Molecular Studies of the *Escherichia coli* K5 Capsule Gene Cluster

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Frederick I Esumeh, BSc (BENSU), MSc (Lagos)

Department of Microbiology and Immunology

University of Leicester

December 1996
Acknowledgements

I am grateful to Professor Ian Roberts for his supervision of this work. I sincerely thank all the members of the Department of Microbiology and Immunology who have given me advice and assistance in the course of this work. In particular I have valued the help and support of the members of Lab 228, among them Mark Stevens, Julie Eastgate, Rowan Pearce, Chantal Petit, Anna Ostrowski, Brad Clarke, Brendan Barrette, Archie Lovatt, Mark Anderson and Gary Griffiths. I am extremely grateful to Gina Manning, Aras Kadioglu, Steve Gordon, Ben Gale, Chris Jones, Gerald Owen son, Hasan Yesilkaya, Ala Ahmed, Austin Asemota, Humphery Ehigie, Ben Emenyeonu and Pedro Da Rocha for their unforgettable friendship. I am indebted to Gordon Rigg for making me knowledgeable in computer analysis of sequence data. My special thanks to Andrew Wallace for his ceaseless words of encouragement. Of the many friends I have made during the course of this work I would like to specially thank Dave Simpson with whom I have shared the uphill struggle of completing a PhD thesis and for being such a golden friend. I am grateful to the Commonwealth Scholarship Commission in the UK for the scholarship awarded to me. This thesis is dedicated to my family for their tireless support and encouragement.
Molecular studies of the *Escherichia coli* K5 capsule gene cluster.

Frederick I Esumeh

Group II K antigens such as the K5 are associated with *Escherichia coli* causing serious extraintestinal infections. Genes for the production of group II capsules (*kps*) are organised into three functional regions 1, 2 and 3. The central region 2 encodes proteins involved in the biosynthesis of type specific polysaccharide. Proteins encoded by the flanking regions 1 and 3 are involved in the export of polysaccharide across the bacterial membranes unto the cell surface. Translocation of polysaccharide across the inner membrane is mediated by two region 3 encoded proteins, KpsM and KpsT which belong to the ATP-binding cassette (ABC) superfamily of transporters. Region 1 encodes six proteins, KpsF-E-D-U-C-S. The KpsE protein was shown to localise in the inner membrane and mutants lacking the encoded protein were unable to export polysaccharide to the cell surface. To understand the export role of the KpsE, its membrane topology was determined using TnphoA mutagenesis and β-lactamase fusions. The topology was confirmed by accessibility of fusion proteins to proteinase K. These results indicated that the KpsE protein adopts a type II bitopic membrane topology consisting of an N-terminus in the cytoplasm followed by a membrane spanning domain, a large periplasmic segment and a C-terminus that may be associated with the outer membrane. On the basis of structural similarity, the KpsE protein is proposed as the membrane fusion protein (MFP) component in the KpsM and T-mediated export of group II capsular polysaccharide and it is postulated to function in the late stages of export across the periplasmic space unto the cell surface. The determined membrane topology contradicts a second C-terminal transmembrane domain in the proposed model based on secondary structure predictions. This points to the importance of experimental data in the establishment of membrane topological models. In addition, the cloning and analysis of a bacteriophage-borne lyase enzyme specific for the *E. coli* K5 capsular polysaccharide is described. The activity of the K5 lyase enzyme was demonstrated *in vitro* and its applications are discussed.
Contents

Chapter 1. Introduction ................................................................. 1

1.1 Bacterial Capsules ................................................................. 1

1.2 Functions of Bacterial Capsules ............................................ 5

1.2.1 Prevention of Desiccation .................................................... 5

1.2.2 Adherence to Surfaces ....................................................... 5

1.2.3 Virulence ............................................................................. 6

1.2.3.1 Resistance to Nonspecific Host Immunity ......................... 7

1.2.3.2 Resistance to Specific Host Immunity ............................... 8

1.2.4 Immunogenicity ................................................................. 9

1.3 Escherichia coli Capsular (K) Antigens ................................. 10

1.3.1 Classification of E. coli Capsular Antigens ......................... 10

1.3.1.1 Group I Capsules .......................................................... 10

1.3.1.2 Group II Capsules ........................................................ 12

1.3.1.3 Group III Capsules ....................................................... 13

1.4 E. coli Group II Capsule Gene Clusters ................................. 15

1.4.1 The E. coli K5 Capsule Gene Cluster ................................. 16

1.5 Genetic Organisation of H. influenzae and N. meningitidis

Capssule Gene Clusters ............................................................... 23

1.6 Polysaccharide and Protein Export in Gram-negative Bacteria .. 27

1.6.1 Protein Export ................................................................. 27

1.6.1.1 Sec-dependent Translocation ........................................... 27
2.3.2 Extraction of Chromosomal DNA ......................................................... 58
2.3.3 Techniques Used in Routine DNA Manipulation ............................ 58
2.3.4 Polymerase Chain Reaction ................................................................. 60
2.3.5 Cloning of Truncated Lengths of PCR-Amplified kpsE DNA Products............................................................................................................ 60
2.4 DNA Hybridisation Procedures .............................................................. 61
2.4.1 Transfer of DNA from Agarose Gels (Southern Blotting) .......... 61
2.4.2 Radiolabelling of DNA for Use as Probe ......................................... 61
2.4.3 Hybridisation and Detection of DNA Immobilised on Filters with Probe................................................................. 62
2.5 DNA Sequencing ..................................................................................... 63
2.5.1 DNA Sequencing Polyacrylamide Gel Electrophoresis ................ 63
2.5.2 Automated DNA Sequencing ............................................................... 65
2.6 Computer Analysis of DNA and Protein Sequences ....................... 65
2.7 Procedures for Protein Analysis .............................................................. 66
2.7.1 Expression of Cloned Genes in E. coli ............................................. 66
2.7.2 In vivo Radiolabelling and Detection of Protein ............................... 66
2.7.3 Detection of Protein Antigen by Western Blotting ......................... 68
2.8 Subcellular Fractionation and Proteolysis with Proteinase K ........ 69
2.9 Assay for Lyase Enzyme Activity ......................................................... 69

Chapter 3 Topology of KpsE, a Protein Involved in the

Export of Polysaccharide in Escherichia coli K5 .................................. 71
3.1 Introduction .......................................................................................... 71
3.2 Results..............................................................................................................73

3.2.1 Prediction of Membrane Topology of the KpsE Protein.............73

3.2.2 Analysis of Membrane Topology of the KpsE Protein

Using Gene Fusion.............................................................................................74

3.2.2.1 TnphoA Mutagenesis.................................................................74

3.2.2.2 Generation of β-lactamase fusion...............................................77

3.2.2.2.1 Ampicillin Resistance of E. coli JM101 Expressing KpsE-BlaM Hybrid Proteins...............................................80

3.2.2.2.2 Visualisation of KpsE-BlaM Fusion Proteins by

Western Blot Analysis......................................................................................82

3.2.2.2.3 Subcellular Fractionation and Proteolysis with

Proteinase K..................................................................................................82

3.2.3 Analysis of the KpsE and Other Proteins Involved in

Bacterial ABC Transport..............................................................................83

3.2.4 Construction of kpsE Null Mutation within the

Capsule Gene Cluster..................................................................................90

3.2.4.1 Phenotype of MSFE100.................................................................98

3.2.4.2 Complementation of kpsE Mutation in MSFE100......................98

3.3 Discussion....................................................................................................99

Chapter 4 Cloning and Sequencing of the K5 Lyase Enzyme from

the K5-Specific Bacteriophage......................................................................103

4.1 Introduction...............................................................................................103

4.2 Results.......................................................................................................103
4.2.1 Construction of a Bacteriophage K5 Genomic DNA Library......103
4.2.2 Screening of Gene Library for Lyase Activity..............................106
4.2.3 Nucleotide Sequence of K5 Lyase Clone, pB4.........................106
4.2.4 Analysis of Deduced Amino Acid Sequence.............................107
4.2.5 T7 Expression of Proteins Encoded by the Lyase Clone..........116
4.2.6 Induced Expression of Lyase Activity........................................116
4.3 Discussion.......................................................................................119

Chapter 5 Discussion ........................................................................123

References..........................................................................................134
List of Abbreviations

Amp  ampicillin
Amp^r  ampicillin resistance
AP  alternative pathway of complement
ATP  adenosine triphosphate
APS  ammonium persulphate
blaM  β-lactamase
bp  base pair
CAT  chloramphenicol acetyltransferase
Cm  chloramphenicol
EDTA  ethylenediamine-tetra acetic acid
GlcA  glucuronic acid
GlcNAc  N-acetylglucosamine
hr  hour
IPTG  isopropylthiogalactoside
kb  kilobase
kDa  kilodalton
KDO  2-keto-3-deoxymanno-octonic acid
Km  kanamycin
LPS  lipopolysaccharide
ManNAc  N-acetylmannosamine
min  minute
NeuNAc  N-acetyl neuraminic acid
orf  open reading frame
PA  phosphatidic acid
PAGE  polyacrylamide gel electrophoresis
PBS  phosphate buffer saline
PCR  polymerase chain reaction
SD  Shine Dalgarno
SDS  sodium dodecyl sulphate
SSC  saline sodium citrate
TEMED  N,N,N',N'-tetramethylethylenediamine
X-Gal  5-bromo-4-chloro-3-indolyl galactopyranoside
XP  5-bromo-4-chloro-3-indolyl phosphate
Chapter 1

Introduction

The production of an extracellular layer of polysaccharide termed capsule is a common feature of many bacteria. Encapsulated bacteria of medical importance are numerous and have been associated with serious invasive infections such as septicaemia, meningitis, pneumonia and urinary tract infections. The capsule as a virulence determinant has been demonstrated in experimental models of infection in which the removal of capsule resulted in loss or reduced virulence of such bacteria. This acidic surface polymer confers on bacteria the ability to evade the various facets of the host defence, including complement-mediated bacteriolysis, uptake and killing by phagocytes. This protective function of the capsule can however be overcome with capsule-specific antibody. It is due therefore to their important role in disease and as likely vaccine candidates that capsules are being studied. In fact the enormous diversity of structure as well as the nature by which these acidic polymers are synthesised and translocated across the bacterial membranes to the cell surface present interesting problems to researchers in the field. This thesis comprises a structure-function study of the KpsE protein and an analysis of a cloned Escherichia coli K5 capsule-specific bacteriophage-borne lyase enzyme.

1.1 Bacterial Capsules

The major surface antigens of many Gram-negative bacteria of diverse genera are the lipopolysaccharide (LPS) and the capsular polysaccharide. The LPS or O antigens consist of lipid A covalently linked to the O-specific polysaccharide (Jann and Jann, 1987, 1990). Distinct from the O antigens are the capsular polysaccharides which are present in an extracellular envelope which may cover the O antigen. These additional antigens on the surface of bacterial outer membranes can be secreted into the surrounding medium as extracellular slime or may be organised into distinct structures termed capsules (Ørskov et al., 1977, 1984; Ørskov and Ørskov, 1992).

Slime polysaccharides are usually released into the growth medium probably due to lack of anchoring structures to the bacterial outer membrane (Ørskov...
et al., 1984; Whitfield and Valvano, 1993) and are easily extracted by washing. Some of the well known examples of this type of polysaccharide are the alginites produced by *Pseudomonas*, the dextrans of *Agrobacterium* and xanthan gum produced by *Xanthomonas campestris* (Sutherland, 1985). On the other hand, capsular polysaccharides exist as an hydrated gel surrounding each bacterial cell and requires a more vigorous extraction procedure by virtue of their association to the bacterial outer membrane. However, it has been noted that encapsulated bacteria do release polysaccharide into the growth medium (Hancock and Cox, 1991). Troy et al. (1971) reported 20% of the capsular polysaccharide of *Aerobacter aerogenes* are easily extracted by washing. The loose binding of some capsular polysaccharides to the bacterial wall is also exemplified by a precipitin halo surrounding a colony grown on serum agar (Petrie, 1932; Kaijser, 1977). Thus, the distinction between slime and capsule is somewhat arbitrary and may be of no functional significance (Boulnois and Roberts, 1990; Ørskov and Ørskov, 1990).

The concept of a bacterial capsule was derived largely from earlier light microscopic studies, where bacterial capsules were identified by negative staining with Indian ink as an exclusion zone around individual cells (Ørskov and Ørskov, 1990). The interpretation of light microscopic results is complicated and ambiguous owing to the distorting effect to capsular polysaccharide structures whenever fixation, dehydration or staining procedures are employed (Kellenberger, 1987; Bayer, 1990). This distortion effect is attributed to the high water content of capsules, which is in excess of 95% (Sutherland, 1972). However, the presence of capsule can now be established in a number of ways including serology (Ørskov and Ørskov, 1990, 1992), the use of DNA probes (Roberts et al., 1986), specific bacteriophages (Gross et al., 1977; Gupta et al., 1985) and improved light and electron microscopy (Bayer, 1990). The diameter of capsule varies considerably ranging from 1μm to 10μm in less and highly expressed conditions respectively (Sutherland, 1977; Bayer, 1990). Capsules surround the organism totally and in a more or less even layer (Bayer, 1990). Capsules are maintained on the cell surface of Gram-negative bacteria through attachment to phospholipid (Gostschlich et al., 1981; Schmidt and Jann, 1982) lipid A (Jann and Jann, 1990) or in the case of Gram-positive bacteria through linkage to peptidoglycan (Yeung and Mattingly, 1986; Sorensen et al., 1990). In addition, ionic interactions may help the cell surface retention (Jann and Jann, 1990).
Capsules are acidic polymers composed of homo- or heteropolysaccharides which can be linear or branched. The repeating mono- or oligosaccharide units of the polysaccharide can be substituted with short side chains. Hence changes in substitution and component monosaccharide residues as well as the nature by which these components are linked together bring about great diversity in capsular polysaccharide. As a result, a wide variety of bacterial polysaccharides have been identified (Kenne and Lindberg, 1983). Some of these polymer types are listed in table 1.1. The components include pentoses, hexoses, heptoses, amino sugars, methylated sugars and uronic acids, and in addition, non-sugar substitutions such as phosphate, formate, ribitol, succinate, pyruvate and acetyl groups (Sutherland, 1977; Kenne and Lindberg, 1983). Also, monomers that were once thought to be unique to particular cell structures have been found in several polysaccharides. For instance, 2-keto-3-deoxymannoctonic acid (KDO) widely considered to be solely associated with LPS is in fact a distinctive feature of several E. coli capsules (Jann and Jann, 1983). The sugar N-acetylmuraminic acid (NeuNAc) that occurs mainly in glycolipids and glycoproteins is a prominent feature of E. coli K1 and Neisseria meningitidis group B capsular polysaccharides (Kasper et al., 1973). The occurrence of amino acids as major constituents of polysaccharides, such as threonine in the K54 antigen of E. coli (Hofmann et al., 1985) and lysine in Proteus mirabilis O polysaccharide (Gromska and Mayer, 1976) is equally uncommon.

Capsular polysaccharide types do result also from differences in linkage of the repeating units. For example, the polysialic acid of E. coli K1 (Dewitt and Rowe, 1961; McGuire and Binckley, 1964), N. meningitidis group B (Bhattacharjee et al., 1975; Ørskov et al., 1979) and Pasturella haemolytica A2 (Adlam et al., 1987) have α(2-8) linkages, whereas α(2-9) linkages are seen in the polysialic acid of N. meningitidis group C (Bhattacharjee et al., 1975). On the other hand, identical polysaccharides can be expressed by different bacterial species. The homopolymer of N-acetylneuraminic acid is found in both E. coli K1 and N. meningitidis group B capsules (Kasper et al., 1973) and the enterobacterial common antigen (ECA) is found in all enteric bacteria except Erwinia chrysanthemi (Kuhn et al., 1988). The presence of identical polysaccharides in diverse bacteria does raise interesting questions concerning the evolution of polysaccharide types.
<table>
<thead>
<tr>
<th>Organism</th>
<th>Antigen/Group</th>
<th>Repeating unit</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>K30</td>
<td>-2)-Man-(1,3)-Gal-(1&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Chakraborty et al. (1980)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>β-GlcA-(1,3)-Gal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K40</td>
<td>4)-β-GlcA-(1,4)-α-GlcNac-(1,6)-α-GlcNac-(1&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Dengler et al. (1986)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CO.NH (serine)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>K1 -8)-α-NeuNac-(2-</td>
<td>McGuire and Binckley (1964)</td>
<td></td>
</tr>
<tr>
<td>K5</td>
<td>-4)-β-GlcA-(1,4)-α-GlcNac-(1-</td>
<td>Vann et al. (1981)</td>
<td></td>
</tr>
<tr>
<td>K10</td>
<td>-3)-α-Rha-(1,3)-β-Qu4NMal-(1-</td>
<td>Sieberth et al. (1993)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OAc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K54</td>
<td>-3)-β-GlcA-(1,3)-α-Rha-(1-</td>
<td>Hofmann et al. (1985)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CO.NH threonine (serine)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-6)-α-ManNac-(1-PO4</td>
<td>Bundle et al. (1974)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OAc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N. B meningitidis</td>
<td>-8)-α-NeuNac-(2-</td>
<td>Bhattacharjee et al. (1975)</td>
<td></td>
</tr>
<tr>
<td>W 135</td>
<td>-6)-α-Gal-(1,4)-α-NeuNac-(2-</td>
<td>Bhattacharjee et al. (1976)</td>
<td></td>
</tr>
<tr>
<td>H. b Influenzas</td>
<td>-3)-β-Ribf-(1,1)-Ribitol-(5-PO4</td>
<td>Crisel et al. (1975)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: Gal, galactose; GlcA, glucuronic acid; GlcNac, N-acetylglucosamine; Man, mannose; ManNac, N-acetylmannosamine; NeuNac, N-acetylneuraminic acid; OAc, O-acetyl group; PO4, phosphate group; Qu4NMal, 4,6-dideoxy-4-malonylaminoglucone; Rha, rhamnose; Ribf, ribofuranose.
Such is the structural diversity and relatedness of capsular polysaccharides among diverse bacterial genera that the focus will henceforth be on those of medical importance. Among this group are the capsules of *Klebsiella* species, *Streptococcus pneumoniae*, *Haemophilus influenzae*, *N. meningitidis* and *E. coli*. These organisms have been associated with serious invasive infections such as sepsis, meningitis and urinary tract infections (Peltola, 1983; Jennings, 1990).

### 1.2 Functions of Bacterial Capsules

Capsules regardless of composition have been shown to affect dehydration, ionic exchange between bacteria and environment, adherence and colonisation of inert and living surfaces, infection by bacteriophages and the overall virulence of the bacterium (Costerton *et al.*, 1981, 1987).

#### 1.2.1 Prevention of Desiccation

The capsule forms a hydrated gel at the cell surface that may prevent desiccation and thus increase the survival of the bacteria (Costerton *et al.*, 1987). This protective effect was recently amply demonstrated with the exopolysaccharide of *Pseudomonas* species (Roberson and Firestone, 1992) and the colanic acid of *E. coli* K12 (Ophir and Gutnick, 1994). Mucoid strains of *E. coli*, *Acinetobacter calcoaceticus* and *Erwinia stewartii* have been shown to be more resistant to drying than the isogenic non-mucoid strains (Ophir and Gutnick, 1994). In *E. coli*, the expression of genes encoding enzymes for the biosynthesis of colanic acid is increased by desiccation (Ophir and Gutnick, 1994). This apparent regulation of capsule expression by desiccation is not fully understood. A possible explanation is that desiccation increases the external osmolarity which triggers increase in capsule biosynthesis. Increased expression of the Vi polysaccharide in *Salmonella typhi* (Pickard *et al.*, 1994) and alginate biosynthesis in *Pseudomonas aeruginosa* (Berry *et al.*, 1989) are known to be induced by high external osmolarity.

#### 1.2.2 Adherence to Surfaces

Capsular polysaccharides may promote adherence and colonisation of inert and living surfaces and thereby facilitate the formation of biofilms (Costerton *et al.*, 1981, 1987). This is attributed to electrostatic attraction between the polysaccharide and surfaces and the intergeneric interactions.
among the microbial population that make the biofilm (Jenkinson, 1994). Electrostatic attraction may particularly benefit bacteria that inhabit oligotrophic environments since nutrients are also attracted to the cell surfaces (Costerton et al., 1981). Microbial biofilms may provide the individual bacteria protection from predation by phagocytic protozoa and slime moulds and from infection by some bacteriophages (Dudman, 1977).

The ability of bacteria to colonise surfaces and form biofilms is important in the development of dental caries and has other far-reaching consequences. These include the colonisation of indwelling catheters which may lead to serious nosocomial infection in hospitalised patients and in the fouling of pipes in industrial processes (Costerton et al., 1987). The formation of biofilm due to overexpression of alginate by Pseudomonas species within the lungs of cystic fibrosis patients is believed to present an impediment to antibiotics during treatment (Govan, 1988). Whilst the majority of capsular polysaccharides may promote adherence to surfaces, the cell surface polysaccharide of P. mirabilis facilitates the swarming of cells over solid surfaces by reducing friction (Gygi et al., 1995).

Capsular polysaccharides may also be involved in bacterial-plant symbiotic interactions. The formation of nitrogen-fixing root nodules on leguminous plants by Rhizobium species involves bacterial polysaccharides (Gray and Rolfe, 1990; Noel, 1992). The major component of capsular polysaccharide of R. meliloti is succinoglycan and mutant strains of R. meliloti that do not synthesise succinoglycan do induce nodulation in alfalfa, but do not penetrate or colonise the nodule, which shows that succinoglycan is required for nodule invasion and development (Noel, 1992).

1.2.3 Virulence

There is considerable evidence that the capsular polysaccharides of both Gram-negative and Gram-positive bacteria play a vital role in virulence by conferring serum resistance and inhibiting phagocytosis (Horwitz and Silverstein, 1980; Kim et al., 1986). The production of alginate by strains of P. aeruginosa causing lung infections is believed to be the major cause of morbidity and mortality among cystic fibrosis patients (Govan, 1988).
1.2.3.1 Resistance to Nonspecific Host Immunity

Capsular polysaccharides often confer resistance to non-specific host immunity through several mechanisms (Cross, 1990; Moxon and Kroll, 1990). During bacteria infection, complement can be activated by either the classical or alternative pathway and in both cases the final product is the activation of the membrane attack complex (MAC), which causes lysis of Gram-negative bacteria by piercing the outer membrane (Joiner, 1985; Joiner et al., 1984). The classical pathway is activated by binding of cell surface components to antibodies, while the alternative pathway (AP) is important in the pre-immune state of infection when specific antibodies are absent. The AP is initiated by the non-specific binding of the serum protein C3b to the bacterial cell surface. The bound C3b is then activated by interacting with factor B to form the C3 convertase, C3bBb. This results in the deposition of more C3b and the formation of MAC (Frank et al., 1987).

Bacterial capsules may resist complement-mediated killing by masking underlying cell surface structures that are likely to activate the alternative pathway (Horwitz and Silverstein, 1980; Loos, 1985; Cross, 1990).

Alternatively, bound C3b can be inactivated by factor I to form iC3b, with factor H acting as a cofactor. Some capsular polysaccharides especially those containing sialic acid, including E. coli K1 and K9 contribute to serum resistance by acting as poor activators of the alternative pathway (Joiner et al., 1984; Michalek et al., 1988). This is achieved by the ability of the sialic acid containing polysaccharides to bind factor H, which breaks the amplification loop of the AP and thereby prevents the formation of MAC (Michalek et al., 1988). E. coli causing serious invasive infections exhibit particular combinations of O and K antigens and it is thought that these polysaccharides act in concert to confer resistance to complement-mediated killing (Kim et al., 1986). In addition, the capsular polysaccharides of E. coli K1 and K92, as well as the types b and c capsular polysaccharides of N. meningitidis exert their serum resistance by exhibiting low affinity for factor B which inhibits the formation of C3 convertase (Stevens et al., 1978; Joiner et al., 1984).

The bacterial capsule may confer protection against complement-mediated opsonophagocytosis. Phagocytes generally have receptors for antibody and complement on their surfaces, and the binding of antibody or complement to the bacterial surface (opsonisation) does promote phagocytosis (Joiner, 1985). Capsule is thought to confer resistance to complement-mediated phagocytosis through steric hindrance, whereby the capsule masks the
underlying C3b deposited on the cell surface from C3b receptors on the phagocyte. The K1 antigen is a potent inhibitor of phagocytosis and this is largely attributed to its inability to activate complement and to bind C3b efficiently (Allen et al., 1987; Michalek et al., 1988). In contrast, the K5 antigen does not appear to confer protection against phagocytosis (Cross et al., 1986, Cross, 1990), despite the frequent isolation of the pathogenic *E. coli* K5 strain from cases of neonatal meningitis, septicaemia, pyelonephritis and urinary tract infections (Orskov and Orskov, 1992). The K5 serotypes involved in these infections are often associated with limited cell surface structures (O75:K5:H5, O6:K5:H1 and O18ac:K5:H1), and their invasive attributes may reflect the function of other surface antigens, including the O antigen. The net negative charge and hydrophobicity conferred on the cell surface by capsular polysaccharide may also serve to inhibit interaction of phagocytes with the bacteria (Allen et al., 1987; Roberts et al., 1989; Cross, 1990). Highly charged K antigens tend to confer greater resistance to phagocytosis (Moxon and Kroll, 1990).

In certain bacteria, the expression of cell surface polysaccharide may be switched on or off from one generation to the next, a term referred to as phase variation (Kimura and Hansen, 1986). This has been observed with the lipopolysaccharide of *H. influenzae* type b and has been suggested as an important mechanism for the bacteria to evade immune responses (Kimura and Hansen, 1986). Phase variation has not been observed with the O and K antigens of *E. coli* (Whitfield et al., 1984), although O-acetylation variants of the K1 polysaccharide have been observed (Orskov et al., 1979). The basis and role of the K1 variation is not known.

### 1.2.3.2 Resistance to Specific Host Immunity

In contrast to most polysaccharide capsules which can elicit immune responses with the development of specific antibodies (Lee, 1987; Jennings, 1990), certain capsular polysaccharides are poorly immunogenic. This is largely attributed to their similarity to particular mammalian structures (Jann et al., 1992). For instance, the poly α-2,8-sialic acid found in *E. coli* K1 and group B meningococcal polysaccharides is a structure similar to the neural cell adhesion molecule, n-CAM in mammalian cells (Finne et al., 1983); and the heteropolymer of β-1,4-linked glucoronic acid and N-acetylgalcosamine of *E. coli* K5 is identical to the first polymeric intermediate of mammalian heparin (Vann et al., 1983). Such structural mimicry by capsular polysaccharides to glycosaminoglycans found on host tissues (Finne et al., 1983; Lindahl et al., 1994) means that infected
individuals are unable to mount a specific immune response to bacterial
strains carrying these capsules.

1.2.4 Immunogenicity

Most capsular polysaccharides are immunogenic and may be used for the
generation of vaccines (Lee, 1987; Jennings, 1990). The potential of capsular
polysaccharides as vaccines has been demonstrated with a number of life-
treating diseases. The vaccine against meningitis caused by *N. meningitidis* consists of the purified polysaccharides from serogroups A, C, W135 and Y, which constitute 90% of the serogroups encountered in all
infections (Cadoz *et al*., 1985). In *H. influenzae* six serological types exist
and of these, the type b *H. influenzae* is the major cause of meningitis in
man (Jennings, 1990) and the type b polysaccharide vaccine was shown to
offer protection against infections caused by this bacterial strain (Moxon
and Rappouli, 1990; Jennings *et al*., 1993). The situation is different with
the *S. pneumoniae* where 23-valent polysaccharide vaccine is needed to
provide effective coverage (Jennings, 1990).

Despite the success obtained with the use of polysaccharide vaccines, their
age-related immunogenicity proved a severe limitation. Infants respond
very poorly to polysaccharide vaccines (Peltola *et al*., 1977; Robbins, 1978;
Jennings, 1983). Purified polysaccharides are T-cell independent antigens,
only capable of inducing immune responses that are mainly of the IgM
isotype which are not boostable by further exposure (Gotschlich *et al*., 1977).
The conjugation of polysaccharides to proteins has been employed to
circumvent this problem (Jennings, 1990). A wide range of antibodies of
the IgG isotype that can be boosted by repeated exposures are produced by
infants in response to polysaccharide-protein conjugates (Robbins and
Schneerson, 1990). The effectiveness of conjugate vaccines in reducing the
incidence of serious infections caused by *H. influenzae* type b has attracted
its wide usage in many countries (Moxon and Rappouli, 1990). The immune
response to polysaccharide vaccines may also be improved by chemical
modification of the polysaccharide (Jennings *et al*., 1993). This
approach has been adopted with some success in the production of *N*
-propionylated B polysaccharide-tetanus toxoid conjugate from the native
polysaccharide moiety (Jennings *et al*., 1986). Neither the replacement of
the *N*-acetyl groups with *N*-propionyl residues nor the conjugation of
this product to protein affected the polysaccharide-specific IgG antibody
responses to *E. coli* K1 and group B meningococci (Jennings, 1990). Thus, a
promising area of study to overcome the limitations of many
polysaccharide vaccines especially with immunogenicity is the design of semi-synthetic or modified polysaccharides for use as vaccines.

In summary, it is clear that bacterial capsules play important roles in the onset and development of disease. The cloning and molecular analysis of several of the capsule gene clusters will obviously be of great benefit in the modification of capsules for therapeutic use.

1.3 *Escherichia coli* Capsular (K) Antigens

*E. coli* is usually a member of the normal gut flora of both humans and animals. However, it is a common cause of opportunistic infection with certain strains associated with severe and life-threatening diseases. Whilst, capsules are expressed by commensals and intestinal etiologic *E. coli* of the gastrointestinal tract, specific *E. coli* K serotypes are frequently associated with extraintestinal isolates (Jann and Jann, 1983; Ørskov et al., 1984). *E. coli* has been shown to produce more than 70 chemically and serologically distinct capsular polysaccharides (Ørskov et al., 1977). Individual isolates can only produce one of these polymers, expression of which is stable and switching of capsular type has not been documented (Boulnois and Roberts, 1990).

1.3.1 Classification of *E. coli* Capsular Antigens

The capsules of *E. coli* were originally divided into two broad groups, I and II on the basis of a number of criteria, including chemical structure, size, mode of expression and chromosomal location of their genetic determinants (Jann and Jann, 1987). The existence of a third group of capsular antigens, designated group I/II (Finke et al., 1990) or group III (Pearce and Roberts, 1995) has now been demonstrated.

1.3.1.1 Group I Capsules

The expression of group I capsular polysaccharides is restricted to *E. coli* serotypes O8, O9 and occasionally O20 and O101 (Jann and Jann, 1990). In general, group I capsules are more stable to heating at 100°C and belong to the low electrophoretic mobility capsular type (Jann and Jann, 1983). Group I polysaccharides have low charge density, high molecular weight and are expressed at all growth temperature (Table 1.2) (Jann and Jann, 1983; 1990).
They are composed mainly of glucuronic acid, galacturonic acid and their ketosidic-pyruvate substitutions. Certain group I capsules have amino sugars as structural compositions and on this basis they have been further divided into groups Ia and Ib, depending on the absence or presence of amino sugars in their structure (Jann and Jann, 1987, 1992). Group Ia K antigens do not contain amino sugars, and they resemble the K antigens found in Klebsiella species. For example, the E. coli K28 and K55 antigens are respectively identical to the K54 and K5 antigens of Klebsiella (Jann and Jann, 1992). Group Ia K antigens show some similarities in structure to colanic acid and the genes for the biosynthesis of colanic acid (cps) and group Ia K antigens are allelic (Keenleyside et al., 1992). It is important to note that colanic acid is loosely associated with the cell surface and therefore should be termed a slime polysaccharide. Group Ib K antigens contain amino sugars and have no counterparts in other bacteria. Strains expressing group Ib capsular polysaccharide are also able to express colanic acid (Keenleyside et al., 1992).

Group I capsular antigens are thought to be attached to the cell surface by core-lipid A and therefore resemble LPS (Jann et al., 1992). It is now known that core-lipid A is not a universal feature of group I K antigens and has been found not to be essential for the surface expression of the K30 polysaccharide (MacLachlan et al., 1993). Analysis of the size of the E. coli K30 antigen has revealed a polysaccharide of primarily high molecular weight that consists of one repeat unit unlinked to core-lipid A (MacLachlan et al., 1993). In addition, E. coli K30 mutants with truncated lipid A core (rfa mutants) still make wild-type capsule as seen on electron microscopy (MacLachlan et al., 1993). This implies that the cell surface expression of such capsules is independent of core-lipid A and such K antigens may be associated at the cell surface through ionic interactions with other attached molecules. For the E. coli K40 group Ib antigen, both low and high molecular weight polysaccharide chains exist. The majority of the capsular polysaccharide exists as high molecular weight species that are linked to lipid A-core which is typical of the LPS molecule (Oodgson et al., 1996). Apart from core-lipid A attachment, the effect of the rol gene product on the expression of group Ib K antigens suggests that the group Ib K antigens should be considered as O antigens. The rol gene product modulates the chain length of heteropolymetric LPS O antigens in a number of Enterobacteriaceae (Batchelor et al., 1992; Bastin et al., 1993; Morona et al., 1995). The rol gene is present on the chromosome of group Ib K antigen expressing strains, and the introduction of a multicopy rol gene was shown to modulate the chain length of the K40 (group Ib) K antigen in a manner identical to the Rol-
dependent O antigen (Dodgeon et al., 1996). In the past, strains expressing the group Ib K antigens have been said to possess two O antigens, one acidic (K antigen) and the other neutral (LPS) (Örskov et al., 1977; Jann and Jann, 1990).

The genetic determinants for group I capsules have been shown to map at 44 minutes on the *E. coli* chromosome near the *his* and *rfb* loci (Örskov et al., 1977). Studies using the K27 antigen indicate that a second *trp*-linked (*rfc*) locus is necessary for surface expression of a complete capsule, although this is not seen in the K30 antigen (Schmidt et al., 1977; Laasko et al., 1988) and has not been demonstrated in other group Ia K antigen expressing strains. The genes for the biosynthesis of group Ib capsular antigens have been assumed to be located at the same *his* and *rfb* loci, but this has not been shown.

1.3.1.2 Group II Capsules

Unlike the group I capsule, group II capsular polysaccharides have a high charge density and are not expressed at growth temperature below 20°C (Örskov et al., 1984; Jann and Jann, 1987). Group II K antigens are heat-labile, usually released from the cell during heating at 100°C. This has been attributed to the labile linkage nature between the group II polysaccharide and its anchor in the outer membrane (Jann and Jann, 1983). Group II polysaccharides are associated with phosphatidyl acid at their reducing termini, and this substitution is thought to serve as a membrane anchor (Schmidt and Jann, 1982; Jann and Jann, 1990). They are of low molecular weight and thus belong to the high electrophoretic mobility group. This group of polysaccharides are co-expressed with many O antigens, excluding those associated with group I (Jann and Jann, 1987). Group II polysaccharides may contain hexuronic acids, N-acetylmuramic acid, or KDO as their acidic components (Jann and Jann, 1987). An important distinguishing factor of group II capsule expressing strains is the high levels of CMP-KDO synthetase (CKS) activity shown at capsule permissive temperature (Schmidt and Jann, 1982; Finke et al., 1989). Genetic determinants for the expression of group II capsules have been shown to map at 64 minutes near *serA* on the *E. coli* chromosome (Örskov et al., 1976; Vimr, 1991) and are termed *kps* (Silver et al., 1984; Vimr et al., 1989). These distinction between groups I and II are summarised in table 1.2. In many respect, the group II K antigens are very similar to capsular polysaccharides of *N. meningitidis* and *H. influenzae* (Jann and Jann, 1987; 1990). This is quite evident in the
serological cross-reactivities of the respective encapsulated strains (Robbins et al., 1974; Ørskov and Ørskov, 1990).

1.3.1.3 Group III Capsules

The serA locus was originally thought to encode a distinct group II capsule gene cluster. Recent studies have shown that some capsules previously assigned as members of group II possess features of group I (Finke et al., 1990; Boulnois et al., 1992; Drake et al., 1993). These capsular antigens include the K2, K3, K10, K11, K19, K54 and K98. It has been shown for the K3, K10, K11, K54 and K98 antigens that despite mapping near serA they are expressed at all growth temperatures (Ørskov et al., 1984). In addition, the K2, K3, K10, K11, K19 and K54 antigen-expressing strains do not have the elevated CKS activity at 37°C characteristic of group II capsule-producing strains (Finke et al., 1990). These former group II capsules were tentatively classed as group I/II (Finke et al., 1990). It should be noted, however that this classification does nothing to distinguish these capsule gene clusters from those clearly assigned to groups I and II.

The cloning and analysis of gene clusters for the expression of groups II and I/II capsules has helped the distinction of these capsules to be further refined. DNA probes from group II capsule gene clusters do not hybridise to DNA from group I/II capsule-expressing strains (Drake et al., 1993). Recently, the cloning and preliminary genetic analysis of the K10 and K54 capsule gene clusters, including hybridisation and complementation studies with the K5 (a group II capsular antigen) have revealed these group I/II capsule gene clusters are quite distinct from the group II capsule gene clusters (Pearce and Roberts, 1995). Thus, despite mapping near serA on the chromosome as the group II gene clusters, Pearce and Roberts (1995) showed that these capsule gene clusters hitherto referred to as group I/II were sufficiently different and were classed as group III. This classification eliminates any ambiguity surrounding the group I/II nomenclature, which does suggest that these capsule gene clusters are in fact hybrid of the clearly classified groups I and II K antigen clusters. Hybridisation studies with the K11 and K19 strains suggest that these capsule gene clusters are not members of group III and may represent a new class of capsule gene clusters (Pearce and Roberts, 1995).
### Table 1.2: Differences Among Groups I, II and III Capsular Polysaccharides of *E. coli*

<table>
<thead>
<tr>
<th>Property</th>
<th>Group I</th>
<th>Group II</th>
<th>Group III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular weight</td>
<td>&gt;100Kd</td>
<td>&lt;50Kd</td>
<td>&lt;50Kd</td>
</tr>
<tr>
<td>State of lipid linkage at 100°C</td>
<td>Stable</td>
<td>Labile</td>
<td>Labile</td>
</tr>
<tr>
<td>Acidic component</td>
<td>GlcA</td>
<td>GlcA</td>
<td>GlcA</td>
</tr>
<tr>
<td></td>
<td>GalA</td>
<td>NeuNAc</td>
<td>NeuNAc</td>
</tr>
<tr>
<td></td>
<td>Pyruvate</td>
<td>KDO</td>
<td>KDO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ManNAc</td>
<td>ManNAc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate</td>
<td>Phosphate</td>
</tr>
<tr>
<td>Expression below 20°C</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Coexpression with O antigens</td>
<td>O8, O9, O20</td>
<td>Many O antigens</td>
<td>NS</td>
</tr>
<tr>
<td>Lipid at reducing end</td>
<td>Core lipid A</td>
<td>PA</td>
<td>PA</td>
</tr>
<tr>
<td>Elevated CKS activity at 37°C</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Genetic determinant on the <em>E. coli</em> chromosome located near</td>
<td><em>rfb</em> (<em>his</em>), <em>rfc</em> (<em>trp</em>)</td>
<td><em>kps</em> (<em>SerA</em>)</td>
<td><em>kps</em> (<em>SerA</em>)</td>
</tr>
<tr>
<td>Intergeneric relationship with</td>
<td><em>Klebsiella</em></td>
<td><em>H. influenzae</em></td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td><em>N. meningitidis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: GalA, galacturonic acid; GlcA, glucuronic acid; KDO, 2-keto-3-deoxy-manno-octulosonic acid; ManNAc, N-acetylmannosaminuronic acid; NeuNAc, N-acetyleneuraminic acid; PA, phosphatidic acid; NS, not stated; CKS, CMP-KDO synthetase
1.4 E. coli  Group II Capsule Gene Clusters

The gene clusters of all group II K antigens studied so far revealed a conserved organisation consisting of three functional regions termed 1, 2 and 3 (Figure 1.1).

Molecular studies on the genes required for the expression of E. coli serotypes K1 (Silver et al., 1981), K5, K7, K12, and K92 (Roberts et al., 1986) and K4 (Drake et al., 1993) show a similar genetic organisation consisting of a central region 2 flanked by regions 1 and 3. Mutations in regions 1 and 3 result in the accumulation of polysaccharide inside the cell, suggesting that the encoded products are needed for translocation of the capsular antigen to the cell surface (Boulnois et al., 1987; Kröncke et al., 1990a,b; Bronner et al., 1993a,b). Regions 1 and 3 are conserved among group II capsule gene clusters and can be interchanged (Roberts et al., 1986). Mutations in either region 1 or 3 can be complemented in trans by the corresponding region from other group II capsule gene clusters (Roberts et al., 1986), which indicates that regions 1 and 3 encode a conserved set of proteins capable of exporting chemically distinct polysaccharides to the cell surface. Thus, genes in regions 1 and 3 are designated kps in keeping with their conservation among group II capsule gene cluster.

The central region 2 is serotype specific and contains genes that encode the proteins necessary for capsule biosynthesis (Boulnois and Jann, 1989; Vimr et al., 1989; Silver and Vimr, 1990). Thus, the region 2 genes encode products that are specific for sugar synthesis, activation and polymerisation of a particular polysaccharide (Roberts et al., 1986; Vimr et al., 1989). Region 2 varies in sizes (Figure 1.1) among E. coli serotypes and these sizes correlate with the complexity of the encoded polysaccharide (Boulnois and Jann, 1989; Drake et al., 1993). For instance, the complex heteropolymer of K4 antigen has a 14 kb region 2 (Drake et al., 1993) in contrast to the 5.8 kb region 2 of the K1 antigen (Silver et al., 1981, 1984; Annunziato et al., 1995), and the 8.0 kb region 2 of the K5 antigen (Petit et al., 1995) (Figure 1.1).
Among the group II capsules, the most extensively studied to date are the *E. coli* K1 and K5 capsule gene clusters. The K5 antigen is of primary interest and will serve as a paradigm for other group II capsules.

### 1.4.1 The E. coli K5 Capsule Gene Cluster

The nucleotide sequence of the entire *E. coli* K5 capsule gene cluster has been determined (Smith *et al.*, 1990; Pazzani *et al.*, 1993a; Petit *et al.*, 1995) and shown to consist of all genes necessary to direct the biosynthesis and export of the K5 capsular polysaccharide (Figure 1.2).

Region 2 of the K5 capsule gene cluster spans an 8.0 kb DNA fragment with an overall G+C content of 33.4% (Petit *et al.*, 1995). This G+C content is lower than the 50% G+C ratio usually observed for *E. coli* (Ørskov, 1984) and also lower than the G+C contents of both region 1 (50.6%) (Pazzani *et al.*, 1993a) and region 3 with 42.3% (Smith *et al.*, 1990) of the K5 capsule gene cluster. In other Gram-negative capsule gene clusters lower G+C contents have been reported (Steenbergen *et al.*, 1992; Edwards *et al.*, 1994; Ganguli *et al.*, 1994; Van Eldere *et al.*, 1995) and was suggested to reflect a common origin of these capsule gene clusters (Frosch *et al.*, 1991). The observed differences in G+C content between region 2 and those of regions 1 and 3.
Figure 1.2: Molecular organisation of the *E. coli* K5 gene cluster

*Bg = Bgl II; Ev = EcoRV; Hc = Hinc II; Sm = Sma I*

↑ : denotes direction of transcript
which are conserved in different group II capsules may indicate the acquisition of different region 2 sequences perhaps through homologous recombination between the flanking regions 1 and 3 of an incoming and resident capsule gene cluster. This view is augmented by the findings of a marked divergence in the amino acid sequences at the C-termini of region 2 flanking KpsS and KpsT proteins from different capsule gene clusters (Drake, 1991; Pavelka et al., 1994). The 3' end of these genes encoding both KpsS and KpsT proteins are located at the junction sites between regions 1/2 and 2/3 respectively. Thus, recombination events may explain the source of capsular diversity in group II capsule gene clusters.

Recent analysis of the K5 capsule gene cluster shows that region 2 contains four genes designated kfiA-D (kfi, for K five antigen) with two large intergenic spaces (Petit et al., 1995) (Figure 1.2). The intergenic regions are between kfiA and B genes and kfiB and C genes. Northern blotting, promoter probe analysis and transcript mapping revealed the presence of three promoters located upstream of kfiA, B and C and three major overlapping transcripts of 8.0, 6.5 and 3.0 kb (Petit et al., 1995). The 8.0 kb transcript originating upstream of kfiA gene is a large polycistronic message that overlaps the other two messages and spans the entire region 2. Similar single polycistronic units have been described for the region 2 capsule gene cluster from E. coli K1 (Silver et al., 1993) and N. meningitidis (Edwards et al., 1994).

The genes described within region 2 encode four proteins KfiA-D that are required for the biosynthesis of the E. coli K5 capsular polysaccharide (Petit et al., 1995). The role of two of the encoded proteins in polysaccharide biosynthesis have been confirmed. The KfiC protein is a glycosyltransferase enzyme that adds both glucuronic acid (GlcA) and N-acetylglucosamine (GlcNAc) in an alternating manner at the non-reducing end of the growing polysaccharide (Petit et al., 1995). Searches of the protein databases and CLUSTER-V alignment of the predicted amino acid sequences of KfiC have identified respectively extensive sequence homology to a number of glycosyltransferase enzymes and three conserved regions that might be of importance in function of these enzymes (Petit et al., 1995). These homologous glycosyltransferase enzymes are involved in the biosynthesis of chemically unrelated polysaccharides. The Exo proteins (ExoM, ExoU, ExoO and ExoW) of R. meliloti are required for the biosynthesis of succinoglycan (Glukmann et al., 1993; Reuber and Walker, 1993). Lsg, the enzyme encoded by the lsg locus of H. influenzae is required for the synthesis of an N-
acetylglucosamine-containing epitope of lipooligosaccharides (LOS) on the cell surface (McLauchlin et al., 1992) and NodC is a glycosyltransferase involved in the biosynthesis of D-glucosamine-containing LOS essential for nodulation by Rhizobium species (Debelle et al., 1992). Others include HetA which is involved in the synthesis of heterocyst-specific envelope polysaccharide in Anabaena species (Holland and Wolk, 1990) and the HasA protein which is required for hyaluronic acid synthesis in Streptococcus pyogenes (Dougherty and Van de Rijn, 1993). It is believed that a polysaccharide biosynthetic complex exists in the inner-face of the cytoplasmic membrane and Western blot analysis using antibodies to the KfiC protein shows its association with the inner-face of the cytoplasmic membrane (G. Rigg, G. Griffiths and I.S. Roberts unpublished). Proteins encoded by regions 1 and 3 are also required for KfiC association with the inner-face of the cytoplasmic membrane (G. Rigg and I.S. Roberts, unpublished). The KfiD protein is a UDP-glucose dehydrogenase enzyme, which catalyses the conversion of UDP-glucose to UDP-glucuronic acid, a component sugar of the K5 polysaccharide (Sieberth et al., 1995). The nucleotide sequence of the kfiD gene shares significant identity to NAD-dependent sugar dehydrogenase encoding genes in other bacteria (Petit et al., 1995). It is 60% identical to the cap3A gene that encode a UDP-glucose dehydrogenase enzyme involved in the expression of the type 3 capsule of S. pneumoniae (Arrecubieta et al., 1994) and also shares 59% identity to the hasB gene required for the synthesis of hyaluronic acid capsule of group A streptococci (Dougherty and Van de Rijn, 1993). The overexpression of the KfiD protein results in elevated UDP-glucose dehydrogenase activity (Petit et al., 1995; Sieberth et al., 1995), which lends further evidence that the KfiD is a UDP-glucose dehydrogenase enzyme. The exact roles of both KfiA and B proteins in the biosynthesis of the K5 polysaccharide are still unknown. The kfiA and kfiB genes encode proteins of 27 kDa and 66 kDa respectively and are believed to be located in the cytoplasm (Petit et al., 1995). The predicted amino acid sequences of both proteins do not share any significant homology to other proteins deposited in the database. A mutant of the kfiA gene has been found to exhibit a low endogenous transferase activity in vitro, but can extend exogenous polysaccharide (Sieberth, 1994). This evidence suggests a role for the encoded KfiA protein in the initiation of polysaccharide biosynthesis, or possibly required for the attachment of the endogenous acceptor on the membrane, onto which polymerisation catalysed by the KfiC occurs. The expression of K5 capsule at 37°C as determined by sensitivity to the capsule-specific bacteriophage was abolished when TnphoA was inserted within the kfiB gene (Stevens, 1995), which also suggests a role in
biosynthesis of the K5 antigen. The polarity effects of this insertion on the expression of the downstream genes \textit{kfiC} and \textit{D} are unlikely as a separate transcript for the \textit{kfiC} and \textit{D} genes has been identified (Petit \textit{et al.}, 1995). Detail analysis of both the \textit{KfIA} and \textit{KfIB} are currently going on to define their exact role in the biosynthesis of the K5 capsular polysaccharide.

Region 3 of the K5 capsule gene cluster consists of two genes \textit{kpsM} and \textit{T} organised in a single transcriptional unit (Smith \textit{et al.}, 1990). A similar organisation was described for the region 3 of K1 capsule gene cluster (Pavelka \textit{et al.}, 1991). The promoter for region 3 of K5 antigen gene cluster has been mapped to 741 bp upstream of the initiation codon for \textit{kpsM} (Stevens, 1995). Transcript originating from the region 3 promoter was temperature regulated, only being detected at capsule-permissive temperature of 37°C (Stevens, 1995). In both the K1 and K5 capsule gene clusters, transcription of region 3 is in the same direction as in region 2 transcript. For the K5 capsule gene cluster, region 2 was shown to be regulated by \textit{RfaH} and a recently described JUMP-start (Hobbs and Reezes, 1994) sequence in region 3 is important for this regulation (Stevens \textit{et al.}, 1995). Analysis of the predicted amino acid composition of the encoded proteins indicated that the KpsM and T probably exist as homodimers (Smith \textit{et al.}, 1990; Pavelka \textit{et al.}, 1991, 1994) that may function in a polysaccharide export system analogous to the periplasmic binding protein-dependent transport system of gram-negative bacteria (Higgins \textit{et al.}, 1985). The functions of these two proteins will be discussed below under the section on polysaccharide export.

The nucleotide sequence for the entire region 1 of the K5 capsule gene cluster has been determined (Pazzani \textit{et al.}, 1993a,b). Region 1 consists of six genes \textit{kpsFEDU} (Figure 1.2) organised as a single transcriptional unit. By Northern blotting and transcript mapping, this transcriptional organisation has been confirmed and the promoter identified to 225 bp upstream from the start codon of the first gene \textit{kpsF} (Simpson \textit{et al.}, 1996). The promoter had -35 and -10 consensus sequence similar to an \textit{E. coli} σ70 promoter. Region 1 is transcribed as an 8.0 kb polycistronic mRNA which was processed to form a separate 1.3 kb transcript encoding the last gene, \textit{kpsS} . The generation of a separate \textit{kpsS} transcript has been postulated to bring about differential expression of KpsS from the other region 1 encoded proteins (Simpson \textit{et al.}, 1996) and thereby enable regulation of capsular polysaccharide expression through a role for KpsS in the modification of the polysaccharide prior to export (Roberts, 1996). Two binding site consensus sequences for integration
host factor (IHF) were identified 110 bp upstream and 130 bp downstream of the region 1 transcription start site. Binding of IHF which is a histone-like DNA-bending protein at these sites was suggested to change the DNA conformation and therefore may alter the efficiency of transcription and/or mediate the action of other regulators on the kps genes (Simpson et al., 1996). By using strains with mutations in himA and himD, which encode the two subunits of IHF, fivefold and sixfold reductions respectively in the expression of KpsE protein at capsule permissive temperature was demonstrated (Simpson et al., 1996). Although, previous results has shown that a functional IHF is not essential for group II capsule expression (Stevens et al., 1995), the reduced expression of KpsE does suggest a role for IHF in mediating the action of other regulatory proteins. A Rho-dependent transcriptional terminator was located within the kpsF gene, and was postulated to modulate the transcription of region 1 in response to the physiological state of the cell such that under conditions of physiological stress expression of region 1 is reduced (Simpson et al., 1996). Like region 3, the region 1 transcript is temperature regulated with no transcription detectable at non-capsule permissive temperature (18°C). The nucleotide sequence of kpsF in the K5 and K1 antigen gene clusters is 98.8% identical, confirming that this gene is conserved between different group II capsule gene clusters (Simpson et al., 1996) Analysis of the predicted amino acid sequence of the KpsF protein has revealed significant homology to the GutQ protein, a hypothetical ORF 328 protein of E. coli and a KpsF homologue in H. influenzae (Simpson et al., 1996). Based on the proposed role of the GutQ protein as a regulator of the glucitol operon in E. coli (Yamada et al., 1990), it has been suggested that the KpsF protein may play a role in the regulation of capsule expression in group II capsule producing E. coli (Cielewicz et al., 1993). It was observed however, that mutants lacking a functional kpsF gene are still able to express the K5 capsule in a temperature-dependent manner (Pazzani et al., 1993a,b). The possibility that GutQ could be acting in trans to complement the deleted kpsF was ruled out by the observation of capsule expression when a plasmid which contains the K5 capsule gene cluster minus the kpsF gene was introduced into a strain with a mutation in the gutQ gene (Simpson et al., 1996). Therefore, KpsF does not appear to have an essential role in the expression of group II capsules.

The kpsE and D genes encode proteins of 43 and 60 kDa respectively. In E. coli K1, the kpsE gene is reported to encode a protein of 39 kDa, and the proposed translational start site of the kpsE gene in E. coli K5 has been disputed (Cieslewicz et al., 1993). The translational start of the kpsE gene in
E. coli K5 was confirmed by purification and N-terminal sequencing of the encoded protein (Rosenow et al., 1995a). The KpsE protein hydropathy plot (Kyte and Doolittle, 1982) which is similar to both the BexC protein of H. influenzae and CtrB protein of N. meningitidis revealed a relatively hydrophilic protein with N- and C-terminal hydrophobic domains (Kroll et al., 1990; Frosch et al., 1991; Pazzani, 1992). These three proteins also share extensive homology. The KpsD protein is a periplasmic protein with a typical N-terminal signal sequence (Pazzani et al., 1993a,b). Both KpsE and D proteins are similar in the K1 and K5 capsule gene cluster (Cieslewicz et al., 1993; Pazzani et al., 1993a,b; Wunder et al., 1994) and have been suggested to play a role in polysaccharide export (see below).

The KpsU protein is a functional CMP-KDO synthetase enzyme (Rosenow et al., 1995b) and the protein is homologous (44.3% identity and 70% similarity) to the KdsB enzyme of E. coli (Pazzani et al., 1993a,b). The KdsB enzyme is also a CMP-KDO synthetase enzyme that catalyses the activation of KDO prior to its linkage to lipid A (Munson et al., 1978; Goldman and Kohlbrenner, 1985). The KpsU enzyme is suggested to function in the same manner in generating activated KDO before linkage of KDO to phospholipid in group II capsules (Jann and Jann, 1983). The identification of a functional CMP-KDO synthetase structural gene within region 1 accounts for the high level of this enzyme activity observed at capsule permissive temperature in group II capsule-expressing E. coli (Finke et al., 1990). Mutations in the kpsU gene of the K5 capsule gene cluster do not abolish capsule production as determined by sensitivity to K5-specific bacteriophage (Pazzani, 1992) and this is thought to be as a result of the compensatory effect of the enzyme encoded by the kdsB gene. An immunoelectron microscopy of this mutant revealed patches of polysaccharide at the cell surface and polysaccharide accumulated as electron dense aggregates in the cytoplasm (Bronner et al., 1993a). Polysaccharide aggregate is a phenotype associated with both KpsC and KpsS mutants (Bronner et al., 1993a), and it is possible that the oligonucleotide linker that contains a stop codon which is inserted in the kpsU gene in this mutant has resulted in a polarity effect on the expression of kpsC and perhaps kpsS genes. The translational start codon for the kpsC gene overlaps the end of kpsU, which shows that these genes may be translationally coupled (Pazzani, 1992).

By minicell analysis, the kpsC and kpsS genes have been shown to encode proteins of 76 kDa and 46 kDa respectively in the K5 capsule gene cluster (Pazzani et al., 1993a). In region 1 of E. coli K1 antigen gene cluster, proteins
of similar sizes are encoded (Silver et al., 1984; Roberts et al., 1986). The KpsC and KpsS proteins are predicted to be located in the cytoplasm in *E. coli* K1 and K5 (Silver et al., 1984; Pazzani et al., 1993a). Mutations in both *kpsC* and *kpsS* genes result in the accumulation of full length intracellular polysaccharide that lack KDO or phospholipid substitution in the cytoplasm as electron dense aggregates (Bronner et al., 1993a,b). It has been reported in both *E. coli* K1 and K5, that membranes from *kpsC* and *kpsS* mutants have reduced endogenous transferase activities *in vitro* (Vimr et al., 1989; Bronner et al., 1993a,b). In the case of *E. coli* K5, the addition of exogenous K5 polysaccharide to membrane preparations of *kpsC* and *kpsS* mutants resulted in increased transferase activity from about 2% to 25% of the wild-type (Bronner et al., 1993b). It is therefore likely that *kpsC* and *kpsS* mutants may reduce the amount of existing endogenous acceptor on the membrane. In support of this notion, is the suggestion that KpsC and KpsS may form a complex with the enzymes for polysaccharide biosynthesis at the cytoplasmic face of the inner membrane (Vimr et al., 1989; Bronner et al., 1993a,b). It has been postulated that the biosynthetic complex at the cytoplasmic face of the inner membrane is linked to enzymes required for the translocation of capsular polysaccharide to the cell surface (Roberts, 1995, 1996). The possible role played by the KpsC and KpsS in the export of polysaccharide will be discussed (see section 1.6.2.2 below).

1.5 Genetic Organisation of *H. influenzae* and *N. meningitidis* Capsule Gene Clusters

*H. influenzae* produces six serologically distinct capsular polysaccharides designated a through f. Of these six serotypes, only the type b strains are associated with invasive diseases in human (Moxon and Kroll, 1990). The genes for the production of *H. influenzae* type b antigen have been cloned and the cluster shows a similar organisation to that of group II capsule gene clusters of *E. coli* (Kroll et al., 1989) (Figure 1.3). The central region 2 of the *H. influenzae* type b capsule gene (*cap*) cluster is serotype specific and encodes proteins necessary for polysaccharide biosynthesis (Ely et al., 1989; Van Eldere et al., 1995). The flanking regions 1 and 3 are common to all of *H. influenzae* serotypes and like the *E. coli* group II capsule genes are involved in the export of polysaccharide to the cell surface (Kroll et al., 1989, 1990). Region 1 of the *H. influenzae* type b *cap* locus has been sequenced and shown to contain four genes, *bexABCD* probably organised in a single transcriptional unit (Kroll et al., 1990).
A striking feature of the majority of type b strains is the duplication of the cap locus, which comprises two directly repeated 17 kb copies of DNA separated by a 1.2 kb 'bridge segment' (Kroll and Moxon, 1988). The duplication is not perfect, since the cap locus contains only one functional copy of bexA (Kroll et al., 1989). The bexA gene is located in the bridge region and is required for polysaccharide export (Kroll et al., 1990). Strains of H. influenzae type b revert to an acapsular phenotype at high frequency. This instability is attributed to recombination events between homologous sites in each repeat which reduce the cap to a single copy and thereby result in the loss of bexA (Kroll and Moxon, 1988).

Strains of H. influenzae type b may contain multiple copies of the cap genes (Kroll and Moxon, 1988). Recombination events mediated by direct repeats of an insertion sequence-like element, IS1016 which flank the cap locus are responsible for this amplification (Kroll et al., 1991). The amplification of cap is accompanied by an increase in capsule synthesis (Kroll and Moxon, 1988). The advantages of the cap instability and gene-dosage phenomena in vivo are not clear. On the other hand, it has been suggested that the down-regulation of capsule expression may promote bacterial adhesion and invasion of host cells mediated by outer membrane opacity proteins (Virji et al., 1994).

The N. meningitidis group B capsule gene cluster (cps) has been analysed and it revealed five regions A-E (Frosch et al., 1989). In all, genes required for the biosynthesis and export of capsular polysaccharide are located in three regions organised in a similar manner to those of group II capsule gene clusters (Frosch et al., 1989, 1991; Hammerschmidt et al., 1994) (Figure 1.3). A central region A encodes all enzymes necessary for biosynthesis of the α-2,8-polysialic acid polymer (Edwards et al., 1994; Ganguli et al., 1994). This region is flanked by region B which encodes two proteins, LipA and LipB needed for phospholipid substitution of the polysaccharide at the reducing end prior to transportation (Frosch and Muller, 1993). Subsequent cell surface translocation is directed by region C encoded proteins on the other side of the central region A. Region C encodes proteins that share strong features with members of the ABC superfamily of transporters (Frosch et al., 1991). The additional two regions, D and E within the cps gene cluster were originally thought to be involved in regulation of capsule expression (Frosch et al., 1989). However, recent studies (Hammerschmidt et al., 1994) have shown that region D contains genes not required for capsule expression, but encodes enzymes involved in meningococcal lipo-oligosaccharide biosynthesis.
Regions analogous to regions D and E have not been identified in E. coli group II capsule gene clusters.

The capsular polysaccharide of N. meningitidis group B is identical to the E. coli K1 antigen (Bhattacharjee et al., 1975). Although initial studies indicated that little homology exists between DNA of these two capsule genes (Echarti et al., 1983), subsequent analysis has shown that the proteins involved in the synthesis of the polysaccharide are similar. For instance, the NeuB protein of E. coli K1, which is postulated to encode an N-acetyleneuraminic acid synthetase, is 57% similar to the CpsB protein of N. meningitidis (Annunziato et al., 1995). In these bacteria, the predicted amino acid sequences of the CMP-N-acetyleneuraminic acid synthetase and the sialytransferase enzymes also share significant homologies (Frosch et al., 1991; Ganguli et al., 1994).

The homologous genetic organisation of the capsule genes of H. influenzae, N. meningitidis and E. coli expressing group II K antigens suggests that they have a common evolutionary origin. Support for this comes from demonstration of DNA and protein homology between the functional regions of the capsule gene clusters of these bacteria. For example, the kpsM and kpsT genes of region 3 in E. coli share extensive sequence homology with the H. influenzae bexB and bexA genes (Kroll et al., 1990; Smith et al., 1990; Pavelka et al., 1991), and with the N. meningitidis ctrC and ctrD genes (Frosch et al., 1991) respectively. In addition, the kpsE gene is homologous to the bexC and ctrB genes of H. influenzae and N. meningitidis respectively (Cieslewicz et al., 1993; Pazzani et al., 1993). These genes encode proteins that are involved in the cell surface expression of capsular polysaccharide in these bacteria (see section 1.6.2.2).
Figure 1.3: Genetic organisation of the *H. influenzae* type b, and *N. meningitidis* group B capsule gene clusters and the *E. coli* K5 group II capsule gene cluster. For clarity, a single copy of the *H. influenzae* capsule gene cluster is shown (see text). The large boxes denote the conserved functional regions and small boxes the specific genes that have been identified within each cluster.
1.6 Polysaccharide and Protein Export in Gram-negative Bacteria

The export of molecules in Gram-negative bacteria has been a subject of intensive research for the past few decades, and the processes involved are yet to be fully elucidated. The majority of what is known today comes from studies of the protein export systems. Some of these systems involved in protein export have been identified in the export of polysaccharide, and therefore represent a common theme in the export of molecules in Gram-negative bacteria. It is also likely that despite the common theme, export processes may differ especially to accommodate differences in substrates being translocated.

Polysaccharides, proteins and many other molecules are synthesised in the bacterial cytoplasm and some of these molecules are located partially or completely outside the cytoplasm. To reach their final destination, these molecules must often cross one or more membranes. For molecules destined for the surface or external milieu of the bacterial cell, they must cross the cytoplasmic membrane in Gram-positive bacteria and both cytoplasmic and outer membranes in Gram-negative bacteria. Many secreted bacterial proteins reach the external milieu or periplasm (in case of Gram-negative bacteria) via the secretory (Sec) pathway (Schatz and Beckwith, 1990; Wickner et al., 1991; Pugsley, 1993). Sec-dependent secretion requires that the protein to be secreted contains an N-terminal signal sequence (Pugsley, 1993). This absolute requirement for N-terminal signal sequences, thus severely limits the number of substrates that can be translocated in this way. Some extracellular proteins and nonprotein products that are translocated across both inner and outer membranes of Gram-negative bacteria cannot use the Sec pathway (Kenny et al., 1991; Faith and Kolter, 1993). For these products, some dedicated export systems generally grouped as ABC transporters exist that facilitate membrane translocations with a large degree of substrate specificity (Higgins, 1992; Lory, 1992; Wandersman, 1992). Certain proteins are also capable of self-promoting their secretion (Lory, 1992; Pugsley, 1993).

1.6.1 Protein Export

1.6.1.1 Sec-dependent Translocation

Proteins transported by the Sec pathway are called secretory proteins or presecretory proteins if they are made as precursors. Secretory proteins are
distinguished from other proteins by the presence of a cleavable secretory signal sequence. Signal sequences are exclusively amino terminal and their characteristic features include a positively charged amino-terminal domain (N) of 1 to 5 residues, a central hydrophobic region (H domain) of 7 to 15 residues and a carboxyl-terminal region (C domain) containing 5 to 7 residues that ends at the signal peptidase cleavage site (von Heijne, 1984; Pugsley and Schwartz, 1985; Wandersman, 1989; Pugsley, 1993). Shortening or introducing charged or strong polar residues into the H domain, or eliminating the basic amino acids from the N domain usually results in reduced export efficiency (Pugsley, 1993). In the majority of cases identified so far, exchange of signal sequences between presecretory proteins does not affect the translocation of such proteins (von Heijne, 1984, 1986; Delpeleaire and Wandersman, 1990). On translocation of presecretory proteins, the signal processing site is recognised and cleaved by signal peptidase. Thus the primary function of the signal sequence is to initiate insertion into the cytoplasmic membrane and translocate an otherwise cytoplasmic protein across the inner membrane.

In the Sec pathway, protein secretion across the inner membrane requires interaction with products of the sec (prl) genes. Six Sec proteins (SecABDEFY) have been identified as components of the secretory pathway (Gardel et al., 1990; de Cock and Tommassen, 1991, 1992; Kumamoto, 1991; Ito, 1992). SecB and A are both cytoplasmic proteins. It has been suggested that SecB exists probably as a tetramer and may be part of a protein complex that binds to presecretory proteins (Watanabe and Blobel, 1989a,b; Kumamoto, 1991). A major role of the SecB protein is to pilot presecretory proteins to the cytoplasmic membrane (de Cock and Tommassen, 1992). The SecB proteins are also referred to as secretory chaperonins and are in many ways similar to the general molecular chaperonins, which bind to exposed polypeptide segments (Pugsley, 1993). The SecA protein is found associated with the ribosomes as well as the membrane (Cabelli et al., 1991). Thus, it could interact with the nascent secretory proteins as they emerge from ribosomes while it is free in the cytosol or when it binds to the inner face of the cytoplasmic membrane (Pugsley, 1993). SecA is essential for cell viability as mutations in the secA gene are lethal (Matsuyama et al., 1990; Akimura et al., 1991; Jarosik and Oliver, 1991). SecA binds and hydrolyses ATP (Matsuyama et al., 1990), and it has a sequence with an amino-terminal region that is identical to nucleotide binding (Walker box A) domain found in proteins such as ATPase and kinases (Schmidt et al., 1988) and has been referred to as secretory ATPase because of the ATP binding capability.
The other four Sec components, Sec D, E, F and Y are integral membrane proteins with several long stretches of hydrophobic amino acids that probably span the cytoplasmic membrane (Gardel et al., 1990; Schatz et al., 1991; Ito, 1992) and are collectively termed translocases (Pugsley, 1993). The membrane organisation of the SecY protein shows that it probably has 10 membrane spanning domains with the N- and C-termini in the cytoplasm (Akiyama and Ito, 1987), a topology that is similar to the LacY lactose permease involved in solute transport in Gram-negative bacteria (Ito, 1992). The SecE protein has three membrane spanning segments, and the C-terminal domain was shown to be sufficient for SecE activity (Schatz et al., 1991). Ito (1992) however reported a single transmembrane domain organisation for the SecE protein from Bacillus subtilis. Mutations in both SecY and E block protein translocation (Ito, 1992). The functions of SecD and SecF proteins have not been fully elucidated. Each appear to span the membrane six times (Gardel et al., 1990). The SecD and SecF proteins are characterised by a large periplasmic domain and mutants lacking both proteins either reduce or block protein translocation where molecular chaperones that interact with newly translocated proteins are absent. Both proteins have been suggested to perform related functions at a late stage in export (Pugsley, 1993). As proposed by Simon and Blobel (1992), the translocases are thought to function in the formation of a translocation channel by one or more of the integral membrane components. This in effect will result in shielding of hydrophilic segments of the presecretory protein from the hydrophobic environment of the lipid bilayer.

Proteins may reach their final location on translocation across the cytoplasmic membrane via the Sec pathway or may rely on the participation of other assembly proteins or complexes like the chaperon and general secretory proteins. The general mechanism involved in the Sec-dependent export process is represented in figure 1.4.
Figure 1.4: General export process in Gram-negative bacteria, showing the Sec-dependent insertion and translocation. The model represented shows from left to right; SecB bound to proteinX (I), which then interact with SecA at the inner face of cytoplasmic membrane (II). The proteinX-SecAB complex then interacts with the integral membrane translocases SecDEFY (III). Both SecAB proteins are released (IV) before the signal peptide shown in thick lines is cleaved, leading to the export of the protein across the cytoplasmic membrane (V). Abbreviation: OMP, outer membrane protein; CA, chaperone assembly; GSP, general secretory pathway.
1.6.1.2 Membrane Topology: Determinants of Protein Secretion

In Gram-negative bacteria the envelope is composed of two membranes (inner and outer) separated by the periplasm, and proteins destined for the cell surface upon translocation across the inner membrane must transverse both the periplasmic space and outer membrane. Within the bacterial membrane, proteins adopt a variety of topological states that may expose different domains of such proteins for interaction with other components of the export system and/or the substrate being exported. The different topological states exhibited by integral membrane proteins have been classed as monotopic, bitopic and polytopic (Blobel, 1980; Pugsley, 1993). The bitopic membrane proteins have at least one transmembrane segment mainly of hydrophobic amino acids embedded in the lipid bilayer with one extremity exposed on each side of the membrane (Blobel, 1980). Bitopic membrane proteins are divided into three classes depending on their orientation and whether they are made as precursors. Type Ia bitopic membrane proteins are made as signal peptide-bearing precursors processed by signal peptidases (Pugsley, 1993) and a second hydrophobic domain that acts as the stop transfer anchors them in the cytoplasmic membrane. The type Ib membrane proteins have a single, unprocessed stretch of hydrophobic amino acids at their amino terminus which acts to insert the proteins in the lipid bilayer (Blobel, 1980; von Heijne and Gavel, 1988). Both types Ia and b membrane proteins have a C-terminus in the cytoplasm. Type II bitopic membrane proteins are similar to periplasmic proteins except that the membrane spanning signal-like sequence at the N-terminus is not cleaved. They are usually anchored by the N-terminal region with the remaining bulk of the protein in the periplasm (von Heijne, 1986; Pugsley, 1993).

Polytopic integral membrane proteins span the membrane at least twice via domains of mainly, but not exclusively, hydrophobic amino acid residues (Blobel, 1980). Each of the transmembrane domain is potentially capable of targeting the nascent polypeptide to the cytoplasmic membrane (Lee et al., 1992; Traxler et al., 1992; Gavel and von Heijne, 1994) and the ‘positive inside rule’ of von Heijne (1986b) proposed that the positively charged amino acids determined the orientation of adjacent transmembrane domains. On the other hand, monotopic proteins are only partially embedded in the lipid bilayer by either the N- or C-termini (Pugsley, 1993). They do not have signal sequence-like regions and their membrane association is thought to be strengthened by interaction with other integral membrane proteins (Ulbrandt et al., 1992).
Some proteins rely for secretion on information encoded within the exported polypeptide (Kuhn, 1988; Koronakis et al., 1989; Boyd and Beckwith, 1990; Schatz and Beckwith, 1990; Svirsky et al., 1995). In these proteins determinants of localisation must be part of the mature polypeptide and they can be located in any part of the polypeptide chain (Lory, 1992). Both N- and C-terminal domains have been implicated as export signals capable of self-promoting the translocation of several bacterial proteins (Koronakis et al., 1989; Calamia and Manoil, 1992; Lory, 1992) and the topology of the protein within the bacterial membrane is important in the disposition of these export signals.

Many integral membrane proteins contain an amino terminal region consisting of a short hydrophilic domain with a net positive charge, followed by a stretch of hydrophobic membrane spanning α-helical domain with no signal processing site. Such N-terminal regions have been shown to promote protein insertion and translocation of the hydrophilic domains across the inner membrane, and thus constitute export signals (Roof et al., 1991; Calamia and Beckwith, 1992; Karlsson et al., 1993; Uhland et al., 1994; Vianney et al., 1994). The positively charged amino terminal region functions as cytoplasmic anchor, determining the orientation of the transmembrane domains and the entire polypeptide (von Heijne, 1986; Boyd and Beckwith, 1990; Nilsson and von Heijne, 1990). The N-terminal sequences that constitute such export signals in the TolA and TonB proteins have been shown to be functionally homologous (Karlsson et al., 1993).

Evidence of C-terminal domains acting as export signals is best exemplified by the hemolysin (HlyA) export. Information sufficient for membrane targeting and export of wild-type hemolysin into the surrounding medium has been shown to be located in the C-terminal 113 amino acid residues (Koronakis et al., 1987; Mackman et al., 1987). Further analysis of HlyA protein from E. coli, Proteus vulgaris and Morganella morganii identified three conserved features within this C-terminal region that are necessary for the export role. These features are the amphiphilic α-helix secondary structure, followed by a cluster of charged residues and the hydrophobic/hydroxylated tail (Koronakis et al., 1989). These three HlyA proteins do not show any primary sequence homology within the C-termini (Koronakis et al., 1987), underlying the importance of secondary consensus features in export functions.
The analysis of protein topology is therefore a step in the study of structure-function relationship.

### 1.6.1.3 Bacterial ABC Transporters

The ATP-binding cassette (ABC) transporters constitute a superfamily of proteins involved in the transport of diverse molecules found in both prokaryotes and eukaryotes. In prokaryotes, the bacterial ABC transporters function as importers and exporters and respectively constitutes two families within the superfamily (Ames et al., 1990; Higgins, 1992; Fath and Kolter, 1993). A third family is the eukaryotic ABC transporters involved in the export of multidrug resistance (MDR) protein (Endicott and Ling, 1989), including antimalarial drugs from *Plasmodium falciparum* (Foote et al., 1989).

The family of bacterial ABC exporters is the largest and fastest-growing among the ABC superfamily of transporters. Over 40 systems have been identified (Fath and Kolter, 1993). In general, ABC transporters have a protein with a conserved ATP-binding motif or 'Walker' motif (Walker et al., 1982) and another with membrane spanning domains (MSDs). The ATP-binding component can be present on the same polypeptide as the MSDs (as in the eukaryotic exporters) or on a polypeptide separate from the membrane spanning unit (as in the ABC importers) (Fath and Kolter, 1993). Many of the bacterial ABC exporters involved in protein export require in addition to the ATP-binding cassette and MSDs units, other protein(s) to form a functional export complex. These additional components include proteins belonging to the newly described membrane fusion protein (MFP) family (Dinh et al., 1994), also referred to as accessory factors (Fath and Kolter, 1993) and the outer membrane proteins (Gilson et al., 1987; Letoffe et al., 1990; Wandersman and Delepelaire, 1990). The accessory proteins have been shown to fractionate mostly to the inner membrane (Delepelaire and Wandersman, 1991). The CvaA protein involved in the export of colicin V belongs to this class of accessory proteins (Fath and Kolter, 1993) or MFP family (Dinh et al., 1994) and it has a topological organisation which consists of a single transmembrane domain near its N-terminus and a large C-terminal region extending into the periplasm (Skvirsky et al., 1995). The accessory factors or membrane fusion proteins are thought to connect the inner and outer membranes and thus function to facilitate export of substrates through both membranes of Gram-negative bacteria (Fath and Kolter, 1993; Dinh et al., 1994). Some systems require the outer membrane
components and in such systems they interact with the MFP components, resulting in the fusion of inner and outer membranes or the formation of an oligomeric protein pore in the outer membrane (Dinh et al., 1994). The processes involved in the ABC-mediated export of α-haemolysin in E. coli is shown in figure 1.5. Genes coding for the outer membrane proteins can be closely linked to the other export genes (Delepelaire and Wandersman, 1990; Letoffe et al., 1990) or physically distant (Wandersman and Delepelaire, 1990). Some of the components of ABC exporters are functionally interchangeable. For instance, the HasD and HasE proteins which normally export the HasA protein (a haem acquisition protein) of Serratia marcescens are capable of secreting the metalloprotease C (PrtC protein) of Erwinia chrysanthemi (Binet and Wandersman, 1995). However, the HasA protein is secreted only by its specific transporters. A summary of prototype bacterial ABC exporters are listed in table 1.3 and they include systems involved in the export of proteins (Felmlee et al., 1985; Koronakis et al., 1987, 1988), peptides (Gilson et al., 1990) and polysaccharides (Stanfield et al., 1988; Smith et al., 1990; Pavelka et al., 1991; Bronner et al., 1994).
Figure 1.5: Structural model of ABC-mediated export of *E. coli* α-haemolysin (HlyA). HlyB is the membrane spanning (MS) and ATP-binding polypeptide, HlyD the accessory protein (AF) and TolC the outer membrane component, which is not encoded within the haemolysin operon. The MS/ATP polypeptide and AF are represented as dimers, consistent with the model of Higgins (1992). The figure is from Fath and Kolter (1993).
Table 1.3: Some Members of Bacterial ABC Transporters

<table>
<thead>
<tr>
<th>ABC homologue</th>
<th>MFP/AF(^a)</th>
<th>OMP(^b)</th>
<th>Transported substrate</th>
<th>Organism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CvaB</td>
<td>CvaA</td>
<td>TolC</td>
<td>Colicin V (CvaC)</td>
<td>E. coli</td>
<td>Gilson et al. (1990)</td>
</tr>
<tr>
<td>HlyB</td>
<td>HlyD</td>
<td>TolC</td>
<td>α-hemolysin (HlyA)</td>
<td>E. coli</td>
<td>Felmlee et al. (1985)</td>
</tr>
<tr>
<td>HlyB</td>
<td>HlyD</td>
<td>TolC</td>
<td>α-hemolysin (HlyA)</td>
<td>Proteus vulgaris</td>
<td>Koronakis et al. (1988)</td>
</tr>
<tr>
<td>HlyB</td>
<td>HlyD</td>
<td>TolC</td>
<td>α-hemolysin (HlyA)</td>
<td>Morganella morganii</td>
<td>Koronakis et al. (1987)</td>
</tr>
<tr>
<td>PrtD</td>
<td>PrtE</td>
<td>PrtF</td>
<td>Protease C (PrtC)</td>
<td>E. chrysanthemi</td>
<td>Letoffe et al. (1990)</td>
</tr>
<tr>
<td>HasD</td>
<td>HasE</td>
<td>TolC</td>
<td>Haem acquisition protein (HasA)</td>
<td>Serratia marcescens</td>
<td>Letoffe et al. (1993, 1994)</td>
</tr>
<tr>
<td>KpsMT</td>
<td>?</td>
<td>?</td>
<td>Capsular polysaccharide</td>
<td>E. coli K5</td>
<td>Smith et al. (1990)</td>
</tr>
<tr>
<td>CtrCD</td>
<td>?</td>
<td>?</td>
<td>Capsular polysaccharide</td>
<td>N. meningitidis</td>
<td>Frosch et al. (1991)</td>
</tr>
<tr>
<td>RfbAB</td>
<td>?</td>
<td>?</td>
<td>O-antigen side chain</td>
<td>Klebsiella pneumoniae O1</td>
<td>Bronner et al. (1994)</td>
</tr>
<tr>
<td>ChvA</td>
<td>?</td>
<td>?</td>
<td>β-1,2-glucan</td>
<td>A. tumefaciens</td>
<td>Cangelosi et al. (1989)</td>
</tr>
</tbody>
</table>

\(^a\): MFP/AF, membrane fusion protein and accessory factor respectively
\(^b\): outer membrane protein
\(c\): not required
\(d\): unknown
1.6.2 Polysaccharide Export

Although the export of polysaccharide may differ significantly from protein export, conserved systems involved in their export processes do exist. The use of molecular genetic and biochemical techniques has begun to shed some light on the mechanisms involved in the export of polysaccharide.

It is generally assumed that assembly or polymerisation of capsular polysaccharide and some O antigen repeating units occurs at the inner face of the cytoplasmic membrane, using precursors synthesised in the cytosol (Boulnois and Jann, 1989; Kroncke et al., 1990a,b; Whitfield and Valvano, 1993; Roberts, 1996). For other O antigens, polymerisation of the repeating unit is thought to be at the periplasmic face of the cytoplasmic membrane. This assembly complex is equally believed to be closely linked to the systems involved in the export of distinct polysaccharides (Roberts, 1996).

1.6.2.1 Export of Lipopolysaccharide (O-antigen)

The lipopolysaccharide of Gram-negative bacteria consists of three components: the lipid A, core oligosaccharide and the O antigen. The lipid A and core oligosaccharide are synthesised together, while the O antigen is independently synthesised (Jann and Jann, 1984). The O antigen component of the LPS consists of short oligosaccharide, O unit repeats. Two known mechanisms have been identified in the assembly of O antigens, which differ in the cellular location of their polymerisation steps and in the direction of chain elongation (Whitfield, 1995), and hence distinct export pathways. For each export pathway, after translocation across the cytoplasmic membrane, the cell surface expression of completed lipopolysaccharide is by a yet unidentified mechanism (Whitfield and Valvano, 1993).

In the chain length, Rol (Cld)- and Rfc-dependent polymerisation pathway, addition of the O units occurs at the reducing terminus (Brown et al., 1992; Klena and Schnaitman, 1993; Wang and Reeves, 1994). The first mechanism involves the synthesis of the O units on a lipid carrier, undecaprenol phosphate (und-P) at the inner face of the cytoplasmic membrane (Jann and Jann, 1984). The galactosyltransferase RfbP enzyme catalyses the initiation reaction through interaction with und-P acceptor, and other Rfb enzymes catalyse the transfer of successive sugars to form lipid-linked O units (Klena and Schnaitman, 1993; Morona et al., 1994; Liu et al., 1995). Lipid-linked O units are then polymerised en block in a reaction requiring the Rfc protein.
The second pathway, Rol- and Rfc-independent polymerisation is currently only known to be involved in the synthesis of homopolymer O antigens of *E. coli* O8 and O9, and *Klebsiella pneumoniae* O1 (Whitfield, 1995). In this pathway Rfe and RfbP function only in initiating O antigen synthesis and this involves the transfer of a non-O-antigen residue to und-P to form the acceptor for monomers of the O antigen (Bronner *et al.*, 1994; Rick *et al.*, 1994; Clarke *et al.*, 1995). In the *E. coli* O8, the acceptor is und-P-P-GlcNAc and GlcNAc is not a component of this O antigen (Rick *et al.*, 1994). Polymerisation in the Rfc-independent pathway occurs at the inner face of the cytoplasmic membrane and it involves the sequential addition of sugar monomers to the non-reducing end of the growing polypeptide chain (Zhang *et al.*, 1993; Bronner *et al.*, 1994). Subsequent translocation of the O antigen to the ligation site at the periplasmic face has been shown to require proteins that belong to the ABC transporter family described by Higgins (1992). In the *rfb* gene cluster of *K. pneumoniae*, two proteins RfbA and B have been identified as members of the ABC family of exporters involved in the translocation of the polymerised O unit across the cytoplasmic membrane (Bronner *et al.*, 1994). A similar mechanism involving ABC transporters is seen in the export of group II polysaccharide (see below).

1.6.2.2 Export of Group II Capsular Polysaccharide (K antigen)

Much of what is known about the biosynthesis and export of group II K antigens comes from studies of group II capsule genes, particularly those from *E. coli* K1 and K5 serotypes. The export of group II capsular polysaccharide is presented therefore in the light of what is known from
Proteins encoded by regions 1 and 3 are involved in the export of group II capsular polysaccharides in *E. coli* (Smith *et al.*, 1990; Pavelka *et al.*, 1991; Pazzani *et al.*, 1993; Wunder *et al.*, 1994). Mutations within region 3 genes result in cytoplasmic polysaccharide of lower molecular weight than the cell surface expressed polymer, which are associated with the inner face of the cytoplasmic membrane (Boulnois *et al.*, 1987; Kroncke *et al.*, 1990; Smith *et al.*, 1990; Pavelka *et al.*, 1994; Pigeon and Silver, 1994). This observation is consistent with the synthesis of polysaccharide occurring on the inner face of the cytoplasmic membrane and the role of encoded proteins in export of polysaccharide across the cytoplasmic membrane (Boulnois and Jann, 1989; Boulnois and Roberts, 1990). Analysis of the predicted amino acid sequence of the encoded KpsM and T proteins revealed that they are members of the ATP-binding cassette (ABC) transporters (Smith *et al.*, 1990; Pavelka *et al.*, 1991) (Table 1.3). The two proteins are homologous to BexAB proteins encoded in *H. influenzae* cap locus and the CtrCD proteins of the *N. meningitidis* capsule gene cluster and constitute an inner membrane polysaccharide export system in these bacteria (Kroll *et al.*, 1990; Smith *et al.*, 1990; Frosch *et al.*, 1991; Pavelka *et al.*, 1991, 1994; Reizer *et al.*, 1992). Mutations in region 3 of the K5 capsule gene cluster can be complemented by cloned genes from *Actinobacillus pleuropneumoniae* (Ward and Inzana, 1995), suggesting that there is conservation in the ABC transport systems involved in group II polysaccharide export in Gram-negative bacteria.

In the ABC-mediated export of group II polysaccharide in *E. coli*, the KpsT protein is the ATP-binding component while the KpsM is the integral membrane-spanning unit. Interaction between KpsT and ATP was demonstrated in structure-function studies of KpsT mutants and binding of photolabelled analogues of ATP (Pavelka *et al.*, 1994). Hydropathy plot of the predicted amino acid sequence indicated that the KpsM protein contains six hydrophobic membrane spanning domains interspersed with short stretches of hydrophilic residues (Smith *et al.*, 1990; Pavelka *et al.*, 1991). Topological analysis using gene fusion approaches confirmed this six membrane spanning organisation, with the N- and C-terminal domains in the cytoplasm (Pigeon and Silver, 1994). Site-directed mutagenesis and linker-insertion identified the N-terminus, the first cytoplasmic loop and a small hydrophobic domain (SV-SVI linker) near the C-terminus as regions that are important for KpsM function (Pigeon and Silver, 1994). Thus, the KpsM and
T structures are consistent with the model of an ABC transporter having a polytopic integral membrane protein and an ATP-binding protein (Higgins, 1992), where KpsT presumably functions in hydrolysing ATP in the KpsM-mediated export of group II polysaccharide (Smith et al., 1990; Pavelka et al., 1991, 1994; Pigeon and Silver, 1994).

Region 1 of the kps gene cluster encodes proteins that are involved in the cell surface expression of group II polysaccharide in *E. coli* (Boulnois and Roberts, 1990; Silver and Vinr, 1990; Pazzani et al., 1993a,b). Mutations within the kpsE and D genes result in periplasmic accumulated polysaccharide (Boulnois et al., 1987; Bronner et al., 1993a) suggesting a role for the KpsE and D proteins in the export of polysaccharide across the periplasmic space. The KpsE protein is homologous and shares similar hydropathy plot profile to the BexC protein of *H. influenzae* and the CtrB protein of *N. meningitidis*, both of which have been implicated in the export of polysaccharide in these bacteria (Kroll et al., 1990; Frosch et al., 1991). Antiserum has been raised against the purified KpsE protein of *E. coli* K5 (Rosenow et al., 1995a). The protein has been shown to localise in the membrane, and preliminary data on the topology indicated that KpsE has an amino terminus in the cytoplasm and a large periplasmic domain of approximately 302 amino acids (Esuehm and Roberts, 1995; Rosenow et al., 1995a). The large periplasmic domain is likely to interact with the KpsD protein and may be functionally important in the export of polysaccharide to the cell surface.

KpsD is a periplasmic protein with a typical N-terminal signal sequence (Pazzani et al., 1993a,b; Wunder et al., 1994). The predicted amino acid sequence of KpsD is not homologous to proteins encoded in either the *H. influenzae* and *N. meningitidis* capsule gene clusters (Pazzani et al., 1993a,b), but is homologous to the ExoF protein which is involved in the expression of succinoglycan in *Rhizobium meliloti* (Muller et al., 1993). Thus the precise role of KpsD in polysaccharide export is not fully elucidated, but mutants lacking the encoded product are not capable of exporting polysaccharide to the cell surface (Bronner et al., 1993a; Wunder et al., 1994). Outer membrane proteins are not encoded in the group II capsule gene clusters of *E. coli* and it is possible that the KpsD protein may be acting in a similar role analogous to outer membrane proteins in the cell surface expression of group II capsule.
The KpsC and S proteins are localised to the cytoplasm and are associated with the inner face of the cytoplasmic membrane (Rigg and Roberts unpublished). Analysis of mutations in either the kpsC or S genes revealed cells with polysaccharide that lack both phospholipid and KDO, located within 'holes' in the cytoplasm (Bronner et al., 1993a,b). Phospholipid and KDO are usually found at the reducing end of cell surface expressed polysaccharide (Kroncke et al., 1990; Finke et al., 1991). The KpsC and S proteins are respectively homologous to LipA and B proteins encoded in the capsule gene cluster of N. meningitidis and both Lip proteins have been proposed to function in lipid substitution (Frosch and Muller, 1993). Based on results obtained from the KpsC and S mutants, it was suggested that phosphatidyl-KDO might form a structural motif that is recognised prior to export of polysaccharide across the cytoplasmic membrane by the KpsM and T proteins (Roberts, 1995). The exact functions of these two proteins are yet to be demonstrated.

The kpsU gene encodes a protein of 27 kDa that have been suggested to provide KDO for the synthesis of 2-phosphatidyl-KDO, needed for ligation to the polysaccharide chain (Bronner et al., 1993a). In the K5 capsule gene cluster, mutations in kpsE and kpsD genes result in the accumulation of phosphatidyl-KDO-linked polysaccharide at the cell surface and in the periplasm, whereas polysaccharide in the cytoplasm lacks phosphatidyl-KDO (Bronner et al., 1993a). This has led to the suggestion that phosphatidyl-KDO substitution of the K5 polysaccharide is a prerequisite for translocation of the polysaccharide (Roberts, 1995, 1996), as is the case with N. meningitidis capsular polysaccharide where phosphatidic acid substitution has been demonstrated (Frosch and Muller, 1993). However, kpsU mutants express some capsular polysaccharide at the cell surface, indicating that this is likely not the case. Therefore, the role of KpsU in the cell surface expression of capsular polysaccharide in E. coli remains poorly defined.

The mechanisms involved in the export of group II polysaccharide are beginning to unfold. The observation of a conserved system involved in the export processes in other Gram-negative bacteria has helped to facilitate research into export of the K5 capsular polysaccharide. Albeit there is a conserved theme in the export process, significant differences do exist. For example, the lack of periplasmic protein in N. meningitidis and H. influenzae, and an outer membrane protein in E. coli and the role of KDO in the transport process raises interesting questions as to the organisation of the processes involved in the last stages of cell surface expression of group II polysaccharide.
capsular antigens. The analysis of proteins encoded by region 1 genes will help to further our understanding on the cell surface expression of the K5 antigen.

1.7 Capsule-specific Bacteriophages

Binding of the phage to the bacterial cell is often the first and crucial step in bacteriophage infection. This action is as a result of specific interaction between the phage and its receptor on the bacterial cell surface (Makela, 1985). Hence, one can test for the presence or absence of such receptors by using a set of phages whose receptor specificities are known. According to Makela (1985) this has proved to be a useful approach in the study of several bacterial cell surface structures and mutants lacking these structures. The receptor specificity of a phage is determined by its binding protein, which usually is the tail fibre or part of it. Some phages especially the large DNA phages have been shown to have two sets of tail fibres which have different binding specificities (Zorzopulos et al., 1982).

Bacteriophage infection and the subsequent killing of the first target bacterium usually means liberation of a burst of new phage particles that bind to other bacteria in the vicinity and the result is a plaque of lysis in a lawn of bacterial growth. However, bacteria can be resistant to the killing action of a phage even after phage binding to specific receptor (Pelkonen et al., 1992). In the absence of specific receptor, the most common cause of such resistance is immunity conferred by a resident prophage (Duckworth et al., 1981). Other mechanisms of resistance which are only but poorly understood include 'superinfection exclusion' at the bacterial cell surface, restriction by endonuclease action and abortive infection (Makela, 1985).

Capsular polysaccharides of bacteria often interact with bacteriophages, either by blocking access to phage receptor sites beneath or by providing specific adsorption sites. A number of bacteriophages specific for E. coli capsular polysaccharides (K phages) have been described (Stirm, 1968; Stirm and Freund-Molbert, 1971; Lindberg, 1977). These K phages are very specific; they do not bind to noncapsulated mutants nor to related strains with different capsular structures and thus of different serotype (Jann and Jann, 1983). Structurally identical capsular polysaccharide from different bacterial strains will be sensitive to the same phage. This is exemplified by the K54 capsular antigens of Klebsiella aerogenes A3 and that of E. coli O8:K27, both a receptor
for the bacteriophage F34 (Sutherland, 1977; Sutherland et al., 1970). The two polysaccharides are very similar with the same backbone structure but substituted with galactose in *E. coli* K27 instead of glucose in *K. aerogenes* K54 antigen (Conrad et al., 1966).

As a result of their binding specificity, the capsule-specific phages can be used as indicators of the presence of certain structures in the polysaccharide. The identification of polysaccharide structures from a number of bacteria including *E. coli* (Gross et al., 1977; Gupta et al., 1982) species of *Klebsiella* (Ravenscroft et al., 1987; Parolis et al., 1988), *Azotobacter* (Pike and Wyss, 1975; Davidson et al., 1977) and *Rhizobium* (Barnet and Humphrey, 1975; Amemura et al., 1983) have benefited from this approach. For polysaccharide structures that are poor immunogens, capsule-specific phages can be used in identification of bacteria responsible for infection. For instance, it is particularly difficult to prepare high titre antisera to the *E. coli* K1 and K5 capsular types (Jann and Jann, 1983), and the K1- or K5-specific phages are of great value in the serotyping of these strains (Gross et al., 1977; Nimmich et al., 1981; Gupta et al., 1982). The K phages have been employed with great success in the epidemiological studies of bacterial etiologic agents, notably in the typing of *Staphylococcus aureus* (Blair and Williams, 1961), *Salmonella* (Wilkinson et al., 1972; Lindberg, 1977; Hickman-Brenner et al., 1991), *Yersinia* (Kawaoka et al., 1983) and *E. coli* (Lindberg, 1977). Capsule-specific phages have also been used to isolate non-capsulated mutants (Stirm, 1968).

Bacteriophage specificity is not restricted to the K antigens. Phages with binding specificities to lipopolysaccharide (O phages) such as the ε15, ε34 and Ω8 phages (Kanegasaki and Wright, 1973; Reske et al., 1973; Prehm and Jann, 1976) and rough (R) antigens like the classical T phages (T3, T4, T7) (Prehm et al., 1976) have been described. The O phages are the most frequently isolated phages from sewage or other sources when using indicator bacteria with O antigen as host (Makela, 1985). The commonly used transducing phage P22 of *Salmonella* is a good example of the specificity of the O phages. It requires smooth type LPS with long O side chains as opposed to semirough (SR) or rough (R) mutants (Makela and Stocker, 1984). Most of the phages used for typing in clinical epidemiological investigations of *Yersinia enterocolitica* are O-specific (Kawaoka et al., 1983). The specificity of phages that bind to the LPS core (R-specific) have been best studied in *Salmonella typhimurium* (Wilkinson et al., 1972) and have been used in the characterisation of R mutants (Makela and Stocker, 1984). In fact, as reported by Makela (1985) the
R-specific phage, FO is useful for differentiation of rfa mutants (with the exception of rfaL) with incomplete LPS core from other R types. Many of the phages that bind to the core LPS do not bind to smooth (S) bacteria due to steric hindrance exerted by the O polysaccharide (Makela, 1985).

1.7.1 Bacteriophage Lyases Specific for Capsules

Capsule phages has been shown to first bind to receptors in the capsule or O polysaccharide remote from the outer membrane surface and then specifically depolymerise the polysaccharide to which they are bound (Stirm and Freund-Molbert, 1971; Lindberg, 1977). The depolymerising enzyme is borne in the spikes or tail fibre of the phage and it is required both for binding and degradation of bacterial polysaccharide for the phage to reach the cell surface (Stirm, 1968; Stirm and Freund-Molbert, 1971). This depolymerisation event of the polysaccharide has been elegantly demonstrated by electron microscopy (Bayer et al., 1979; Bayer and Bayer, 1994). A large number of bacteriophage-associated enzymes which degrade carbohydrate-containing polymers on the bacterial cell surface have been described and a few of these are polysaccharide lyases (Sutherland, 1995).

In general, all lyase enzymes that cleave anionic polysaccharide polymers exhibit a common mode of action, by a β-elimination mechanism. The specificity of these lyases is related to this eliminase mechanism (Linhardt et al., 1986; Desai et al., 1993), by which certain glycosidic (usually 1,4 β- or 1,4 α-linked) bonds between hexosamine and uronic acid are cleaved (Jandik et al., 1994; Sutherland, 1995). A product of this cleavage is a modified non-reducing terminus to form unsaturated uronic acid and this UV chromophore formed in the non-reducing termini of the small oligosaccharide products is useful for their detection (Jandik et al., 1994). Most of the enzymes, with the exception of xanthan lyases (Lesley, 1961) are randomly endolytic in their action, cleaving the main chain of the polysaccharide structure (Jandik et al., 1994; Sutherland, 1995). The products of lyase action may in a few cases be monosaccharides but are more commonly oligosaccharides ranging in size from a degree of polymerisation (DP) of 2 to 5. Polysaccharide lyases act on various substrates which may be found in either eukaryotes or prokaryotes, the later source according to Sutherland (1995) providing most diverse range of substrates. A few of the polysaccharide substrates are common or similar to both types of organisms and thus will be cleaved to some degree of specificity by the same enzyme. Notable among these substrates are the polysialic acid which is a polymer
present in both humans and the bacteria *E. coli* K1 and *N. meningitidis* serogroup B strains (Livingston et al., 1988; Troy, 1992). Heparin and desulphatoheparin which are common in eukaryotes has similar structure to *E. coli* K5 polymer (Vann et al., 1981; Casu et al., 1994). Likewise, the *E. coli* K4 antigen has the structure of a chondroitin backbone to which D-fructosyl side chain residues are attached (Rodriguez et al., 1988).

Among the polysaccharide lyases found in phage preparations are alginate lyases (Davidson et al., 1977a) and those cleaving various linkages in both homo- and hetero-polymers produced by several bacterial species including *S. aureus* (Farrel et al., 1995), *Streptococcus* (Niemann et al., 1976; Hynes and Feretti, 1989), *Rhizobium trifolii* (Barnet and Humphrey, 1975; Higashi and Abe, 1978; Hollingsworth et al., 1984), *Klebsiella* (Ravenscroft et al., 1987; Parolis et al., 1988), *Pseudomonas* (Davidson et al., 1977a,b) and *E. coli* (Stirm and Freund-Molbert, 1971; Hallenbeck et al., 1987; Petter and Vimr, 1993; Long et al., 1993, 1995; Gerardy-Schahn et al., 1995). The *Azotobacter vinelandii* phage-borne lyase enzyme was shown to degrade alginate rich in mannuronate residues (Davidson et al., 1977a) yielding oligosaccharide products containing mannuronate and unsaturated uronic acid and guluronate residues (Davidson et al., 1977b). The hyaluronate lyase borne by streptococcal bacteriophage described by Niemann et al. (1976) yielded a tetrasaccharide and an octasaccharide along with two higher oligomers from its substrate. Each oligosaccharide were characterised by the unsaturated uronic acid at the non-reducing end and N-acetyl glucosamine at the reducing terminus. Lyases found in virulent phage lysates of *R. trifolii* have been used for structural analysis of the repeating unit of the *R. trifolii* polysaccharide (Amemura et al., 1983; Hingsworth et al., 1984). The enzyme cleaves a 1,4 β-D-glucosyl-D-glucuronic acid linkages with resultant oligosaccharide products composed of D-glucose: D-galactose: D-glucuronic acid: 4-deoxy-L-threo-hex-4-enopyranosyluronic acid in the ratio 5:1:1:1. These oligosaccharides have been found to be biologically active products capable of binding trifolin A and stimulating root hair infection in clover (Abe et al., 1984).

A number of polysaccharide lyases from bacteriophages useful in the structural analysis of capsular polysaccharides from *E. coli* have been identified. The enzyme from *E. coli* K28 phage cleaves the α-D-glucosyl (1,4)-β-D-glucuronic acid linkage in the acetylated tetrasaccharide repeat unit of the polysaccharide (Altman et al., 1986). The lyase from *Klebsiella* phage 45 was capable of cleaving the β-D-mannosyl (1,4)-β-D-glucuronic acid linkage
in the *E. coli* K55 polysaccharide (Anderson and Parolis, 1989). In the trisaccharide repeat unit of the K55 antigen structure, the mannose carried 4,6-linked pyruvate ketals (on up to 40% of residues) and O-2-acetyl groups, whereas the *Klebsiella* K5 polymer (which is also degraded by the same phage) carried only an O-2-acetyl group on the glucose residue. Thus the specificity of this enzyme is on the linkage in the backbone structure.

In all, the action of lyase enzymes provides oligosaccharide products that are more amenable to structural analysis by physical methods. Of special interest are the phage-borne lyases specific for the low molecular weight K antigens, particularly the K1 and K5 antigens of *E. coli*.

### 1.7.1.1 Bacteriophage-borne Lyase Specific for *E. coli* K1 Capsule

Several bacteriophages specific for polysialic acid capsule of *E. coli* K1 have been described (Gross et al., 1977; Kwiatkowski et al., 1982, 1983; Hallenbeck et al., 1987). The phages designated members of the PK1 family recognise the capsular polysialic acid as a primary cell surface receptor by their capsule-degrading endosialidases, cleaving their substrate by a common endo-eliminase mechanism (Kwiatkowski et al., 1982; Fine and Makela, 1985; Hallenbeck et al., 1987; Pelkonen et al., 1989). *E. coli* K1 has been implicated in a number of infections (Robbins et al., 1974) with high mortality rates in cases of meningitis (Robbins et al., 1974; Ørskov et al., 1977). As cited by Long et al. (1995), it was proposed by Taylor (1987) that the PK1 phage-induced endosialidase could be used in the diagnosis and therapy of *E. coli* K1 meningitis, septicaemia or bacteraemia due to the enzyme’s high specificity for hydrolysing α-2,8-sialosyl linkages of the K1 polymer. Polysialic acid is an onco-developmental antigen in human kidney and brain and may contribute to the invasive and metastatic potential of some tumors (Livingstone et al., 1988; Roth et al., 1988). Thus as reported by Tomlinson and Taylor (1985) bacteriophage E endosialidase has the potential of being used as probe for neoplastic tissues displaying the polysialic acid marker.

To date phages that recognise the K1 capsule are the only known source of polysialic acid lyase (Kwiatkowski et al., 1983; Finne and Makela, 1985). Most of the phage-borne enzymes have been purified and shown to depolymerise both O-acetylated and non-O-acetylated α-2,8-linked homopolymers of N-acetyleneuraminic acid from a variety of sources (Tomlinson and Taylor, 1985; Hallenbeck et al., 1987; Long et al., 1993). Limited lyase activity was observed when the alternating α-2,8 and α-2,9-linked N-acetyleneuraminic
acid homopolymer of *E. coli* K92 antigen was used as substrate (Tomlinson and Taylor, 1985). In the studies reported by Hallenbeck *et al.* (1987) both oligo- and poly-sialosyl units were substrates and examination of the intermediate products showed that cleavage occurred at random sites on the poly(sialosyl) chains, as opposed to progressive depolymerisation from an initiated cleavage site at one end. Some of the genes encoding these enzymes have been cloned and the amino acid sequence determined (Petter and Vimr, 1993; Gerardy-Schahn *et al.*, 1995; Long *et al.*, 1995), and detail analysis of the encoded proteins will further our understanding on the structural-functional relationship of this group of polysaccharide lyases.

1.7.1.2 Bacteriophage-borne Lyase Specific for *E. coli* K5 Capsule

Gupta *et al.* (1982) reported the isolation of a phage specific for the K5 antigen, to aid in the clinical diagnosis of infections caused by *E. coli* K5 strains. They also found that this phage was capable of degrading the *E. coli* K5 polysaccharide. The capsular polysaccharide from *E. coli* K5 strain is formed from disaccharide repeating units of 4-β-D-glucuronosyl-1,4-α-N-acetyl-D-glucosamine and essentially represents a structural intermediate of heparin biosynthesis (Vann *et al.*, 1981) and is similar to other intermediates required as substrates to generate products with high affinity for antithrombin activity (Casu *et al.*, 1994). The K5 polymer and those of heparin are substrates for polysaccharide lyases (Linhardt *et al.*, 1986; Lohse and Linhardt, 1992; Jandik *et al.*, 1994).

The K5 phage-borne degrading enzyme was recently characterised and shown to be a lyase, cleaving its substrate through a β-eliminative mechanism, and yielding oligosaccharides characterised mainly of hexa-, octa- and deca(saccharides with 4,5-unsaturated glucuronic acid (Δ4,5GlcA) at their nonreducing end (Hänfling *et al.*, 1996). The few documented reports on lyases that are specific for heparin or heparin-like compounds come from the use of the bacterium *Flavobacterium heparinum*. Three distinct lyases have been purified from this bacterium (Lohse and Linhardt, 1992). These are heparin lyase I (EC 4.2.2.7), II (no EC number) and III (heparitinase, EC 4.2.2.8) with corresponding molecular weights of 42.8, 84.1 and 70.8 KDa. All had high alkaline pH optima in the range of 9.0-10.0 and isoelectric points of 8.9 to 10.1.

The heparin lyases from *F. heparinum* act in a random endolytic fashion and their specificity is related to cleavage of the glycosidic linkages between
hexosamine and uronic acid to yield oligosaccharide products characterised by unsaturated uronic acid residues at the non-reducing end (Jandik et al., 1994). The three heparin lyases act on heparan sulphate, with heparin lyase I degrading the substrate much more slowly and all but heparin lyase III act on heparin (Linhardt et al., 1982; Jandik et al., 1994) (Figure 1.6). Lyases with specificity for heparin and heparan sulphate have found many applications including the determination of structure of such polymers (Linhardt et al., 1991, 1992), preparation of new therapeutic agents (Linhardt et al., 1988; Casu et al., 1994) and the analysis of glycosaminoglycans found in tissues and biological fluids (Linhardt et al., 1992).

Figure 1.6: Action pattern of heparin lyases on both heparin and heparan sulphate. Arrows indicate site of cleavage by the polysaccharide lyases, and the asterik denotes the slow action of heparin lyase I on heparan sulphate. Figure was adapted from Jandik et al. (1994).
Unlike in *E. coli K*1, where phages represent the only source of lyase enzyme specific for the capsule, certain strains of *E. coli K*5 are known to express lyase enzyme that are specific for the K5 polysaccharide. These strains include *E. coli* O4:K5, O10:K5 and O117:K5 (K. Jann, personal communication). The encoded enzymes in these strains have not been studied in detail, but it is known that these K5 strains express low molecular weight capsular polysaccharide (K. Jann, personal communication).
Aims

The aims of this thesis are two fold. First, to determine the topology of the KpsE protein. Secondly, to clone, determine the nucleotide sequence, and overexpress the K5-specific bacteriophage-borne lyase enzyme. The usefulness of this enzyme in the study of K5 polysaccharide will be highlighted.
Chapter 2

Materials and Methods

2.1 Bacterial Strains and Plasmids

The bacterial strains and plasmids that were used in this study are listed in tables 2.1 and 2.2 respectively.

Table 2.1: Bacterial Strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli LE392</td>
<td>F', hsdR514(r'k,m'm,k), SupE44, SupF58, lacY1, galK2, galT22, metB1, trp R55,λ⁺</td>
<td>Murray et al., 1977</td>
</tr>
<tr>
<td>E. coli JM101</td>
<td>supE, thi Δ(lac -proAB), F'[traD36, proAB⁺, lac19, lacZAM15]</td>
<td>Yanisch-Perron et al, 1985</td>
</tr>
<tr>
<td>E. coli JM109(DE3)</td>
<td>ε14(mcr A), recA1, endA1, gyr A96, hisd R17(r'k,m'm,k), supE44, thi Δ(lac -proAB), F'[traD36, proAB⁺, lacZAM15, λ(DE3)</td>
<td>Yanisch-Perron et al., 1985</td>
</tr>
<tr>
<td>E. coli SURE™</td>
<td>mcr A, (Δmcr BC-hsd RMS-mcr Y71, sup E44, thi 1, λ, gyr A96, rel A1, lac , rec B, rec J, shc C, umu C::Tn5 (lac I²), wvr C, [F, pro AB⁺, lac ZAM15, Tn10 (tet)]</td>
<td>Stratagen® Cloning Systems</td>
</tr>
<tr>
<td>E. coli CC118</td>
<td>araD139, (ara, leu Y7697, lac X74, phoA Δ20, galE, galK, thi , rpsE, rpoB, arg F42, rec A1</td>
<td>Michaelis et al., 1983</td>
</tr>
<tr>
<td>E. coli SM10pir</td>
<td>F', thi -1, thr -1, lev B6, supE44, tonA 21, lac Y 1, rec A4::RP4-2-Tc::Mu Km'</td>
<td>Simon et al., 1983</td>
</tr>
</tbody>
</table>
Table 2.1 (continued)

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> SY327Xpir</td>
<td><em>F</em>, ara D, Δlac-proA, arg Eam, rec A56, rifA, nal A</td>
<td>Miller and Mekalanos, 1988</td>
</tr>
<tr>
<td><em>E. coli</em> MS101</td>
<td><em>rpsL</em>, serA, K5*</td>
<td>Stevens et al., 1994</td>
</tr>
<tr>
<td><em>E. coli</em> MSPE100</td>
<td><em>rpsL</em>, serA, ΔkpsE, K5*</td>
<td>This study</td>
</tr>
</tbody>
</table>

Table 2.2: Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Characteristic</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pYZ4</td>
<td>Km*, Cloning and expression vector</td>
<td>Zhang and Broome-Smith, 1990</td>
</tr>
<tr>
<td>pLH21</td>
<td>Cm*, Contains BlaM cassette</td>
<td>Broome-Smith</td>
</tr>
<tr>
<td>pUC19</td>
<td>Ap*, High copy number cloning vector</td>
<td>Yanisch-Perron et al., 1985</td>
</tr>
<tr>
<td>pACYC184</td>
<td>Cm*, Cloning vector</td>
<td>Chang and Cohen, 1987</td>
</tr>
<tr>
<td>pRT733</td>
<td>Km*, Suicide vector, λpir, TnphoA</td>
<td>Taylor et al., 1989</td>
</tr>
<tr>
<td>pPC6</td>
<td>Cm*, Cloned <em>E. coli</em> K5 capsule gene cluster except kpsF in pACYC184</td>
<td>Pazzani, 1993</td>
</tr>
<tr>
<td>pH18</td>
<td>Ap*, Hind III fragment encoding region 1 minus kpsF of <em>E. coli</em> K5 capsule gene cluster in pUC18</td>
<td>Pazzani, 1993</td>
</tr>
<tr>
<td>pCR6</td>
<td>Ap*, Sme I-Hinc II kpsE fragment in pCE30</td>
<td>Rosenow et al., 1995a</td>
</tr>
</tbody>
</table>

52
2.1.1 Growth Conditions and Media

Bacteria were routinely grown at 37°C in Luria broth (LB; 0.5% NaCl, 0.5% yeast extract, and 1% tryptase peptone) with 1.5% bacteriological agar (BBL) added as required. M9 minimal medium (0.6% Na2HPO4·7H2O, 0.3% KH2PO4, 0.05% NaCl, 0.1% NH4Cl, 0.4% glucose, 10mM MgCl2, 10μgml⁻¹ thiamine) with 1.5% agar was used where necessary. For soft top agar 0.6% agar was used. Where necessary media were supplemented with appropriate antibiotics at the following concentrations: ampicillin 100μgml⁻¹, chloramphenicol 25μg/ml⁻¹, kanamycin 25μg/ml⁻¹, streptomycin 25μg/ml⁻¹, and tetracycline 25μg/ml⁻¹. Antibiotics and amino acids were purchased from Sigma Chemical Company Ltd. Solid agar media were supplemented with 0.2mg/ml⁻¹ X-gal and 1mM IPTG (Novabiochem), and XP (Sigma) where required.

Bacterial cells were harvested by centrifugation at 3300g for 5 mins at 4°C in Sorval centrifuge (Dupont). Small volumes of culture were centrifuged in a bench top minifuge at 13,400g for 5 mins at room temperature.

2.1.2 Phage Assay

Cells were grown to mid-logarithmic phase at 37°C with shaking at 200rpm, harvested and then resuspended in 10mM MgSO₄ solution. Resuspended cells were incubated with dilutions of capsule-specific phage at ambient temperature for 20 mins. Soft top agar was added and then overlaid onto L-agar plates containing the appropriate antibiotics. Plates were allowed to set, then incubated at 37°C and examined after 5 hours and overnight for plaques.

2.1.3 Determination of Ampicillin Resistance of E. coli JM101 Producing KpsE-BlaM Hybrid Proteins

The minimum inhibitory concentration (MIC) required to prevent colony formation by single cell was determined based on the method described by Broome-Smith and Pratt (1986). Overnight cultures of E. coli JM101
harbouring different *kpsE-blaM* clones were serially diluted in L-broth to provide $10^{-1}$-$10^{-9}$ dilutions. A drop corresponding to 5µl from these dilutions were spotted on different segments of L-agar plates containing Kanamycin (25µgml$^{-1}$), IPTG (1mM final concentration) and various range (5µgml$^{-1}$-200µgml$^{-1}$) of Ampicillin. The formation of a single colony was examined after overnight incubation at 37°C.

2.2 Transfer of Plasmid DNA into Bacterial Cells

Bacterial cells were transformed using variations of either the Calcium chloride (CaCl$_2$) (Mandel and Higa, 1970) or the high efficiency electroporation (Dower et al., 1988) methods. DNA was also delivered into bacterial cells through conjugation.

2.2.1 Calcium Chloride Method

In the CaCl$_2$ method, 100µl of an overnight culture grown at 37°C was diluted 1:100 in 10ml of L-broth and grown at 37°C with shaking to mid-exponential phase (OD$_{600}$~0.5). Cells were harvested (3300g at 4°C for 5 mins), washed once in 10ml of ice-cold 10mM NaCl solution, pelleted and resuspended in 4ml ice-cold 100mM CaCl$_2$. The cells were made competent by incubation on ice for 30 mins and collected by gentle centrifugation (1800g for 5 mins at 4°C). Pelleted cells were resuspended in 1ml ice-cold 100mM CaCl$_2$ and then used immediately in transformation. For transformation, 100µl aliquot of competent cells was mixed with 5-20µl of plasmid DNA and placed on ice for 1 hr. Cells were heat shocked at 42°C for 3 mins and then immediately mixed with 0.5ml L-broth. The transformed cells were incubated at 37°C for 1 hr, after which 100µl aliquot was plated unto each L-agar plates containing the appropriate antibiotics. Plates were then incubated overnight at 37°C.

2.2.2 Electrotransformation Method

For electrotransformation 100µl of an overnight culture was back-diluted 1:100 with 10ml of L-broth and grown to mid-exponential phase (OD$_{600}$~0.5). Cells were pelleted by centrifugation at 3300g for 10 mins at
4°C. Pelleted cells were washed three times in 10ml of ice-cold nanopure water and then once in 10% v/v glycerol with centrifugation between washes. The cell pellet was resuspended in 80μl of 10% glycerol and used immediately in transformation. Aliquots of 1-2μl of DNA was then mixed with 40μl of the competent cells and transferred into an ice-cold 2mm Gene Pulser™ cuvette (Biorad). A high voltage was delivered through the sample using a Biorad Gene Pulser™ with pulse control parameters at 25μF capacitance, 200 Ω resistance and 2.4KV voltage. Time constants were typically in the range of 4.6-4.8 msec, depending on the purity (salt content) of the DNA samples. Immediately after pulsing a ml of ice-cold SOC recovery medium (20gl⁻¹ tryptase peptone, 5gl⁻¹ yeast extract, 0.57gl⁻¹ 10mM NaCl, 0.175gl⁻¹ 2.5mM KCl, 2.025gl⁻¹ 10mM MgCl₂, 2.45gl⁻¹ 10mM MgSO₄, 3.6gl⁻¹ 20mM glucose) was mixed with cells, followed by incubation for 1 hr at 37°C before plating out 100μl aliquots unto each L-agar plates containing the appropriate antibiotics.

2.2.3 Conjugation Procedures

2.2.3.1 TnphoA Mutagenesis

Random phoA insertions were generated in plasmid DNA using filter mating experiments between the recipient (E. coli LE392pCR6) and the donor (E. coli SM10λpir pRT733) strains. Plasmid pRT733 is a suicide vector (Table 2.2) based upon the λpir system and thus can only replicate in a Pir+ background and have been described (Manoil and Beckwith, 1985). Briefly, the donor and recipient were grown to OD₆₀₀~0.4. Cells from 1ml of each culture were pelleted together and resuspended in 50μl of L-broth. The suspension was then spread over a microcellulose filter (Hybond-C, Amersham) on L-agar plates. Plates were incubated at 37°C and cells washed off the filter after 4 hrs with 1x M9 salt solution. Conjugants resulting from the mating were selected on M9 salt minimal medium containing the required antibiotics. Plasmid DNA extracted from conjugants was used to transform E. coli CC118 and PhoA+ mutants were identified as blue colonies on L-agar plates containing the chromogenic substrate, 5-bromo-4-chloro-3-indonyl phosphate (XP; 40μgml⁻¹).
2.2.3.2 Homologous Gene Recombination

Homologous gene recombination using a positive selection suicide vector, pCVD422 was carried out as previously described (Donnenberg and Kasper, 1991). The suicide vector pCVD422 is also based upon the \( \lambda pir \) system (see above) and therefore ideal for gene replacement or introduction of mutations into the chromosome of \( \lambda pir^+ \) strains. pCVD422 contains the \( sacB \) gene, which confers sucrose sensitivity and thus allow the positive selection for the loss of vector sequences after homologous recombination. This system was used in filter mating experiments as described earlier to introduce mutations in the chromosome.

2.3 Procedures for Recombinant DNA Manipulation

2.3.1 Extraction and Purification of Plasmid DNA

Small scale preparation of plasmid DNA was carried out using the alkaline lysis method. Cells pelleted from 1.5ml of an overnight culture were resuspended in 100\( \mu l \) ice-cold solution I (50mM glucose, 25mM Tris-HCl pH 8.0, 10mM EDTA and 5mgml\(^{-1}\) lysozyme) and left on ice for 30 mins, followed by the addition of 200\( \mu l \) of freshly made solution II (0.2M NaOH and 1% SDS). The content of the tube were then mixed by inverting the tube five times and placed on ice. After 5 mins on ice 150\( \mu l \) of ice-cold solution III (5M potassium acetate buffer pH 4.8, which consists of 3M potassium and 5M acetate) was added, again followed by gentle mixing and a precipitate was allowed to form by standing the tube on ice for 5 mins. The supernatant was recovered after centrifugation and extracted once with equal volume (about 400\( \mu l \)) of phenol:chloroform (1:1), vortexed for 5 secs and centrifuged for 1 min. The upper aqueous phase was collected and mixed with 1ml of ethanol at room temperature. After 5 mins the DNA precipitate was sedimented by centrifugation in a bench top minifuge for 5 mins. The ethanol was aspirated and discarded, leaving the DNA pellet which was then dried for 5 mins using a vacuum desicator. DNA pellet was resuspended in 50\( \mu l \) of sterile nanopure water and stored at -20°C.

Large scale preparation of plasmid DNA was carried out using either the alkaline lysis method described by Birnboim and Doly (1979) followed by caesium chloride-ethidium bromide density gradient centrifugation or the Maxi prep kit (Qiagen). For the caesium chloride-ethidium bromide
density gradient method, cells from a 400ml overnight culture were harvested in two large pots by centrifugation. Each cell pellet was resuspended in 5ml of ice-cold solution I, transferred to a 30ml Oakridge tube and left to stand on ice for 30 mins. Freshly made solution II (10ml) at room temperature was added, gently mixed and left for a further 10 mins on ice. Ice-cold solution III (7.5ml) was added, again mixed gently and incubated for a further 10 mins on ice. Cell debris was removed from the plasmid preparation by centrifugation at 45°C for 20 mins at 35000g. The supernatant (18ml) was mixed with 12ml of isopropyl alcohol and left to stand at room temperature for a minimum of 15 mins. The resultant DNA precipitate was pelleted by centrifugation in 30ml Corex tubes at 3500g for 30 mins at 20°C. The DNA pellet was air dried for 15 mins and resuspended in sterile nanopure water to a final volume of 17ml. Caesium chloride (17g) was added (ie to a final concentration of 1mg/ml⁻¹), allowed to dissolve and the solution was transferred to a 30ml Sorval vertical rotor centrifuge tube. Ethidium bromide was added to a final concentration of 50μg/ml⁻¹ and the tube filled with mineral oil, balanced and crimp-sealed. Chromosomal and plasmid DNA were separated by centrifugation at 40000 rpm using a Sorval TV850 fixed-angle rotor in a Sorval OTD 60 ultracentrifuge for 20 hrs at 20°C. DNA was visualised under UV light and the lower plasmid DNA band was extracted. Ethidium bromide was extracted by equilibration with caesium chloride-saturated isopropanol. Caesium chloride was removed by exhaustive dialysis against distilled water at room temperature. The DNA solution was aliquoted and stored at -20°C.

The maxi prep was performed according to manufacturers instructions (Qiagen). Briefly, 150ml (for high copy number plasmid) or 500ml (for low copy number plasmid) of an overnight grown culture was harvested by centrifugation at 4°C for 15 mins at 6000xg. The resulting pellet was resuspended in 10ml of ice-cold buffer P1 (100μg/ml⁻¹ RNase A, 50mM Tris-HCl, 10mM EDTA pH 8.0). Ten millilitre (10ml) of buffer P2 (200mM NaOH, 1% SDS) was added, gently mixed and left to stand at room temperature for 5 mins. A 10ml aliquot of ice-cold buffer P3 (3M KAc pH 5.5) was added, again gently mixed by inverting the tube five times and then incubated on ice for 20 mins. Cell debris were removed by centrifugation at 4°C for 30 mins at 30000xg and the clear supernatant was promptly recovered. QIAGEN-tip 500 column was equilibrated with 10ml buffer QBT (750mM NaCl, 50mM MOPS, 15% ethanol pH 7.0, 0.15% Triton
X-100) and the supernatant was applied to the column. Possible contaminants were removed by applying a total of 60ml wash buffer QC (1.0 M NaCl, 50mM MOPS, 15% ethanol pH 7.0) and allowing the buffer to flow through the column. DNA was eluted with 15ml of buffer QF (1.25M NaCl, 50mM Tris-HCl, 1.5% ethanol pH 8.5) and precipitated with 0.7 volumes of isopropanol, previously equilibrated to room temperature. The DNA was immediately pelleted by centrifugation at 15000xg for 30 mins at 4°C, washed once with 15ml of cold 70% ethanol and then air dried for 5 mins. DNA pellet was resuspended in 200μl of nanopure water and stored at -20°C.

2.3.2 Extraction of Chromosomal DNA

The extraction of chromosomal DNA was based on the method described by Siato and Muir (1963). Bacterial cells from a 10ml overnight broth culture were resuspended in 5ml of solution I and left to stand on ice for 30 mins. SDS and EDTA were added to final concentrations of 1% and 50mM respectively. The suspension was left to stand at room temperature until the turbidity was clear (usually within 20 mins). The cell lysate was then extracted repeatedly to remove the proteins with equal volume of phenol:chloroform (1:1), followed by a single chloroform:isoamyl alcohol (24:1). After centrifugation at 4000g for 20 mins at 4°C the upper aqueous phase was collected carefully using a truncated plastic pipette to avoid mechanical shearing of the DNA. The clear aqueous phase was then mixed with 0.1 volume of 3M sodium acetate pH 5.2 and 3 volumes of cold absolute ethanol (-20°C) was slowly poured down the side of the tube. Chromosomal DNA precipitated at the interface was then collected by spooling around the end of a sterile glass pasteur pipette and resuspended in sterile nanopure water.

2.3.3 Techniques Used in Routine DNA Manipulation

Restriction endonucleases and DNA modifying enzymes were purchased from Life Technologies Ltd (Gibco BRL) or Pharmacia Biochemicals Inc and were used according to manufacturers recommendations. Restriction endonuclease cleavage of DNA was performed typically in 20μl reaction volume with one unit of enzyme per μg of DNA at 37°C. T4 DNA ligase
was used at 4°C overnight in T4 ligase buffer (50mM Tris-HCl pH 7.5, 10mM MgCl₂, 1mM ATP). For ligation the molarity ratio of vector to insert used depended on the DNA ends after cleavage with restriction enzyme; for sticky ends the ratio was 1:1 and for blunt ends a ratio of 1:2 was used.

Standard analysis and sizing of DNA were estimated in kilobases using the 1 kb ladder (Gibco BRL) in agarose gel electrophoresis. DNA samples were mixed with the appropriate volumes of 6x loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 15% ficoll). DNA fragments were separated by agarose gel electrophoresis using 0.7-1.2% TAE agarose in TAE buffer (40mM Tris-acetate pH 7.7, 1mM EDTA) with 0.5μgml⁻¹ ethidium bromide and the DNA visualised using a longwave UV transilluminator. DNA in agarose gel slices was routinely purified using a Sephaglas™ bandprep kit (Pharmacia) according to the manufacturers instructions.

The concentration of DNA was routinely determined from the A₂₆₀ of the DNA in water using a Philip UV/VIS spectrophotometer and was calculated on the assumption that an A₂₆₀ of 1 represents a concentration of 50μgml⁻¹ for a double stranded DNA and 33μgml⁻¹ for a linear DNA. Alternatively DNA concentrations were estimated by comparing band intensities to those of known standards after agarose gel electrophoresis.

Dephosphorylation of plasmid vector by the removal of 5' terminal phosphate groups from cleaved vector DNA fragments was carried out to achieve high frequency recombination between vector and insert DNA. Basically, 100μg of vector DNA was cleaved to completion in a reaction volume of 50μl. A total of 45μl of the digested DNA was mixed with 5μl of 10x calf intestinal phosphatase (CIP) buffer (10x CIP buffer; 50mM Tris-Cl pH 9.0, 10mM MgCl₂, 1mM ZnCl₂), followed by 1μl of 5U μl⁻¹ CIP (New England Biological, NEB). After 30 mins incubation at 37°C a further 0.5μl of CIP was added and the reaction allowed to proceed for a further 30 mins at 37°C. The reaction mixture was then made up to 200μl and the DNA extracted twice with an equal volume of phenol:chloroform (1:1). The phosphatase DNA was ethanol precipitated and the pellet washed with 70% ethanol as previously described (see section 2.3.1). DNA was resuspended in 50μl sterile nanopure water. Phosphatase efficacy was assessed by the transformation efficiency of self-ligated phosphatase vector.
2.3.4 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed based on the method of Saiki et al. (1988). PCR reactions were typically in 100µl reaction volume consisting of oligonucleotide primers (0.25µM each), dNTPs (200µM final for each), template DNA (10ng), reaction buffer (1x; 10mM Tris-HCl pH 8.8, 50mM KCl, 1.5mM MgCl2, 0.1% Triton X-100) and 2.5U of Taq DNA polymerase (NEB). Reaction mixtures were overlaid with 20µl of molecular biology grade mineral oil (Sigma) prior to DNA amplification through 30 cycles using a Perkin Elmer Cetus thermal cycler. Cycles consisted of a minute denaturation step at 95°C, a minute annealing at 55°C (temperature usually chosen to be at least 5°C lower than the Tm of the primer with the lowest melting point) and 2 mins at 72°C extension step. The amplified DNA was recovered by pipetting from under the oil overlay and 5-10µl aliquot analysed on agarose gel.

DNA was also amplified directly from bacterial cells in colony PCR. Basically the template DNA was replaced in the above reaction mixture either by using 10µl of an overnight broth culture or introducing a colony with a sterile toothpick directly into the reaction mixture given above and the DNA amplified through 30 cycles as previously described.

2.3.5 Cloning of Truncated Lengths of PCR-Amplified kpsE DNA Products

In order to generate site specific fusions of blaM to the kpsE gene, different lengths of the kpsE gene were amplified by polymerase chain reaction (PCR). The oligonucleotides for the PCR were designed to include restriction enzyme sites (shown in bold below), to facilitate cloning of the amplified DNA products and blaM insertion to the constructs. The oligonucleotides designated e1 to e7 include, 5'-GGGGTACCCGCCCACTGAAA CCTTGAT-3' (e1), 5'-CGGAATTTCGCCGGGCACTGAAA CTTTGAT-3' (e2), 5'-CGGAATTTCGCCGGGCACTGAAA CTTTGAT-3' (e3), 5'-CGGAATTTCGCCGGGCACTGAAA CTTTGAT-3' (e4), 5'-CGGAATTTCGCCGGGCACTGAAA CTTTGAT-3' (e5), 5'-CGGAATTTCGCCGGGCACTGAAA CTTTGAT-3' (e6), and 5'-CGGAATTTCGCCGGGCACTGAAA CTTTGAT-3' (e7). The 5' oligonucleotide (e1) located immediately upstream of the kpsE gene had a Kpn I site introduced for cloning and the 3' oligonucleotides each had an Eco RI site for cloning and a Sma I (Pvu II in the case of e4) site to generate blaM fusion within the cloned DNA products. The PCR-
amplified *kpsE* fragments were cloned into a *Kpn I-Eco RI* linearised expression vector pYZ4 (Zhang and Broome-Smith, 1990) to generate a series of clones designated pFE3-8 (see section 3.2.2.2).

2.4 DNA Hybridisation Procedures

2.4.1 Transfer of DNA from Agarose Gels (Southern Blotting)

DNA was transferred to filters using the capillary blotting technique described by Southern (1975). DNA sample was separated by agarose gel electrophoresis and the gel photographed along side a ruler to serve as reference point. The DNA was depurinated by soaking the gel in 0.25M HCl for 7 mins. The gel was then rinsed in distilled water and soaked in denaturing solution (0.5M NaOH, 1.5M NaCl) for 30 mins with gentle shaking. The gel was again rinsed in distilled water, followed by soaking for 30 mins in neutralising solution (0.5M Tris-HCl pH 7.5, 3M NaCl) with gentle shaking. The gel was rinsed again and then laid unto six sheets of pre-wet (20x SSC, 3M NaCl, 0.3M sodium citrate, pH 7.0) Whatman 3mm filter paper without trapping any air bubbles. A pre-soaked (3x SSC) sheet of Hybond-N nylon membrane (Amersham) was placed on the gel with a 3x SSC pre-wet Whatman paper on top, followed by four sheets of dry Whatman paper and a stack of dry paper towels. Finally, the transfer sandwich was evenly compressed by placing a glass plate and a 500g weight on top of the paper towels. The lower sheets of Whatman paper were constantly kept soaked in 20x SSC and the paper towels changed once before dismantling the apparatus after overnight of DNA transfer. The nylon membrane was air dried, wrapped in Saran wrap and the DNA cross-linked to the membrane by exposure to UV from a long wave transilluminator for 5 mins.

2.4.2 Radiolabelling of DNA for Use as Probe

Plasmid DNA was cleaved with the appropriate restriction endonucleases and the fragments separated by agarose gel electrophoresis on a 1% low melting point agarose gel (BRL). The DNA fragment(s) of interest were excised from the gel and added to sterile nanopure water (1.5ml/gram of agarose). The sample was boiled in a water bath for 7 mins before use or stored at -20°C and reboiled for 3 mins immediately before use.
Labelling reactions were performed as described by Feinberg and Vogelstein (1983). Approximately 10ng of DNA was radiolabelled using random hexanucleotide primers. Nucleotide and hexanucleotides were obtained from Pharmacia and $^{32}$P-$\alpha$-dCTP from DuPont.

2.4.3 Hybridisation and Detection of DNA Immobilised on Filters with Probe

Hybridisation and washing steps were carried out in a rotary hybridisation oven (Hybaid) using cylindrical canisters. Southern blot membrane filters were incubated in prehybridisation solution (Table 2.3) at 65°C for 2 hrs, followed by incubation overnight at 65°C in hybridisation solution (Table 2.3) containing the radiolabelled probe.

After hybridisation steps, the filters were washed twice in high stringent solutions of 0.1% SDS and 0.1x SSC at 65°C for 15 mins each and twice in 0.1% SDS and 0.5x SSC at 65°C for further 15 mins each. Filters were air dried, wrapped in Saran wrap and placed in an autoradiography cassette with the DNA side up. Kodak X-Omat AR film was exposed to the filters at -70°C. Films were developed in an Agfa-Geveart automatic film processor.

<table>
<thead>
<tr>
<th>Table 2.3: Composition of Hybridisation Solutions$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prehybridisation solution</td>
</tr>
<tr>
<td>3x SSC</td>
</tr>
<tr>
<td>0.1% SDS</td>
</tr>
<tr>
<td>5x Denhardt's solution$^b$</td>
</tr>
<tr>
<td>6% PEG 600</td>
</tr>
<tr>
<td>200μg/ml$^c$ Denatured Salmon Sperm DNA</td>
</tr>
<tr>
<td>Hybridisation solution</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Radiolabelled Probe

a, solutions were stored at -20°C

b, 100x Denhardt's solution contains 2% Ficoll, 2% BSA and 2% Polyvinol-pyrollinidine

c, Salmon sperm DNA was denatured by boiling and forcing the solution through a narrow gauge syringe needle

2.5 DNA Sequencing

Nucleotide sequence was determined based upon the chain termination method described by Sanger et al. (1977), in which DNA synthesis from deoxynucleotide triphosphates is terminated by the addition of dideoxynucleotide triphosphates. Double stranded sequencing was carried out using plasmid DNA. Sequence reactions were performed using the Sequenase™ Version 2.0 Kit (United States Biological, USB) with [$\alpha$-35S] dATP radiolabelling incorporated in the extension reaction and the ABI PRISM™ Ready Reaction Dye Deoxy™ Terminator Cycle Sequencing Kit (Perkin Elmer) according to manufacturers recommendations. Reactions relied on multiple annealing/denaturing cycles to amplify second strand yield.

2.5.1 DNA Sequencing Polyacrylamide Gel Electrophoresis

The radiolabelled chain termination products were resolved using Tris/Borate/EDTA (TBE) gradient gel electrophoresis (Biggs et al., 1983).
Gel solution 1
7ml 5x TBE acrylamide/urea mix
45μl 10% ammonium persulphate (APS)
2.5μl TEMED

Gel solution 2
40ml 0.5x TBE acrylamide/urea mix
180μl 10% ammonium persulphate (APS)
7.5μl TEMED

Electrophoresis grade APS and TEMED were purchased from Sigma, and 30% Accugel acrylamide solution was from Flowgen. 10x TBE buffer consists of 121.1g/l Tris-HCl, 55g/l Boric acid and 40ml of 0.5M EDTA adjusted to pH 8.3. The 0.5x TBE/acrylamide urea mix consists of 430g/l urea, 150ml 30% acrylamide solution and 50ml/l 10x TBE. The 5x mix is similar except that it contains 150ml/l 10x TBE, with 50g/l sucrose and 50mg/l bromophenol blue. To prepare the gel, two glass plates (20cm by 50cm) were cleaned free of grease using teepol and water. One of the plates was siliconised by applying a smear of dimethylchlorosaline evenly on a surface and allowing to dry. Plates were then tapped together separated by 0.4mm spacers. In a 25ml pipette, 10ml of gel solution 2 followed by 14ml of gel solution 1 were drawn and a rough gradient was formed within the pipette by drawing up some air bubbles. This mixture was run down between the glass plates avoiding formation of air bubbles and the cavity filled with the remaining of gel solution 2. The comb was positioned and plates clamped along each side. Gel was allowed to set at room temperature for 1 hr and the tape was removed from the bottom of the gel. The gel was clamped in vertical position with aluminium sheets of a similar dimension as the gel plates on either side for even heat distribution. The upper buffer reservoir was filled with 0.5x TBE and the lower with 1x TBE. The gel was pre-run for 30 mins at a constant power of 40W and the wells rinsed with running buffer prior to sample loading. Electrophoresis was performed at constant power of 40W for 3 or 6 hrs. After electrophoresis the gel plates were prised apart with the non-siliconised plate down to retain the gel, which was then soaked in gel fix (10% methanol, 10% acetic acid) for 15 mins and rinsed with distilled water. The gel was transferred to a pre-wet Whatman filter paper, covered with Saran wrap and dried under vacuum for 1 hr at 80°C. Autoradiography used the Dupont Cronex film and exposure at room temperature.
2.5.2 Automated DNA Sequencing

DNA sequencing is based on the chain termination method, where four dye-labelled dideoxy nucleotides, GATC DyeDeoxy™ terminators replaces the standard dideoxy nucleotides in the enzymatic sequencing reaction. Double stranded DNA sequencing was performed using the PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit (Perkin Elmer) according to manufacturers recommendations. The supplied terminator premix consists of 1.58μM A-DyeDeoxy, 94.74μM T-DyeDeoxy, 0.42μM G-DyeDeoxy, 47.37μM C-DyeDeoxy, 78.95μM dITP, 15.79μM dATP, 15.79μM dCTP, 15.79μM dTTP, 168.42mM Tris-HCl pH 9.0, 4.21μM (NH₄)₂SO₄, 42.10mM MgCl₂ and 0.42U/μl AmpliTaq DNA polymerase. Typically reactions consisted of 9.5μl terminator premix, 0.6μg DNA template, 0.0pmol primer in a reaction volume of 20μl. Reaction mixture was overlaid with a drop of mineral oil. Reaction was for 25 cycles in a Perkin Elmer Model 9600 and consisted of a rapid thermal ramp and a hold at 96°C for 10 secs, annealing at 50°C for 5 secs and extension at 60°C for 4 mins. Extended products were extracted with sequencing grade chloroform once and with phenol:water:chloroform (68:18:14) solution twice. Products were precipitated with 2M sodium acetate pH 4.5 and 100% ethanol, washed once with 70% ethanol and the pellet dried. The dried product was sent to the Protein and Nucleic Acid Laboratory, University of Leicester for gel analysis.

2.6 Computer Analysis of DNA and Protein Sequences

DNA sequence was assembled using programs of the Genetics Computer Group (GCG) suite of the University of Wisconsin (Devereaux et al., 1984) running on a Silicon Graphic Crimson mainframe (Daresbury Seqnet Facility, UK) and (Irix, Leicester University) under a UNIX platform. The program MAP was used to identify ORFs and restriction enzyme sites. TRANSLATE was used in creating protein sequence from the DNA sequence. DNA and protein homologies were identified from database searches using either FASTA (Lipman and Pearson, 1985) or the BLAST series (Altschul et al., 1990). Databases used included the Genbank/European Molecular Biology Laboratory (Gen_EMBL) and the Swissprot/Protein Information Resource (OWL).
The hydropathy profile of Kyte and Doolittle (1982) was used for the analysis of predicted amino acid sequences on the PEPLOT or the DNA strider software. Attempts at structure prediction for proteins were performed using the PREDICTPROTEIN version 1.0 PHD neural network facility at EMBL Heidelberg (Sander and Schneider, 1991). Multiple alignments were obtained using either CLUSTER-W or ALIEN using a PAM 250 matrix (Higgins et al., 1994).

2.7 Procedures for Protein Analysis

2.7.1 Expression of Cloned Genes in *E. coli*

Protein expression from genes cloned into the pTTQ18* and pGEM-3Zf+ vectors were performed using the *lac* and the T7 promoters respectively. For total cell lysate, cells in 10ml broth were grown to an OD$_{600}$ of 0.5 and induced with 1mM IPTG final concentration. Cells were allowed to grow to an OD$_{600}$ of 1.0 and were harvested from 1.5ml culture by centrifugation. Cells were then resuspended in nanopure water and the optical densities of cell suspension were adjusted to equivalence by diluting the higher density cell suspensions. To 200μl of cell suspension, 20μl of 10% SDS was added, boiled for 3 mins and 5-15μl of this lysate was separated by SDS-PAGE (see section 2.7.2). Gels were stained in 0.2% Coomassie blue (Sigma) solubilised in the destain solution (10% acetic acid, 40% methanol) for 1 hr and destained to achieve the desired contrast.

2.7.2 *In vivo* Radiolabelling and Detection of Protein

Proteins were expressed in pGEM-3Zf+ using the T7 expression system based on the method of Tabor and Richardson (1985). The gene of interest was cloned downstream from the T7 promoter and the pGEM-3Zf+ derivative transformed into *E. coli* JM109DE3. This strain contains the structural gene for bacteriophage T7 RNA polymerase on the chromosome. The T7 RNA polymerase is transcribed from the *lac* promoter on induction with IPTG. *E. coli* JM109DE3 harbours an F' factor containing *lacI* to repress the *lac* promoter under non-induced conditions.
To express plasmid encoded genes, a single colony from an agar plate of Davis medium (7gl⁻¹ K2HPO4, 2gl⁻¹ KH2PO4, 0.5gl⁻¹ Na citrate, 0.1gl⁻¹ MgSO4, 1gl⁻¹ (NH4)2SO4) supplemented with glucose (0.4%) and thiamine (1µg⁻¹) was used to inoculate 10ml of L-broth and cells grown at 37°C with aeration overnight. The cell culture was diluted 1:10 in 10ml of fresh L-broth and grown further at 37°C with aeration to an OD600 of 0.6. Cells were harvested from 1ml of this culture by centrifugation and washed twice in unsupplemented Davis medium. The final pellet was resuspended in 1ml of Davis medium supplemented with 0.4% glucose and 1µg⁻¹ thiamine-HCl. This cell suspension was then incubated in a 37°C water bath for 90 mins. T7 RNA polymerase was induced by the addition of 0.5mM IPTG final concentration to the cell suspension followed by incubation at 37°C for 20 mins. After induction, E. coli RNA polymerase was inhibited by the addition of 40µl of rifampicin (10mgml⁻¹ in methanol) and incubation at 37°C for 30 mins. Products of the encoded genes transcribed from the T7 promoter were then labelled by the addition of 10µCi [³⁵S] methionine (Amersham), followed by incubation at 37°C for 1 hr. The radiolabelled cells were harvested by centrifugation. Cells were then lysed by resuspending the pellet in 30µl of lysis buffer (60mM Tris-HCl pH 6.8, 1% SDS, 1% 2-mercaptoethanol, 10% glycerol, 0.01% bromophenol blue) and by boiling at 100°C for 5 mins. A 5µl aliquot of cell lysate was separated by SDS-PAGE.

SDS-PAGE was performed as described by Laemmli (1970). Gels were run using the mini PROTEAN™ II system (Biorad) according to the manufacturers instructions. The resolving gel (4.65ml of 1.5M Tris-HCl pH 8.8/0.2% SDS, 4.95ml of 30% Protogel acrylamide, 348µl of 1% APS, 40µl of TEMED) was poured first, overlaid with butanol and allowed to set. On setting the butanol was removed and the top of the gel was washed with distilled water. The resolving gel was then overlaid with stacking gel (3ml of 0.5M Tris-HCl pH 6.8/0.2% SDS, 1.04 ml of 30% Protogel acrylamide, 2.12ml distilled water, 150µl of 1% APS, 20µl of TEMED) and the comb immediately put in place. Gel was allowed to set and the comb removed before samples were loaded. Electrophoresis was performed in a Tris/Glycine/SDS buffer (25mM Tris base, 193mM Glycine, 0.1% SDS) at 150V for 1-2 hrs until the bromophenol blue dye front has migrated to the bottom of the gel. Broad range prestationed SDS-PAGE markers (Biorad) were used as standards. After electrophoresis the gel was soaked in a fixing solution (10% acetic acid and 25% isopropanol) for 30 mins and treated
with Amplify (Amersham) according to the manufacturers recommendations. Gels were dried under vacuum at 80°C and autoradiographed at room temperature using DuPont Cronex film.

2.7.3 Detection of Protein Antigen by Western Blotting

For Western blotting, proteins resolved by SDS-PAGE were transferred to PVDF membrane (Millipore Corporation) using the electroblotting technique of Towbin et al. (1979). Electrotransfer was performed in 25mM Tris base, 193mM glycine, 20% methanol (transfer buffer) at 120V for 1 hr. Protein detection was performed using the ECL™ western blotting protocol (Amersham Life Science). Essentially immediately after protein transfer, non-specific binding sites on membranes were blocked for overnight at 4°C in blocking solution (5% skimmed milk in phosphate buffer saline, PBS; 137mM NaCl, 1.5mM KH2PO4, 8mM Na2HPO4, 6.7mM KCl, containing 0.1% Tween 20™). Membranes were then washed three times for 15 mins each in PBS-Tween 20™ wash buffer at room temperature. The membranes were incubated with the appropriate primary antibody dilutions for 1 hr at room temperature with constant agitation on a rocking platform (Hybaid). The membranes were washed as previously described and incubated with the HRP-conjugated second antibody for 1 hr at room temperature. This was followed by washing the membranes five times, which was then wrapped in Saran wrap. Detection was through hydrogen peroxide, H2O2/HRP catalysed oxidation of luminal in alkaline conditions, achieved by incubating membranes in pre-mixed equal volumes of the detection solutions 1 and 2. Film exposure was for 5-30 secs. Antibodies were routinely diluted in blocking solution. The KpsE antibody was used at 1:1000 dilutions and has been described (Rosenow et al., 1995). Antibody to β-lactamase was obtained from 5'-3' Inc, and was used at a dilution of 1:10000. The secondary antibody was anti rabbit IgG-horseradish peroxidase (HRP) conjugate (5'-3' Inc) and was used at 1:2500 dilutions.

Membranes were routinely stripped and reprobed several times with antibodies. Stripping was achieved by incubating membranes in stripping buffer (10mM 2-mercaptoethanol, 2% SDS, 62.5mM Tris-HCL pH 6.7) for 30 mins at 65°C.
2.8 Subcellular Fractionation and Proteolysis with Proteinase K

Subcellular fractionation was carried out based on the method described by Kampfenkel and Braun (1993). A 10ml culture of the appropriate strain was grown at 37°C with aeration to an OD600 of 0.8 and cells were induced with 1mM IPTG final concentration. The cells were harvested at an OD600 of 1.2 by centrifugation at 4000g for 5 mins at 4°C, followed by resuspension in 1ml ice-cold solution of 0.2M Tris-HCl, 0.5M sucrose, 0.5mM EDTA pH 8.0 and 0.5mgml⁻¹ lysozyme, and incubation on ice for 20-30 mins. Spheroplast formation and stability were observed by phase contrast microscopy. The spheroplasts were collected by centrifugation at 3000g for 5 mins at 4°C and the supernatant fraction containing the periplasm was precipitated with 0.5ml of 30% trichloroacetic acid. The pelleted spheroplasts were resuspended in 1ml of ice-cold PBS and 20% sucrose, and 0.5ml of this suspension was left on ice as control for untreated sample. To the remaining 0.5ml of spheroplast suspension, 10µl of 50mgml⁻¹ of proteinase K was added and incubated at 37°C for 15 mins. Proteolysis of the spheroplasts was stopped by the addition of 2mM final concentration of phenylmethylsulfonyl fluoride (PMSF) in isopropyl alcohol. Both proteinase K treated and untreated spheroplast were collected by centrifugation and the resulting pellet resuspended in 150µl of PBS pH 7.6. To this 150µl suspension, 15µl of 10x SDS-PAGE loading buffer was added and then boiled for 5 mins. The boiled samples (10-15µl) were analysed by SDS-PAGE and the proteins were probed with specific antibodies as described in Western blot procedure.

2.9 Assay for Lyase Enzyme Activity

Enzymatic activity was determined by incubating cell lysate of strain harbouring the lyase clone with the extracted polysaccharide of E. coli K5. Lyophilised K5 polysaccharide was a generous gift from Professor K Jann Laboratory in Freiburg, Germany. The K5 polysaccharide was essentially extracted by harvesting cells from 400ml overnight culture at 10000xg for 10 mins at 4°C. The resulting pellet was washed once with 50ml PBS. The pellet was then incubated in 50mM Tris-HCl, 5mM EDTA pH 7.3 for 30 mins at 37°C, followed by centrifugation, with this process repeated three times. The resulting polysaccharide extract in the supernatant from the three centrifugation steps was precipitated by the addition to the
supernatant of 10% cetyl-3-ethyl ammonium bromide sodium salt to saturation and incubation at room temperature overnight. The precipitate was recovered by centrifugation at 10000xg for 20 mins at 20°C which was then dissolved in 1M NaCl and 80% ethanol solution, followed by centrifugation. The pellet was resuspended in 1-5ml of nanopure water and dialysed against water at 4°C for 2 days. The suspension was freeze dried, followed by resuspension in water and then ultracentrifugation with fix angle rotor at 100000xg for 4 hrs at 4°C. The supernatant containing the polysaccharide was collected and freeze dried.

The enzyme lysate was obtained from a 10ml culture of strain harbouring the appropriate clone. Cells were grown to OD_{600} of 0.6, followed by induction for 1 hr with 1mM IPTG final concentration. Cells were collected by centrifugation at 4000g for 5 mins at 4°C, washed once and resuspended in ice-cold sonication buffer (100mM Tris-HCl pH 7.5). The suspension was sonicated to clarity and the supernatant (lysate) obtained after centrifugation.

The enzyme assay involved incubating 20μl of K5 polysaccharide substrate (5mgml^{-1} in Tris buffer pH 7.5) with 5-20μl of lysate for 5 mins (pFE50 clone) or 6 hrs (B4 clone) at 37°C. At the end of reaction, sample buffer (10x; 2M sucrose in 10x TBE and bromophenol blue) was added to 1x final concentration. Samples were then resolved by PAGE in 1x TBE at 350-400V for 2-3 hrs. The gel was fixed for 30 mins in acetic acid, water and ethanol (10:110:80) solution, stained for 3 hrs with Alcian Blue (0.5% in 2% acetic acid) (Sigma) and destained with 2% acetic acid in water. Samples were resolved in 25% acrylamide gel, which consists (for 40ml) of 4ml 10x TBE, 15ml water, 20ml 50% acrylamide stock (49% or 495.9g^{-1} acrylamide and 1% or 4.1g^{-1} bisacrylamide), 150μl 10% APS and 15μl TEMED.
Chapter 3

Topology of KpsE, a Protein Involved in the Export of Polysaccharide in *Escherichia coli* K5

3.1 Introduction

Capsular polysaccharides play a critical role in the interaction of a microorganism and its environment. Capsules have been recognised as important virulence determinants conferring on bacteria the ability to evade the human host defence. The capsular polysaccharide of *E. coli* have been divided into three groups on the basis of chemical structure, size, mode of expression and genetic determinants (Jann and Jann, 1990; Pearce and Roberts, 1995). Group II capsule gene clusters studied so far show a common organisation consisting of three functional regions, 1, 2 and 3. Region 2 encodes proteins necessary for the biosynthesis and polymerisation of type specific polysaccharide. The flanking regions 1 and 3 are homologous among different group II capsules and are thought to encode products involved in the export of polysaccharide across the bacterial membrane and its surface expression. The *E. coli* K5 KpsE protein encoded in region 1 has been shown to localise in the inner membrane and mutants lacking the encoded product were unable to export polysaccharide to the cell surface (Bronner *et al.*., 1993a; Rosenow *et al.*, 1995a). The predicted amino acid sequence of the KpsE protein of *E. coli* K5 and K1 is homologous to those of the CtrB and BexC proteins encoded by the capsule gene clusters of *N. meningitidis* and *H. influenzae* respectively. A distinct subfamily (ABC-2) of the ABC-transporters involved in polysaccharide export has been proposed on the basis of sequence comparison studies (Reizer *et al.*, 1992). Analysis of the amino acid sequences of DrrAB and NodIJ proteins involved in drug resistance transport in *Streptomyces peucetius* and nodulation in *Rhizobium leguminosarum* respectively showed that they are homologous to three sets of proteins (KpsMT of *E. coli*, BexABC of *H. influenzae* and CtrDCB of *N. meningitidis*) comprising capsular polysaccharide export systems in gram-negative bacteria (Reizer *et al.*, 1992). The Bex and Ctr systems of this ABC-2 subfamily are postulated to have two integral membrane components, the BexC protein is thought to interact with BexAB transporters and the CtrB with CtrCD transporters in the inner membrane. On the basis of their likely function and location
within the cytoplasmic membrane, both the CtrB and BexC proteins have been suggested to form the third components of the ABC-2 transporters involved in the export of polysaccharide across the cytoplasmic membrane (Reizer et al., 1992). By analogy, the KpsE protein which is homologous to the BexC and CtrB proteins, might interact with the KpsMT ABC-transport components in the export of polysaccharide in E. coli. In addition, the ABC exporters have been identified in many bacteria and they constitute a family of proteins involved in the export of several unrelated polypeptides in these bacteria. In the best studied cases, they have been postulated to comprise three components (Fath and Kolter, 1993; Dinh et al., 1994; Binet and Wandersman, 1996). These include the inner membrane ATP-binding cassette (ABC) protein(s), in which the membrane spanning and the ATP-binding segments may be present on the same polypeptide or as distinct polypeptides, and a second inner membrane protein with a large periplasmic domain and a C-terminus that may interact with the outer membrane. Because of their membrane topology these proteins are classified into the membrane fusion protein (MFP) family (Dinh et al., 1994). The third component is an outer membrane protein (Wandersman, 1992; Binet and Wandersman, 1996). The inner membrane location of the BexC, CtrB and KpsE proteins suggests that they represent the MFP protein in the ABC-mediated export of polysaccharide. To understand the possible role played in the export of polysaccharide by the KpsE, it is required to define the topology within the cytoplasmic membrane.

The ability to predict protein topology from the primary sequence and the use of gene fusions to study membrane protein topology are both well established techniques (von Heijne, 1986, 1992; von Heijne and Gavel, 1988; Boyd and Beckwith, 1990; Broome-Smith et al., 1990; Gafvelin and von Heijne, 1994; Prinz and Beckwith, 1994). Topological models for a membrane protein can be generated from its predicted amino acid sequence, using the hydropathy profile (Kyte and Doolittle, 1982) and the 'positive inside rule' (von Heijne, 1992). Such models can be confirmed using gene fusion approach. TnphoA have been used to generate alkaline phosphatase (AP) fusions in determining the topology of membrane proteins (Manoil and Beckwith, 1985, 1986; Boyd et al., 1993; Prinz and Beckwith, 1994). AP is enzymatically active when localised to the periplasm but is inactive in the cytoplasm. Thus, when the mature AP moiety is fused to a membrane protein, its location is determined by the topogenic information encoded in that protein and can be analysed by measuring or detecting the AP enzymatic activity. This system relies on random phoA
gene fusions and it is therefore generally inefficient in that even distribution of \textit{phoA} fusions is not a certainty and many fusions will have to be isolated to obtain a complete topological analysis. A similar gene fusion approach for studying protein topology has been developed using \(\beta\)-lactamase (\textit{blaM}) gene (Broome-Smith and Spratt, 1986; Broome-Smith \textit{et al.}, 1990). \textit{BlaM} confers resistance to \(\beta\)-lactam drugs such as ampicillin when exported to the periplasm. When \textit{BlaM}, which lacks its N-terminal signal sequence is fused to various lengths of a membrane protein, its localisation can be analysed from the levels of ampicillin resistance conferred, thus acting as topological reporter protein. \(\beta\)-lactamase is only active in the periplasm and from the level of ampicillin resistance conferred both periplasmic and cytoplasmic fusions can be selected and has been successfully used to analyse the topology of a number of membrane proteins (Wang \textit{et al.}, 1991; Kampfenkel and Braun, 1993; Reeves \textit{et al.}, 1994; Gunn \textit{et al.}, 1995). The membrane topology of the KpsE protein was investigated using both Tn\textit{phoA} mutagenesis and \(\beta\)-lactamase fusions. The topology was confirmed by accessibility of fusion proteins to proteinase K.

Previous mutational analysis of the KpsE protein involved the use of deletion and insertion mutants, where the entire gene had been deleted or an oligonucleotide inserted within the gene. These mutations have resulted in a polarity effect on the expression of downstream region 1 genes and are therefore unsuitable for structure-function studies. In order to avoid this polarity effect, it is necessary to construct a null \textit{kpsE} mutant. The construction of a strain (MS101) with a single copy of the K5 capsule gene cluster in its normal location in the chromosome of a laboratory strain of \textit{E. coli} has been described (Stevens \textit{et al.}, 1994) and was exploited in making a null \textit{kpsE} mutant.

### 3.2 Results

#### 3.2.1 Prediction of Membrane Topology of the KpsE Protein

The hydropathy profile of the KpsE protein was determined by the Kyte and Doolittle algorithm (Kyte and Doolittle, 1982) to expose potential \(\alpha\)-helices forming regions. Based on the hydropathy plot (Figure 3.1) of the predicted amino acid sequence and the use of von Heijne’s ‘positive inside rule’, a working model (Figure 3.2) consisting of two transmembrane spanning regions for the membrane topology of KpsE was proposed. The
membrane spanning regions consisted of residues 29-49 and 351-371. These chosen membrane spanning regions were strengthened by the prediction of helical transmembrane segments (PHDhtm) for the KpsE protein using a profile fed neural network system from Heidelberg (Rost et al., 1994).

3.2.2 Analysis of Membrane Topology of the KpsE Protein Using Gene Fusion

3.2.2.1 TnphoA mutagenesis

To confirm the proposed model of the membrane topology, TnphoA mutagenesis was used to generate phoA fusions to the KpsE protein. Random mutagenesis were generated in filter mating experiments (see section 2.2.3) using E. coli LE392 carrying the plasmid pCR6 as recipient and E. coli SM10pirpRT733 as donor strains. Conjugants resulting from the mating were selected on M9 salt minimal medium containing the appropriate antibiotics. Plasmid DNA extracted from kanamycin resistant transconjugants was used to transform E. coli CC118 and PhoA+ mutants were identified as blue colonies on plates containing the chromogenic substrate, 5-bromo-4-chloro-3-indonyl phosphate (XP, 40μg/ml). TnphoA insertions were mapped within the kpsE gene in plasmid pCR6 using restriction endonucleases. A total of 128 plasmid insertion mutants representing 2 blues and 126 whites were isolated. Of these 128 insertion mutants, 15 mapped within the kpsE gene and the remaining 113 were within the vector sequence. Thirteen out of the 15 insertions within the kpsE were in the wrong orientation to the transcribed gene. The two blue insertion mutants represented PhoA+ insertions at amino acid positions 54 and 56 of the KpsE protein. The precise insertion site within the kpsE gene was determined by nucleotide sequence with oligonucleotide primers to phoA. Two oligonucleotide primers, 5'-CTGAGCCACGCG-3' and 5'-GTTAGGGCTACATG-3', that bind respectively within 63 bp from the start and 78 bp to the end of TnphoA sequence were used to sequence from the phoA into the kpsE gene.

The two phoA fusions within the kpsE gene represented a low isolation frequency and more fusions were needed to cover the remaining parts of the KpsE protein to determine its membrane topology.
Figure 3.1. Hydropathicity plot of the KpsE protein: by the method of Kyte & Doolittle (1982) with a window of 20. Boxes denote positions of the two α-helices. The amino acid sequence of KpsE is shown below the hydropathy plot and positions of the two α-helices are underlined. Residue numbers are shown on the left.
Figure 3.2. Model of the membrane organisation of the KpsE protein. The net positive and negative charges are indicated. Single letter amino acids are shown within circles.
3.2.2.2 Generation of β-lactamase fusion

The isolated PhoA fusions were close together at amino acid positions 54 and 56 of the KpsE protein and because of the random insertion associated with the use of the TnphoA system, site-specific fusions were needed in other domains of the protein to obtain better topological information. The β-lactamase gene fusion system described by Broome-Smith and Spratt (1986) was used to generate site-specific fusions to other parts of the KpsE protein.

Site specific inframere fusions of blaM to the kpsE gene were made using a unique EcoRV site in kpsE and to truncated lengths of PCR-amplified kpsE DNA products. In order to utilise the EcoRV site in kpsE for a translational inframere fusion to BlaM, a 2.1 kb Sma I-Hinc II kpsE fragment from plasmid pH18 was ligated into an EcoRV linearised expression vector pYZ4 (Zhang and Broome-Smith, 1990) to generate plasmid pFE2. To cover other regions, truncated lengths of the kpsE gene were amplified by polymerase chain reaction (PCR). Oligonucleotides for the PCR were made with restriction enzyme sites within, to facilitate cloning of the amplified DNA products and blaM insertion to the constructs. The 5’ oligonucleotide (5’-GGGGTACCCGCCCAGCTCTGATGAT-3’) had a Kpn I site (shown in bold) introduced for cloning and the 3’ oligonucleotides each had an EcoRI site for cloning and a Sma I site to generate blaM fusion within the cloned DNA product (Figure 3.3) (see section 2.3.5). The PCR amplified kpsE fragments were cloned in the Kpn I-EcoRI linearised expression vector pYZ4 (Figure 3.3) to generate a series of constructs designated pFE3-8 (Figure 3.4). A 0.9 kb Sma I DNA fragment encoding the matured BlaM protein in plasmid pLH21 was ligated into the EcoRV site in pFE2 and the Sma I sites in pFE3-8 to generate BlaM fusions at amino acid positions 24, 41, 51, 211, 345, 361 and 381 yielding respectively constructs pFE2::24, pFE7::41, pFE8::51, pFE3::211, pFE4::345, pFE5::361 and pFE6::381 (Figure 3.4). The KpsE-BlaM fusion junctions were confirmed by nucleotide sequence using the primer 5’-CACTCGTGCACCCGACTG-3’, which is complementary to codons 15-19 of blaM.
Figure 3.3. Cloning of different lengths of PCR-amplified kpsE DNA products. PCR products each with unique restriction enzyme sites are shown (A). DNA fragments were cloned in the expression vector pYZ4 (B), before ligating the blaM fragment from plasmid pLH21 (C) in the cloned gene. Abbreviations: B= BamHI, E= EcoRI, Ev= EcoRV, H= Hind III, K= Kpn I, Nc= Nco I, Sa= Sal I, Sm= Sma I, Sp= Sph I, Sst= Sst I, Xb= Xba I. Sizes are in kilobase.
Figure 3.4. Construction of the pFE::blaM gene fusion series. The precise site of blaM insertions are indicated in parenthesis.
3.2.2.2.1 Ampicillin Resistance of *E. coli* JM101 Expressing KpsE-BlaM Hybrid Proteins

Strains expressing KpsE-BlaM fusion proteins were selected by patching colonies onto L-agar plates that contained 25μgml⁻¹ kanamycin, 1mM IPTG and 30μgml⁻¹ ampicillin. This initial screening ensures that *E. coli* colonies carrying BlaM fusions to both cytoplasmic and periplasmic domains of the KpsE protein were selected, when inoculated at high cell density (cell patches) (Broome-Smith and Spratt, 1986). Cell patches containing plasmids with putative inframe BlaM fusions to KpsE (Ampicillin resistant) were then tested for their ability to form single colonies on L-agar containing various concentrations of ampicillin when grown at low cell density (see section 2.1.3). The minimum inhibitory concentration, MIC (see section 2.1.3) determined for the KpsE-BlaM fusions are listed table 3.1, and these Amp⁷ levels are superimposed on the topological model (Figure 3.5). Fusions were tested in the same experiment.

Table 3.1. The Amp⁷ Values of *E. coli* JM101 Expressing KpsE-BlaM Fusion Proteins

<table>
<thead>
<tr>
<th>KpsE-BlaM fusion</th>
<th>Amp⁷ (MIC values) in μgml⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>pFE2::BlaM,24</td>
<td>&lt;5</td>
</tr>
<tr>
<td>pFE7::BlaM,41</td>
<td>50</td>
</tr>
<tr>
<td>pFE8::BlaM,51</td>
<td>100</td>
</tr>
<tr>
<td>pFE3::BlaM,211</td>
<td>150</td>
</tr>
<tr>
<td>pFE4::BlaM,345</td>
<td>100</td>
</tr>
<tr>
<td>pFE5::BlaM,361</td>
<td>10</td>
</tr>
<tr>
<td>pFE6::BlaM,381</td>
<td>15</td>
</tr>
</tbody>
</table>

Amp⁷, ampicillin resistance; MIC, minimum inhibitory concentration.
Figure 3.5. The MIC values of KpsE-BlaM fusion proteins, superimposed on the working model. MIC values are shown in parenthesis, next to the amino acid residues to which BlaM was fused. The two PhoA+ insertions are represented with a plus (+) sign.
3.2.2.2.2 Visualisation of KpsE-BlaM Fusion Proteins by Western Blot Analysis

To ensure that the MIC values (Figure 3.5) were not affected by the quantity and/or the stability of fusion proteins expressed, total protein lysate of all KpsE-BlaM fusions were visualised in Western blot analysis using anti-β-lactamase antibody (Figure 3.6A). The results revealed similar levels of expression by the KpsE-BlaM fusions despite the varying levels of Amp\(^R\) (≤ 5 to 150 μg/ml), indicating a lack of correlation between the MIC values and the amount of fusion protein expressed. This suggests that the high MIC values determined for fusions to amino acid positions 51, 211 and 345 are reflections of BlaM location in the periplasm.

The membrane used in the Western blot analysis using anti-BlaM antibody was stripped (see section 2.7.3) and then reprobed with anti-KpsE antibody. The results showed lack of hybridisation to KpsE-BlaM fusions, pFE7::41, pFE8::51 and pFE3::211, while fusions pFE2::24, pFE4::345, pFE5::361 and pFE6::381 were reactive to the KpsE antibody (Figure 3.6B). The pCR6 construct (Rosenow et al., 1995a) was used as a positive control for the KpsE antibody reaction, and it is essentially a 2.1 kb Sau I-Hinc II fragment containing the entire \(kpsE\) gene placed under the bacteriophage \(\lambda\) promoter in plasmid pCE30 (Elvin et al., 1990). The BlaM fusion at amino acid position 24 (pFE2::24) is between the first 24 N-terminal amino acids and the remaining residues of the KpsE protein and the reactivity with the KpsE antibody is indicative of a second translational start downstream from the fusion point (see figure 3.4). The BlaM fusions to residues 41, 51 and 211 lack respectively the last 341, 331 and 171 amino acid residues of the KpsE protein and their inability to react with the KpsE antibody shows that the epitope recognised by the KpsE antisera lie within the last 170 (212-382) amino acid residues of the KpsE protein.

3.2.2.2.3 Subcellular Fractionation and Proteolysis with Proteinase K

The cytoplasmic membrane (spheroplast) and the periplasm of strain carrying all KpsE-BlaM fusions were separated and Western blot analysis of the cell fractions revealed that the KpsE-BlaM fusion proteins were entirely membrane associated (Figure 3.7). The location of the β-lactamase in each hybrid protein was verified by proteinase K accessibility to spheroplasts made from KpsE-BlaM hybrid proteins. The results showed that the fusion
at position 24 was protected indicative of a cytoplasmic location while the
remaining fusions were cleaved, indicating that these fusions must be
exposed to proteinase K at the periplasmic face of the cytoplasmic
membrane (Figure 3.8).

Whilst from the topological model, the BlaM fusion at position 381 was
expected to be cytoplasmic, it was found accessible to proteinase K. One
reason for this unexpected result could be due to disruption of the inner
membrane integrity during spheroplast preparation thereby allowing
proteinase K to enter the cytoplasm. To ascertain if this were the case,
plasmid pACYC184 which encodes a cytoplasmic chloramphenicol acetyl
transferase (CAT) protein was transformed into a strain carrying the KpsE-
BlaM fusion at position 381. Western blot analysis of proteinase K treated
and untreated spheroplast obtained from this strain using antibody to CAT
(Figure 3.9) showed that the CAT protein was exclusively protected,
indicative of an intact spheroplast. This also means that the C-terminus of
the KpsE protein is not cytoplasmic but rather within the inner membrane,
a result that is in agreement with that of the Amp\(^R\) level conferred by the
strain carrying BlaM at position 381.

3.2.3 Analysis of the KpsE and Other Proteins Involved in Bacterial ABC
Transport

Comparing sequences of the KpsE and some members of the recently
described membrane fusion protein (MFP) family (Dinh et al., 1994)
revealed some similarities. The majority of the MFP family are fairly
uniform in size, typically in the range of 367-478 residues, with the
exception of a larger (520 residues) CzcB protein as reported by Dinh et al.
(1994). The KpsE protein with 382 residues is well within this size range.
Most striking is the similarity of predicted topology of these proteins to that
determined for KpsE, namely a short hydrophilic N-terminal region,
followed by a segment of 20 hydrophobic membrane spanning residues and
a large periplasmic domain. Comparisons using the sequence alignment
program of Lipman and Pearson (1985) between the predicted amino acid
sequences of the KpsE protein and four members of the MFPs, CvaA
involved in colicin V export in \textit{E. coli}, HasE required for heme acquisition
in \textit{Serratia marcescens}, HlyD involved in haemolysin export in \textit{E. coli},
and PrtE needed for protease export in \textit{Erwinia chrysanthemi} revealed
respectively 17.5%, 19.2%, 18.9% and 20.7% identities. Taking conservative
Figure 3.6. Western blot of KpsE-BlaM fusion proteins, using antibodies to both BlaM (A) and KpsE (B). Lane order: 1= JM101, 2= JM101pUC18, 3= JM101pCR6, 4= JM101pFE2::24, 5= JM101pFE7::41, 6= JM101pFE8::51, 7= JM101pFE3::211, 8= JM101pFE4::345, 9= JM101pFE5::361, 10= JM101pFE6::381. Molecular weight markers are in kDa.
Figure 3.7. Western blot of cell fractions obtained from KpsE-BlaM fusion proteins. Lane order (left to right): cytoplasmic and periplasmic fractions; 1/2 = JM101pBR, 3/4 = JM101pFE2::24, 5/6 = JM101pFE7::41, 7/8 = JM101pFE8::51, 9/10 = JM101pFE3::211, 11/12 = JM101pBR, 13/14 = JM101pFE4::345, 15/16 = JM101pFE5::361, 17/18 = JM101pFE6::381. Molecular weight markers are in kDa.
Figure 3.8. Western blot of spheroplast (untreated and proteinase K treated) made from KpsE-BlaM fusion proteins. Lane order (left to right): untreated and proteinase K treated; 1/2 = JM101pFE2::24, 3/4 = JM101pFE7::41, 5/6 = JM101pFE8::51, 7/8 = JM101pFE3::211, 9/10 = JM101pFE4::345, 11/12 = JM101pFE5::361, 13/14 = JM101pFE6::381. Molecular weight markers are in kDa.
Figure 3.9. Western blot of spheroplast obtained from KpsE-BlaM fusion protein, pFE6::381 probed with antibody to CAT. Lane order (left to right): untreated and proteinase K treated; 1/2 = JM101pACYC, 3/4 = JM101pFE6::381, 5/6 = JM101pFE6/pACYC. The position of the CAT protein is indicated with an arrow head.
amino acid substitutions into account, the similarity is extended to 65.0%, 59.5%, 65.6% and 58.0% respectively. A similar analysis between the KpsE and the BexC protein of *H. influenzae* and the CtrB protein of *N. meningitidis* capsule gene clusters revealed 28.1% identity (73.2% similarity) and 26.8% identity (73.2% similarity) respectively. Multiple alignment of the predicted amino acid sequences of these membrane fusion proteins and the KpsE (Figure 3.10) revealed similar aligned residues that was observed for the MFP family by Dinh *et al.* (1994), that of mainly hydrophobic (L, V and I) with a few structural (G) and strongly hydrophilic (E and Q), immediately adjacent to the transmembrane hydrophobic segment and before the hydrophobic stretch towards the C-terminus. A phylogenetic tree plot for the five sequences analysed and those of BexC and CtrB (proteins that share great similarity with the KpsE) showed that the polysaccharide exporting BexC and CtrB proteins constitute a subcluster, while the HasE and PrtE constitute another. The colicin V (CvaA), haemolysin (HlyD) and polysaccharide (KpsE) transporters each forming a distinct sub-branch (Figure 3.11).

| KpsE (56) | SKV---AIKRSDDLNSRGLFLGGLGASNPSAEADLYLKEYINSPIIQMA |
| CVAA (63) | GPPVQPVHBEQQLISTKDDVKLVLYLDDTVKDHDRDNLQKLYVVD |
| Hase (72) | GYVQGQVRHDVKEQVAGQVLLLHQKVDQERTQQRDARSHQNSAQQQHQLQ |
| HLYD (195) | GIVQVEEIVKDQVEQAVGVLLKLTQALGADVLKQSLQARLEQIQY |
| Prte (77) | GIVQVQXKQRDVRAGQVLLTLNAVADTSSERLOGSQQYDQLAARLELL |

*G V G L S*  

| KPSE (222) | VLDPOAQQA---ASTLVMQTMDQKLQMEADLEMLTLYL8EDAPVVSAR |
| CVAA (237) | ELTVYQDQVDS---IIRALSDGVRDSL5-TVQWMVNTGDSLQVIERP |
| Hase (272) | GADNREKAEADLGHTQVKEAPVAGTVGKVT3GQ3AGQFGMSIVPPSD |
| HLYD (313) | LIKLELEGNEERQGAGVIRAPVSS3QVQLFVHRQVTATLMQVIVRED |
| Prte (283) | EIVQREARADFLAVRQPAPAVQ8VMSIPTKVIGVAPQ3QGDSYVPE |

**G V G L S**  

| KPSE (274) | LQAQID---EEXKITAQPGDFKNRHALVADVEATQKACKE-KVEFNYEKLKLTTSIK |
| CVAA (287) | LILNVPNDAVIFASA---CDKN1RVEAQPFEGFSATTVKTSRPGPSA |
| Hase (327) | VBEARIQHELAVQV---GLPVQELPAFQMSTTPRFBVNBTVLGAQRETEDK |
| HLYD (368) | TVALQVWNQKFG2IN---GQNAI1KVBFAPPYTVGGYLUQXVXNPLDAEDQK |
| Prte (338) | VDC1HPVWMTKSW6---GLPVBEQQPTAPSEQSTRFPPKGTVYLLSADNLDEXK |

**G V I**  

Figure 3.10. Multiple alignment of the KpsE protein and four members of the MFPs. Alignment positions in the complete multiple alignment of the five proteins are shown above the alignment. The figures in bracket showing residue numbers in each protein are provided at the beginning of each line. Identical residues (*) between the KpsE and any of the MFPs, and similar residues are shown below the alignment.
Figure 3.11. Phylogenetic tree for the KpsE, BexC, CtrB proteins and members of the MFP family. Relative evolutionary distance is proportional to branch length, which is given in arbitrary unit adjacent to the branch. Tree was drawn from the dendrogram data in the multiple alignment output using the Phylogenetic Inference Package (PHYLIP 3.4) of Joseph Felsensteen and the University of Washington (Copyright 1986-1993).
3.2.4 Construction of *kpsE* Null Mutation Within the Capsule Gene Cluster

The need for the construction of a *kpsE* null mutation within the capsule gene cluster on the chromosome arose from the fact that previous *kpsE* mutants constructed through insertion of an oligonucleotide within the gene and through the deletion of the entire gene (Pazzani, 1993) resulted in polarity effect on the expression of downstream genes and were therefore unsuitable for structure-function studies. The construction of a null *kpsE* mutant involved making a translational in-frame internal deletion within the KpsE protein encoded in a plasmid, which was then delivered unto the capsule gene cluster in the chromosome by allelic exchange. The strategy is outlined in figure 3.12.

The suicide vector pCVD442 (Donnenberg and Kaper, 1991) was chosen. Plasmid pCVD442 contains the R6K origin of replication, the *mob* region from plasmid RP4, the *sacB* gene from *Bacillus subtilis* and a *bla* gene conferring resistance to ampicillin (Donnenberg and Kaper, 1991). Replication at the R6K origin requires *π* protein (Kolter et al., 1978) which is not encoded in the pCVD442 and thus can only replicate in strains that supply the *π* protein *in trans*. The pCVD442 derivative carrying truncated *kpsE* gene is unable to replicate in the *pir~* MS101 strain, but on ampicillin resistance selection, merodiploid strains in which pCVD442 derivative has integrated into the chromosome can be isolated. Integration occurs through homologous recombination. A second recombination event can be selected in which the truncated *kpsE* gene replaces the wild-type copy within the capsule gene cluster and the loss of the suicide vector. This selection involves the *sacB* gene, which encodes the enzyme levan sucrase. The expression of this enzyme is toxic for Gram negative bacteria when grown in the presence of 5% sucrose (Blomfield et al., 1991). It means therefore, that when the merodiploid is cultured in the absence of ampicillin selection and plated onto sucrose containing medium, only colonies which have lost the suicide vector will be able to grow (Figure 3.12).
Figure 3.12. Recombination events required for the transfer of a plasmid borne allele onto the chromosome. Integration can occur by recombination on either side of the plasmid allele (a or b). For excision to lead to successful allelic exchange, recombination must occur in the second region of homology. Adapted from Blomfield et al. (1991).
The \( kpsE \) gene was subcloned from plasmid pFE2 on a \( Xba \) I-Sac I fragment ligated into \( Xba \) I-Sac I linearised pUC19 to generate plasmid pFE9. The use of \( Xba \) I-Sac I sites eliminates the \( BamH \) I in pUC19. Plasmid pFE9 was then linearised using \( BamH \) I and \( Dra \) III that cut within the \( kpsE \) gene resulting in a 200 bp deletion (Figure 3.13). The 5' overhang left from the \( BamH \) I digest was end-filled with dNTPs and the 3' overhang produced from the \( Dra \) III digest was end repaired using T4 DNA polymerase (Figure 3.13B). The resulting blunt ends were ligated together to produce plasmid pFE10 (Figure 3.14). This plasmid encodes a truncated KpsE protein which has a deletion of 65 amino acid residues (from 225 to 290). The deleted \( kpsE \) gene in pFE10 was then subcloned as a \( Sph \) I-Sac I fragment in pCVD442 to yield plasmid pFE11 (Figure 3.14). \( E. coli \) SY327\( \Delta \)pir was used as a host to isolate the recombinant plasmid. Plasmid pFE11 was then transformed into \( E. coli \) SM10\( \Delta \)pir and introduced into MS101 by conjugation with selection for ampicillin and streptomycin resistant colonies. Strain SY327\( \Delta \)pir was not used to mobilise pFE11 (pCVD442 derivative) since it lacks the RP4 transfer genes.

The integration of plasmid pFE11 into the correct site within the capsule genes in the chromosome (Figure 3.15) was confirmed by Southern blot of \( BamH \) I digested chromosomal DNA prepared from ampicillin resistant transconjugants (Figure 3.16) using the 1.2 kb PCR amplified \( kpsE \) DNA (see figure 3.3A) as a probe. Chromosomal DNA from \( E. coli \) strain MS101, which has the K5 capsule gene cluster in its normal place on the chromosome was used as a control (Lane 1, figure 3.16). Plasmid pFE11 has lost the \( BamH \) I site within the \( kpsE \) gene, therefore integration occurring on either side of the plasmid allele (a and b in figure 3.12) resulted in a change of the native 11.0 and 5.2 kb \( BamH \) I fragments to produce two sets of fragment sizes (11.6 and 1.0 kb, and 11.6, 5.5 and 1.1 kb, figure 3.16 lanes 2 and 3 respectively) (see figure 3.15). Strains in which the pFE11 had integrated in the correct site were found to be K5 phage sensitive, however this could be due to the presence of a functional wild-type copy of the \( kpsE \) gene (see figure 3.12).

The selection for double recombinants was carried out by growing a merodiploid of MS101 in which the pFE11 had integrated correctly within the capsule gene to late logarithmic phase in L-broth in the absence of ampicillin. The culture dilutions (10\(^{-1}\)-10\(^{-4}\)) were plated onto L-agar (lacking sodium chloride) containing 6% sucrose and streptomycin, followed by incubation overnight at 30°C, as reported by Blomfield et al.
(1991) to be optimal for selection of double recombinants. Sucrose resistant colonies were plated in duplicate onto media containing streptomycin alone and streptomycin with ampicillin, to check for the loss of suicide vector. Sucrose resistant, ampicillin sensitive isolates were checked for the presence of the truncated kpsE gene. Chromosomal DNA prepared from these isolates was digested with BamH I and a Southern blot carried out using the PCR amplified kpsE DNA as probe. The truncated gene in plasmid pFE11 has lost the BamH I site, therefore the native 11.0 kb and 5.2 kb BamH I fragments in MS101 produced a band shift to 16.0 kb in the isolated mutants (Figure 3.17). The replacement of the normal copy by the truncated kpsE gene in the chromosome was confirmed by colony PCR using primers 5'-GGGGTACCCGCCCACGTGAAACTTTTGAT-3' and 5'-CGGAATTCGCCGGATCCTCAATAACAGC-3', that amplify the kpsE gene. The result showed a 1.1 kb PCR product for MS101 chromosomal DNA and a 0.9 kb amplified product for the mutants (Figure 3.18).

Figure 3.13. Construction of translational in-frame deletion in KpsE. The BamHI I and Dru III sites are shown in bold (A). The 5' and 3' overhangs resulting from restriction digest are shown (B). The resulting amino acid residues after end-filling/exonuclease and ligation are shown in bold below the nucleotide sequence (C).
Figure 3.14. Construction of plasmids pFE9, pFE10 and pFE11. Plasmid pFE11 was used for integration into the chromosome of strain MS101. Abbreviations: B, Bam HI; D, Dra III; Nc, Nco I; Sa, Sac I; Sp, Sph I; X, Xmn I and Xb, Xba I.
Figure 3.15 Recombination events showing the integration of plasmid pFE11 into the capsule gene cluster in MS101. L and R denotes respectively left and right sides of the deleted kpsE segment. Sizes are in kilobase. Abbreviation: B, Bam HI; D, Dra III; Hc, Hinc II; Sm, Sma I.
Figure 3.16. Southern blot of \textit{Bam} H I digested chromosomal DNA from MS101(lane 1) and MS101 merodiploids containing pFE11 (lanes 2 and 3). A 1.2 kb PCR-amplified \textit{kpsE} fragment was used as probe. The sizes shown are in kb and were determined by comparison to \textit{\lambda}DNA fragments obtained from \textit{Hind} III digest.

Figure 3.17. Southern blot analysis of \textit{BamH} I digested chromosomal DNA from MS101 (track 1) and MSFE100 (tracks 2-6). A 1.2 kb PCR-amplified \textit{kpsE} DNA product was used as probe. This is represented below the gel photograph. Dark box denotes the deleted \textit{kpsE} segment.
Figure 3.18. Confirmation of the retention of truncated \textit{kpse} gene in MSFE100 by PCR. Chromosomal DNA from MS101 (lane 2) and from 6 MSFE100 strains (lanes 3-7). DNA ladder (lane 1) was used for size markers.
3.2.4.1 Phenotype of MSFE100

The mutant carrying the mutated \textit{kpsE} gene in the chromosome is designated MSFE100 and the encoded truncated protein of 36 kDa was demonstrated in Western blotting using antibody to the KpsE protein (Figure 3.19). The phenotype of MSFE100 as assessed by K5-specific bacteriophage assay showed that the mutant is resistant to the phage. This result means that the mutant has lost its ability to export capsular polysaccharide to the cell surface.

![Western blot analysis of the KpsE protein in MS101 (lane 1) and MSFE100 (lane 2) using anti-KpsE antibody.](image)

**Figure 3.19.** Western blot analysis of the KpsE protein in MS101 (lane 1) and MSFE100 (lane 2) using anti-KpsE antibody.

3.2.4.2 Complementation of \textit{kpsE} Mutation in MSFE100

To ascertain if the mutation in MSFE100 can be complemented by a functional copy of the \textit{kpsE} gene, several plasmids carrying the \textit{kpsE} gene were transformed into this strain and the resulting transformants were assayed for sensitivity to K5-specific bacteriophage. Plasmids pH18 and pFE9 in which the \textit{kpsE} gene is cloned in the high copy number vectors pUC18 (Table 2.2) and pUC19 (Figure 3.14) respectively were able to complement the \textit{kpsE} mutation in MSFE100 and restore K5-phage sensitivity (Table 3.2). Neither plasmid pFE2, which contains an intact \textit{kpsE} gene cloned in pYZ4, nor plasmids pFE3-9 containing truncated derivatives of the \textit{kpsE} gene,
were able to complement MSFE100 (Table 3.2). The introduction of plasmid pPC6, which contains the entire K5 capsule gene cluster (Pazzani et al., 1993) into MSFE100 resulted in turbid K5-phage plaques. This is indicative of the profound effect on the cell surface expression of capsular polysaccharide exerted by the null kpsE mutant.

Table 3.2. Complementation of kpsE Mutation in MSFE100

<table>
<thead>
<tr>
<th>Strain/Plasmid</th>
<th>Sensitivity to K5-specific phage</th>
</tr>
</thead>
<tbody>
<tr>
<td>LE392</td>
<td>R</td>
</tr>
<tr>
<td>MSFE100</td>
<td>R</td>
</tr>
<tr>
<td>MSFE100pFE2-pFE8</td>
<td>R</td>
</tr>
<tr>
<td>MSFE100pH18</td>
<td>S</td>
</tr>
<tr>
<td>MSFE100pFE9</td>
<td>S</td>
</tr>
<tr>
<td>MSFE100pPC6</td>
<td>S*</td>
</tr>
<tr>
<td>LE392pPC6</td>
<td>S</td>
</tr>
<tr>
<td>MS101</td>
<td>S</td>
</tr>
</tbody>
</table>

Abbreviation: R, resistance; S, sensitive; S*, turbid plaque morphology.

3.3 Discussion

The KpsE protein is known to be involved in the cell surface expression of the E. coli K5 capsular polysaccharide. To further gain an insight into the possible role in the export of polysaccharide, the organisation of this protein within the bacterial cytoplasmic membrane was investigated.

From the hydropathic analysis of the predicted amino acid sequence and use of von Heijne's 'positive inside rule', a working model consisting of two transmembrane domains (residues 29-49 and 351-371) was proposed. The boundaries of transmembrane domains were chosen to minimise intramembrane charged residues. These transmembrane domains are in agreement with those predicted for the KpsE protein using the PHDhtm profile neural network system of Rost et al. (1994). This model was analysed using gene fusion techniques. Whilst the low level of Amp^R conferred by a BlaM fusion at amino acid position 24 suggests that this position of KpsE is cytoplasmic, the high levels conferred by BlaM fusion to positions 51, 211 and 345, and the isolation of PhoA^+ insertions at residues 54 and 56 are clear indications of the periplasmic locations of these positions. A relatively high Amp^R was recorded for fusion at amino acid
position 41 compared to fusion at position 361 (two inner membrane locations, from the model). This observation is consistent with the finding of N-terminal hydrophobic domains that may function in translocating the protein in question across the cytoplasmic membrane (Roof et al., 1991; Calamia and Beckwith, 1992; Uhland et al., 1994; Vianney et al., 1994). The 15μg/ml Amp\(^\text{R}\) conferred by BlaM fusion at position 381 as opposed to the <5μg/ml\(^\text{R}\) resistance by fusion at position 24 suggests that the C-terminus of KpsE is membrane associated rather than inserting into the cytoplasm. The accessibility of BlaM at position 381 to proteinase K of spheroplast and the lack of net positive charge at the C-terminus is in keeping with this model. Cell fractionation results showed that the KpsE-BlaM fusion proteins are membrane associated and thus in keeping with the prediction that the KpsE protein is anchored in the cytoplasmic membrane probably via the N-terminus.

The visualisation of KpsE-BlaM fusion proteins using anti-β-lactamase antibodies showed that there was no correlation between the amount of protein expressed and the deduced localisation of the β-lactamase protein. The conferred Amp\(^\text{R}\) by different fusion proteins are therefore true reflection of the position of these segments within the membrane. Thus the KpsE protein has a membrane topology consisting of a short hydrophilic N-terminus in the cytoplasm, followed by a single transmembrane domain, a large periplasmic region of about 302 amino acid residues and a C-terminus that is membrane associated. While it is tempting from the topological model and the low Amp\(^\text{R}\) results of BlaM fusions at amino acid positions 361 and 381 to conclude that the most C-terminal 31 amino acid residues are associated with the cytoplasmic membrane, it is possible that this region of the KpsE protein could be located in the outer membrane. The lack of data on Amp\(^\text{R}\) values of β-lactamase fusions to outer membrane locations makes interpretations of the possible membrane association of the KpsE C-terminus the more difficult. A second limitation that could account for this problem, which can be attributed to β-lactamase and other commonly used (alkaline phosphatase and β-galactosidase) gene fusion techniques is that these enzyme moieties are too large and may interfere with the folding and thus the localisation of the protein under study. More recently, small reporter epitope (C3 neutralisation epitope from poliovirus) tagging, recognised by specific monoclonal antibody was used in determining the topology of the outer membrane protein LamB, and of the 23 hybrid LamB-C3 proteins generated, nine of them did not lead to topological conclusions (Newton et
al., 1996), thus underlining the limitations of genetic foreign epitope insertion for the study of integral membrane proteins. Regardless, of the location of the C-terminus, the KpsE topology within the bacterial membrane is that of a type II bitopic membrane protein, with a single N-terminal membrane spanning domain preceded by a large amino acid segment that is capable of spanning the periplasm.

Although the precise functional mechanism of the KpsE and other members of the MFP proteins has not been elucidated, the result presented here shows that KpsE shares structural features with the HlyD (Mackman et al., 1998), PrtE (Delepelaire and Wandersman, 1991), HasE (Lefort et al., 1994) and CvaA (Skvirsky et al., 1995), proteins involved in the export of substrate across both membranes. The alignment of the predicted amino acid sequence of KpsE protein with the HlyD (required for export of α-haemolysin in E. coli), PrtE (protease export in E. chrysanthemi), HasE (heme acquisition protein in S. marcescens) and CvaA (involved in colicin V export in E. coli) revealed conserved residues that have been observed in members of the MFP family. As has been noted (Dinh et al., 1994), most of these conserved residues are structural (G and P), hydrophobic (L, V, I and A) and a few strongly hydrophilic residues (E and Q), although the functional significance of this residue conservation is not known. In addition, the KpsE like the CvaA, HlyD and PrtE fractionate primarily with the inner membrane. The membrane topology of KpsE presented is consistent with the proposed function of the MFP proteins, which are thought to act in connecting the inner and outer membranes and hence promote transport to the cell exterior.

Whilst the feature of the MFP proteins points to conserved functional mechanism, results of the phylogenetic tree showed that the KpsE as well as CvaA and HlyD each represents distinct sub-branches, while the HasE and PrtE form a subcluster. Both the BexC and CtrB proteins involved in the export of polysaccharide in H. influenzae and N. meningitidis respectively are homologous to the KpsE protein, however the phylogenetic analysis showed that both proteins form a subcluster, indicative of evolutional divergence between KpsE and both proteins. Indeed, the KpsE protein was found to be relatively closer to the CvaA protein, with evolutional distance of 3 units as oppose to the 4 units between the KpsE and the BexC/CtrB cluster. This evolutional divergence of the KpsE from the other proteins, BexC and CtrB also involved in polysaccharide export might reflect specificity in substrate being exported or
differences in organisation of the protein complex involved in the cell surface expression of capsular polysaccharide in these bacteria. The lack of an outer membrane protein in *E. coli*, and a periplasmic protein in *N. meningitidis* and *H. influenzae* capsule gene clusters is in keeping with the later. The phylogenetic grouping of MFP is said to correlate with types of cytoplasmic membrane transport systems with which they function and with the substrate being transported (Dinh et al., 1994). This observation is particularly in agreement with the phylogenetic findings of the polysaccharide (KpsE, BexC and CtrB) and the protein (HasE, HlyD and PrtE) exporters each clustering at different sides of the tree while the peptide exporter (CvaA) represents a distinct branch.

The construction of the *KpsE* null mutant, with an inframe deletion of 65 amino acid within the last 160 amino acid residues of the encoded protein proved sufficient to disrupt the export of capsular polysaccharide to the cell surface. The deleted residues are within the large periplasmic spanning domain of the KpsE protein, that is likely to interact with the periplasmic KpsD protein and may be functionally significant in the role of KpsE in the export of polysaccharide to the cell surface. The last 170 amino acid residues are also essential for the recognition of the KpsE protein by the KpsE antisera as demonstrated by Western blot analysis of proteins encoded by the truncated *kpsE* gene products (Figure 3.6B). The *kpsE* mutant (MSFE100) was complemented by a functional copy of the *kpsE* gene, which indicates that the deletion within *kpsE* did not result in any polarity effect on the expression of the downstream genes. In addition, the truncated KpsE product encoded in strain MSFE100 has been confirmed in Western blot analysis using antibody to the KpsE protein (Figure 3.19). The complementation of the *kpsE* mutant (MSFE100) using plasmid pPC6 resulted in K5-phage turbid plaque in contrast to clear 'normal' plaque morphology observed when plasmids pH18 and pFE9 in which the *kpsE* gene is cloned in high copy number vectors (see section 3.2.4.2) were used. This observation suggests that the *KpsE* mutant exhibits a dominant negative effect on the biosynthetic/export complex postulated to be involved in the expression of group II capsules (Roberts, 1995, 1996; Bliss and Silver, 1996) Therefore, the construction of a non-polar *KpsE* mutant in the chromosome has provided a base to conduct detail structure-function analysis of the last 160 amino acid residues to ascertain their importance for the KpsE role in the cell surface expression of capsular polysaccharide.
Chapter 4
Cloning and Sequencing of the K5 Lyase Enzyme from the K5-specific bacteriophage.

4.1 Introduction

The K5 capsular antigen is a polymer with the structure 4)-β-GlcA-(1,4)-α-GlcNAc-(1 (Vann et al., 1981) and essentially represents desulfo heparin, an intermediate in heparin biosynthesis. Oligosaccharides chemically derived from E. coli K5 capsular polysaccharide has been used as acceptor for both GlcA and GlcNAc transferase reaction implicated in heparin biosynthesis (Lidholt and Lindahl, 1992). The isolation of a coliphage specific for E. coli expressing the K5 capsular antigen was described by Gupta et al. (1982b). This phage has been used in typing of E. coli K5 strains (Gupta et al., 1982a,b). On further analysis, the K5-specific phage was found to posses a lyase enzyme that hydrolysed the K5 capsular polysaccharide. The K5 lyase enzyme is highly specific for the 1,4 linkage between the glucoronic acid and the N-acetyleneuraminic acid, yielding oligosaccharides that could serve as substrates for the study of both heparin and K5 polymer biosynthesis. In addition, the high specificity of the lyase enzyme for the K5 capsular polysaccharide could be exploited in developing rationales for the diagnosis and treatment of infections caused by strains of E. coli K5.

In order to produce large amounts of the K5 lyase enzyme it was decided to clone and overexpress the lyase encoding gene from the K5-specific bacteriophage in E. coli.

4.2 Results

4.2.1 Construction of a Bacteriophage K5 Genomic Library

In order to increase the expression of the cloned gene(s) in E. coli, the plasmid expression vector pTTQ18 (Stark, 1987) was chosen. This vector
contains a tac promoter the activity of which can be induced by the addition of IPTG. In addition, pTTQ18 contains a lac repressor gene to ensure that there is no transcription in the absence of IPTG (Figure 4.1B). This vector is therefore suitable for the expression of cloned gene(s) that may be lethal for cell growth (Stark, 1987). The genes of interest are cloned 3' to the tac promoter such that addition of IPTG leads to an increase in transcription. To ensure that the genomic DNA library will contain a high proportion of recombinants, the BamHI cleaved vector DNA was treated with calf intestinal phosphatase to remove the 5' terminal phosphate. The removal of 5' terminal phosphate was confirmed by the visualisation in agarose gel electrophoresis of predominantly linear vector DNA band in the self-ligated vector mixture. The cloning into the BamHI site results in insertional inactivation of the lacZ' gene. This allows the identification of recombinants which are unable to breakdown X-Gal thus appearing as white colonies, while the non-recombinants appear as blue colonies.

The 45 kb genomic DNA of the K5 bacteriophage was partially cleaved with the restriction endonuclease Sau3A to generate DNA fragments in the range of 0.1 to >12 kb. Fragments were analysed by agarose gel electrophoresis and fragment ranges of 0.9 to 2.0 kb and 2.1 to 7.0 kb were isolated from the gel (Figure 4.1A). Approximately 0.5μg of dephosphorylated vector DNA was self-ligated as a control or with approximately 0.5μg of each isolated DNA fragment range in a final volume of 10μl. E. coli SURE™ was electrotansformed with 2μl of each ligation mixture, selecting for ampicillin resistance colonies. The bacteriophage K5 gene library generated a total of 804 transformants, which were screened for the presence of insert DNA by plating transformants on L-agar containing ampicillin, 1mM IPTG and 0.004% X-Gal, for blue or white colony screening. Of these 88% were white (Lac') colonies, containing a cloned DNA insert. Plasmid DNA was isolated from 50 recombinants and insert DNA confirmed by restriction endonucleases. These recombinants were then screened for lyase activity.
Figure 4.1. Size selected bacteriophage DNA fragments (lanes 2 and 3) (A) cloned in the expression vector pTTQ18 (B). Lanes 1 and 4 are kilobase molecular weight markers. Insert DNA were ligated into the BamHI cleaved and dephosphorylated vector. Abbreviations: E; EcoRI, Sa; Sac I, Kp; Kpn I, Sm; Sma I, B; BamHI, Xb; Xba I, Pst; Pst I, Sph; Sph I, H; Hind III.
4.2.2 Screening of Gene Library for Lyase Activity

The screening of the library was performed by Dr. Peter Hänfling in the laboratory of Professor Dr. K. Jann at the Max-Planck-Institute für Immunbiologie, Freiburg, Germany. Essentially, this involved growing pools of ten recombinants to mid logarithmic phase, then inducing the cells with 1mM IPTG for 2 hrs, followed by sonication to obtain the enzyme lysate. Equal volumes (20µl) of lysate was incubated with the K5 polysaccharide substrate (5mgml⁻¹ in Tris buffer) at 37°C for 6 hrs. Samples were then resolved by PAGE in 1XTBE and gel stained for 3 hrs with Alcian Blue (see section 2.9). The cleavage of polysaccharide substrate was observed after destaining gel with 2% acetic acid in water. Positive pools of ten recombinants were individually screened in a similar manner to obtain a positive clone. A total of 244 recombinants were screened and of these, a positive clone which contains the largest insert DNA fragment was chosen for further analysis.

4.2.3 Nucleotide Sequence of K5 Lyase Clone, pB4

An isolated clone with lyase activity and containing approximately 2.8 kb bacteriophage K5 insert DNA was designated clone pB4. The nucleotide sequence of clone pB4 was determined using double stranded DNA template obtained from Maxiprep procedure (Qiagen). Sequence reactions were performed using the PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing (Perkin Elmer). Initial 800 bp sequence was obtained using the -40 universal (5'-GGTTTCCAGTCACGAC-3') and reverse (5'-AACAGCTATGACCATTG-3') primers, which bind on either sides of the polylinker in pTTQ18. Oligonucleotide primers were successively generated based on the obtained sequences to cover the remaining 2.0 kb and both strands of the cloned DNA fragment. The DNA sequence of clone pB4 was found to be 2872 bp in length (Figure 4.2). Nucleotide sequence searches of the composite GenBank/EMBL database revealed significant DNA homology (83.0% identity) over the first 571 bp of the K5 lyase sequence and those of K1E endosialidase (Figure 4.3). Contained within this region is a putative bacteriophage SP6-like promoter region (302-351 nucleotides) 87 bp from the start of the first open
reading frame (ORF) (Figure 4.2). The promoter was identified on the basis of sequence similarities to those reported by Long et al. (1995) and Melton et al. (1984). It has the same AT bases located at position -2 and -3 of the guanine (G) residue at which transcription initiates in the SP6 promoter (see figure 4.2). These AT bases has been shown to be important functional elements of the SP6 promoter (Melton et al., 1984).

4.2.4 Analysis of Deduced Amino Acid Sequence

Analysis of the reading frame in all six phases revealed two large ORFs, one of which exceeded the range of cloned DNA fragment (Figure 4.2). The two ORFs translated in opposite orientation to the direction of the LacZ' expression in plasmid pTTQ18. The largest complete open reading frame spans a DNA fragment of 1905 bp and encoded a protein of 635 amino acid residues, designated ORF 635. ORF 635 is immediately downstream of a putative SP6-like promoter region and encodes for a predicted protein of 66.9 kDa. The predicted amino acid sequence of this protein shares 50.8% and 50.7% similarities over the entire 635 residues with the bacteriophage KIE endosialidase and the PKIE encoded Endo NE proteins respectively (Figure 4.4). In addition, the predicted amino acid sequence of the putative 66.9 kDa protein is 54% identical (89.4% similarity) to a protein of 820 amino acid residues deposited in the database as Sequence 2 from patent US 5480800 (Figure 4.5). The putative 66.9 kDa protein has a predicted isoelectric point of 7.54. The second and incomplete ORF is located downstream of the putative 66.9 kDa protein and searches with the deduced amino acid sequence did not reveal any significant identity to sequences in the protein databases.
CTTCTGCTGATAAGTACAGAAT(3CTT(3GGGATGGTCCTC3TATCATTAGCTGTAGGTGGGG  

S AD K Y R MLG D GA V S L A V G G  

T S S Q V R L F T S D G T S R T V S L T  

NG N V R L S T S S T G Y L G L G A D A  

T S S Q V R L F T S D G T S R T V S L T  

M T P D S T G T Y A L G S A S R A W S G  

G P T Q A A P T V T S D A R C K T E P L  

T I S D A L L D A W S E V D F V Q F Q Y  

G C G G T T T A C T C A G A G T T G A T T G  

M T P D S T G T Y A L G S A S R A W S G  

G F T Q A A F T V T S D A R C K T E P L  

T I S D A L L D A W S E V D F V Q F Q Y  

L D R V E E K G A D S A R W N F G I A  

Q R A K G G F R N V T G I D A H R Y A F  

L I T P A G S R Y G I R Y E E V L I L E  

A A L M R T I E R M O E A L A A L P K  


110
Figure 4.2. Nucleotide sequence and deduced amino acid sequence of bacteriophage K5 lyase coding region. An SP6-like promoter region and putative Shine-Dalgarno sites (SD) are underlined. The guanine (G) residue at which transcription initiates in the SP6 promoter (Melton et al., 1984) is shown in bold. Nucleotide and amino acid positions are given on either side of the sequence. The K5 lyase nucleotide sequence data have been deposited in the EMBL nucleotide sequence databases under Accession Number Y10025.
Figure 4.3. Nucleotide homology between the K5 lyase sequence and the KIE endosialidase (KILYASE) sequence, published by Long et al. (1995). Positions (bp) in each sequence are shown on the right.
Figure 4.4. Amino acid sequence comparison of the K5 lyase, the K1E endosialidase (KILYASE above) and the PKIE Endo NE protein (KIENDON above). Alignment of the amino acid sequences using CLUSTER W (1.5). Gaps introduced in the alignment by computer are shown with hyphens. The asterisks and dots represent identical and similar amino acid residues respectively.
Figure 4.5. Amino acid sequence comparison of the K5 lyase and Sequence 2 (SEQ2) from patent US 5480800 (Legoux et al., unpublished). Alignment of the amino acid sequence using CLUSTER W (1.5). Gaps introduced in the alignment by the computer are shown in hyphens. The asterisks and dots represent identical and similar amino acid residues respectively.
4.2.5 T7 Expression of Proteins Encoded by the Lyase Clone

The potential ORFs in clone pB4 translated in opposite orientation to the \textit{tac} promoter in the expression vector pTTQ18 and the encoded proteins could not be overexpressed from the native promoter located upstream of the two ORFs. The entire 2.8 kb insert DNA was then excised from clone pB4 in a \textit{Sma} I-\textit{Xba} I fragment, which was then ligated to an \textit{Xba} I-\textit{Hinc} II linearised pGEM-3Zf + to generate plasmid pFE50 (Figure 4.6), thus placing the cloned genes 3' to the T7 promoter. \textit{E. coli} JM109DE3 strain carrying either pFES0 or pGEM3Zf + were induced with 1mM IPTG for protein expression (see section 2.7) and proteins were separated by SDS-PAGE. A coomassie blue stained gel of the total protein lysate from \textit{E. coli} JM109[DE3]pFES0 did not reveal any overexpressed protein band. The encoded proteins were then radiolabelled with $^{35}$S methionine, before separation on an SDS-PAGE. This revealed an over-expressed polypeptide migrating at approximately 62 kDa in SDS-PAGE (Figure 4.7). This polypeptide band was found more intense in the IPTG induced \textit{E. coli} JM109DE3pFES0 cells (Figure 4.7, lane 4), when compared with the uninduced \textit{E. coli} JM109DE3pFES0 cells (Figure 4.7, lane 3). A vector band running at approximately 62 kDa was observed in the induced \textit{E. coli} JM109DE3pGEM-3Zf + (Figure 4.7, lane 2), although not of the same intensity as the over-expressed ORF 635 product.

4.2.6 Induced Expression of Lyase Activity

To determine the activity of the lyase enzyme, \textit{E. coli} JM101 carrying plasmid pB4 and its derivatives (pB4Apst I and pFES0) (Figure 4.6) were grown to mid logarithmic phase (OD$_{600}$ of 0.6) and then induced with 1mM IPTG for 2 hrs. Cells were resuspended in 5 ml of sonication buffer and sonicated to obtain a lysate. This lysate (5\mu l for pFES0 and 30\mu l for pB4/pB4Apst I) was incubated with 20\mu l of 25mgmL$^{-1}$ K5 polysaccharide extract at 37$^\circ$C for 10 mins (pFES0) and for overnight in the case of pB4 and pB4Apst I. Five microlitre (5X10^{6}pfu) of the K5 bacteriophage lysate was similarly incubated as control. Samples were separated by polyacrylamide gel electrophoresis to visualise cleavage of the polysaccharide substrate (Figure 4.8). Polysaccharide cleavage was observed after 10 mins incubation with enzyme lysate obtained from pFES0, 1 hr with the K5 phage and overnight with the pB4 lysate. A slight substrate cleavage was observed
after overnight incubation with the pB4ΔPst I lysate (Figure 4.8, lane 6). Plasmid pB4ΔPst I was constructed by re-ligating a Pst I linearised pB4 clone. A Pst I digest of clone pB4 resulted in the deletion of the first 545 bp sequences in the cloned bacteriophage K5 lyase DNA fragment, which included the first 36 amino acid residues of the encoded lyase protein. This deleted segment of the K5 lyase sequence also contains the putative SP6-like promoter region.

Figure 4.6. Physical map of cloned DNA fragment in clone pB4 (A) and its derivatives, pB4ΔPst I (B) and pFE50 (C). The polycloning sites in the vector are shown in boxes. The thick and thin horizontal lines represent cloned DNA and vector DNA fragments respectively. Thin line arrows indicate the direction cloned genes are transcribed. The position of T7 and tac promoters are shown with shaded arrows. Enzyme abbreviations: Av; Ava I, Ac; Acc I, B; Bam HI, E; EcoRI, H; Hind III, Hc; Hinc II, Kp; Kpn I, P; Pst I, Pv; Pvu II, S; Sal I, Sa; Sac I, Sm; Sma I, Sp; Sph I, Xb; Xba I.
Figure 4.7. Autoradiograph of $^{35}$S methionine labelled polypeptides analysed by SDS-PAGE. Lane order: 1; E. coli JM109DE3 induced, 2; JM109DE3pGEM-3Zf+ induced, 3; JM109DE3pFE50 uninduced, 4; JM109DE3pFE50 induced.

Figure 4.8. The K5 lyase activity gel. Lane order: 1/7; uncleaved capsular polysaccharide (PS), 2; pFE50 after 10 mins incubation with PS, 3; pFE50 after 1 hr, 4; K5 phage lysate after 1 hr, 5; pB4 after overnight incubation, 6; pB4ΔPst I after overnight incubation. Sizes are in kDa.
4.3 Discussion

An isolated recombinant from the bacteriophage K5 DNA gene library has been shown to express K5 capsule-specific lyase activity. The isolated recombinant clone pB4 contains a bacteriophage K5 insert DNA of approximately 2.8 kb. The DNA sequence of clone pB4 was found to be 2872 bp in length and nucleotide sequence searches of the database revealed significant DNA homology (83.0%) over the first 571 bp of the K5 lyase sequence and the DNA sequence of the bacteriophage K1E endosialidase described by Long et al. (1995). This homologous region extends from the non-coding region to the first 43 amino acid residues encoded in the putative 66.9 kDa protein of K5 lyase and a potential open reading frame capable of encoding a 12 kDa protein found upstream of the K1E endosialidase protein (Figure 4.9A). Identified within the non-coding homologous region in both the K5 lyase and the K1E endosialidase sequences is a putative promoter region that shares 75% identity with the bacteriophage SP6 promoter described by Melton et al. (1984) and Long et al. (1995). This DNA homology between the K5 lyase and the K1E endosialidase sequences will suggest that the enzyme-bearing coliphages may be similar. Comparison of nucleotide sequence of this region of the K5 lyase and the sequence 5' to the encoded PK1E Endo NE protein (Gerardy-Schahn et al., 1995) did not reveal any nucleotide sequence identity. Whereas, a putative 12 kDa protein (found upstream of the encoded K1E endosialidase protein) is not present in the K5 lyase sequence, comparison of the predicted amino acid sequences of the encoded K5 lyase protein and the 12 kDa protein shows that they share 41.3% identity over the first 46 amino acid (Figure 4.9B). This observation suggests that the K5 lyase protein may have fused with the N-terminus of the 12 kDa protein, thus further lending support to the earlier view that the K5 and the K1E coliphages are related.

Analysis of the open reading frames (ORFs) has identified two ORFs, one of which exceeded the range of the clone DNA fragment. The first and complete ORF (ORF 635) is immediately downstream of the promoter region and encodes a putative 66.9 kDa protein with 635 amino acid residues. Database searches with ORF 635 sequence revealed very high
Figure 4.9. Sequence comparison of the bacteriophage K5 lyase (K5) and the K1E endosialidase (K1). The translated regions of the sequence are shown as unfilled boxes and the non-coding regions as single horizontal lines. Shaded boxes represent regions of homology between the two sequences. The identified promoter region is shown (P). The N and C denote N- and C-termini respectively.
amino acid similarity (89.0%) with a protein of 820 amino acid residues deposited under a US patent as sequence 2 (pat US 5480800). The putative 66.9 kDa protein encoded by ORF 635 is also homologous (50.8% similarity) to bacteriophage KIE endosialidase protein described by Long et al. (1995) and the bacteriophage K1 endoneuraminidase protein (Gerardy-Schahn et al., 1995). The amino acid sequence of the K5 lyase protein and the two KIE encoded proteins show a high degree of identity within the last 33 C-terminal residues (see figure 4.4). The last 32 C-terminal amino acid residues have been implicated in bacteriophage PK1-endoneuraminidase (Endo NE) activity (Gerardy-Schahn et al., 1995) and it is likely that this region will equally be essential for the enzymatic activity of the bacteriophage K5-encoded lyase enzyme. Sequence 2 from patent US 5480800 encodes an enzyme that fragments n-acetylheparosan (Legoux et al., unpublished) and this substrate is essentially identical to the K5 capsular polysaccharide (Vann et al., 1981). Although not yet demonstrated, it is possible that K5 capsular polysaccharide, heparin and heparin-like compounds including n-acetylheparosan can be substrates for both the K5 lyase and the enzyme encoded by Sequence 2. In such a scenario, the K5 lyase enzyme may be functionally important in fragmenting heparin and its precursor molecules to yield substances that could be of biotechnological value. The chemical modifications of the K5 capsular polysaccharide to yield products with antithrombin activity has been reported (Casu et al., 1994). In addition, the K5 lyase enzyme has been characterised and shown to randomly cleave the K5 polysaccharide, yielding oligosaccharide products that are mainly hexa-, octa-, and decasaccharides with unsaturated 4,5-glucuronic acids at their nonreducing ends (Hänfling et al., 1996). The oligosaccharide products of the K5 lyase action could therefore serve as acceptors in the study of polymerisation of the K5 polysaccharide, especially the enzyme involved in the initiation of such process. Searches with the deduced amino acid sequence of the second and incomplete ORF did not reveal any significant identity to sequences in the database, and it is assumed not to be required for the function of the encoded K5 lyase enzyme.

Expression of the K5 lyase protein in *E. coli* revealed a protein migrating at approximately 62 kDa in SDS-PAGE, which differs from the predicted molecular mass of 66.9 kDa. Such discrepancy in the apparent molecular weight in SDS-PAGE and the predicted molecular weight has been observed with the Endo NE protein which has a predicted molecular mass
of 90.6 kDa, but migrates at -74 kDa in SDS-PAGE (Gerardy-Schahn et al., 1995). Although the discrepancy in molecular weight observed with the Endo NE protein has been attributed to post translational processing (Gerardy-Schahn et al., 1995), this is not likely to be the case with the encoded K5 protein as no peptidase processing site is present within the protein. A likely explanation to the discrepancy in molecular weight observed with the K5 lyase protein can be due to a translational start located downstream from that proposed in figure 4.2. Though, such potential translational start that will result in a protein of 62 kDa was not seen, it will be necessary to establish the N-terminal amino acid composition of the K5 lyase protein through N-terminal sequencing, to ascertain the translational start site.

In vitro determination of lyase activity from the plasmid encoded enzyme showed that the lyase was successfully over-expressed using the T7 promoter, with K5 capsular polysaccharide cleavage within 10 mins. The enzyme encoded in clone pB4 translated in the opposite orientation to the tac promoter in plasmid vector pTTQ18 and needed over 8 hrs to cleave the polysaccharide substrate. The reduced activity observed for clone pB4 could be attributed to low level of protein expression from the native SP6-like promoter located upstream of the K5 lyase open reading frame. A slight substrate degradation was observed with the enzyme expressed from plasmid pB4ΔPst I which lacks the native promoter region, including the first 36 amino acid residues of the encoded lyase protein and with the tac promoter in the opposite orientation to the translated protein (see figure 4.6). This observed activity could have resulted from a read-through expression from unknown promoter sequence(s) in the vector. In addition, the reduction in activity observed with clone pB4ΔPst I could have resulted from the deletion of the native promoter sequences and the first 36 amino acid residues of the K5 lyase protein. One way to establish the effect of this deletion will be to clone the K5 lyase DNA fragments in plasmids pB4 and pB4ΔPst I 3' to the tac promoter in pTTQ18 and then to determine the enzyme activity obtained from these plasmid derivatives.
Chapter 5

Discussion

The expression of group II capsular polysaccharide in E. coli is believed to be determined by a membrane-bound hetero-oligomeric biosynthetic/translocation complex (Roberts, 1995, 1996; Bliss and Silver, 1996). A general picture emerging from the processes involved in this complex is viewed to proceed in three stages of biosynthesis, translocation across the inner membrane and the cell surface expression. It should be noted that this categorisation is only for the sake of convenience, as biosynthesis and translocation appear to be in a concurrent process with no intermediate(s). From studies of the E. coli K5 and K1 capsule gene clusters, it has been established that biosynthesis of group II capsular polysaccharides is achieved through a series of reactions mediated by region 2 encoded proteins. Translocation of polysaccharide to the cell surface is mediated by proteins encoded by regions 1 and 3. The region 3 encoded proteins, KpsM and KpsT are members of the ATP-binding cassette (ABC) superfamily, involved in the export of capsular polysaccharide across the cytoplasmic membrane (Smith et al., 1990; Pavelka et al., 1991; Pigeon and Silver, 1994). After the KpsMT-mediated transport across the cytoplasmic membrane, the final stage involves the export across the periplasmic space to the cell surface. This stage in the export of capsular polysaccharide is the least-characterised. Mutational analysis of region 1 genes have provided some insights into the encoded proteins that may be involved. Mutations in kpsE and kpsD abolished capsule expression on the cell surface (Bronner et al., 1993a). In addition, the localisation of both KpsD and KpsE suggests a role in the final stage of polysaccharide export. The KpsD is a periplasmic protein that is secreted via the classical sec-dependent pathway, and its cleavable leader sequence has been established (Wunder et al., 1994). Our earlier data has shown that the KpsE protein is localised in the inner membrane (Rosenow et al., 1995). KpsE is homologous to BexC and CtrB proteins from the capsule gene clusters of H. influenzae and N. meningitidis respectively, and on the basis of their likely function and localisation they have been suggested to form a component of the ABC-mediated transport of capsular polysaccharide in these bacteria (Reizer et al., 1992). Thus, the membrane organisation of the KpsE protein
was investigated to ascertain its possible export role in this model for the cell surface expression of group II capsular polysaccharide.

A topological model consisting of two transmembrane domains with the N- and C-termini located in the cytoplasm was proposed for the KpsE, based on the hydrophatic profile of the predicted amino acid sequence and by using the von Heijne's 'positive inside rule'. The predicted transmembrane domains (Rost et al., 1994) and those proposed in the topological model for the KpsE were in agreement. The proposed membrane topology was investigated using both TnphoA mutagenesis and β-lactamase gene fusion approach. TnphoA mutagenesis generated only two PhoA+ fusions at residues 54 and 56 within the KpsE protein. The PhoA fusions within the KpsE represented a low isolation frequency (2 out of 15 insertions within the KpsE in a total of 128 plasmid insertion mutants analysed). This low frequency of isolation can be attributed to the inefficient random mutagenesis approach in defining protein topology, as KpsE will only constitute a small target area relative to other genes within the recipient in the mating experiment. Secondly, it could be due to some hot-spot for TnphoA within the recipient chromosome or to some yet unknown intrinsic properties of the kpsE gene, as previous attempt by Cieslewicz et al. (1995) to generate TnphoA fusions within the kpsE gene proved unsuccessful. Further β-lactamase (blaM) gene fusion analysis provided additional insight into the membrane organisation of the KpsE protein. A BlaM fusion at amino acid position 24 of the KpsE gave a low level of ampicillin resistance (Amp'), suggestive of the cytoplasmic location of this position of the KpsE. The periplasmic locations of residues 51, 211 and 345 were indicated by the high levels of Amp' conferred by BlaM fusions at these positions and by the isolation of PhoA+ insertions at residues 54 and 56. BlaM fusion at position 381 gave a high Amp' value compared to the fusion at the cytoplasmic position 24, which suggests that the C-terminus of the KpsE protein is not located in the cytoplasm. Confirming this view is the proteolysis of BlaM fused to position 381, and the lack of net positive charge at the C-terminus is in keeping with this finding. Whilst, it is certain that the C-terminus of the KpsE is not inserted into the cytoplasm thus confirming the absence of a second transmembrane domain, the exact localisation with regards to the inner or outer membrane remains uncertain. It was not possible to draw a topological conclusion from the Amp' values of BlaM fused to residues 361 (10μg/ml-1) and 381 (15μg/ml-1) which represent two most C-terminal
fusions to the KpsE protein. Hence, with the C-terminus not located within the cytoplasm, the KpsE can adopt two possible membrane topology (Figure 5.1) in which the C-terminus may be associated either with the inner or outer membrane (Figure 5.1, panels A and B respectively). The "coiled coil" conformation of the large periplasmic segment as suggested from the solvent accessibility prediction using the neural network profile of Rost et al. (1994), indicates that this segment of the KpsE may not loop back into the inner membrane after spanning the periplasmic space. This therefore favours a membrane topology in which the C-terminus is associated with the outer membrane (Panel B, figure 5.1).

![Diagram showing membrane topology of KpsE protein](image)

Figure 5.1. Membrane topology of the KpsE protein, showing the association of the C-terminus either with the inner membrane (A) or with the outer membrane (B). The inner (IM) and outer (OM) membranes are indicated.

One way of confirming the C-terminal location of the KpsE protein will be through carboxyl peptidase and mapping. In addition, epitope tagging has been used to study the topology of outer membrane proteins (Newton et al., 1996). This approach is particularly useful in providing topological information on the external or periplasmic locations of segments of protein under study. Thus, in the case of the KpsE protein the C-terminus association with the outer rather than the inner membrane can only be confirmed using epitope tagging when such fusion to the KpsE C-terminus is exposed to the cell surface, which can then be detected through cytofluorometric (FACS) analysis with the whole cell. Also, the deletion of
the transmembrane segment may result in the KpsE being anchored to the outer membrane which can then be probed for the presence of KpsE protein or an epitope tagged to its C-terminus.

The membrane organisation of KpsE, with an N-terminal transmembrane domain followed by a large periplasmic domain and a C-terminus that may be associated with the outer membrane is identical to the topology adopted by the CvaA (Skvirsky et al., 1995), a protein that belongs to the membrane fusion protein (MFP) family. Similar membrane topology was established for LcnD, a protein implicated in the transport of bacteriocins from *Lactococcus lactis* and this protein was proposed as the MFP homologue in the transport process (Frank et al., 1996). Although analysis of the predicted amino acid sequences between the KpsE and some members of the MFP family has revealed identities (17.5 - 20.7%) which are within the "twilight zone" of 15-25% identity described by Doolittle (1986), the membrane topology of the KpsE with a large periplasmic domain and a C-terminus that may be associated with the outer membrane is consistent with the proposed function of the MFPs (Fath and Kolter, 1993; Dinh et al., 1994). This observation is in line with the view of related proteins with similar function that have diverge over the years (Doolittle, 1986). Consistent with this view is the finding that ABC exporters of unrelated polypeptides show poor sequence identity (Delepelaire and Wandersman, 1990) and the observation of the polysaccharide exporters (KpsE, BexC, and CtrB) clustering to one side in the phylogenetic tree analysis of these proteins and some members of the MFPs.

Thus, in the model for the cell surface expression of group II capsular polysaccharide, the KpsE protein is proposed as the MFP component (on the basis of structural similarity) in the KpsMT-mediated export. In this role, the KpsE may interact with the KpsD protein to effectively export capsular polysaccharide unto the cell surface. Though the function of KpsD remains enigmatic with no known homologues even in the very similar polysaccharide export systems like hex and ctr which have dedicated outer membrane components, the KpsD protein has been proposed to function in the "recruitment" of a porin (protein K) to the transport machinery complex (Bliss and Silver, 1996). This porin was then suggested to complete the assembly of capsular polysaccharide on the cell surface (Bliss and Silver, 1996). Alternatively, the finding that the predicted amino acid sequence of KpsD shares a 20.1% identity with the TolC protein, an outer
membrane component in haemolysin secretion (Figure 5.2) suggests that KpsD may function in a role analogous to outer membrane proteins. In addition, the predicted amino acid sequences of the KpsD and the GumB protein of Xanthomonas campestris is 25.3% identical over 170 amino acid residues. The GumB protein has been implicated in the expression of xanthan gum polysaccharide in this bacterium (Harding et al., 1987). Therefore, in line with the proposed function of the MFPs it is possible that both KpsE and KpsD may interact to effectively allow fusion of the outer and inner membranes or the formation of an oligomeric pore, that will enable the passage of polysaccharide directly to the cell surface. In keeping with this view is the observation that cell surface expression of capsular polysaccharide may be achieved through sites (Bayer junctions) where the inner and outer membranes come in close aperture (Bayer and Thurow, 1977). Whereas the existence of such adhesion zones between the inner and outer membrane has been disputed (Kellenberger, 1990), their involvement in the expression of capsular polysaccharide have been suggested by electron microscopy (Bayer, 1990; Krönke et al., 1990b). The model (Figure 5.3) for the expression of capsular polysaccharide in E. coli K5 serotype therefore, involves the biosynthesis of the K5 polymer at the inner face of the cytoplasmic membrane by proteins encoded by region 2 genes (Petit et al., 1995; Roberts, 1995, 1996). The lipid-linked and KDO substituted polysaccharide is then translocated across the inner membrane by the ABC exporters, KpsM and KpsT proteins encoded by region 3 genes (Smith et al., 1990) and the cell surface assembly of the capsular polysaccharide determined by the KpsE and KpsD proteins (Figure 5.3).

In the light of existing knowledge on the mechanism involved in the cell surface expression of group II capsular polysaccharides, it is possible that the KpsE, KpsD, KpsM and KpsT may form a hetero-oligomeric complex involved in the export of capsular polysaccharide in E. coli. However, it is not known whether the KpsE interacts physically with the KpsM and T or directly with the polysaccharide on crossing the cytoplasmic membrane. Thus, further work will be to understand the KpsE function and interaction with other kps encoded proteins. This will require some protein-protein interaction analysis, through cross-linking and immuno-
Figure 5.2. Amino acid sequence comparison of the KpsD and the TolC proteins. Alignment of the amino acid sequences using CLUSTER W (1.5). Gaps introduced by the computer are shown with hyphens. The asterisks and dots denotes identical and similar amino acid residues respectively.
Figure 5.3. The schematic diagram showing the biosynthesis and export of polysaccharide in *E. coli* K5. The proteins believed to be involved in each process are represented.
precipitation studies. Such cross-linking studies can be achieved through the use of several cross-linking agents like glutaraldehyde treatment of whole cells, dimethylsuberimate (DMS) and acylazides treatments of proteins under study, followed by two dimensional SDS-PAGE analysis. Likewise, evidence of protein interaction can be obtained through immunoprecipitation of a cross-linked protein product in the presence or absence of a cross-linking reagent, using monoclonal antibody to one of the proteins under study. In addition, site-directed mutagenesis especially within the periplasmic domain may reveal conserved residues that may be necessary for the KpsE function. Furthermore, complementation studies between the KpsE and the MFPs may confirm the functional conservation of these proteins in the ABC transport system. Also, since there are no KpsD homologues in the bex and ctr capsule gene clusters of H. influenzae and N. meningitidis respectively, it will be interesting to see if kpsE mutation can be complemented by its homologues, BexC and CtrB proteins. In addition, it will be important to show if the BexC or CtrB proteins can function in a kpsE/D double mutant. These complementation studies will provide an insight as to whether the KpsD functions in a role analogous to outer membrane proteins, which are involved in the export process in the bex and ctr systems.

The cell surface expression of capsular polysaccharide was abolished by an inframe deletion of 65 amino acid residues of the KpsE protein encoded within the kps gene locus. The kpsE mutant can be complemented by a functional kpsE gene. The complementation of this mutant using the plasmid construct, pPC6 which encodes the kps genes excluding the kpsF (see table 2.2) resulted in turbid plaque morphology as oppose to clear 'normal' plaque morphology observed with clone pFE9 (see table 3.2). One interpretation of the result is that the mutated kpsE gene encodes a protein (KpsE*) which forms some form of dysfunctional complex disrupting the biosynthetic/export complex. This suggests that the kpsE mutant exerts a dominant negative effect on this complex. In support of this hypothesis is the finding that this dysfunctional complex can be titrated out using multiple copy of the kpsE gene (see section 3.2.4.2). This observation is consistent with the existence of a biosynthetic/export complex and points to an important role for the KpsE protein in this complex.
The polysaccharide lyases are finding increasing application in a number of biological processes, including the structural determination and analysis of glycosaminoglycans found in tissues and body fluids (Linhardt et al., 1991, 1992) and the preparation of novel therapeutic agents (Jandik et al., 1994). The isolation and characterisation at molecular level could therefore provide some useful information on this least characterised and important group of enzymes.

The isolation and sequencing of the bacteriophage K5 lyase gene has identified nucleotide sequence homology (83.0%) over the first 571 bp of the K5 lyase sequence to the nucleotide sequence of the bacteriophage K1E endosialidase described by Long et al. (1995). A promoter region that shares 75% identity with the bacteriophage SP6 promoter described by Melton et al. (1984) and Long et al. (1995) was identified within the non-coding homologous region. This nucleotide homology may reflect a common origin for these bacteriophage genes. Whilst, an open reading frame capable of encoding a 12 kDa protein was found between the promoter region and the encoded K1E endosialidase protein, no potential open reading frame was observed upstream of the encoded K5 lyase protein. However, comparison of the predicted amino acid sequence between the K5 lyase and the 12 kDa revealed a 41.3% identity over the first 46 amino acid residues of the two proteins. Although the 12 kDa has been postulated not to play a role in the K1E endosialidase activity, if the suggestion of a binding role for a 38 kDa protein that co-purifies with the Endo NE protein (Gerardy-Schah et al., 1995) is true for the 12 kDa protein, then it is likely that the N-terminus of the encoded K5 lyase protein may play a similar role in binding of the enzyme to the phage. A similar binding function for the N-terminus of a bacteriophage encoded Endo NF protein has been postulated (Petter and Vimr, 1993). The deduced amino acid sequences of the two bacteriophage K1 proteins (Endo NE and K1E endosialidase) and that of the K5 lyase protein are identical in the last C-terminal 32 residues. This region have been implicated in the Endo NE activity and was suggested to function in the stabilisation of the secondary or tertiary structures of the Endo NE (Gerardy-Schah et al., 1995). Likewise, this region in the K5 lyase protein may play a similar role and may therefore be required for the insertion of the enzyme into the phage head assembly. As a component of the phage tail structure, the K5 lyase specificity for K5 capsular polysaccharide defines host range interaction, which explains the low identity observed over the central region between
the predicted amino acid sequences of the K5 lyase and the K1E endosialidase. This view together with the nucleotide homology suggest that the K5 and the K1E phages may belong to the same family of DNA bacteriophages but with different capsule specificities.

Furthermore, database searches with the deduced amino acid sequence of the K5 lyase protein revealed high amino acid identity (54.0%) (89.4% similarity) to a protein of 820 amino acid residues deposited under a US patent as sequence 2 (pat:US15480800). This homology does imply that these two proteins are not the same but are very similar. The 820 amino acid (sequence 2) protein functions in fragmenting n-acetylheparosan (Legoux et al., unpublished), a substrate with similar structure to the K5 capsular polysaccharide. Therefore, the similarity between this enzyme and the K5 lyase suggest that the K5 lyase enzyme may be capable of utilising n-acetylheparosan as substrate.

The activity of the plasmid encoded K5 lyase enzyme has been demonstrated in vitro and shown to effectively cleave the K5 polysaccharide substrate. The specificity of the bacteriophage-borne K5 lyase on K5 capsular polysaccharide which is a polymer with similar structure to an intermediate of heparin biosynthesis could therefore prove to be a valuable tool for the study of the complex reactions involved in the biosynthesis of both heparin and K5 capsular polysaccharide. Hence, future studies on the K5 lyase will include the determination of substrate specificity against heparin or heparin-like compounds found in tissues and biological fluids and its applications.

In one of such likely applications, the gene encoding the K5 lyase can be introduced into the kps gene cluster where it can be induced in a controlled fermentation process to yield defined K5 oligosaccharides that can be processed into heparin-like substances with anti-thrombin activity as have been reported (Casu et al., 1994) using chemically modified E. coli K5 polysaccharide. In addition, a specificity of the K5 lyase for heparin and heparin-like substances means that this enzyme can be employed for the removal of heparin from circulation, especially during extra-corporeal therapy (Langer et al., 1982). Furthermore, enzymatically derived oligosaccharides could serve as sugar acceptors for transferase reactions involved in both heparin and K5 polysaccharide biosynthesis. Such GlcA and GlcNAc transfer reactions involved in heparin/heparan sulphate biosynthesis have been demonstrated using chemically derived
oligosaccharides from the *E. coli* K5 polysaccharide (Lind *et al.*, 1993). This suggests a possible prospect for the commercial production of heparin.

The K5 lyase can also be used *in vitro* to study the biosynthesis and export of the K5 capsular polysaccharide by different *kps* mutants. Here, polysaccharide produced by these mutants are radiolabelled, and spheroplast or inverted membrane vesicles made from the mutant(s) under study. The K5 lyase enzyme can then be used to cleave radiolabelled polysaccharides exposed to the periplasmic side (in the case of spheroplasts) and the cytoplasmic side (for reverse membrane vesicles) of the inner membrane, before assaying for the amount of radiolabelled polymer within the cytoplasm or exported across the inner membrane to the periplasm. The difference in radioactivity between the spheroplast for instance, and the intact cells will give an indication of the amount of polysaccharide exported across the inner membrane. In a pulse-chase experiment where cells are collected at different time points, such data generated can give an indication of the efficiency in biosynthesis and export of capsular polysaccharide by different mutants.

Thus, the purification of the K5 lyase enzyme will be of great importance in providing a homogenous enzyme that could find useful biotechnological applications. Structural studies including site-directed mutagenesis will be required to address some of the structure-function properties of the K5 lyase enzyme.

As a result of my studies, we have gained significant insight into the mechanisms involved in the cell surface expression of group II capsular polysaccharides in *E. coli*. In addition, the cloning and expression of an active K5 lyase may provide an enzyme for use in a number of commercial biotechnological ventures.
References


