



**CHARACTERISATION OF THE  
CAPSULAR POLYSACCHARIDE  
BIOSYNTHESIS LOCI OF  
*STREPTOCOCCUS PNEUMONIAE*  
SEROGROUP 19**

by

**Judy Kay Morona B. Sc. (Adelaide)**

A thesis submitted for the degree of Doctor of Philosophy

Department of Microbiology and Immunology  
University of Adelaide

November, 1998

# ABSTRACT

In this thesis, the genetic loci encoding capsular polysaccharide synthesis (*cps*) have been characterised for all members of *Streptococcus pneumoniae* serogroup (19F, 19A, 19B and 19C). In each serotype, the *cps* locus is located in the *S. pneumoniae* chromosome between *dexB* and *aliA* and appears to be arranged as a single transcriptional unit. The arrangement of the genes within the *cps19* loci is highly conserved with 13 genes (*cps19A-H*, and *K-O*) common to all four serogroup 19 members. These genes encode functions required for the synthesis of the shared trisaccharide component of the group 19 capsular polysaccharide (CPS) repeat unit structures. Furthermore, the genetic differences between the group 19 *cps* loci identified are consistent with the differences in the CPS structures of individual serotypes. Functions have been assigned to nearly all the *cps19* gene products, based on either gene complementation or similarity to other proteins with known functions. This has enabled biosynthetic pathways for production of all four group 19 CPSs to be proposed.

Nearly all of the common genes from types 19F, 19B and 19C are >95% identical to each other. However, closely related homologues of *cps19fl* and *J*, which encode the type 19F polysaccharide polymerase and repeat unit transporter, respectively, are not found in the type 19B and 19C *cps* loci. In type 19B and 19C this region of the *cps* locus (between *cps19bH* and *cps19bK*) contains five genes which encode a unique polysaccharide polymerase and repeat unit transporter, as well as two additional putative glycosyl transferases and a protein which may be involved in synthesis of an activated ribose precursor. Transformation studies indicated that these five genes encode all of the functions required to convert a type 19F pneumococcus to type 19B. The type 19C *cps*

locus differs from the 19B *cps* locus only in the insertion of a glucosyl transferase gene (*cps19cS*) between *cps19cK* and *cps19cL*. Transformation studies have shown that the presence of this gene accounts for the additional glucose side chain in the otherwise identical repeat unit structures. The type 19C *cps* locus contains 19 genes, and at 21 kb it is the largest pneumococcal capsule gene cluster characterised to date.

Although the *cps19a* and *cps19f* loci are identical in the number and arrangement of the genes present, the similarity between individual genes varies from 70% to 99% identity (for both the nucleotide and the deduced amino acid sequences). This sequence divergence is surprising given that the only difference between their CPS repeat units is the glycosidic linkage which joins the repeat units together ( $\alpha(1\rightarrow 2)$  for 19F and  $\alpha(1\rightarrow 3)$  for 19A). Theoretically, only a difference in the *cps19aI* gene, which presumably encodes the polysaccharide polymerase responsible for this linkage, is required to change a type 19F pneumococcus into type 19A. Indeed, this was demonstrated by a transformation event in which the region of the *cps19a* locus encoding Cps19aH and Cps19aI replaced the homologous portion of the *cps19f* locus was sufficient to convert CPS type from 19F to 19A. Given that Cps19fH and Cps19aH are >95% identical, it seems probable that Cps19aI (79% identity) is solely responsible for the observed alteration in CPS type.

The serotype specificity of the *cps19f* genes was investigated by Southern hybridisation analysis of chromosomal DNA from other *S. pneumoniae* serotypes. Large variations in the hybridisation patterns were obtained with the different gene-specific probes. Probes specific for sequences flanking *cps19f* hybridised with all the serotypes tested. However, within the *cps* loci, only *cps19fA* and *cps19fB* were common to all serotypes. Based on the Southern hybridisation analysis a protocol for PCR amplification of *cps* loci was developed and used to amplify the *cps* regions from a variety of pneumococcal serotypes. Direct sequencing of the 5' end of the PCR products was undertaken and identified two classes of *cpsC* gene. Southern hybridisation studies with

*cps19aC*- and *D*-specific gene probes demonstrated that homologues of the first four genes in the *cps* locus, *cpsA-D*, are present in all serotypes and that all the *cps* loci tested evolved from one of two clonal origins which contained either class I or class II *cpsC* and *D* genes. The *cpsE* gene, which encodes a glucose-1-phosphate transferase, is also conserved (in the two distinct classes) in the *cps* loci of all serotypes tested which contain glucose in their CPS, except type 3.

The sequence analysis of the various *cps* loci presented in this thesis provides further evidence that in nature frequent recombination occurs between different *cps* loci resulting in either complete exchange of the *cps* locus or exchange of only part of the *cps* locus, and could potentially result in the expression of new capsular serotypes.

# TABLE OF CONTENTS

<b>1. INTRODUCTION</b>	<b>1</b>
1.1 <i>Streptococcus pneumoniae</i> , a Distinguished History	1
1.1.1 Development of the Gram stain	1
1.1.2 Humoral immunity and phagocytosis	2
1.1.3 The concept of serotyping	2
1.1.4 The observation of antibiotic resistance in bacteria	2
1.1.5 Recognition of the first non-protein antigen	3
1.1.6 Bacterial transformation	3
1.1.7 Discovery of DNA as the genetic material	4
1.2 Pneumococcal Disease	5
1.2.1 Colonisation of the nasopharynx	5
1.2.2 Risk factors for pneumococcal disease	6
1.2.3 Incidence of pneumococcal disease	7
1.2.4 Pathogenesis of pneumococcal infections	8
1.2.5 Pneumococcal virulence factors	8
1.2.5.1 Capsule	8
1.2.5.2 Cell wall	9
1.2.5.3 Autolysin	9
1.2.5.4 Pneumolysin	9
1.2.5.5 Neuraminidase	10
1.2.5.6 Hyaluronidase	11
1.2.5.7 Pneumococcal surface protein A (PspA)	11
1.2.5.8 PsaA	11
1.2.5.9 Choline binding protein A (CbpA)	12
1.3 The Cell Surface of the Pneumococcus	12
1.3.1 The cell wall	13
1.3.2 Teichoic acid	14
1.3.3 The plasma membrane	15
1.3.4 Proteins associated with the cell surface	16
1.4 The Pneumococcal Capsule	16
1.4.1 Capsule serotyping	17
1.4.2 Chemical structure of the capsular polysaccharide repeat units	18
1.4.3 The chemical structure of group 19 CPS	20
1.4.4 The relationship between serotype and virulence	21

1.5	Distribution of Capsule Types	21
1.5.1	Capsule types associated with paediatric pneumococcal disease	22
1.5.2	Capsule types associated with invasive pneumococcal disease in adults	24
1.5.3	Geographic distribution of invasive pneumococci	24
1.5.4	Changes in serotype distribution over time	25
1.6	Antimicrobial Resistance in the Pneumococcus	26
1.6.1	Geographic distribution of antimicrobial resistance	26
1.6.2	The spread of antibiotic resistance	27
1.6.3	Antibiotic resistance among different pneumococcal serotypes	27
1.7	The Pneumococcal Polysaccharide Vaccine	28
1.7.1	Antibody response	28
1.7.2	Conjugate polysaccharide-protein vaccines	29
1.8	Polysaccharide Biosynthesis	30
1.8.1	Some common features in polysaccharide biosynthesis	31
1.8.2	Polysaccharide biosynthesis via lipid-linked repeat unit intermediates	32
1.8.3	Polysaccharide biosynthesis involving ABC transporters	34
1.8.4	Polysaccharide synthesis via processive glycosyl transferases	34
1.8.5	Polysaccharide synthesis not involving lipid-linked intermediates	35
1.8.5	Capsule gene clusters	36
1.9	The Capsule Locus of <i>S. pneumoniae</i>	37
1.9.1	Pneumococcal capsular transformation	37
1.9.2	Localisation of the type 19F capsule locus	38
1.9.3	Partial cloning and sequencing of the type 19F capsule locus	39
1.9.4	Characterisation of <i>cps19fA-G</i>	40
1.9.4.1	Cps19fA	40
1.9.4.2	Cps19fB	41
1.9.4.3	Cps19fC	41
1.9.4.4	Cps19fD	41
1.9.4.5	Cps19fE	42
1.9.4.6	Cps19fF	43
1.9.4.7	Cps19fG	43
1.10	Aims of this Thesis	43
<b>2.</b>	<b>MATERIALS AND METHODS</b>	<b>45</b>
2.1	Bacterial strains and cloning vectors	45
2.1.1	Bacterial strains	45
2.1.2	Cloning vectors	46
2.1.3	Growth media	47

2.2	Chemicals and reagents	48
2.2.1	General chemicals	48
2.2.2	Antibiotics	48
2.2.3	Enzymes	48
2.2.4	Oligodeoxynucleotides	49
2.3	Serotyping of pneumococcal strains	51
2.4	Bacterial transformation	51
2.4.1	Preparation of competent <i>E. coli</i> strains	51
2.4.1.1	RbCl <sub>2</sub> method	51
2.4.1.2	Electroporation method	52
2.4.1.3	CaCl <sub>2</sub> method	52
2.4.2	Transformation of competent <i>E. coli</i> strains	53
2.4.3	Electroporation of <i>E. coli</i> strains	53
2.4.4	Transformation of pneumococcal strains	54
2.4.5	Preparation of competence factor for pneumococcal transformation	54
2.5	DNA extraction procedures	55
2.5.1	Plasmid isolation	55
2.5.2	Plasmid isolation for pneumococcal transformation	56
2.5.3	Preparation of pneumococcal chromosomal DNA	57
2.6	Analysis and manipulation of DNA	58
2.6.1	Restriction endonuclease digestion of DNA	58
2.6.2	Calculation of sizes of restriction fragments	58
2.6.3	Analytical and preparative separation of restriction fragments	59
2.6.4	Isolation of restriction fragments from agarose gels	59
2.6.5	<i>In vitro</i> cloning	59
2.7	Southern hybridisation	60
2.7.1	Labelling of DNA probes	60
2.7.2	DNA hybridisation by Southern blotting	61
2.7.2.1	Transfer of DNA to nylon membrane	61
2.7.2.2	High stringency hybridisation	61
2.7.2.3	Low stringency hybridisation	62
2.7.2.4	Immunological detection	62
2.8	DNA sequencing and analysis	62
2.8.1	Nested deletions	62
2.8.2	DNA sequencing using dye-labelled primers	64
2.8.3	Sequencing with dye-labelled terminators	65
2.8.4	DNA sequencing	65
2.8.5	DNA sequence analysis	65

2.9 PCR Amplification	66
2.9.1 PCR amplification using Taq polymerase	66
2.9.2 Long range PCR	66
2.9.3 Inverse PCR	67
2.9.4 Purification of PCR products	67
2.10 Insertion-duplication mutagenesis	68
2.11 Plasmid insertion/rescue	69
2.12 Protein and LPS analysis	70
2.12.1 T7 expression	70
2.12.2 LPS preparation	70
2.12.3 SDS-PAGE	71
2.12.4 Western blotting	71
<b>3. CHARACTERISATION OF THE COMPLETE TYPE 19F CAPSULE LOCUS</b>	<b>73</b>
3.1 Introduction	73
3.2 Results	73
3.2.1 Isolation, cloning and DNA sequencing of the type 19F <i>cps</i> locus	73
3.2.2 Analysis of the <i>cps19f</i> DNA sequence	76
3.2.3 Characterisation of <i>cps19fG-O</i>	77
3.2.3.1 <i>cps19fG</i>	77
3.2.3.2 <i>cps19fH</i>	79
3.2.3.3 <i>cps19fI</i>	80
3.2.3.4 <i>cps19fJ</i>	83
3.2.3.5 <i>cps19fK</i>	85
3.2.3.6 <i>cps19fLMNO</i>	87
3.2.4 Insertion-duplication mutagenesis of <i>cps19fG-O</i> genes	96
3.2.5 T7 expression of several Cps19f proteins	98
3.2.6 Complementation of an <i>E. coli rffE</i> mutant with Cps19fK	100
3.2.7 Complementation of <i>S. flexneri rfbBDAC</i> with <i>cps19fLMNO</i>	100
3.2.8 Putative biosynthetic pathway for <i>S. pneumoniae</i> type 19F CPS	102
3.3 Conclusions	104
<b>4. ANALYSIS OF CAPSULE LOCI FROM VARIOUS PNEUMOCOCCAL SEROTYPES</b>	<b>106</b>
4.1 Introduction	106



4.2 Results	107
4.2.1 Serotype specificity of the <i>cps19f</i> genes and flanking DNA sequences	107
4.2.1.1 Presence of <i>cps19f</i> homologues in members of serogroup 19	107
4.2.1.2 Presence of <i>cps19f</i> homologues in other serotypes	109
4.2.1.3 Analysis of the intergenic regions	110
4.2.2 Amplification of capsule loci by LR-PCR	111
4.2.3 Southern hybridisation analysis of LR-PCR products	113
4.2.4 Partial DNA sequencing of the LR-PCR products	116
4.2.5 Phylogenetic analysis of <i>cpsC</i> sequences	120
4.3 Conclusions	120
<b>5. CHARACTERISATION OF THE <i>S. PNEUMONIAE</i> TYPE 19B CAPSULE LOCUS</b>	<b>124</b>
5.1 Introduction	124
5.2 Results	125
5.2.1 Isolation of the type 19B specific <i>cps</i> genes	125
5.2.2 Characterisation of the <i>cps19b</i> genes	128
5.2.2.1 <i>cps19bP</i>	129
5.2.2.2 <i>cps19bI</i>	129
5.2.2.3 <i>cps19bQ</i>	129
5.2.2.4 <i>cps19bR</i>	132
5.2.2.5 <i>cps19bJ</i>	133
5.2.3 Serotype specificity of the <i>cps19b</i> genes	134
5.2.4 Capsule type switching by transformation	135
5.3 Conclusions	136
<b>6. MOLECULAR BASIS FOR PNEUMOCOCCUS TYPE 19A</b>	<b>139</b>
6.1 Introduction	139
6.2 Results	140
6.2.1 PCR amplification and sequencing of the type 19A locus	140
6.2.2 Analysis of the <i>cps19a</i> locus	141
6.2.3 Comparison of the <i>cps19a</i> locus from <i>S. pneumoniae</i> strains 19A1 and 19A2	148
6.2.4 Serotype specificity of the <i>cps19a</i> genes	149
6.2.5 Capsule transformation from 19F to 19A	153
6.2.6 Sequence variation in the 5' intergenic region of serogroup 19	155
6.3 Conclusions	159

<b>7. CHARACTERISATION OF THE <i>S. PNEUMONIAE</i> TYPE 19C SPECIFIC CPS REGION</b>	<b>163</b>
7.1 Introduction	163
7.2 Results	164
7.2.1 Isolation of the type 19C specific <i>cps</i> region	164
7.2.2 Characterisation of Cps19cS	168
7.2.3 Serotype specificity of <i>cps19cS</i>	171
7.2.4 Transformation of <i>S. pneumoniae</i> type 19B to type 19C	171
7.2.5 Characterisation of IS19C, located in the 5' intergenic region	172
7.3 Conclusions	174
<b>8. DISCUSSION</b>	<b>176</b>
8.1 Introduction	176
8.2 Analysis of the <i>S. pneumoniae</i> group 19 <i>cps</i> loci	176
8.2.1 Comparison of the group 19 <i>cps</i> loci	176
8.2.2 Distribution of the <i>cps19</i> genes in other pneumococcal serotypes	181
8.2.3 G+C content	181
8.2.4 Transcription and translation of the <i>cps19</i> loci	182
8.3 The <i>cps</i> loci from other <i>S. pneumoniae</i> serotypes	184
8.3.1 The <i>S. pneumoniae</i> type 1 <i>cps</i> locus	184
8.3.2 The <i>S. pneumoniae</i> type 2 <i>cps</i> locus	185
8.3.3 The <i>S. pneumoniae</i> type 3 <i>cps</i> locus	187
8.3.4 The <i>S. pneumoniae</i> type 4 <i>cps</i> locus	187
8.3.5 The <i>S. pneumoniae</i> type 14 <i>cps</i> locus	188
8.3.6 The <i>S. pneumoniae</i> type 23F <i>cps</i> locus	189
8.3.7 The <i>S. pneumoniae</i> type 33F <i>cps</i> locus	190
8.4 Analysis of the <i>cps</i> loci from <i>S. pneumoniae</i>	191
8.4.1 Organisation of pneumococcal <i>cps</i> loci	191
8.4.2 Comparison of the <i>cps</i> genes and their protein products	194
8.4.3 Capsular transformation <i>in vivo</i>	198
8.4.4 The presence of IS elements in the 5' and 3' intergenic regions	200
8.4.5 Transcription of the pneumococcal <i>cps</i> locus	200
8.4.6 Regulation of CPS production in <i>S. pneumoniae</i>	202
8.4.7 The function of CpsC and CpsD	205
8.4.8 Relationship of the <i>S. pneumoniae cps</i> locus to other Gram-positive <i>cps</i> loci	206
8.5 Future studies	209