Child* **76**: 518-521.
Newstead, S.L., Potter, J.A., Wilson, J.C., Xu, G., Chien, C.H., Watts, A.G., Withers, S.G., and Taylor, G.L. (2008) The structure of Clostridium perfringens NanI sialidase and its catalytic intermediates. *J Biol Chem* **283**: 9080-9088.
Ng, E.W., Samiy, N., Rubins, J.B., Cousins, F.V., Ruoff, K.L., Baker, A.S., and D'Amico, D.J. (1997) Implication of pneumolysin as a virulence factor in *Streptococcus pneumoniae* endophthalmitis. *Retina* **17**: 521-529.
Ng, E.W., Costa, J.R., Samiy, N., Ruoff, K.L., Connolly, E., Cousins, F.V., and D'Amico, D.J. (2002) Contribution of pneumolysin and autolysin to the pathogenesis of experimental pneumococcal endophthalmitis. *Retina* **22**: 622-632.
Novak, R.W., Martin, C.R., and Orsini, E.N. (1983) Hemolytic-uremic syndrome and T-cryptantigen exposure by neuraminidase-producing pneumococci: an emerging problem? *Pediatr Pathol* **1**: 409-413.
Nurkka, A., Joensuu, J., Henckaerts, I., Peeters, P., Poolman, J., Kilpi, T., and Kayhty, H. (2004) Immunogenicity and safety of the eleven valent pneumococcal polysaccharide-protein D conjugate vaccine in infants. *Pediatr Infect Dis J* **23**: 1008-1014.
O'Brien, A.D., Newland, J.W., Miller, S.F., Holmes, R.K., Smith, H.W., and Formal, S.B. (1984) Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**: 694-696.
Obaro, S., and Adegbola, R. (2002) The pneumococcus: carriage, disease and conjugate vaccines. *J Med Microbiol* **51**: 98-104.
Obaro, S.K. (2000) Prospects for pneumococcal vaccination in African children. *Acta Trop* **75**: 141-153.
Obert, C., Sublett, J., Kaushal, D., Hinojosa, E., Barton, T., Tuomanen, E.I., and Orihuela, C.J. (2006) Identification of a Candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. *Infect Immun* **74**: 4766-4777.
Obregon, V., Garcia, P., Garcia, E., Fenoll, A., Lopez, R., and Garcia, J.L. (2002) Molecular peculiarities of the lytA gene isolated from clinical pneumococcal strains that are bile insoluble. *J Clin Microbiol* **40**: 2545-2554.
Ogunniyi, A.D., Folland, R.L., Briles, D.E., Hollingshead, S.K., and Paton, J.C. (2000) Immunization of mice with combinations of pneumococcal

virulence proteins elicits enhanced protection against challenge with *Streptococcus pneumoniae*. *Infect Immun* **68**: 3028-3033.

Ogunniyi, A.D., Woodrow, M.C., Poolman, J.T., and Paton, J.C. (2001) Protection against *Streptococcus pneumoniae* elicited by immunization with pneumolysin and CbpA. *Infect Immun* **69**: 5997-6003.

Ogunniyi, A.D., Grabowicz, M., Briles, D.E., Cook, J., and Paton, J.C. (2007a) Development of a vaccine against invasive pneumococcal disease based on combinations of virulence proteins of *Streptococcus pneumoniae*. *Infect Immun* **75**: 350-357.

Ogunniyi, A.D., LeMessurier, K.S., Graham, R.M., Watt, J.M., Briles, D.E., Stroeher, U.H., and Paton, J.C. (2007b) Contributions of pneumolysin, pneumococcal surface protein A (PspA), and PspC to pathogenicity of *Streptococcus pneumoniae* D39 in a mouse model. *Infect Immun* **75**: 1843-1851.

Ohno-Iwashita, Y., Iwamoto, M., Ando, S., and Iwashita, S. (1992) Effect of lipidic factors on membrane cholesterol topology--mode of binding of theta-toxin to cholesterol in liposomes. *Biochim Biophys Acta* **1109**: 81-90.

Orihuela, C.J., Gao, G., Francis, K.P., Yu, J., and Tuomanen, E.I. (2004) Tissue-specific contributions of pneumococcal virulence factors to pathogenesis. *J Infect Dis* **190**: 1661-1669.

Orihuela, C. J., Fillon, S., Smith-Sielicki, S. H., El Kasmi, K. C., Gao, G., Soulis, K., Patil, A., Murray, P. J., Tuomanen, E. I. (2006) Cell wall-mediated neuronal damage in early sepsis. *Infect Immun* **74(7)**: 3783-9.

Ortqvist, A., Hedlund, J., Burman, L.A., Elbel, E., Hofer, M., Leinonen, M., Lindblad, I., Sundelof, B., and Kalin, M. (1998) Randomised trial of 23-valent pneumococcal capsular polysaccharide vaccine in prevention of pneumonia in middle-aged and elderly people. Swedish Pneumococcal Vaccination Study Group. *Lancet* **351**: 399-403.

Ostergaard, C., Konradsen, H.B., and Samuelsson, S. (2005) Clinical presentation and prognostic factors of *Streptococcus pneumoniae* meningitis according to the focus of infection. *BMC Infect Dis* **5**: 93.

Overweg, K., Pericone, C.D., Verhoef, G.G., Weiser, J.N., Meiring, H.D., De Jong, A.P., De Groot, R., and Hermans, P.W. (2000) Differential protein expression in phenotypic variants of *Streptococcus pneumoniae*. *Infect Immun* **68**: 4604-4610.

Pai, R., Limor, J., and Beall, B. (2005a) Use of pyrosequencing to differentiate *Streptococcus pneumoniae* serotypes 6A and 6B. *J Clin Microbiol* **43**: 4820-4822.

Pai, R., Moore, M.R., Pilishvili, T., Gertz, R.E., Whitney, C.G., and Beall, B. (2005b) Postvaccine genetic structure of *Streptococcus pneumoniae* serotype 19A from children in the United States. *J Infect Dis* **192**: 1988-1995.

Pai, R., Gertz, R.E., and Beall, B. (2006) Sequential Multiplex PCR Approach for Determining Capsular Serotypes of *Streptococcus pneumoniae* Isolates. *J Clin Microbiol* **44**: 124-131.

Palaniappan, R., Singh, S., Singh, U.P., Sakthivel, S.K., Ades, E.W., Briles, D.E., Hollingshead, S.K., Paton, J.C., Sampson, J.S., and Lillard, J.W., Jr. (2005) Differential PsaA-, PspA-, PspC-, and PdB-specific immune responses in a mouse model of pneumococcal carriage. *Infect Immun* **73**: 1006-1013.

Palmer, M., Vulicevic, I., Saweljew, P., Valeva, A., Kehoe, M., and Bhakdi, S. (1998) Streptolysin O: a proposed model of allosteric interaction between

a pore-forming protein and its target lipid bilayer. *Biochemistry* **37**: 2378-2383.
Palmer, M. (2001) The family of thiol-activated, cholesterol-binding cytolysins. *Toxicon* **39**: 1681-1689.
Park, I.H., Pritchard, D.G., Cartee, R., Brandao, A., Brandileone, M.C., and Nahm, M.H. (2007) Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* **45**: 1225-1233.
Pasteur, L. (1881) Note sur la maladie nouvelle provoquee par la salive d'un enfant mort de la rage. *Bulletin de l'Academie de Medicine* **10**: 94-103
Paterson, G.K., and Mitchell, T.J. (2006) The role of *Streptococcus pneumoniae* sortase A in colonisation and pathogenesis. *Microbes Infect* **8**: 145-153.
Paton, J.C., and Ferrante, A. (1983) Inhibition of human polymorphonuclear leukocyte respiratory burst, bactericidal activity, and migration by pneumolysin. *Infect Immun* **41**: 1212-1216.
Paton, J.C., Lock, R.A., and Hansman, D.J. (1983) Effect of immunization with pneumolysin on survival time of mice challenged with *Streptococcus pneumoniae*. *Infect Immun* **40**: 548-552.
Paton, J.C., Rowan-Kelly, B., and Ferrante, A. (1984) Activation of human complement by the pneumococcal toxin pneumolysin. *Infect Immun* **43**: 1085-1087.
Paton, J.C., Andrew, P.W., Boulnois, G.J., and Mitchell, T.J. (1993) Molecular analysis of the pathogenicity of *Streptococcus pneumoniae*: the role of pneumococcal proteins. *Annu Rev Microbiol* **47**: 89-115.
Paudyal, B.P. (2007) Twelve episodes of meningitis in the same patient! *Kathmandu Univ Med J (KUMJ)* **5**: 243-246.
Pebody, R.G., Leino, T., Nohynek, H., Hellenbrand, W., Salmaso, S., and Ruutu, P. (2005) Pneumococcal vaccination policy in Europe. *Euro Surveill* **10**: 174-178.
Peltola, V.T., and McCullers, J.A. (2004) Respiratory viruses predisposing to bacterial infections: role of neuraminidase. *Pediatr Infect Dis J* **23**: S87-97.
Peltola, V.T., Murti, K.G., and McCullers, J.A. (2005) Influenza virus neuraminidase contributes to secondary bacterial pneumonia. *J Infect Dis* **192**: 249-257.
Pericone, C.D., Overweg, K., Hermans, P.W., and Weiser, J.N. (2000) Inhibitory and bactericidal effects of hydrogen peroxide production by *Streptococcus pneumoniae* on other inhabitants of the upper respiratory tract. *Infect Immun* **68**: 3990-3997.
Pettigrew, M.M., Fennie, K.P., York, M.P., Daniels, J., and Ghaffar, F. (2006) Variation in the presence of neuraminidase genes among *Streptococcus pneumoniae* isolates with identical sequence types. *Infect Immun* **74**: 3360-3365.
Phillips, G., Barker, R., and Brogan, O. (1988) Optochin-resistant *Streptococcus pneumoniae*. *Lancet* **2**: 281.
Pichichero, M.E., and Casey, J.R. (2007) Emergence of a multiresistant serotype 19A pneumococcal strain not included in the 7-valent conjugate vaccine as an otopathogen in children. *Jama* **298**: 1772-1778.
Plotkowski, M.C., Puchelle, E., Beck, G., Jacquot, J., and Hannoun, C. (1986) Adherence of type I *Streptococcus pneumoniae* to tracheal epithelium of mice infected with influenza A/PR8 virus. *Am Rev Respir Dis* **134**: 1040-1044.

Polekhina, G., Giddings, K.S., Tweten, R.K., and Parker, M.W. (2005) Insights into the action of the superfamily of cholesterol-dependent cytolysins from studies of intermedilysin. *Proc Natl Acad Sci U S A* **102**: 600-605.
Polissi, A., Pontiggia, A., Feger, G., Altieri, M., Mottl, H., Ferrari, L., and Simon, D. (1998) Large-scale identification of virulence genes from *Streptococcus pneumoniae*. *Infect Immun* **66**: 5620-5629.
Porat, N., Greenberg, D., Givon-Lavi, N., Shuval, D.S., Trefler, R., Segev, O., Hanage, W.P., and Dagan, R. (2006) The important role of nontypable *Streptococcus pneumoniae* international clones in acute conjunctivitis. *J Infect Dis* **194**: 689-696.
Poulsen, K., Reinholdt, J., and Kilian, M. (1996) Characterization of the *Streptococcus pneumoniae* immunoglobulin A1 protease gene (iga) and its translation product. *Infect Immun* **64**: 3957-3966.
Poulsen, K., Reinholdt, J., Jespersgaard, C., Boye, K., Brown, T.A., Hauge, M., and Kilian, M. (1998) A comprehensive genetic study of streptococcal immunoglobulin A1 proteases: evidence for recombination within and between species. *Infect Immun* **66**: 181-190.
Prellner, K., Hermansson, A., White, P., Melhus, A., and Briles, D. (1999) Immunization and protection in pneumococcal otitis media studied in a rat model. *Microb Drug Resist* **5**: 73-82.
Proulx, J.F., Dery, S., Jette, L.P., Ismael, J., Libman, M., and De Wals, P. (2002) Pneumonia epidemic caused by a virulent strain of *Streptococcus pneumoniae* serotype 1 in Nunavik, Quebec. *Can Commun Dis Rep* **28**: 129-131.
Quin, L.R., Moore, Q.C., 3rd, and McDaniel, L.S. (2007) Pneumolysin, PspA, and PspC contribute to pneumococcal evasion of early innate immune responses during bacteremia in mice. *Infect Immun* **75**: 2067-2070.
Ramachandran, R., Tweten, R.K., and Johnson, A.E. (2005) The domains of a cholesterol-dependent cytolysin undergo a major FRET-detected rearrangement during pore formation. *Proc Natl Acad Sci U S A* **102**: 7139-7144.
Rapola, S., Jantti, V., Haikala, R., Syrjanen, R., Carlone, G.M., Sampson, J.S., Briles, D.E., Paton, J.C., Takala, A.K., Kilpi, T.M., and Kayhty, H. (2000) Natural development of antibodies to pneumococcal surface protein A, pneumococcal surface adhesin A, and pneumolysin in relation to pneumococcal carriage and acute otitis media. *J Infect Dis* **182**: 1146-1152.
Rayner, C.F., Jackson, A.D., Rutman, A., Dewar, A., Mitchell, T.J., Andrew, P.W., Cole, P.J., and Wilson, R. (1995) Interaction of pneumolysin-sufficient and -deficient isogenic variants of *Streptococcus pneumoniae* with human respiratory mucosa. *Infect Immun* **63**: 442-447.
Regev-Yochay, G., Trzcinski, K., Thompson, C.M., Malley, R., and Lipsitch, M. (2006) Interference between *Streptococcus pneumoniae* and *Staphylococcus aureus*: In vitro hydrogen peroxide-mediated killing by *Streptococcus pneumoniae*. *J Bacteriol* **188**: 4996-5001.
Reichmann, P., Konig, A., Linares, J., Alcaide, F., Tenover, F.C., McDougal, L., Swidsinski, S., and Hakenbeck, R. (1997) A global gene pool for high-level cephalosporin resistance in commensal Streptococcus species and *Streptococcus pneumoniae*. *J Infect Dis* **176**: 1001-1012.
Ren, B., McCrory, M.A., Pass, C., Bullard, D.C., Ballantyne, C.M., Xu, Y., Briles, D.E., and Szalai, A.J. (2004) The virulence function of *Streptococcus pneumoniae* surface protein A involves inhibition of complement

activation and impairment of complement receptor-mediated protection. *J Immunol* **173**: 7506-7512.
Robbins, J.B., and Schneerson, R. (1983) Planning for a second (23 valent) generation pneumococcal vaccine. With special reference to new developments in our understanding of the structure and biology of polysaccharides. *Bull Eur Physiopathol Respir* **19**: 215-226.
Roberts, P., Jeffery, P.K., Mitchell, T.J., Andrew, P.W., Boulnois, G.J., Feldman, C., Cole, P.J., and Wilson, R. (1992) Effect of immunization with Freund's adjuvant and pneumolysin on histologic features of pneumococcal infection in the rat lung in vivo. *Infect Immun* **60**: 4969-4972.
Rogers, P.D., Thornton, J., Barker, K.S., McDaniel, D.O., Sacks, G.S., Swiatlo, E., and McDaniel, L.S. (2003) Pneumolysin-dependent and -independent gene expression identified by cDNA microarray analysis of THP-1 human mononuclear cells stimulated by *Streptococcus pneumoniae*. *Infect Immun* **71**: 2087-2094.
Roggentin, P., Rothe, B., Kaper, J.B., Galen, J., Lawrisuk, L., Vimr, E.R., and Schauer, R. (1989) Conserved sequences in bacterial and viral sialidases. *Glycoconj J* **6**: 349-353.
Rosenow, C., Ryan, P., Weiser, J.N., Johnson, S., Fontan, P., Ortqvist, A., and Masure, H.R. (1997) Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol Microbiol* **25**: 819-829.
Rossjohn, J., Feil, S.C., McKinstry, W.J., Tweten, R.K., and Parker, M.W. (1997) Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form. *Cell* **89**: 685-692.
Rubins, J.B., Duane, P.G., Charboneau, D., and Janoff, E.N. (1992) Toxicity of pneumolysin to pulmonary endothelial cells in vitro. *Infect Immun* **60**: 1740-1746.
Rubins, J.B., Duane, P.G., Clawson, D., Charboneau, D., Young, J., and Niewoehner, D.E. (1993) Toxicity of pneumolysin to pulmonary alveolar epithelial cells. *Infect Immun* **61**: 1352-1358.
Rubins, J.B., Charboneau, D., Paton, J.C., Mitchell, T.J., Andrew, P.W., and Janoff, E.N. (1995) Dual function of pneumolysin in the early pathogenesis of murine pneumococcal pneumonia. *J Clin Invest* **95**: 142-150.
Rubins, J.B., Charboneau, D., Fasching, C., Berry, A.M., Paton, J.C., Alexander, J.E., Andrew, P.W., Mitchell, T.J., and Janoff, E.N. (1996) Distinct roles for pneumolysin's cytotoxic and complement activities in the pathogenesis of pneumococcal pneumonia. *Am J Respir Crit Care Med* **153**: 1339-1346.
Rubins, J.B., Paddock, A.H., Charboneau, D., Berry, A.M., Paton, J.C., and Janoff, E.N. (1998) Pneumolysin in pneumococcal adherence and colonization. *Microb Pathog* **25**: 337-342.
Sa-Leao, R., Nunes, S., Brito-Avo, A., Alves, C.R., Carrico, J.A., Saldanha, J., Almeida, J.S., Santos-Sanches, I., and de Lencastre, H. (2008) High rates of transmission of and colonization by *Streptococcus pneumoniae* and *Haemophilus influenzae* within a day care center revealed in a longitudinal study. *J Clin Microbiol* **46**: 225-234.
Sampson, J.S., Furlow, Z., Whitney, A.M., Williams, D., Facklam, R., and Carlone, G.M. (1997) Limited diversity of *Streptococcus pneumoniae* psaA among pneumococcal vaccine serotypes. *Infect Immun* **65**: 1967-1971.
Sato, K., Quartey, M.K., Liebeler, C.L., Le, C.T., and Giebink, G.S. (1996) Roles of autolysin and pneumolysin in middle ear inflammation caused by a type

3 *Streptococcus pneumoniae* strain in the chinchilla otitis media model. *Infect Immun* **64**: 1140-1145.
Saunders, F.K., Mitchell, T.J., Walker, J.A., Andrew, P.W., and Boulnois, G.J. (1989) Pneumolysin, the thiol-activated toxin of *Streptococcus pneumoniae*, does not require a thiol group for in vitro activity. *Infect Immun* **57**: 2547-2552.
Scanlon, K.L., Diven, W.F., and Glew, R.H. (1989) Purification and properties of *Streptococcus pneumoniae* neuraminidase. *Enzyme* **41**: 143-150.
Scott, J.A. (2007) The preventable burden of pneumococcal disease in the developing world. *Vaccine* **25**: 2398-2405.
Seger, R., Joller, P., Baerlocher, K., Kenny, A., Dulake, C., Leumann, E., Spierig, M., and Hitzig, W.H. (1980) Hemolytic-uremic syndrome associated with neuraminidase-producing microorganisms: treatment by exchange transfusion. *Helv Paediatr Acta* **35**: 359-367.
Shakhnovich, E.A., King, S.J., and Weiser, J.N. (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**: 7161-7164.
Shannon, J.G., Ross, C.L., Koehler, T.M., and Rest, R.F. (2003) Characterization of anthrolysin O, the *Bacillus anthracis* cholesterol-dependent cytolysin. *Infect Immun* **71**: 3183-3189.
Shaper, M., Hollingshead, S.K., Benjamin, W.H., Jr., and Briles, D.E. (2004) PspA protects *Streptococcus pneumoniae* from killing by apolactoferrin, and antibody to PspA enhances killing of pneumococci by apolactoferrin [corrected]. *Infect Immun* **72**: 5031-5040.
Shepard, L.A., Shatursky, O., Johnson, A.E., and Tweten, R.K. (2000) The mechanism of pore assembly for a cholesterol-dependent cytolysin: formation of a large prepore complex precedes the insertion of the transmembrane beta-hairpins. *Biochemistry* **39**: 10284-10293.
Shivers, R.P., and Sonenshein, A.L. (2005) *Bacillus subtilis* ilvB operon: an intersection of global regulons. *Mol Microbiol* **56**: 1549-1559.
Sibold, C., Henrichsen, J., Konig, A., Martin, C., Chalkley, L., and Hakenbeck, R. (1994) Mosaic pbpX genes of major clones of penicillin-resistant *Streptococcus pneumoniae* have evolved from pbpX genes of a penicillin-sensitive *Streptococcus oralis*. *Mol Microbiol* **12**: 1013-1023.
Silva, N.A., McCluskey, J., Jefferies, J.M., Hinds, J., Smith, A., Clarke, S.C., Mitchell, T.J., and Paterson, G.K. (2006) Genomic diversity between strains of the same serotype and multilocus sequence type among pneumococcal clinical isolates. *Infect Immun* **74**: 3513-3518.
Simonsen, L. (1999) The global impact of influenza on morbidity and mortality. *Vaccine* **17 Suppl 1**: S3-10.
Sjostrom, K., Spindler, C., Ortqvist, A., Kalin, M., Sandgren, A., Kuhlmann-Berenzon, S., and Henriques-Normark, B. (2006) Clonal and capsular types decide whether pneumococci will act as a primary or opportunistic pathogen. *Clin Infect Dis* **42**: 451-459.
Sjostrom, K., Blomberg, C., Fernebro, J., Dagerhamn, J., Morfeldt, E., Barocchi, M.A., Browall, S., Moschioni, M., Andersson, M., Henriques, F., Albiger, B., Rappuoli, R., Normark, S., and Henriques-Normark, B. (2007) Clonal success of piliated penicillin nonsusceptible pneumococci. *Proc Natl Acad Sci U S A* **104**: 12907-12912.
Skinner, L.J., Beurg, M., Mitchell, T.J., Darrouzet, V., Aran, J.M., and Dulon, D. (2004) Intracochlear Perfusion of Pneumolysin, a Pneumococcal Protein,

Rapidly Abolishes Auditory Potentials in the Guinea Pig Cochlea. *Acta Otolaryngol* **124**: 1000-1007.

Smart, L.E. (1986) An alternative approach to typing of *Streptococcus pneumoniae* strains by coagglutination. *Acta Pathology Microbiology and Immunology Scandinavia (B)* **94**: 409-413.

Smit, P., Oberholzer, D., Hayden-Smith, S., Koornhof, H.J., and Hilleman, M.R. (1977) Protective efficacy of pneumococcal polysaccharide vaccines. *Jama* **238**: 2613-2616.

Smith, A., Jackson, M.S., and Kennedy, H. (2004) Antimicrobial susceptibility of viridans group streptococcal blood isolates to eight antimicrobial agents. *Scand J Infect Dis* **36**: 259-263.

Smith, M.W., Schmidt, J.E., Rehg, J.E., Orihuela, C.J., and McCullers, J.A. (2007) Induction of pro- and anti-inflammatory molecules in a mouse model of pneumococcal pneumonia after influenza. *Comp Med* **57**: 82-89.

Srivastava, A., Henneke, P., Visintin, A., Morse, S.C., Martin, V., Watkins, C., Paton, J.C., Wessels, M.R., Golenbock, D.T., and Malley, R. (2005) The apoptotic response to pneumolysin is Toll-like receptor 4 dependent and protects against pneumococcal disease. *Infect Immun* **73**: 6479-6487.

Steinfort, C., Wilson, R., Mitchell, T., Feldman, C., Rutman, A., Todd, H., Sykes, D., Walker, J., Saunders, K., and Andrew, P.W. (1989) Effect of *Streptococcus pneumoniae* on human respiratory epithelium in vitro. *Infect Immun* **57**: 2006-2013.

Sternberg, G. M. (1881) A fatal form of septicaemia in the rabbit, produced by the subcutaneous injection of human serum. *Annual reports to the National Board of Health* **3**: 87-108

Stone, W.J., and Swift, G.W. (1919) Influenza and influenzal pneumonia at Fort Riley, Kansas. *Jama*: 487-493.

Sung, C.K., Li, H., Claverys, J.P., and Morrison, D.A. (2001) An rpsL cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. *Appl Environ Microbiol* **67**: 5190-5196.

Swiatlo, E., King, J., Nabors, G.S., Mathews, B., and Briles, D.E. (2003) Pneumococcal surface protein A is expressed in vivo, and antibodies to PspA are effective for therapy in a murine model of pneumococcal sepsis. *Infect Immun* **71**: 7149-7153.

Tai, S.S. (2006) *Streptococcus pneumoniae* protein vaccine candidates: properties, activities and animal studies. *Crit Rev Microbiol* **32**: 139-153.

Talkington, D.F., Brown, B.G., Tharpe, J.A., Koenig, A., and Russell, H. (1996) Protection of mice against fatal pneumococcal challenge by immunization with pneumococcal surface adhesin A (PsaA). *Microb Pathog* **21**: 17-22.

Tangsinmankong, N., Kamchaisatian, W., Day, N.K., Sleasman, J.W., and Emmanuel, P.J. (2004) Immunogenicity of 23-valent pneumococcal polysaccharide vaccine in children with human immunodeficiency virus undergoing highly active antiretroviral therapy. *Ann Allergy Asthma Immunol* **92**: 558-564.

Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickinson, T., Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K., and Fraser, C.M. (2001) Complete

Genome Sequence of a Virulent Isolate of *Streptococcus pneumoniae*. *Science* **293**: 498-506.
Thornton, J., and McDaniel, L.S. (2005) THP-1 monocytes up-regulate intercellular adhesion molecule 1 in response to pneumolysin from *Streptococcus pneumoniae*. *Infect Immun* **73**: 6493-6498.
Tilley, S.J., Orlova, E.V., Gilbert, R.J., Andrew, P.W., and Saibil, H.R. (2005) Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell* **121**: 247-256.
Tong, H.H., McIver, M.A., Fisher, L.M., and DeMaria, T.F. (1999) Effect of lacto-N-neotetraose, asialoganglioside-GM1 and neuraminidase on adherence of otitis media-associated serotypes of *Streptococcus pneumoniae* to chinchilla tracheal epithelium. *Microb Pathog* **26**: 111-119.
Tong, H.H., Blue, L.E., James, M.A., and DeMaria, T.F. (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**: 921-924.
Tong, H.H., Liu, X., Chen, Y., James, M., and Demaria, T. (2002) Effect of neuraminidase on receptor-mediated adherence of *Streptococcus pneumoniae* to chinchilla tracheal epithelium. *Acta Otolaryngol* **122**: 413-419.
Tong, H.H., Li, D., Chen, S., Long, J.P., and DeMaria, T.F. (2005) Immunization with recombinant *Streptococcus pneumoniae* neuraminidase NanA protects chinchillas against nasopharyngeal colonization. *Infect Immun* **73**: 7775-7778.
Tseng, H.J., McEwan, A.G., Paton, J.C., and Jennings, M.P. (2002) Virulence of *Streptococcus pneumoniae*: PsaA mutants are hypersensitive to oxidative stress. *Infect Immun* **70**: 1635-1639.
Tuomanen, E. (1999) Molecular and cellular biology of pneumococcal infection. *Curr Opin Microbiol* **2**: 35-39.
Tuomanen, E., Mitchell, T. J., Morrison, D. A. and Spratt, B. G. (2006) The pneumococcus - ISBN 1-55581-297-X
Tweten, R.K., Parker, M.W., and Johnson, A.E. (2001) The cholesterol-dependent cytolysins. *Curr Top Microbiol Immunol* **257**: 15-33.
van Ginkel, F.W., McGhee, J.R., Watt, J.M., Campos-Torres, A., Parish, L.A., and Briles, D.E. (2003) Pneumococcal carriage results in ganglioside-mediated olfactory tissue infection. *Proc Natl Acad Sci U S A* **100**: 14363-14367.
van Rossum, A.M., Lysenko, E.S., and Weiser, J.N. (2005) Host and bacterial factors contributing to the clearance of colonization by *Streptococcus pneumoniae* in a murine model. *Infect Immun* **73**: 7718-7726.
Vanderkooi, O.G., Kellner, J.D., Wade, A.W., Jadavji, T., Midgley, J.P., Louie, T., and Tyrell, G.J. (2003) Invasive *Streptococcus pneumoniae* infection causing hemolytic uremic syndrome in children: Two recent cases. *Can J Infect Dis* **14**: 339-343.
Virolainen, A., Jero, J., Kayhty, H., Karma, P., Eskola, J., and Leinonen, M. (1995a) Nasopharyngeal antibodies to pneumococcal pneumolysin in children with acute otitis media. *Clin Diagn Lab Immunol* **2**: 704-707.
Virolainen, A., Jero, J., Kayhty, H., Karma, P., Leinonen, M., and Eskola, J. (1995b) Antibodies to pneumolysin and pneumococcal capsular polysaccharides in middle ear fluid of children with acute otitis media. *Acta Otolaryngol* **115**: 796-803.
Virolainen, A., Jero, J., Chattopadhyay, P., Karma, P., Eskola, J., and Leinonen, M. (1996) Comparison of serum antibodies to pneumolysin with those to

pneumococcal capsular polysaccharides in children with acute otitis media. *Pediatr Infect Dis J* **15**: 128-133.
Waldor, M.K., and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910-1914.
Walker, J.A., Allen, R.L., Falmagne, P., Johnson, M.K., and Boulnois, G.J. (1987) Molecular cloning, characterization, and complete nucleotide sequence of the gene for pneumolysin, the sulfhydryl-activated toxin of *Streptococcus pneumoniae*. *Infect Immun* **55**: 1184-1189.
Wartha, F., Beiter, K., Albiger, B., Fernebro, J., Zychlinsky, A., Normark, S., and Henriques-Normark, B. (2007) Capsule and D-alanylated lipoteichoic acids protect *Streptococcus pneumoniae* against neutrophil extracellular traps. *Cell Microbiol* **9**: 1162-1171.
Waters, A.M., Kerecuk, L., Luk, D., Haq, M.R., Fitzpatrick, M.M., Gilbert, R.D., Inward, C., Jones, C., Pichon, B., Reid, C., Slack, M.P., Van't Hoff, W., Dillon, M.J., Taylor, C.M., and Tullus, K. (2007) Hemolytic uremic syndrome associated with invasive pneumococcal disease: the United kingdom experience. *J Pediatr* **151**: 140-144.
Weintraub, A. (2003) Immunology of bacterial polysaccharide antigens. *Carbohydr Res* **338**: 2539-2547.
Weis, S., and Palmer, M. (2001) Streptolysin O: the C-terminal, tryptophan-rich domain carries functional sites for both membrane binding and self-interaction but not for stable oligomerization. *Biochim Biophys Acta* **1510**: 292-299.
Weiser, J.N., Austrian, R., Sreenivasan, P.K., and Masure, H.R. (1994) Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. *Infect Immun* **62**: 2582-2589.
Weiser, J.N., Bae, D., Fasching, C., Scamurra, R.W., Ratner, A.J., and Janoff, E.N. (2003) Antibody-enhanced pneumococcal adherence requires IgA1 protease. *Proc Natl Acad Sci U S A* **100**: 4215-4220.
Wellmer, A., Zysk, G., Gerber, J., Kunst, T., Von Mering, M., Bunkowski, S., Eiffert, H., and Nau, R. (2002) Decreased virulence of a pneumolysin-deficient strain of *Streptococcus pneumoniae* in murine meningitis. *Infect Immun* **70**: 6504-6508.
Whalan, R.H., Funnell, S.G., Bowler, L.D., Hudson, M.J., Robinson, A., and Dowson, C.G. (2006) Distribution and genetic diversity of the ABC transporter lipoproteins PiuA and PiaA within *Streptococcus pneumoniae* and related *streptococci*. *J Bacteriol* **188**: 1031-1038.
Whatmore, A.M., Barcus, V.A., and Dowson, C.G. (1999) Genetic diversity of the streptococcal competence (com) gene locus. *J Bacteriol* **181**: 3144-3154.
Whatmore, A.M., and Dowson, C.G. (1999) The autolysin-encoding gene (lytA) of *Streptococcus pneumoniae* displays restricted allelic variation despite localized recombination events with genes of pneumococcal bacteriophage encoding cell wall lytic enzymes. *Infect Immun* **67**: 4551-4556.
Whatmore, A. M., King, S. J., Doherty, N. C., Sturgeon, D., Chanter, N., Dowson, C. G. (1999) Molecular characterisation of equine isolates of *Streptococcus pneumoniae*: natural disruption of genes encoding the virulence factors pneumolysin and autolysin. *Infect Immun* **67(6)** 2776-82.
Whatmore, A.M., Efstratiou, A., Pickerill, A.P., Broughton, K., Woodard, G., Sturgeon, D., George, R., and Dowson, C.G. (2000) Genetic relationships between clinical isolates of *Streptococcus pneumoniae, Streptococcus oralis,* and *Streptococcus mitis*: characterization of "Atypical"

pneumococci and organisms allied to *S. mitis* harboring *S. pneumoniae* virulence factor-encoding genes. *Infect Immun* **68**: 1374-1382.
Whitney, C.G. (2005) Impact of conjugate pneumococcal vaccines. *Pediatr Infect Dis J* **24**: 729-730.
Wickremesinghe, R., and Russell, C. (1976) Viridans *streptococci* associated with periapical dental abscesses. *Infection* **4**: 196-203.
Williamson, Y.M., Gowrisankar, R., Longo, D.L., Facklam, R., Gipson, I.K., Ades, E.P., Carlone, G.M., and Sampson, J.S. (2008) Adherence of nontypeable *Streptococcus pneumoniae* to human conjunctival epithelial cells. *Microb Pathog* **44**: 175-185.
Winkelstein, J.A. (1981) The role of complement in the host's defense against *Streptococcus pneumoniae*. *Rev Infect Dis* **3**: 289-298.
Winter, A.J., Comis, S.D., Osborne, M.P., Tarlow, M.J., Stephen, J., Andrew, P.W., Hill, J., and Mitchell, T.J. (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**: 4411-4418.
Wizemann, T.M., Heinrichs, J.H., Adamou, J.E., Erwin, A.L., Kunsch, C., Choi, G.H., Barash, S.C., Rosen, C.A., Masure, H.R., Tuomanen, E., Gayle, A., Brewah, Y.A., Walsh, W., Barren, P., Lathigra, R., Hanson, M., Langermann, S., Johnson, S., and Koenig, S. (2001) Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect Immun* **69**: 1593-1598.
Wright, A.E., Parry Morgan, W., Colebrook, L., and Dodgson, R.W. (1914) Prophylactic inoculation against pneumococcal infection. *Lancet*: 87-95.
Wu, H.Y., Nahm, M.H., Guo, Y., Russell, M.W., and Briles, D.E. (1997) Intranasal immunization of mice with PspA (pneumococcal surface protein A) can prevent intranasal carriage, pulmonary infection, and sepsis with *Streptococcus pneumoniae*. *J Infect Dis* **175**: 839-846.
Wu, T.T., Kabat, E.A., and Bilofsky, H. (1975) Similarities among hypervariable segments of immunoglobulin chains. *Proc Natl Acad Sci U S A* **72**: 5107-5110.
Yesilkaya, H., Soma-Haddrick, S., Crennell, S.J., and Andrew, P.W. (2006) Identification of amino acids essential for catalytic activity of pneumococcal neuraminidase A. *Res Microbiol* **157**: 569-574.
Yesilkaya, H., Manco, S., Kadioglu, A., Terra, V.S., and Andrew, P.W. (2008) The ability to utilize mucin affects the regulation of virulence gene expression in *Streptococcus pneumoniae*. *FEMS Microbiol Lett* **278**: 231-235.
Zhang, Y., Masi, A.W., Barniak, V., Mountzouros, K., Hostetter, M.K., and Green, B.A. (2001) Recombinant PhpA protein, a unique histidine motif-containing protein from *Streptococcus pneumoniae*, protects mice against intranasal pneumococcal challenge. *Infect Immun* **69**: 3827-3836.
Zhu, B., Cai, G., Hall, E.O., and Freeman, G.J. (2007) In-fusion assembly: seamless engineering of multidomain fusion proteins, modular vectors, and mutations. *Biotechniques* **43**: 354-359.
Zysk, G., Bejo, L., Schneider-Wald, B.K., Nau, R., and Heinz, H. (2000) Induction of necrosis and apoptosis of neutrophil granulocytes by *Streptococcus pneumoniae*. *Clin Exp Immunol* **122**: 61-66.
Zysk, G., Schneider-Wald, B.K., Hwang, J.H., Bejo, L., Kim, K.S., Mitchell, T.J., Hakenbeck, R., and Heinz, H.P. (2001) Pneumolysin is the main inducer of cytotoxicity to brain microvascular endothelial cells caused by *Streptococcus pneumoniae*. *Infect Immun* **69**: 845-852.