Identification and characterisation of a novel gene *algK* from the alginate biosynthetic cluster of *Pseudomonas aeruginosa.*

Simon J. Aarons

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

May 1995

DECLARATION

I declare that this thesis was composed by myself and the research presented is my own except where otherwise stated.

Simon J. Aarons

May 1995.

ACKNOWLEDGEMENTS

There are a number of people I have to thank for various contributions and assistance throughout my Ph.D., particularly my Supervisors Dr. M. P. Gallagher and Professor I. W. Sutherland; all the people in the laboratory past and present (in no particular order) Karen Scott, Dr. Jane Craig, Patrick Taylor, Moira Priestley, Dr. Patricia McCabe, Dr. Kevin Francis, Dr. Yian-Biao Zhang, Lisa Foster, and all the honours students; Aileen Greig and Joan Davidson for their washing and autoclaving service; Graham Brown and Frank Johnston for their photographic service and the SERC for some financial support. I would also like to thank the rest of the 4th floor for generating such a unique working environment and their contribution towards making the last 4 years such an experience.

CONTENTS

FIGURES AND TABLES

Chapter 3:

Chapter 4:

Chapter 5:

ABBREVIATIONS

Ab	antibody
ABC	ATP binding cassette
cAMP	adenosine 3', 5'-cyclic monophosphate
Amp^R	ampicillin resistant
ADP	adenosine diphosphate
ATP	adenosine triphosphate
dATP	deoxyadenosine triphosphate
ddATP	dideoxyadenosine triphosphate
bp	basepairs
°C	degrees Celsius
CF	cystic fibrosis
CFTR	cystic fibrosis transductance regulator
Ci	Curies
CIP	calf intestinal phosphatase
Cml^R	chloramphenicol resistant
CPS	capsular polysaccharide
dCTP	deoxycytidine triphosphate
ddCTP	dideoxycytidine triphosphate
Da	Daltons
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetate
EM	electron microscopy
EMS	ethylmethanesulphonate
EPS	exopolysaccharide
ExoIII	exonuclease III
g	grams
g	standard acceleration of gravity
GDP	guanosine diphosphate
GMD	GDP-mannose dehydrogenase
GMP	GDP-mannose pyrophosphorylase
dGTP	deoxyguanosine triphosphate
ddGTP	dideoxyguanosine triphosphate
dH_2O	distilled water
ddH_2O	deionised distilled water
IgG	immunoglobulin G
IPTG	isopropyl-β-D-thiogalactoside
kb	kilobases
Kan^R	kanamycin resistant
$lb./in^2$	pounds per square inch
LB	Luria-Bertani
LPS	lipopolysaccharide

LSB	Laemmli sample buffer
M	molar
mA	milliamps
MIC	minimum inhibitory concentration
mg	milligrams
MW	molecular weight
µl	microlitres
NAD	nicotinamide adenine dinucleotide
dNTP	deoxynucleotide triphosphate
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PEG	polyethyleneglycol
pg	picograms
PMI	phosphomannoisomerase
PMM	phosphomannomutase
RBS	ribosome binding site
RNA	ribonucleic acid
RNAse	ribonuclease
SDS	sodium dodecyl sulphate
TE	Tris/EDTA buffer
TEM	specifies the source of β-lactamase (i.e. from *E. coli* TEM, Datta and Kontomichalou, 1965).
TEMED	N, N, N'-tetramethylethylenediamine
Tet^R	tetracycline resistant
dTTP	deoxythymidine triphosphate
ddTTP	dideoxythymidine triphosphate
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet

ABSTRACT

Pseudomonas aeruginosa is an opportunistic pathogen that can cause chronic pulmonary infections in patients with cystic fibrosis (CF), the most common, lethal, hereditary disease amongst the Caucasian population. Colonisation of the CF lung by *P. aeruginosa* is greatly facilitated by the production of an exopolysaccharide called alginate and the formation of mucoid microcolonies. Such microcolonies have enhanced survival in the host respiratory system, with the alginate coating preventing phagocytosis, impeding antibiotic penetration and facilitating bacterial cell adherence to host tissue. Once *P. aeruginosa* is established in the airway it releases proteases which cause tissue damage and impair the humoral response further. The alginate-coated microcolony is also thought to elicit a hyper-responsiveness reaction within the lung and this together with protease action is considered to contribute substantially to the immune-mediated damage which ensues. Chronic pulmonary infection by mucoid *P. aeruginosa* frequently leads to mortality amongst CF patients as a consequence of the pulmonary stress which inevitably follows.

In *P. aeruginosa*, the majority of genes involved in alginate synthesis and export are clustered as a polycistronic operon at 34 minutes on the chromosome. Within this operon are two uncharacterised regions, thought to encode the final unidentified steps in alginate exopolysaccharide production. The characterisation of one of these two regions is reported here.

Initially, Southern blot analysis was undertaken on plasmids bearing fragments from the 34 minute region to identify a putative ATP binding cassette transporter. However, this proved unsuccessful. As an alternative approach, the region between *alg44* (which encodes a protein of unknown function) and *algE* (which encodes an outer membrane, alginate-permeable pore) was subcloned and sequenced. Sequence analysis revealed a continuous open reading frame of 1475 base pairs which encodes a putative polypeptide of 475 amino acids with a predicted

molecular weight of 52,470Da. This genetic locus was termed *algK* and a computer search through the available databases did not reveal any obvious homologous proteins to the predicted AlgK polypeptide. However, computer analysis did predict that AlgK may have an N-terminal lipoprotein signal sequence for extracytoplasmic secretion. Such signal sequences are generally removed by signal peptidase II before anchoring the mature polypeptide, by the N-terminal addition of fatty acids, to either the inner or outer membrane.

To confirm the sequence data, protein expression surveys in both *E. coli* and *P. aeruginosa* strains was performed. Results from these experiments confirmed, both the presence of the open reading frame for *algK* and the approximate molecular weight of the AlgK polypeptide. Expression studies of *algK* within *E. coli* and *P. aeruginosa* in the presence of [^{35}S]methionine also suggested that the protein was post-translationally modified, consistent with the view that cleavage and removal of a signal peptide was occurring.

The topological organisation and subcellular localisation of AlgK within *E. coli* has been analysed experimentally with the aid of 3' deletions of *algK* which result in in-frame fusions to mature β-lactamase (the level of ampicillin resistance conveyed by the fusion proteins reflects on their subcellular locations). These data predicted that mature AlgK had an entirely periplasmic location (although potentially anchored to a membrane) and also provides supporting evidence for the cleavage of the AlgK pre-protein.

As a further part of the study, the *algK* gene within the *P. aeruginosa* chromosome of a mucoid *P. aeruginosa* strain was insertionally inactivated using a suicide vector. The resulting mutant was found to display a non-mucoid phenotype. Complementation analysis of this mutant failed to reveal whether the AlgK protein is obligately required for alginate exopolysaccharide production in *P. aeruginosa*.

Chapter 1

Introduction.

1.1. GENERAL INTRODUCTION

Pseudomonas aeruginosa is extremely widespread in nature and in the majority of environments, is effectively harmless. Despite this finding, *P. aeruginosa* is extremely adept as an opportunistic pathogen and can cause life threatening infections in immunosuppressed hosts, as well as in cancer patients undergoing chemotherapy. However, of most significance perhaps, is the common association of pathogenic *P. aeruginosa* with lung infections in patients suffering from cystic fibrosis (Govan *et al*, 1986).

Cystic fibrosis (CF) is the most common inheritable lethal disorder found amongst Caucasians. An estimated 5% of the Caucasian population are carriers, with the disease being inherited as an autosomal recessive trait in 1:2000 live births (Bye *et al.*, 1994). The leading cause of mortality and morbidity in CF is chronic respiratory disease complicated by intractable infections with mucoid *P. aeruginosa* and other bacterial pathogens, and together these account for over 90% of fatalities (Govan, 1988).

Cystic fibrosis results from aberrations in the gene encoding the cystic fibrosis transductance regulator (CFTR), which normally functions as a chloride channel. In the CF lung, the deficiency in CFTR impairs chloride and water secretion by the respiratory epithelial cells, resulting in the formation of a layer of thick, dehydrated mucus, which renders the respiratory mucosa excessively susceptible to bacterial infection. Antibiotic therapy clears the majority of pathogenic microbial infections but *P. aeruginosa* is extremely recalcitrant to even the most severe treatment and thus becomes the predominant coloniser of the CF lung (May *et al.*, 1991). The lung environment of CF individuals also induces *P. aeruginosa* to produce a highly viscous alginate exopolysaccharide. Alginate is a linear copolymer of β(1-4)-linked D-mannuronate (partially *O*-acetylated) and its C-5

epimer, L-guluronic acid (Evans and Linker, 1973); the molecular weight of the polysaccharide may be in excess of 1 million daltons. Alginate clogs the airways of CF patients further and also allows the *P. aeruginosa* in the CF lung to form a polyanionic biofilm, consisting of exopolysaccharide-embedded cells (Lam *et al.*, 1980). The pattern of microcolony formation has profound implications for the course of the disease, as the alginate biofilm protects *P. aeruginosa* from phagocytosis (Govan and Harris, 1986), impedes antibiotic penetration (Govan and Fyfe, 1978) and also, facilitates bacterial adherence to the respiratory tract (Rhamphal and Pier, 1985; Saiman *et al.*, 1992). Emergence of mucoid strains allows the establishment of chronic infection within the CF lung, ultimately leading to irreversible lung damage and respiratory failure. Indeed the overwhelming majority of CF patients die due to respiratory failure associated with pulmonary hypertension and *cor pulmonale* (Hodson and Warner., 1992).

The biosynthetic pathway of *P. aeruginosa* for alginate production shares many common genetic and enzymatic similarities with the exopolysaccharide biosynthetic pathways of other Gram-negative bacteria. The enzymatic steps and proteins involved in the conversion of the initial precursor, fructose 6-phosphate to GDP-mannuronic acid, the monomeric unit providing the activated monosaccharide that is polymerised to form alginate, are well characterised. However, the subsequent steps of polymerisation, acetylation and export of the alginate exopolysaccharide of *P. aeruginosa* have yet to be completely defined.

Understanding the biosynthesis and regulation of alginate production could lead to the identification of non-toxic inhibitors that could act within the lungs of CF patients, thus allowing control of pulmonary infections by *P. aeruginosa.* However, as well as its medical importance, alginate also has important commercial uses as a gelling agent and a viscosity modifier in both the food and pharmaceutical industry. At present, alginate is extracted solely from marine algae. Like most crops, the

quality and yield of alginate is subject to seasonal variations and clearly the ability to harvest alginate from a more consistent bacterial producer, such as *Pseudomonas* or *Azotobacter vinelandii*, as required, throughout the year would be preferable. While *P. aeruginosa* is a human pathogen and therefore of little industrial value, technology derived from the study of alginate synthesis in this organism could possibly be applied to other non-pathogenic alginate-producing pseudomonal species such as *P. putida*, *P. syringae*, *P. cicheri* or *P. viridiflava* (Fett *et al.*, 1992). Therefore, studies on the expression of the alginate biosynthetic and modification enzymes of *P. aeruginosa* may lead to both the synthesis of bacterial alginate on an industrial scale and also, the subtle manipulation of alginate size, water binding capacity and gel fluidity.

1.2. EXOPOLYSACCHARIDE PRODUCTION IN GRAM-NEGATIVE BACTERIA

1.2.1. Introduction

The architecture of a composite Gram-negative cell surface is depicted in figure 1.1. The Gram-negative cell surface, as shown, can be divided into three domains, the cytoplasmic membrane, the periplasm and the outer membrane.

The cytoplasmic membrane in Gram-negative bacteria (also referred to as the inner membrane), like other membranes, consists primarily of a mosaic of proteins embedded in a bilayer of phospholipids. The polar groups of the phospholipid are exposed at the two membrane surfaces (either the cytoplasm or periplasm) and the hydrophobic tails of the fatty acids are oriented to the interior of the membrane. The cytoplasmic membrane is stabilised by both the electrostatic forces of the polar head

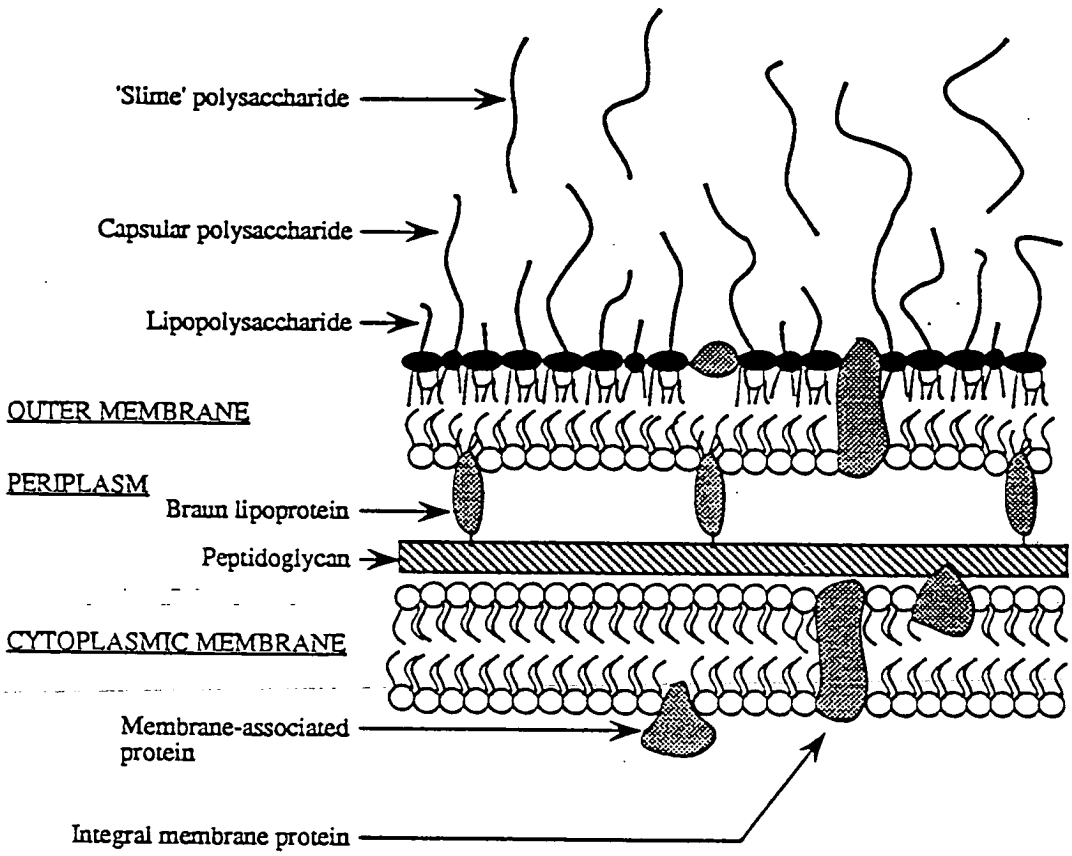


Figure 1.1. Diagram showing a stylised representation of the cell-surface architecture of Gram-negative bacteria. Adapted from Whitfield and Valvano, (1993).

groups as well as the hydrophobic interactions of the fatty acid tails. Cytoplasmic membrane proteins may be associated with the inner membrane by traversing the lipid bilayer (integral membrane proteins) or by being only partially associated with the membrane (peripheral membrane proteins). In general, the bilayer acts as a thin osmotic barrier, allowing the retention of metabolites and ions in the cytoplasm but excluding larger external compounds (larger than glycerol). However, the bacterial cytoplasmic membrane also allows the active concentration of nutrients and ions, and is the site for lipid and exopolysaccharide (EPS) synthesis, protein secretion, electron transport, oxidative phosphorylation and specific solute transport systems.

Immediately external to the cytoplasm is the periplasm, this layer (or sometimes mistermed 'space') is approximately 10-25 nm wide, depending on the bacterial species, and contains a gel-like solution which incorporates the peptidoglycan layer (Graham *et al.*, 1991). Within this gel matrix are an abundant array of proteins including substrate binding proteins, degradative enzymes, biosynthetic enzymes and detoxifying enzymes.

Bacterial peptidoglycan is a heteropolymer of sugars and amino acids, only the outer portion of which is sufficiently cross-linked to be isolated as an intact 'wall', the inner region is mostly non-cross-linked polymer. The peptidoglycan layer of Gram-negative bacteria is required for both cell shape and rigidity, the inner membrane is held in place against the peptidoglycan layer by the turgor pressure of the protoplast. The outer membrane however, requires structural attachments to the peptidoglycan layer. Binding of the outer membrane to the underlying peptidoglycan layer is via a small lipoprotein (Regue and Wu, 1988) also known as the Braun lipoprotein. The Braun lipoprotein is covalently linked to approximately one tenth of the tetrapeptides located in the polypeptide component of the peptidoglycan. The reaction links the D-ala of the peptidoglycan to the carboxy-terminal lysine of the protein. The N-terminal end of the lipoprotein is attached to the inner leaflet of the

outer membrane by an embedded hydrophobic substituent (fatty acid) linked to its amino terminal cysteine residue (Regue and Wu, 1988).

The inner leaflet of the outer membrane as in both leaflets of the cytoplasmic membrane is composed principally of phosphatidic lipids, however, in the outer leaflet of the outer membrane the predominant lipophilic macromolecule is the lipopolysaccharide (LPS). The LPS units, cohere more closely than the phosphatides. They also contain saturated fatty acids (as opposed to the unsaturated fatty acids in phospholipids), negatively charged groups and proximal sugar residues thus resulting in an atypical lipid bilayer which is markedly less fluid than the cytoplasmic membrane. Nearly half the mass of the outer membrane is protein. Most of these proteins are exclusive to the outer membrane but some are found in both the outer and cytoplasmic membranes. Amongst its proteins, the outer membrane has channels that permit passive diffusion of hydrophilic compounds such as amino acids, sugars and certain ions. Larger compounds such as vitamin B12, polysaccharides and iron chelate may only pass through the outer membrane via specific permeation systems. Thus, the major function of the outer membrane is to act as a molecular sieve (Nikaido and Vaara, 1985).

The presence of LPS in the outer leaflet of the outer membrane also adds a hydrophilic boundary to the cell surface, and this hydrophilicity may be further augmented by a layer of exopolysaccharide. The EPS can take the form of an adherent (often covalently anchored) cohesive layer, forming the morphological entity termed the capsule (Jann and Jann, 1990). Alternatively, EPS can consist of a 'slime' exopolysaccharide, with little or no firm cell association. However, the 'slime' polysaccharide definition is mainly an operational one and is based on the lack of retention of the EPS, to the cell surface, during centrifugation. By virtue of their location, cell surface polysaccharides influence the manner in which bacterial cells interact with their environment and are especially important as virulence

determinants in pathogenic infections (review Moxon and Kroll, 1990 and Lee, 1987).

1.2.2. **Exopolysaccharide structure**

Lipopolysaccharides and capsular polysaccharides are important virulence determinants, which interfere with the host's defence against pathogenic bacteria. Capsular exopolysaccharides also act as poor immunogens (Moxon and Kroll, 1990) and as such have particular importance in the virulence of extraintestinal and invasive bacteria. For this reason, bacterial capsules and exopolysaccharides have been the subject of extensive studies involving clinical, biological, biochemical and genetic investigations (Jann and Jann, 1990; Sutherland, 1992; Whitfield and Valvano, 1993). However, encapsulation also results in the impairment of phagocytosis and complement interaction.

Microbial exopolysaccharides may exist as homopolysaccharides, formed from chains of a single sugar, or heteropolysaccharides, composed of two or more sugars. A large number of different monosaccharides have been identified as being incorporated into bacterial exopolysaccharides, including the neutral hexose sugars such as D-glucose, D-galactose and D-mannose, the 6-deoxyhexoses such as L-fucose and L-rhamnose, the amino sugars, such as N-acetyl D-glucosamine and N-acetyl D-galactosamine and the uronic acids, D-glucuronic acid, D-galacturonic acid and D-mannuronic acid. In addition, exopolysaccharides may contain phosphate, ester-linked acetate or succinate or ketal-linked pyruvate. The heteropolysaccharides are normally composed of regular sequences of repeating sugar units, but may vary in size, from disaccharides to octasaccharides depending on the EPS and overall, this results in an extracellular polyanionic, encapsulating, polysaccharide (Sutherland, 1992).

There are several forms in which exopolysaccharides can be organised at the cell surface (figure 1.2). At the simplest level, exopolysaccharides may be perpetually released directly from the cell, without any terminal cell surface anchor molecule thereby forming a 'slime' polysaccharide (Whitfield and Valvano, 1993). This situation is assumed to be the case for the alginate exopolysaccharide of *Azotobacter vinelandii* and *P. aeruginosa* and for the capsule polysaccharide (CPS) of *Klebsiella aerogenes*, where no EPS anchors have been discovered (Whitfield and Valvano, 1993). However, it remains to be firmly established whether ionic or other interactions are sufficient to maintain association of 'slime' exopolysaccharides to the cell surface. For 'capsule' exopolysaccharides, two modes of covalent surface attachment have been proposed. These involve a hydrophobic anchor, provided either by phospholipid or alternatively, by lipid A. *E. coli* strains show both modes of interaction depending on the type of capsule (K-antigen) they produce. *E. coli* CPS's are divided into two groups (I and II) on the basis of a variety of physical, chemical and genetic criteria (Jann and Jann, 1990). Group II capsules typically use a phospholipid anchor as do the CPS's of *Neisseria meningitidis* and *Haemophilus influenzae*, whereby the polysaccharide is directly linked to the phospholipid. However, some group II K-antigens of *E. coli*, exhibit a variation, whereby the L-glycerophosphatidyl residue (phospholipid) is instead, linked to the polysaccharide through a 3-deoxy-D-manno-octulosonic acid (KDO) residue (this difference is reflected in a slight variation in the method of polysaccharide assembly, see later). The second and better characterised membrane anchor is that of lipid A. Lipid A is involved in attaching a few of the group I CPSs of *E. coli*, as well as attaching the lipopolysaccharide O-antigen. Indeed, *E. coli* with group I polysaccharides may thus have two different LPSs, a neutral polymer being the O antigen and an acidic macromolecule being the capsular K-antigen (Jann and Jann, 1990). The polysaccharides, both of the K-antigens and LPS are attached to the lipid by a

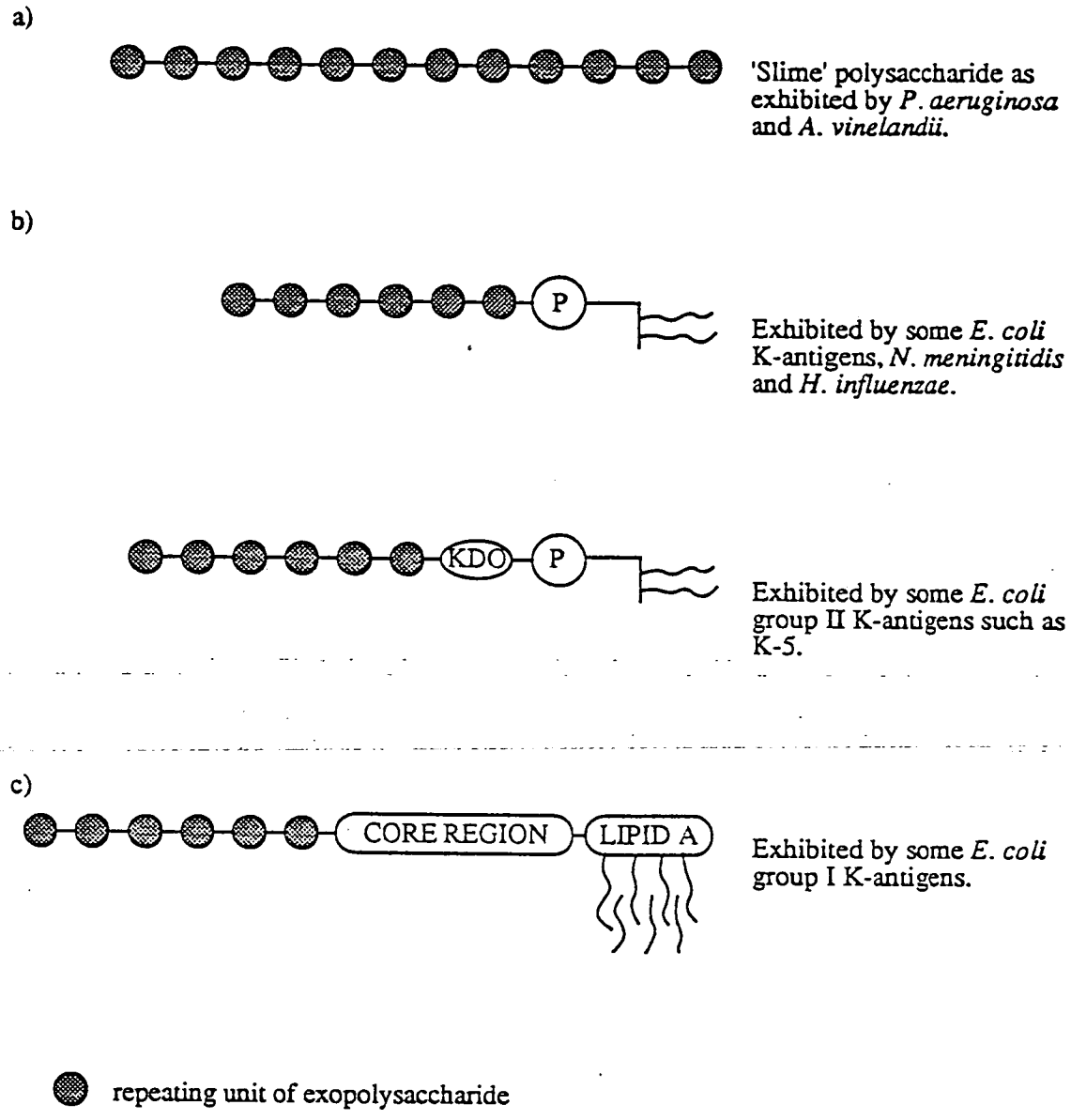


Figure 1.2. Cell-surface association of polysaccharides in Gram-negative bacteria. (a) No membrane anchor. (b) Attachment through diacylglycerolphosphate. The diacylglycerol moiety is linked to the exopolysaccharide through a phosphodiester linkage (represented by 'P') and in some instances 3-deoxy-D-manno-octulosonic acid (KDO) is required as well. (c) Attachment through lipid A. Adapted from Whitfield and Valvano (1993).

defined core region consisting of KDO and other specific sugars and this core region may vary in content between bacterial species (Jann and Jann, 1990).

1.2.3. **Biosynthesis of exopolysaccharides**

a) Carbohydrate precursors

Biosynthesis of bacterial cell-surface polysaccharides is a sequential process and the individual steps have been characterised to various degrees. The initial step in the production of exopolysaccharide is the biosynthesis of activated precursors within the cytoplasm. Such, activated precursors exist as nucleoside diphosphate or less frequently nucleoside monophosphate sugars. Activated sugar nucleotides function in the synthesis of all carbohydrate-containing polymers found outside the microbial cell membrane (except dextran and levans). Sugar nucleotides are also required for the synthesis of intracellular storage products such as glycogen as well as the formation of the peptidoglycan layer and membrane derived (MD) oligosaccharides. The remaining steps in exopolysaccharide synthesis from these activated sugar precursors involves the formation of repeating sugar units, the polymerisation of multiple repeating sugar units and the export of the polysaccharide to the cell surface.

b) Formation of repeating sugar units

The requirement for a lipid acceptor in the cytoplasmic membrane on which to assemble the repeating sugar units (containing 1-8 sugar residues) of the exopolysaccharide is common to the majority of Gram-negative bacteria. The discovery of the involvement of isoprenoid lipids in the formation of LPS and peptidoglycan, and the structural similarity of these polymers to EPS indicated that isoprenoid lipids might also have a role in EPS synthesis. Early work by Troy *et al.*

(1971), on the biosynthesis of *K. aerogenes* (strain DD45) EPS, revealed a process whereby the tetrasaccharide repeating sugar unit of the CPS is produced sequentially by the addition of D-galactose from UDP-D-galactose, D-mannose from GDP-mannose and glucuronic acid (glcA) from UDP-glcA. The lipid involved in the formation of the repeating unit has been extracted and characterised and was found to be a C_{55} isoprenyl phosphate (also known as undecaprenol phosphate). This assembly is summarised in figure 1.3.

A series of studies on *Xanthomonas campestris* has confirmed that the synthesis of xanthan follows much the same pattern as found for the EPS synthesis of *K. aerogenes*, thus *X. campestris* forms an undecaprenol linked pentasaccharide unit which is subsequently polymerised to give the EPS (Ielpi *et al.*, 1993). A similar process is probably involved in the assembly of EPS in other plant-associated bacteria including *Rhizobium trifolii* and *R. meliloti* (review Leigh and Walker, 1994), as well as the majority of bacteria synthesising group II K-antigens (Jann and Jann, 1990).

However, in the synthesis of *E. coli* K5 (a group II CPS) no directly linked lipid-sugar intermediates have been observed (Finke *et al.*, 1991). Instead it was thought that the biosynthesis of the K5 polysaccharide was initiated by substitution of a carrier (possibly undecaprenol phosphate), with KDO from CMP-KDO, which then functioned as an acceptor for subsequent addition of glucose and N-acetyl glucosamine (glcNac) to the non-reducing end directly from their UDP derivatives. Thus, KDO from CMP-KDO was added at the start of CPS synthesis rather than being added with the phospholipid anchor in a post-polymerisation process, to give the typical cell surface attachment (see figure 1.2.).

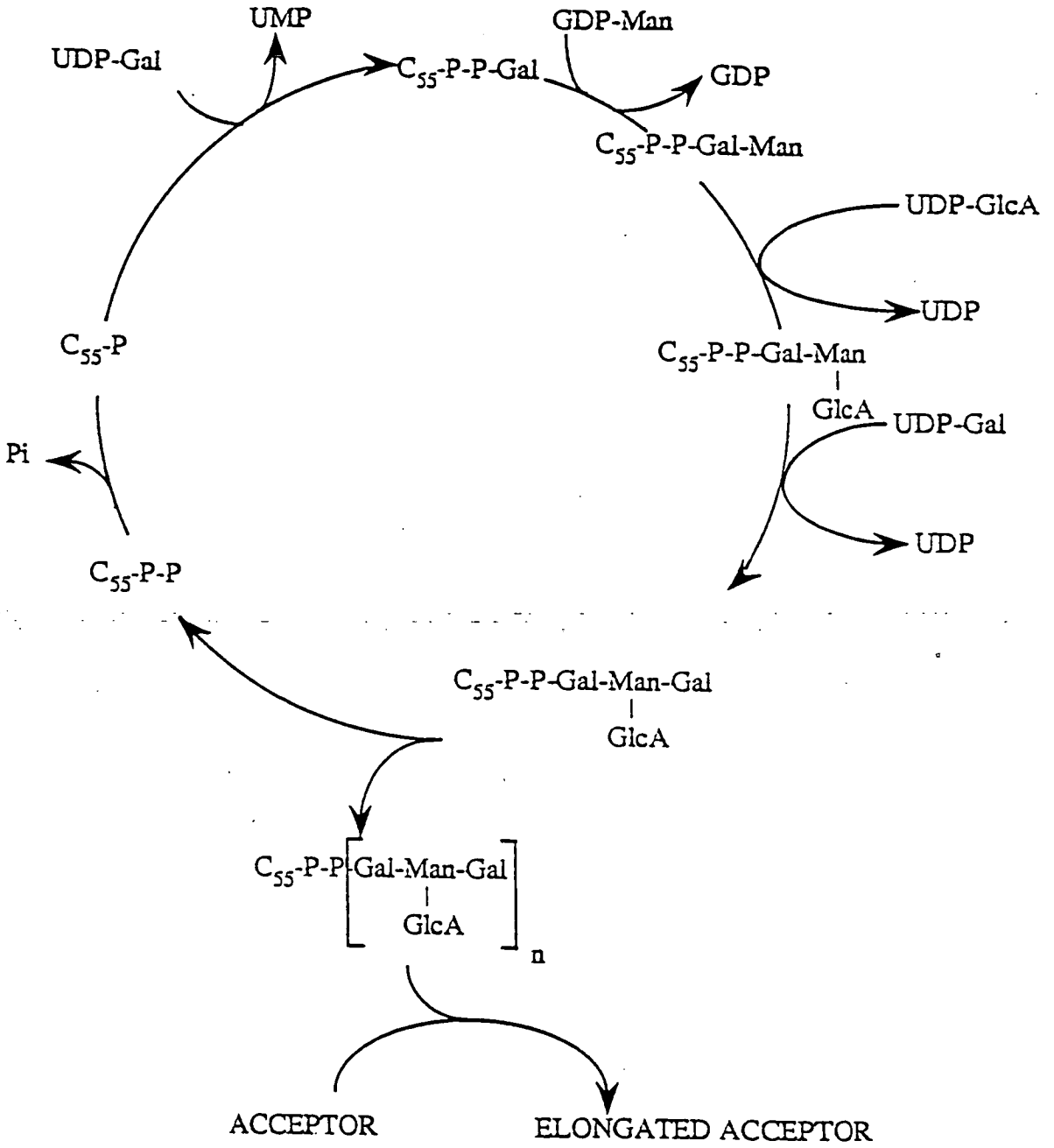


Figure 1.3. Pathway for the assembly of undecaprenol-linked oligosaccharide intermediates in the biosynthesis of the capsular polysaccharide of *Klebsiella aerogenes* DD45. The tetrasccharide unit is polymerised to give the polysaccharide of n (number of) repeating units (see text for fuller details). Undecaprenol and phosphate are represented by C_{55} and P, respectively (adapted from Whitfield and Valvano, 1993).

c) Polymerisation reactions

As previously mentioned, the majority of EPSs are synthesised through undecaprenol dependent systems, although there is no such requirement for a lipid-linked sugar intermediate in bacterial cellulose synthesis.

Two mechanisms have been described for polymerisation of undecaprenol dependent cell-surface polysaccharides. These involve either polymerisation of preformed undecaprenol-linked units in a blockwise manner, or alternatively, the processive, sequential transfer of individual sugar residues to a growing polysaccharide attached to undecaprenol. The prototypes for these models are shown in figure 1.4i and 1.4ii.

Polymerisation of preformed undecaprenol-linked blocks is typically shown in the biosynthesis of *K. aerogenes* (see figure 1.3) and *X. campestris* EPS whereby tetrasaccharide or pentasaccharide units, on multiple undecaprenol lipid carriers are polymerised (polymer growth at the reducing terminus of the polysaccharide) to give the EPS.

The sequential transfer of sugars to a growing polysaccharide is exhibited in the formation of the polysialic capsule of *E. coli* K1 and *Neisseria meningitidis* B (Frosch and Müller, 1993). In both these systems, polysialyltransferase activity elongates an undecaprenol lipid acceptor (one undecaprenol molecule for the growth of each chain) within the membrane, adding sialic acid units (from CMP-sialic acid) in a processive manner to form α(2-8)-linked polysialylic acid (Annunziato *et al.*, 1995).

Undecaprenol-dependent polymerisation mechanisms have not been reported in the production of alginate exopolysaccharide from *A. vinelandii* or *P. aeruginosa*. However, although no undecaprenol-linked intermediates have been identified for the biosynthesis of alginate in either of these bacteria (Sutherland, 1982; May *et al.*, 1991) it is still thought that polymerisation proceeds by the sequential addition of

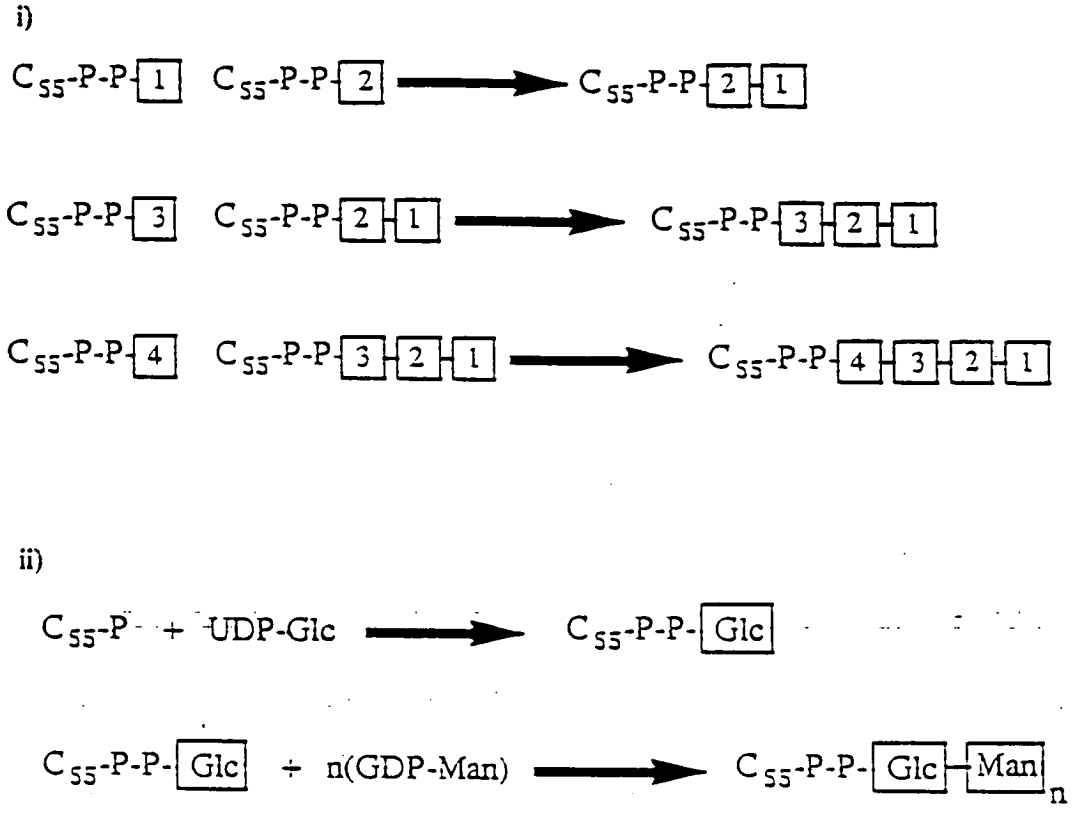


Figure 1.4. Mechanisms involved in the polymerisation of exopolysaccharides: i) Polymerisation of preformed undecaprenol-linked repeating units in a blockwise manner at the reducing terminus, as exhibited by *K. aerogenes*. The boxes represent the repeating units and the internal number defines their linkage position. ii) Sequential transfer of sugars to a growing polysaccharide at the non-reducing terminus of the polysaccharide which is attached to undecaprenol, as exhibited by *E. coli* group II exopolysaccharide producers. Undecaprenol and phosphate are represented by C_{55} and P, respectively (adapted from Whitfield and Valvano, 1993).

mannuronic acid (from GDP-mannuronic acid) on undecaprenol lipids to form the polymannuronate polysaccharide (Sutherland, 1992).

Alternatively, the polymerisation reaction can proceed directly, without the requirement for undecaprenol-linked sugar intermediates, and such a system was observed in the synthesis of bacterial cellulose in *Acinetobacter xylinum*. *A. xylinum* produces a single loosely wound, extracellular ribbon of cellulose fibrils containing ca. 1000 β-1,4 glucan chains in total, each chain being between 2000 and 18000 residues long (Ross *et al.*, 1991). The synthesis of this biopolymer takes place directly, via a single polymerisation step, utilising UDP-glucose as the substrate, and is catalysed by the enzyme cellulose synthetase (UDP-glucose 1,4 β-D-glucosyl transferase) (Ross *et al.*, 1991). Homologues of cellulose synthase, NdvB (*R. meliloti*) and ChvB (*Agrobacterium* sp) utilise a similar undecaprenol independent mechanism to that operating in synthesis of bacterial cellulose (β-glucan).

d) Export of polysaccharides to the cell surface

i) Location of biosynthetic complexes at the cytoplasmic membrane:

It is generally assumed that assembly of polysaccharide repeating units and subsequent polymerisation reactions occur at the cytoplasmic side of the cytoplasmic membrane, using activated sugar precursors synthesised in the cytoplasm. In many systems the involvement of undecaprenol provides an obligatory requirement for the cytoplasmic membrane. Sequential assembly of bacterial polysaccharides has led to the belief that the enzymes function as a co-ordinated complex with undecaprenol sequestered in the active complex (Anderson *et al.*, 1972). A co-ordinated complex is also predicted by the limited amount of undecaprenol molecules available in the cell (approximately 10^5 molecules per cell, Sutherland, 1992). The ability to solubilise glycosyltransferases has produced additional evidence for such a complex

(Danilev *et al.*, 1991). Solubilised glycosyltransferases retain the ability to catalyse individual reactions in the formation of undecaprenol-linked intermediates but fail to polymerise the repeating units. This is thought to be a result of the solubilising process affecting the integrity of the complex thus preventing the formation of polymerised repeating units.

Most transferases characterised to date do not show properties expected of transmembrane or inner membrane proteins. The polysialyltransferase enzymes involved in biosynthesis of polysialyic acid-containing CPS (capsular polysaccharides) in *E. coli* K1 and K92 and *N. meningitidis* B also show no significant membrane-spanning domains (Frosch *et al.*, 1991, Vimr *et al.*, 1992). Thus, association of the transferases with the membrane may either be due to direct interaction (as they are basic proteins) or interaction with other membrane associated proteins.

ii) Transport across the cytoplasmic membrane:

Completion of polymerisation of exopolysaccharides at the cytoplasmic face of the cytoplasmic membrane results in a requirement for a post polymerisation export mechanism. Putative polysaccharide inner membrane export components have been identified in a diverse range of bacteria, to transfer different polysaccharides which are assembled by different polymerisation mechanisms, across the cytoplasmic membrane.

Initial studies in *E. coli* have defined components that are required for translocation of polysaccharides across the cytoplasmic membrane. Both K1 and K5 serotypes (Smith *et al.*, 1990; Pavelka *et al.*, 1991) were found to contain an operon of two genes *kpsM* and *kpsT*. KpsT and KpsM proteins are highly conserved between these *E. coli* serotypes and KpsT shows amino acid similarity and KpsM shows general structural homology with proteins belonging to the ATP Binding

Cassette (ABC) transport superfamily, as defined by Higgins *et al.* (1990). ABC transporters are responsible for passage of a diverse range of substrates into and out of prokaryotic and eukaryotic cells. The complex typically contains a transporter comprised of two hydrophobic membrane proteins with 5 or 6 membrane-spanning segments, and two hydrophilic membrane associated proteins that couple ATP to transport. A typical example (of an ABC-transporter), the oligopeptide permease (Opp) of *S. typhimurium* as shown in figure 1.5, where OppB and OppC are integral membrane proteins and OppD and OppF bind ATP (OppA is a periplasmic substrate binding protein and is additional to the typical superfamily structure).

The properties of KpsM and KpsT are consistent with the basic organisational model of these transport systems. KpsT is a peripheral cytoplasmic membrane protein with a consensus ABC binding domain. This polypeptide is thought to couple ATP hydrolysis with polysaccharide secretion through KpsM (Smith *et al.*, 1990). KpsM has a hydrophobicity profile typical of an inner membrane protein and is predicted to be involved in transport of polysaccharide across this barrier. KpsT and KpsM are thought to dimerise to give an active four sub-unit ABC exporter. Proteins which are homologous to KpsM and KpsT are encoded by the CPS-biosynthesis gene clusters from *Haemophilus influenzae* type B (*bexA/B*) and *N. meningitidis* (*ctrC/D*) (Kroll *et al.*, 1990; Frosch *et al.*, 1991, respectively). Analysis of polysaccharide export genes have also been performed on *Agrobacterium tumefaciens and R. meliloti*. Both these organisms utilise a large single protein, ChvA (Cangelosi *et al.*, 1989) and NdvA (Stanfield *et al.*, 1988) respectively, to transport β-1-2-linked glucan across the inner membrane. Both proteins contain an ATP binding domain and a hydrophobic domain and thus the protein is thought to exist as a homodimer to give the typical four domain model.

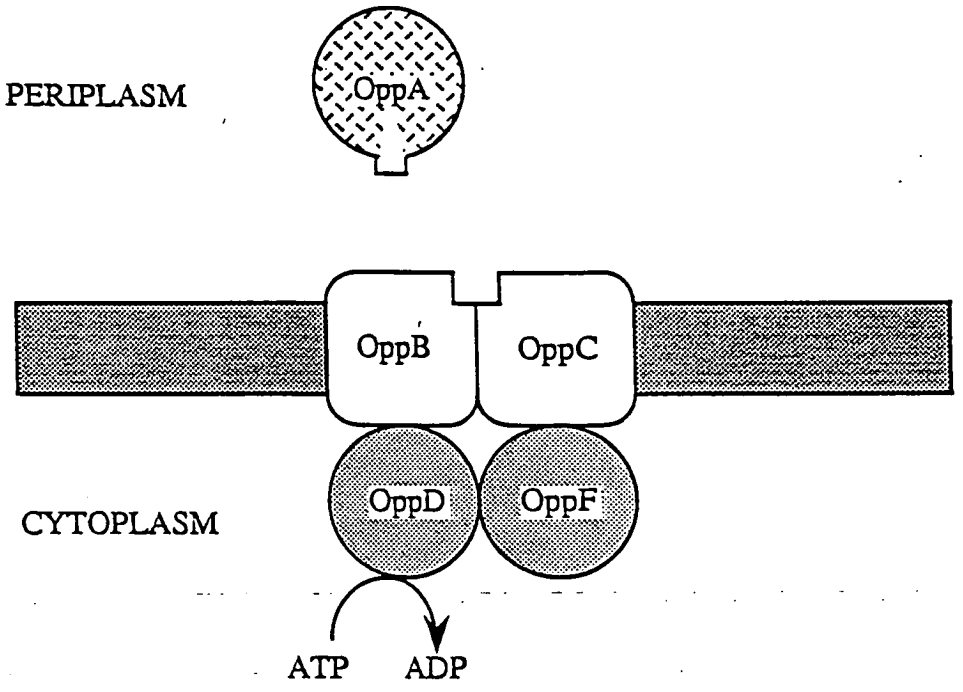


Figure 1.5. Oligopeptide permease of *S. typhimurium* showing the two integral membrane proteins, OppB and OppC, and the two ATP binding proteins OppD and OppF. OppA is a periplasmic substrate binding protein (adapted from Higgins *et al.*, 1990).

iii) Polysaccharide transport across the periplasm and outer membrane:

The proteins involved in the transport of the polymerised polysaccharide across the periplasm and outer membrane to the bacterial cell surface are still poorly defined for both undecaprenol dependent and independent Gram-negative EPS biosynthesis. However, Wunder *et al.* (1994) identified in *E. coli* K1, a 60-kDa periplasmically oriented protein, KpsD, involved in the transport of polysialic acid across the periplasm, this protein had homologues in K2, K5, K7, K12, K13 and K92 (all are undecaprenol-dependent systems). Inactivation of the gene encoding KpsD, results in periplasmic accumulation of polysialic acid (EPS). However, the specific role of KpsD in the translocation of polymer to the cell surface is unclear. Wunder *et al.* (1994) postulated two possible functions for KpsD. Firstly, KpsD may act as the periplasmic binding element of the polysialic acid transport system, in which KpsD transiently interacts with the inner membrane component of the transporter (KpsM), binds polysaccharide and transports the polymer to components in the outer membrane. Alternatively, KpsD along with other accessory proteins (encoded in the region 1, biosynthetic cluster, see below) directly connects the export machinery of the inner membrane (KpsMT) to a protein in the outer membrane, allowing the polysaccharide to bypass the periplasmic space. The second model thus predicts the occurrence of a large biosynthetic-translocation complex that would span both inner and outer membranes. The protein complex would thus generate a pore or channel through which the polymer passes, allowing the polysaccharide to move from the cytosol to the surface of the bacterial cell. The outer membrane porin of this putative system has yet to be definitively identified, although the outer membrane porin, protein K, may be involved (Wunder *et al.*, 1994). Interestingly, ExoF, a periplasmic protein involved in the succinoglycan exopolysaccharide biosynthetic pathway of *R. meliloti* (the *exoF* gene is located in the succinoglycan biosynthetic operon) is considered to be a functional homologue of KpsD (Leigh and Walker, 1994). More

recently, Rosenow *et al.* (1995) further characterised the KpsE protein, which is encoded on the same operon as KpsD. KpsE was found to have a periplasmic orientation, but is anchored to the cytoplasmic membrane by N and C-terminal membrane-spanning α-helices. The subcellular orientation of KpsE led Rosenow *et al.* (1995) to postulate that the protein may also have a role in periplasmic polysaccharide transport.

The outer membrane porin involved in polysaccharide export has been identified for *N. meningitidis* (Frosch *et al.*, 1992), *H. influenzae* (Kroll *et al.*, 1990), and *P. aeruginosa* (Rehm *et al.*, 1994a). These outer membrane porins (CtrA, BexD and AlgE, respectively) share a similar secondary structure and exist as a β-barrel like arrangement of β–strands in the outer membrane. However, CtrA is predicted to have just 8 membrane spanning strands compared to the 18 of AlgE. This may be a reflection in the size difference between the respective exopolysaccharides (<50,000-Da for *N. meningitidis* and *H. influenzae* compared to 1 million-Da for *P. aeruginosa*). Both *N. meningitidis* and *H. influenzae* also express a periplasmic located protein involved in polysaccharide export, CtrB and BexC respectively. These proteins are encoded in the same operon as the outer porin and the inner membrane polysaccharide transporter and show structural similarity to the KpsE protein (Cieslewicz *et al.*, 1993). Both CtrB and BexC are considered to be candidates for part of a pore or carrier mechanism involved in transferring growing polysaccharide from the cytosol to the outer membrane (Kroll *et al.*, 1990).

Saxena *et al.* (1994) identified a potential outer membrane pore/channel encoded by *acsC*, and a further periplasmic protein encoded by *acsD*, located in the cellulose biosynthetic operon (undecaprenol-independent synthesis) of *A. xylinum*. AcsD is thought to be involved either in the pore structure itself or in the organisation of pores within the terminal synthesising complexes (sites of cellulose synthesis and extrusion). However, neither AcsC or AcsD share any obvious amino

acid similarity with the outer membrane and periplasmic proteins described in other polysaccharide biosynthetic systems (Saxena *et al.*, 1994).

Thus, Gram-negative bacterial polysaccharide export systems, appear to employ several common features. Evidence so far gathered predicts either the presence of a large biosynthetic-translocation complex for export of polysaccharide from the cytosol to the bacterial surface or inner and outer membrane polysaccharide transporters/porins linked by a periplasmic polysaccharide shuttle protein.

1.2.4. Common genetic features in the biosynthesis of exopolysaccharide

The biosynthetic genes so far characterised for EPS synthesis, appear to be arranged in clusters and predominately as operons within the bacterial chromosome. This is true for bacteria that exhibit either undecaprenol-dependent or undecaprenol-independent polymerisation mechanisms. Regulatory genes however, tend not to be linked to clusters of biosynthetic genes.

Study of the genes required for the synthesis of different K-antigens in *E. coli*, revealed a common gene cluster organisation (Boulnois *et al.*, 1989; Pazzani *et al.*, 1993). A region of approximately 17-kb of DNA formed three functional segments. The first region (~9-kb) encodes genes responsible for the translocation of the completed polysaccharide across the periplasm and outer membrane to the bacterial surface. A second region (~5-kb) consists of the genes responsible for the enzymes involved in biosynthesis of sugar nucleotides specific to the polymer including specific sugar transferases and the polymerase (Frosch *et al.*, 1991). Region 3 (~2.5-kb) contains just two genes (in an operon), *kpsM* and *kpsT*, which form the ABC inner membrane polysaccharide transporter. Regions 1 and 3 are highly conserved and functionally interchangeable between different *E. coli* strains, suggesting that post-polymerisation modification of the exopolysaccharide and transport of the mature polysaccharide from the site of synthesis are independent of

polysaccharide structure (Roberts *et al.*, 1988). A similar gene cluster organisation is observed in *N. meningitidis* B (Frosch *et al.*, 1992) and *H. influenzae* (Kroll *et al.*, 1989) and is thought to indicate a common evolutionary origin of capsule production in Gram-negative bacteria expressing group II capsular polysaccharides (Frosch *et al.*, 1991). The genes responsible for succinoglycan (EPS) and EPS1 production in *R. meliloti* (Glucksmann *et al.*, 1993 (a, b) and Becker *et al.*, 1993 (a, b)) and cellulose production in *A. xylinium* (Saxena *et al.*, 1994) are also found in biosynthetic clusters.

1.3. ALGINATES

Alginates have been used widely for a broad range of applications in industry and many biochemical properties of this polysaccharide have been exploited. Most alginate is used in the food and pharmaceutical industries where alginates tend to be utilised in many forms but mainly as stabilisers, gelling agents, and films (see table 1.1) and this reflects on their ability to form a gel in the presence of Ca^{2+} ions. Indeed, the calcium salt may be formed into a thread and spun using weaving technology. The widespread use of alginates in industry has resulted in the steady growth in the use of seaweed as a source of supply such that by the end of the 1980's approximately 15,000 tonnes dry weight of polysaccharide was processed from 400,000 tonnes wet weight of seaweed per year (Gacesa, 1988).

The extraction of alginate from marine algae, is subject to seasonal variation and recent research has centred on determining whether a regular and reproducible bacterial source may be possible. Several microbial producers of alginates have been identified although none at present are used commercially. Originally two species of bacteria, *A. vinelandii* (Gorin and Spencer, 1966) and *P. aeruginosa* (Linker and Jones, 1966) were identified as alginate producers. Subsequently, alginate producing

Area of application	Function	Specific examples
Food and drinks industry	Stabiliser	Foam stabiliser (beer). Phase-separation retardant (ice cream).
	Viscosifier	Suspension of fruit pulp. Reconstitution of foods (stoneless fruit, onion rings).
	Film	Coating of fish.
Pharmaceutical industry	Stabiliser	Emulsions in cosmetic preparations. Binder for tablets and lozenges.
	Gelling agent	Moulds for dental impressions.
	Film/fibres	Gastroenteric coatings for tablets. Haemostatic bandages.
	Therapeutic agents	Anti-acid and anti-ulcer compound.
Other uses	Viscosifier	Printing inks.
	Gelling agent	Enzyme/cell immobilisation.

Table 1.1. A selection of the major industrial applications of alginates (from Gacesa, 1988).

strains of *P. florescens*, *P. mendocina* and *P. putida* (Govan *et al.*, 1981) *P. syringae* (Fett *et al.*, 1986), *P. cicheri* and *P. viridiflava* (Fett *et al.*, 1992) have been isolated. The majority of research into the expression of biosynthetic and modification enzymes for bacterial alginate has centred on its role as a major virulence determinant of *P. aeruginosa* infection of cystic fibrosis patients (see later). However, information derived from this model system may eventually be utilised with a non-pathogenic pseudomonal species and allow industrial level synthesis of alginate from a bacterial source.

1.4. *PSEUDOMONAS AERUGINOSA* AND CYSTIC FIBROSIS

Cystic fibrosis prevalence and genetics

In most parts of the world CF is the most common lethal inherited disorder among Caucasian populations. In the US, the heterozygote frequency in Caucasians is approximately 1 in 25, with a disease incidence of approximately 1 in 2500 live births. The incidence of disease in African-Americans is ~1 in 17000 live births. Whilst, in Europe, variability in frequency is common, being as high as 1 in 500 in areas of Scotland (Bye *et al.*, 1994).

CF is transmitted through autosomal recessive inheritance. Carriers have no detectable symptoms and no known electrophysiological abnormalities. However, May *et al.* (1991) suggested that the high incidence of CF in Caucasian populations may be due to heterozygotes having a selective advantage by being more resistant to the bacterial-toxin-mediated diarrhoea once prevalent in Europe.

The CF gene was isolated in 1989 by the groups of Lap-Chee Tsui in Toronto, Canada, and Francis Collins in Michigan, USA. The gene causing CF is encoded on 250-kb of the long arm of chromosome 7 (7q31). The coding region of CFTR is made up of 27 exons which together codes for a 1480 amino acid protein

known as the CF transmembrane conductance regulator or alternatively the CF transductance regulator (CFTR). 70% of CF patients have a 3-bp deletion in exon 10 leading to the deletion of phenylalanine at position 508. 300 alternative mutations have also been found, many of which pertain to only small numbers of people. These include deletions, missense mutations and frameshift mutations (Bye *et al.*, 1994).

The 1480 amino acid, CFTR protein predicted from the 6.2-kb cDNA sequence appears to be a member of the ABC-transport superfamily (Hyde *et al.*, 1990). It is made up of two very similar halves that each contain a nucleotide-binding motif (ABC) and an integral membrane domain. In addition the two halves appear to be separated by the regulatory 'R' domain (Harris, 1992) as shown in figure 1.6 below.

Studies on chloride channels of epithelial cells from airways, sweat glands and pancreas showed that, in particular, the activation of chloride channels in response to cAMP was impaired in CF patients. Several independently regulated chloride channels are found in the apical membrane of epithelial cells. One of these is activated by changes in cell volume, one by Ca^{2+}-dependent protein kinases and one by cAMP. CFTR was identified as the cAMP-activated channel (Anderson *et al.*, 1991; Bear *et al.*, 1992). The movement of chloride across the membrane, through CFTR occurs by diffusion and does not require active transport. The requirement for ATP-binding by the nucleotide binding domains of CFTR appears to be for channel opening, allowing Cl^- diffusion. ATP hydrolysis may well induce a conformational change in the CFTR protein necessary for opening the channel, although this has yet to be shown. The R-domain is known to be highly charged and contains four sites which are phosphorylated by protein kinases *in vivo*. Phosphorylation events appear to play a regulatory role, opening the channel in response to changes in cAMP levels (Cheng *et al.*, 1991). The CF defect is due to the impaired activation of chloride secretion in response to cAMP.

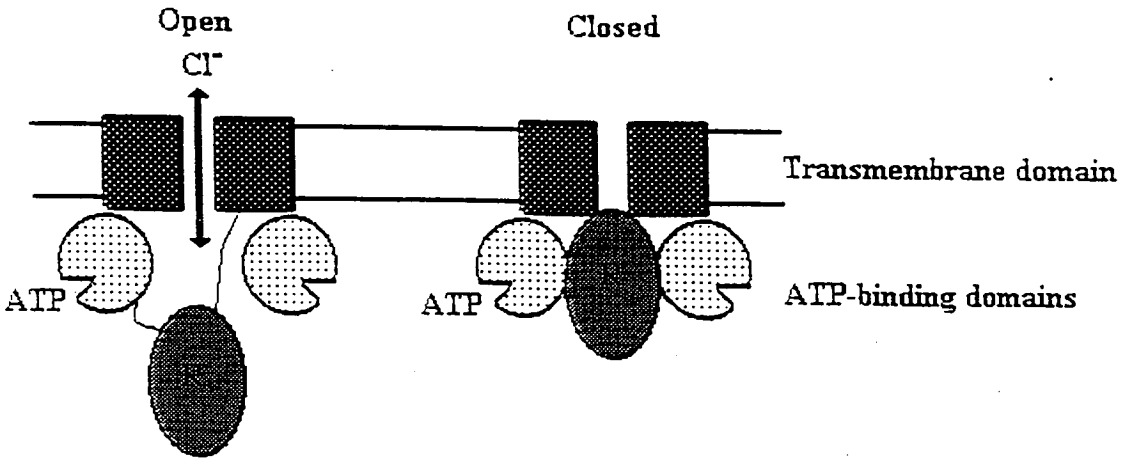


Figure 1.6. Schematic model for CFTR function and regulation. Two transmembrane domains form a pathway for chloride transport (diffusion) across the lipid bilayer. The R-domain opens and closes in response to cAMP-activated protein kinases. The role of the ATP-binding domains in channel function remains to be firmly determined (adapted from Higgins, 1992).

In CF the physiological function of the epithelia of the airways, intestinal tract and exocrine glands (pancreas, salivary glands and exocrine sweat glands) is disrupted. All these tissues have components involved in secretion. They produce a protein-containing fluid which serves either in a digestive, lubricative or protective capacity. The correct ionic, protein and water composition of these epithelial secretions is crucial for their normal function. The secretion or reabsorbtion of water depends on the movement of ions across the cell and water follows the ions by osmosis. In CF, the abnormality in ion transport due to inactivation of CFTR causes an alteration in composition of these epithelial secretions, resulting in an abnormally viscous mucus which leads to the main clinical manifestations of the disease.

Cystic fibrosis pathophysiology

i) pancreatic insufficiency

Because the pancreatic exocrine secretions are abnormal, the ducts are generally obstructed and enzymes and bicarbonate are poorly secreted into the duodenum. This results in malabsorbtion of fats and proteins, resulting in poor weight gain. Malabsorbtion of fat-soluble vitamins, such as vitamins A, E and K, may also occur and result in vitamin deficiency. Pancreatic insufficiency may also predispose CF patients to mechanical problems in the gastrointestinal tract. This along with the additionally viscous mucus in the intestine may result in intestinal obstruction.

ii) sweat gland involvement

CF affects the sweat glands in over 95% of CF patients. An abnormality in electrolyte reabsorbtion in the sweat duct leads to increased concentrations of Na^+ and Cl^- in the sweat (an abnormality that is used as a diagnostic test for CF). The sweat abnormality causes clinical manifestations during hot weather resulting in

excess chloride loss and hyperchloronic metabolic alkalosis in infants and young children.

iii) other features

Expression of mutant CFTR is also found in sweat ducts, salivary glands, intestine and liver and in the latter it leads to focal biliary cirrhosis in 25% of CF patients. Genitourinary abnormalities are found in both genders and results in male sterility due to mucus obstruction of the vas deferens, epididymis and seminal vesicles. Fertility in females is also decreased as increased amounts of extra-thick mucus in the cervical canal acts as a partial barrier to conception. Mutant CFTR is also expressed at low levels in the respiratory epithelia which is the site of the most severe clinical symptoms.

Respiratory pathophysiology

The pulmonary pathophysiology begins in the peripheral airways where increased thick mucus secretions cause airway obstruction and promote bacterial infection and colonisation (Govan and Glass., 1990). Chronic bronchopulmonary infection is the leading cause of lung damage in CF patients and leads to pulmonary dysfunction and death (Govan and Harrison, 1986). The initial infecting organisms are often *S. aureus* and *H. influenzae*. However, as the disease progresses, colonisation with mucoid strains of *P. aeruginosa* frequently occurs. Some patients have also been known to be colonised with mucoid strains of *E. coli*, *Burkholderi cepacia* as well as mycobacteria and fungi (Govan and Nelson, 1992).

In general, microbial colonisation tends to follow a chronological sequence during a patients lifetime. Colonisation usually begins with *S. aureus* infection at infancy followed by *H. influenzae* with *P. aeruginosa* infections developing in early

adolescence, and the highest incidence of *B. cepacia* and mycobacteria occurring in the late teens (Govan and Nelson, 1992).

The first organism recognised to cause chronic lung infection in the majority of young CF patients was *S. aureus*. In the pre-antibiotic era, few CF patients survived beyond infancy mainly because of the *S. aureus* infection. The marked improvement in life expectancy of CF patients is a result of the effective control of *S. aureus* infection with anti-staphylococcal antibiotics (Govan and Nelson, 1992).

S. aureus, however is thought to predispose the CF-affected lung to *P. aeruginosa* colonisation in numerous ways. Indeed, *S. aureus* infection has been shown to contribute to airway inflammation, epithelial damage and altered mucous production and therefore may create an environment favourable to *P. aeruginosa* adherence and colonisation. *H. influenzae* may also assist in *P. aeruginosa* colonisation by disturbing cilliary function. *H. influenzae* initially coexists with *P. aeruginosa*, but *P. aeruginosa* being resistant to even the most aggressive antibiotic therapy, gradually becomes the predominant pathogenic microorganism within the CF lung.

In the early 1980's, several North American clinics reported a disturbing increase in the isolation of *B. cepacia* from CF patients. *B. cepacia* is a relatively little known phytopathogen (first isolated in 1950 as the cause of bulb rot in onions) and was not described as a human pathogen until the 1970's. *B. cepacia* is an especially virulent and antibiotic resistant organism and causes serious infections in immunocompromised and CF patients. It is not, however, mucoid (Fett *et al.*, 1992). Three distinct patterns of *B. cepacia* infection are seen: 1) chronic asymptomatic carriage; 2) progressive deterioration over many months, with recurrent fever, weight loss and frequent hospital admissions; and 3) rapid, usually fatal deterioration of previously mildly affected patients. *B. cepacia* generally causes a rapid and

terminal course of microbial infection in contrast with the slow deterioration associated with the more common CF lung colonisers (Govan and Nelson, 1992).

Colonisation with *P. aeruginosa*

Early bacteriological investigations of CF patients showed that colonisation with *P. aeruginosa* played a minor role in relation to *S. aureus*. In 14 post mortem studies by Di Sant' Agnese in 1948 only one showed the presence of *P. aeruginosa*. However, development of antistaphylococcal antibiotics has led to an increase in population of older CF patients and the emergence of *P. aeruginosa* as the predominant bacterial pathogen responsible for lung disease in CF patients (Govan and Glass, 1990).

Primary colonisation of the CF airways by *P. aeruginosa* is usually by non-mucoid forms. However, since the interval between primary colonisation and emergence of mucoid strains can be less than three months, mucoid strains may be observed on primary bacterial investigation (Govan and Nelson, 1992). In most cases, chronic colonisation of the CF lung occurs with a single strain of *P. aeruginosa* which undergoes phenotypic variation to the mucoid form over a period of time. Transient carriage of more than one strain of *P. aeruginosa* may occur, however. Initially *P. aeruginosa* colonises the upper respiratory tract, infecting surfaces of the maxillary sinuses, tongue, buccal mucosa, saliva and dental plaque. *P. aeruginosa* also adheres to respiratory mucins and evidence for the chemotaxis of the organism towards mucin and various amino acid and sugar components of mucin has also been shown (Nelson *et al.*, 1990). The particular affinity of *P. aeruginosa* for respiratory mucus may explain the propensity of the organism to colonise and persist in the mucosal secretions of CF patients. Nevertheless, the biosynthesis of alginate by mucoid *P. aeruginosa* is probably the key factor in its persistence. Molecular evidence suggests that the mucoid phenotype results from the derepression

of a normal regulatory system resulting in a shift from low, barely detectable, levels of alginate in non-mucoid strains to high levels in the mucoid phenotype in response to various environmental stimuli (see later).

The alginate exopolysaccharide produced by *P. aeruginosa* in CF patients lungs is highly viscid in aqueous solutions and has the capacity to bind both water and divalent cations to form flexible gels. Lam *et al.* (1980) found in Gram films of sputum and direct electron micrographs of post-mortem lung material from CF patients that *P. aeruginosa* exists as microcolonies wherein the alginate matrix occupies considerably greater space than the enclosed microbes. Furthermore, the growth of *P. aeruginosa* in biofilms has grave implications for the course of CF respiratory disease. Although there is an apparent presence of opsonins (Doggett and Harrison., 1972) that can be deposited on the bacterial surface, their interactions with receptors on phagocytic cells are inefficient due to the presence of an EPS barrier on *P. aeruginosa*, preventing ligand accessibility (Pier *et al.*, 1990). The alginate biofilm also suppresses leukocyte function, and non-opsonic phagocytosis and also, promotes adhesion to respiratory surfaces (Mai *et al.*, 1993; May *et al.*, 1991). *P. aeruginosa* isolated from the CF lung is found to be susceptible to a large range of antibiotics. However, in *vivo*, the high electrolyte content of the microfilm is antagonistic to the actions of β-lactams and aminoglycoside inhibitors and, as such, prevents eradication of the organism by antibiotic treatment.

Once *P. aeruginosa* is established within the airway therefore, it is rarely totally eliminated. The organism produces cilio-toxins which further impair ciliary clearance. It also produces proteases (Vasil *et al.*, 1990) which cause tissue damage and cleave immunoglobulins which further impairs the humoral host response to the organism (May *et al*, 1991). These proteases also inhibit the function of phagocytic cells, natural killer cells and T-lymphocytes by affecting the CD4 molecule on T-helper cells and inhibiting interleukin I and interleukin 2 activity (Warner, 1992).

Phagocytes stimulated by immune complexes and immunoglobulins fail to engulf the microcolonies due to their large size (up to 60μm), but still release elastase and toxic oxygen radicals which fail to penetrate the alginate exopolysaccharide and instead act on the lung epithelial surface, thus contributing to the considerable immune mediated damage observed in the CF lung. Indeed, much of the key damage observed in CF may well result from the host immune hyper-responsiveness to infection by mucoid *P. aeruginosa.*

1.5. THE ALGINATE BIOSYNTHETIC PATHWAY OF *P. AERUGINOSA*

Alginate is a highly negatively charged linear-copolymer of β-1-4 D mannuronic acid and its C-5 epimer, L-guluronic acid. Evans and Linker showed that bacterial alginates differ from the algal polymer as mannuronate residues may be esterified with *O*-acetyl groups (Skjåk-Bræk *et al.*, 1986) at positions *O*-2, *O*-3 or both, see below, figure 1.7.

A. vinelandii alginate more closely resembles the algal form with the EPS alginate consisting of both polymannuronate and polyguluronate blocks. Detailed analysis of alginate from *P. aeruginosa* showed acetylated poly-D-mannuronic acid with only individual residues of L-guluronate interspersed through the polymer (Sutherland, 1989).

The biosynthetic route of alginate was first established by Lin and Hassid (1966) for the marine brown alga *Fucus gardneri.* Later, a similar pathway for alginate biosynthesis in *A. vinelandii* was proposed by Pindar and Bucke (1975). Their proposed biosynthetic pathway was based on the detection of all the enzymes except the epimerase and *O*-acetyl transferase. Subsequent workers have assumed that a similar pathway exists in *P. aeruginosa.* The pathway for the biosynthesis of alginate is shown below in figure 1.8.

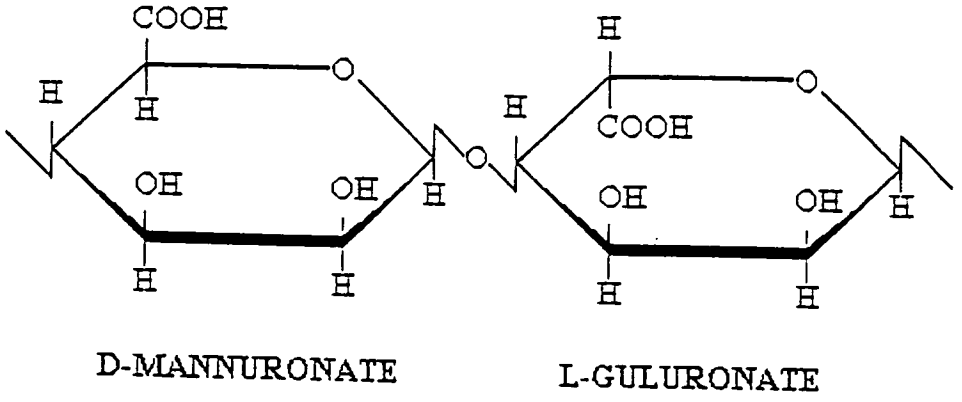


Figure 1.7. Alginate structure showing β-1-4-linked D-mannuronic acid and L-guluronic acid. Mannuronic acid residues may be modified with *O*-acetyl groups at position *O*-2, *O*-3 or both. The *O*-2 position is favoured (Skjåk-Bræck *et al.*, 1986).

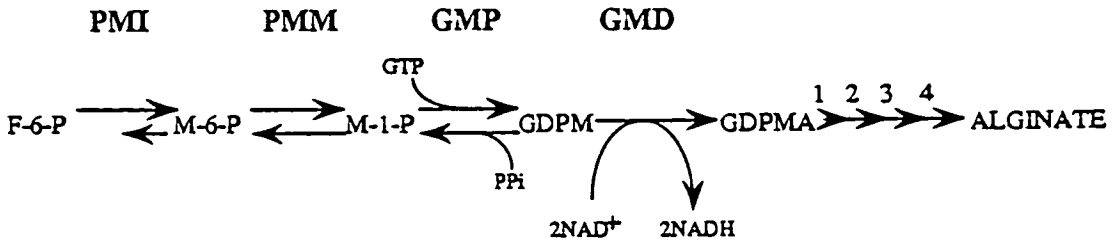


Figure 1.8. The *Azotobacter vinelandii* alginate biosynthetic pathway as proposed by Pindar and Bucke (1975). Arrows 1, 2, 3, 4 indicate the steps of polymerisation, acetylation, export and epimerisation. PMI, phosphomannose isomerase; PMM, phophomannomutase; GMP, GDP-mannose pyrophosphorylase; GMD, GDP-mannose dehydrogenase; F-6-P, fructose 6-phosphate; M-6-P, mannose 6-phosphate; M-1-P, mannose 1-phosphate; GDPM, GDP-mannose; GDPMA, GDP-mannuronic acid.

In the scheme of Pinder and Bucke (1975) the starting precursor was regarded as being fructose 6-phosphate which was isomerised to mannose 6-phosphate by phosphomannose isomerase (PMI). Transfer of the phosphate group via phosphomannomutase (PMM) produces mannose 1-phosphate. Guanosine diphosphate-mannopyrophosphorylase (GMP) catalysed the formation of GDP mannose from mannose 1-phosphate and GTP. The subsequent oxidation via GDP-mannose dehydrogenase (GMD) results in the formation of GDP mannuronic acid. The product of polymerisation of GDP mannuronic acid was assumed to be polymannuronate, which is secreted from the bacterial cell and becomes the substrate for an extracellular epimerase (Ertesvåg *et al.*, 1994) to produce the final polymer (although the epimerase in *P. aeruginosa* was found to have a periplasmic location, see later).

A study of alginate biosynthesis in *P. aeruginosa* mutants that were defective in carbohydrate metabolism, demonstrated the primary role of the Entner-Douderoff pathway enzymes in alginate biosynthesis generating the required intermediate from glucose, mannitol and gluconate (Bannerjee *et al.*, 1983). Thus carbon atoms 1, 2 and 3 of the glucose molecules are lost as pyruvate, whilst carbon atoms 4, 5 and 6 were channelled via the glyceraldehyde 3-phosphate into alginate.

1.5.1. **Genetic organisation of the genes involved in alginate biosynthesis in *P. aeruginosa***

The alginate synthesis and regulatory genes are clustered at four locations in the *P. aeruginosa* chromosome. Except for the *algC* gene, which is located at 10 minutes, all the known alginate biosynthetic genes are located at 34 minutes. The regulatory genes map at 10 minutes and 13 minutes and the genes responsible for the genetic switch to mucoidy (Flynn *et al.*, 1988) are located at 68 minutes on the *P.*

aeruginosa chromosome (Zielinski *et al.*, 1990). The genetic organisation of the genes involved in alginate production is represented schematically in figure 1.9.

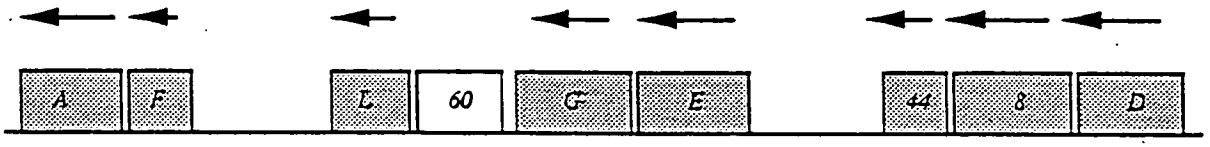
1.5.2. **Enzymology of the alginate biosynthetic pathway**

Phosphomannose isomerase-GDP mannose pyrophosphorylase (PMI-GMP)

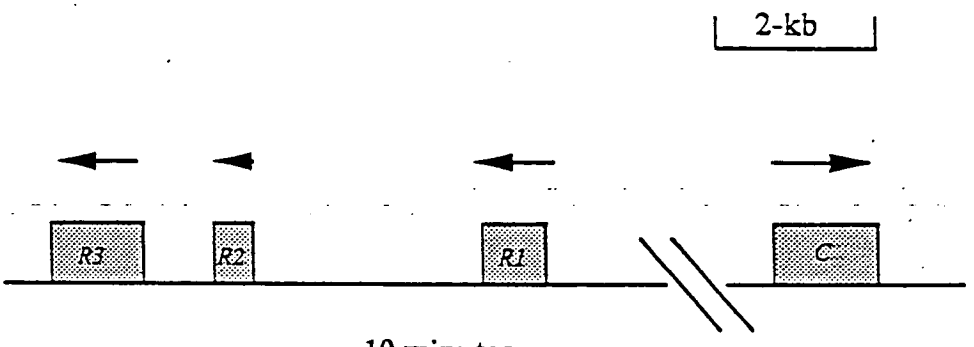
The bifunctional enzyme PMI-GMP is encoded by the *algA* gene, which was sequenced and initially characterised by Sá-Correia *et al.* (1987). This 56-kDa protein catalyses the first and third steps in the alginate pathway (isomerisation of fructose 6-phosphate and the formation of GDP mannose from mannose 1-phosphate and GTP). The PMI reaction strongly favours the forward direction and is crucial to the alginate pathway as it keeps fructose 6-phosphate from the metabolic pool. The GDP mannose pyrophosphorylase reaction (mannose 6-phosphate to mannose 1-phosphate), usually prefers the reverse direction, it is pushed in favour of alginate synthesis by the GMD reaction (GDP-mannose to GDP mannuronic acid). This arrangement conserves the supply of GTP in the absence of the GMD enzyme. No significant areas of amino acid similarity were identified in a comparison of PMI-GMP with the PMI enzymes of *E. coli*, *S. typhimurium* or *R. meliloti*. However, PMI-GMP was found to have a high degree of amino acid identity with the GMP enzymes RfbM and CpsP of *S. typhimurium* and *E. coli* (49.5% and 53.4% identity respectively, May *et al.*, 1994). XanB of *X. campestris*, a bifunctional PMI-GMP enzyme involved in xanthan gum synthesis, was found to be 59% similar (Köplin *et al.*, 1992).

Three regions of the PMI-GMP protein important for catalytic activity have been identified (review May *et al.*, 1994). Site-directed mutagenesis studies have shown that Arg 19 and Lys 20 facilitate GTP binding, probably by bringing the neighbouring glycine-rich region closer to the nucleotide binding site. Site-directed

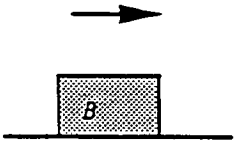
Figure 1.9. The organisation of the alginate gene clusters. The alginate genes are clustered at three locations in the *Pseudomonas aeruginosa* chromosome. All the alginate biosynthetic (structural) genes, except for *algC*, are located at 34 minutes. The regulatory genes map at 10 and 13 minutes, and the genes responsible for the genotypic switch to mucoidy are located at 68 minutes. The arrows above the genes represent the direction of transcription (modified from May and Chakrabarty, 1994).



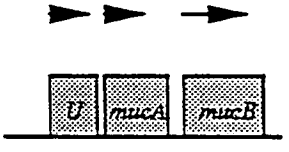
34 minutes



10 minutes



13 minutes



68 minutes

mutagenesis has also shown that Lys 175 is critical for binding the GMP substrate mannose 1-phosphate. PMI-GMP has a phosphate binding sequence (FVEKP) which is similar to the phosphosugar-binding sites of other pyrophosphorylases, suggesting that the active lysine of these proteins binds its substrate via the phosphate moiety (Sá-Correia *et al.*, 1987). Proteolytic cleavage of the *P. aeruginosa* PMI-GMP has shown that the carboxyl terminus is essential for PMI activity, but not for GMP activity. Similar conclusions have come from studies of the *X. campestris* XanB (PMI-GMP) protein. The PMI and GMP enzymatic activities are thought to reside within different catalytic domains since the substrates or products of one activity do not inhibit the other activity. In addition, these two enzymatic activities showed different requirements with respect to metal cofactors and reducing agents (May *et al.*, 1991).

Phosphomannomutase (PMM)

Zielinski *et al.* (1991) initially characterised *algC* and determined that it encodes the 51-kDa enzyme phosphomannomutase, PMM. This enzyme interconverts mannose 6-phosphate to mannose 1-phosphate, the second step of alginate synthesis and is an obligate enzyme of the pathway leading to the biosynthesis of alginate. PMM has also been shown to be required for the synthesis of mannuronic acid residues which are present as part of the O antigen side chain of *P. aeruginosa*, PAO1 (serotype O5) (Coyne *et al.*, 1994). The AlgC protein also appears to have an efficient phophoglucomutase activity that is required for the synthesis of glucose 1-phosphate (from glucose 6-phosphate), a necessary intermediate in the pathway leading to the addition of glucose residues in the LPS core. XanA of *X. campestris* has both PMM and PGM activity and shows a 56.5% similarity and 34.5% identity to the *P. aeruginosa* PMM along the full length of the polypeptide (Coyne *et al.*, 1994).

PMM and PGM catalyse intramolecular phosphate transfer reactions by way of a phosphorylated enzyme intermediate. The *P. aeruginosa algC* appears to provide the only source of PMM and PGM (*algC* mutants show neither activity). These findings suggest that *P. aeruginosa* minimises redundancy by utilising the *algC* gene in two biosynthetic pathways (LPS and alginate) and as the single enzymatic source for two functions, those of PMM and PGM (Coyne *et al.*, 1994).

GDP-mannose dehydrogenase (GMD)

GMD, belongs to the class of four electron transfer dehydrogenases and catalyses the nearly unidirectional, four electron oxidation of GDP-mannose to GDP-mannuronic acid (Pugashetti *et al.*, 1983). GMD is the commitment enzyme of the alginate pathway since the product GDP mannuronic acid is known to be utilised solely as a precursor of alginate synthesis (Roychoudhury, 1989). The *algD* gene, encoding GMD, has been cloned and sequenced by Deretic *et al.* (1987) and was found to encode a polypeptide of 48-kDa. The enzyme has been produced in large quantities and purified for enzymatic characterisation (Roychoudhury, 1989).

The effects of proteolysis on substrate binding and catalytic activity of the enzyme were examined in order to elucidate the structural and functional properties of GMD. The proteolytic studies indicated that the enzyme may fold into two domains, a 25-26-kDa amino-terminal domain and a 16-17-kDa carboxyl terminal domain. The amino terminal domain contains the substrate and cofactor binding sites, but does not possess catalytic activity, whereas the carboxyl domain is essential for catalytic activity. The Cys 268 site has been implicated in the catalytic activity of GMD due to its sequence homology with the active cysteine sites of two other, four electron transfer dehydrogenases (UDP-D-glucose dehydrogenase from bovine liver and L-histidinol dehydrogenase from *E. coli*). Replacement of Cys 268 by serine via

oligonucleotide-directed mutagenesis, abolished 95% of GMD activity confirming the importance of this Cys residue in enzyme catalysis.

Examination of the amino acid sequence of GMD revealed two copies of a consensus sequence for binding of the terminal phosphate group of GTP and a third consensus which forms part of the guanine binding site in the amino terminal domain. The spacing between each consensus is such that it allows folding to bring the regions together in the 3D structure to form the GDP-mannose binding site. Differences between the consensus sequence for other proteins that bind GTP and these for GMD appear to be due to the enzyme binding GDP rather than GTP. The primary sequence of GMD also reveals a sequence at the amino terminal region that bears homology to the β α β (Rossman) fold structure required for NAD^+ binding (Roychoudhury *et al.*, 1992). Consequently, GMD oxidises GDP-D-mannose to GDP-D-mannuronic acid via a four electron transfer redox reaction using NAD^+ as a co-factor.

Sá-Correia *et al.* (1987) showed that even in highly mucoid *P. aeruginosa* isolates from CF patients the key biosynthetic enzymes are still at low levels. This may suggest that PMI-GMP, PMM and GMP exist as an enzyme complex. Many metabolic enzymes are thought to exist in complexes known as 'metabolons' to allow efficient coupling of biosynthetic reactions. Such a complex was shown by the work of Matthews (1993) on T4 dNTP synthetase which suggested that most of T4 phage encoded enzymes of dNTP synthesis interact to form a specific multienzyme complex.

1.5.3 **Alginate modification enzymes and export proteins**

AlgL, alginate lyase

Alginates are enzymatically depolymerised by alginate lyases, which cleave the 1-4 glycosidic linkage by β elimination, resulting in an unsaturated non-reducing terminus. Linker and Evans (1984) found that strains of *P. aeruginosa* which produce alginate also synthesise an intracellular enzyme which can degrade these polysaccharides as well as the seaweed derived, alginic acid. This alginase appeared to be a polymannuronate lyase which degrades the polysaccharides, (depending on their uronic acid composition) to a series of oligosaccharides. Sequencing and characterisation of *algL* within the alginate biosynthetic cluster (Schiller *et al.*, 1993; Boyd *et al.*, 1993) predicts an AlgL (alginate lyase) protein of 41-kDa molecular weight, with a signal peptide that directs the lyase to the periplasmic space. Although *algL* has been shown not to be required for alginate synthesis it appears to be co-transcribed with the alginate biosynthetic genes and there is a correlation between the synthesis of alginate and the increased production of alginate lyase. AlgL does not appear to degrade alginate for use as a carbon source since *P. aeruginosa* expressing AlgL from an inducible plasmid cannot grow when alginate is the sole carbon source (Boyd *et al.*, 1993). The presence of the *algL* gene in the alginate gene cluster suggests that AlgL may serve in a biosynthetic capacity. It has been postulated that AlgL may cleave the polymer during export, thereby controlling the length of the polysaccharide chain. However, hyperexpression of *algL* (with the aid of a plasmid carrying the *tac* promoter) through early growth phase allows considerable sloughing of cells from the growth film, suggesting that alginate lyase may be involved in cell detachment from the biofilm under unfavourable conditions or in extending colonisation within the lung (May and Chakrabarty, 1994). The *algL* gene is transcriptionally regulated through the *algD* promoter, however, since lyase

activity appears well into the stationary phase, long after the alginate transcripts are produced it is probable that the synthesis of the protein is also regulated at the translational stage (Schiller *et al.*, 1993).

AlgF; acetyltransferase

Shinabarger *et al.* (1993) characterised the *algF* gene. Sequence analysis did not reveal any significant homology to other genes but biochemical characterisation of alginate produced by an *algF* mutant revealed that the gene controls the addition of *O*-acetyl groups to the mannuronate residues of alginate (approximately 30 to 60% of the mannuronate residues are acetylated in *P. aeruginosa* alginate, Skjåk-Bræk *et al.*, 1986). Complementation analysis demonstrated that *algF* is not obligately required for synthesis of the alginate polymer.

Nucleotide sequence analysis of *algF* predicted a 23-kDa polypeptide. This polypeptide appears to have a 28 amino acid signal peptide sequence which can mediate in export of the protein through the cytoplasmic membrane, giving it a periplasmic location. Since AlgF is a periplasmic protein, it is unlikely that acetyl CoA, which cannot cross the cytoplasmic membrane, is a substrate for this enzyme. A potential acyl carrier protein has been identified upstream of *algF*, and it seems plausible that the acetylated acyl carrier protein, is the donor of acetyl groups for the acetylation of polymannuronate.

The function of the *O*-acetyl groups in alginate biosynthesis is not well understood, but is has been suggested that acetylation protects the mannuronate residues from being converted to guluronate residues. Since acetyl groups protect against epimerisation, it is thought likely that the epimerisation event occurs after acetylation. The presence of acetyl groups within alginate is believed to increase the water-binding capacity of the polymer - it is probable that the alginate which

encapsulates *P. aeruginosa* would protect the organism better from the dehydrated environment within the lungs of CF patients if acetylated (Shinabarger *et al.*, 1993).

AlgG, C-5 epimerase

Mutants of *P. aeruginosa* which produce polymannuronate exopolysaccharide without any L-guluronate residues were found to be due to a defect in an alginate C-5 epimerase (Chitnis and Ohman, 1990). The defect in an alginate C-5 epimerase was due to a mutation in *algG* which encodes the enzyme. Franklin *et al.* (1994) sequenced *algG* and deduced the amino acid sequence of its protein product, AlgG. Observation on expression of AlgG in *P. aeruginosa* and *E. coli* suggest that AlgG has a signal sequence which is removed during export from the cytoplasm to the periplasm, implying that AlgG acts upon alginate at the polymer level (as polymerisation of activated mannuronate units is assumed to occur on the cytoplasmic side of the inner membrane, prior to transport to the periplasm). Studies with deacetylated polymannuronate and AlgG revealed epimerisation to the C-5 epimer, L-guluronate. Thus, AlgG encodes a mannuronate-C-5 epimerase, which is inhibited by the presence of acetyl groups on the substrate.

AlgE, outer membrane porin

AlgE was first sequenced by Chu *et al.* (1991). The gene encoded an outer membrane protein (Grabert *et al.*, 1990) of 54-kDa. Further functional and structural characterisation of AlgE (Rehm *et al.*, 1994a) found AlgE to be a minor outer membrane protein of *P. aeruginosa*. AlgE is anchored in the outer membrane by 28 membrane spanning β strands, that probably form a β-barrel. Recombinant AlgE was spontaneously incorporated into planar lipid bilayers, forming an ion channel. Single channel current measurements in the presence of various salts and reversal potential measurements in salt gradients revealed that the AlgE channel was strongly

anion selective. AlgE shows significantly lower conductance and open-state lifetime values compared to those of other porin channels, such as OmpF and OmpC and this predicts that AlgE does not function as a typical porin (Rehm *et al.*, 1994a).

Other enzymes involved in alginate synthesis

The incorporation of mannuronic acid residues into the alginate polymer is not entirely understood at present. However the DNA sequences of *alg8* and *alg44* genes predict that their gene products are membrane bound proteins, which may form components of a polymerisation complex (Maharaj *et al.*, 1993). The *alg60* gene was found to encode a 43-kDa polypeptide (Wang *et al.*, 1987), however the function of this protein is entirely unknown (May *et al.*, 1991; May and Chakrabarty, 1994) and the nucleic acid sequence of *alg60* has yet to be published.

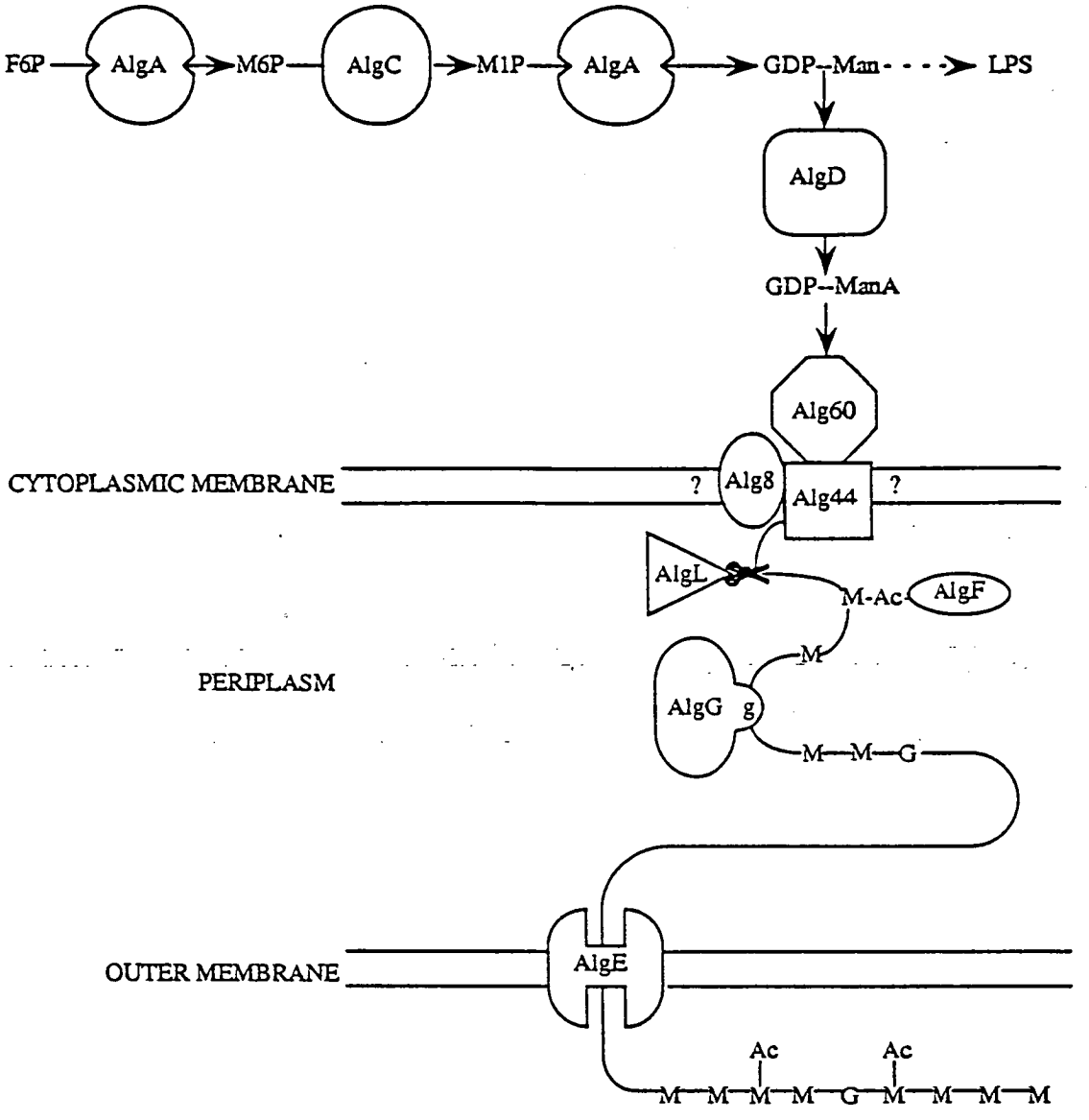
A model for the way the known alginate biosynthetic enzymes interact in the synthesis of the exopolysaccharide is shown in figure 1.10.

1.6. REGULATION OF ALGINATE SYNTHESIS AND CONVERSION TO MUCOIDY

Operonic structure

Most of the genes encoding enzymes in the pathway of alginate biosynthesis are clustered in about 18-kb of DNA near *argF* at 34 minutes on the *P. aeruginosa* chromosome. Results from expression studies using DNA fragments and genes from this region, suggest that all the genes for alginate production are transcribed in the same direction. Chitnis and Ohman (1993) used various genetic and Northern hybridisation experiments to test the hypothesis that the large alginate biosynthetic cluster had an operonic structure (which would be in keeping with other Gram-negative bacterial EPS biosynthetic clusters). Mutants of the Alg^+ strain FRD1 were

Figure 1.10. The biosynthesis of alginate by *Pseudomonas aeruginosa*. Fructose 6-phosphate (F6P) obtained from the metabolic pool is converted to GDP-mannuronic acid (GDP-ManA), which provides mannuronate residues (M) for polymerisation. Occasionally guluronate residues (G) are formed from epimerisation of mannuronate residues by a C-5 epimerase (AlgG); g represents the transition stage of mannuronate epimerisation to guluronate. Mannuronic acid residues of bacterial alginates are partially *O*-acetylated by the AlgF protein. The question marks indicate that the functions of the *alg8*, *alg44* and *alg60* gene products and the putative interactions of these proteins are hypothetical. The dashed arrow indicates the enzymatic steps leading to lipopolysaccharide (LPS) synthesis. Abbreviations used: M6P, mannose 6-phosphate; M1P, mannose 1-phosphate; GDP-Man, GDP-mannose; and Ac, *O*-acetyl groups (diagram as modified from May and Chakrabarty, 1994).


F6P
AlgA
M6P
AlgC
M1P
AlgA
GDP–Man
LPS
AlgD
GDP–ManA
Alg60
CYTOPLASMIC MEMBRANE
?
Alg8
Alg44
?
AlgL
M-Ac
AlgF
M
PERIPLASM
AlgG
g
M—M—G
AlgE
OUTER MEMBRANE
Ac
Ac
M M M M G M M M M

constructed by gene replacement with defined Tn*501* (8.2-kb) insertions in the alginate biosynthetic gene cluster, resulting in an Alg^- phenotype. The Alg^+ phenotype of these mutants was restored by integration of narrow host range plasmids containing DNA fragments from *P. aeruginosa* that reconstructed a continuous alginate gene cluster. A broad host range plasmid containing the entire alginate gene cluster except for the terminal gene *algA*, was unable to complement Tn*501* mutants unless *algA* was transcribed from a second plasmid. This indicated that any Tn*501* insertion in the cluster was polar on downstream alginate genes. Northern blot hybridisation experiments also showed that a transposon insertion downstream of *algD* adversely affects both *algG* and *algA* transcription. These results provided evidence that the alginate biosynthetic gene cluster has an operonic structure and the genes are co-transcribed from the *algD* promoter.

Insertional inactivation of *algL* (Boyd *et al.*, 1994) and *algF* (Shinabarger *et al.*, 1993) also resulted in an Alg^- phenotype. However, conversion to the Alg^- phenotype was shown to be entirely due to a polar effect on expression of *algA*, again predicting an operonic structure for the alginate biosynthetic genes.

Regulation of alginate synthesis

An intriguing aspect of *P. aeruginosa* within the CF lung is the genotypic transition to a mucoid form. Although *P. aeruginosa* causes a wide range of infections in immunocompromised patients (Clark, 1990), the transition to mucoidy (the state of constitutive alginate production) is rare, so that prolonged exposure to the unique environment within the CF-affected lung would seem to be important in triggering non-mucoid *P. aeruginosa* to produce alginate (Bayer *et al.*, 1990; Terry *et al.*, 1991). Once this transformation to a mucoid form has occurred, the cells become increasingly less susceptible to the host immune system, allowing chronic colonisation of the pulmonary tract (May *et al.*, 1991). The transition to mucoidy is

accompanied by reduced levels of several other virulence factors produced by *P. aeruginosa* (e.g. proteases, exotoxins, phospholipase C and pyochilin). In addition, a partial to complete loss of the LPS O-antigen occurs in mucoid strains of *P. aeruginosa*; this may be due to activated *algD* draining the supply of GDP-mannose to produce alginate (May *et al.*, 1991).

The transition to mucoidy is accompanied by the activation of the *algC* and the *algD* promoters. The *algD* promoter controls the alginate biosynthetic cluster at 34 minutes, presumably as an operon, whereas the *algC* gene which is located at 10 minutes on the biosynthetic gene cluster is encoded separately (Zielinski *et al.*, 1992). Both of these promoters have been found to be activated by specific environmental factors. The defective chloride channel in the CF-affected lung leads to the accumulation of a sticky, dehydrated mucus and a salty cellular environment and these are thought to be the primary signals for *algC* and *algD* promoter activation. It is thought likely that the initial impact of the CF lung environment is exerted through the activation of these alginate promoters. In order to deduce the validity of this supposition a variety of experiments have been performed to test the environmental activation of both *algD* and *algC* genes by mimicking various aspects of the CF lung environment and evaluating their affect by the use of transcriptionally fused *algD* or *algC* promoters to a reporter gene. The two reporter systems utilised in these experiments were *xylE* and *lacZ*. *xylE* encodes catechol 2,3-dioxygenase (C23O), which catalyses the formation of an intensely yellow product (2-hydroxymuconic semialdehyde) from catechol and thus transcription of the fusion can be tracked spectrophometrically. The β-galactosidase (*lacZ*) reporter system also allows the following of transcription spectrophometrically as β-galactosidase hydrolyses ONPG (O-nitrophenyl-β-D-galactopyranoside) to a yellow product. The results of these experiments are summarised in table 1.2.

Environmental signal	Promoter-reporter gene fusion	Host bacterium (*P. aeruginosa*)[i]	Level of accumulation
Osmolarity	*algC-lacZ*	8821 8822	5-6 fold 5-6 fold
	algD-xylE	8821 8822	4-5 fold 3 fold
Ethanol	*algD-xylE*	8821 PA01	3-4 fold 12 fold
N_2 limitation	*algD-xylE*	8821 PAO1	2 fold 3-4 fold
PO_4 limitation	*algD-xylE*	8821 PAO1	2 fold 2 fold
Adherence[ii]	*algC-lacZ* *algD-lacZ*	8830 PAO579	3-5 fold 2-5 fold

[i] *P. aeruginosa* strains 8821, 8830 and PAO579 are mucoid, while strains 8822, 8830 and PAO1 are non-mucoid.

[ii] Adherence to a solid surface reflects biofilm formation.

Table 1.2. Environmental factors that cause the activation of *algC* and *algD* genes of the alginate biosynthetic pathway of *P. aeruginosa* (from May and Chakrabarty, 1994).

The respiratory tract fluid of CF patients, due to the defect in CFTR, exhibits abnormally high levels of Na^+ and Cl^- (90 and 80 mM respectively). The effect of high osmolarity on *algD* activation was studied by Berry *et al.* (1989). *P. aeruginosa* strains containing a plasmid bearing a *algD-xylE* transcriptional fusion showed a 4-5 fold increase in expression of *xylE* in the presence of 0.5 M NaCl. Nalidixic acid and novobiocin, which reduce supercoiling by inhibition of DNA gyrase activity, abolished the activation of *algD* under high osmolarity, suggesting that DNA supercoiling is required for high level activation of *algD*.

The genetically determined CF defect also leads to a viscous dehydrated mucus and an overall desiccated lung environment. DeVault *et. al.* (1990) incorporated ethanol into agar plates to simulate this *in-situ* scenario. Using an *algD-xylE* transcriptional fusion plasmid, induction was from 3-4 fold in a mucoid strain and 10-12 fold in a non mucoid strain. Again, nalidixic acid prevented any activation of *algD*.

Formation of *P. aeruginosa* biofilms due to the production of alginate within the CF lung was shown by Lam *et al.* (1980). Production of exopolysaccharide by planktonic and adherent *P. aeruginosa* was monitored using an *algD-lacZ* promoter-reporter gene construct and showed a 2-5 fold increase in expression of *algD* on adherence of the bacteria to a solid surface (Hoyle *et al.*, 1993).

Nutrient deprivation was also found to induce *algD* transcription. DeVault *et al.* (1989) found that nitrogen and phosphate starvation resulted in a 2-3 fold increase in *algD* transcription, also determined by an *algD-xylE* transcriptional fusion.

Thus the transcription of *algD* (and thus the rest of the operon) has been found to be induced by high osmolarity, dehydration, biofilm formation, and nutrient deprivation, all respiratory conditions common to CF.

Zielinski *et al.* (1992) studied the effect of osmolarity on *algC* expression using an *algC-lacZ* system. Maximal activation occurred with 0.3M NaCl giving a

4-5 fold increase in expression. However, Zielinski *et al.* (1992) also showed that *algC* promoter activation was not appreciably modulated by other environmental conditions peculiar to the CF lung, such as oxygen tension, pH, nutrient-limitation or ethanol induced dehydration, although Davies *et al.* (1993) found that adherence also caused activation of *algC* by 3-5 fold levels.

The role of regulatory proteins which affect alginate production

Activation of the *algC* and *algD* promoters is mediated by the regulator proteins AlgR1 (Deretic *et al.*, 1989; Kato and Chakrabarty, 1989), AlgR2 (Konyecsni and Deretic, 1990; Kato *et al.*, 1989), and AlgR3 (Konyecsni and Deretic, 1990; Kato *et al.*, 1990) also referred to as AlgR, AlgQ, and AlgP, respectively.

The main regulatory protein in mucoidy appears to be AlgR1. DNA sequence analysis suggests that the gene belongs to a class of two-component, environmentally responsive, bacterial regulatory genes (Deretic *et al.*, 1989). Two-component sensory transduction systems consist of a sensory (transmitter) protein that detects a certain environmental stimulus and relays the information to a regulatory (receiver) protein that in turn activates the genes necessary for growth in the presence of the specified environmental condition. The *algR1* gene product appears to be of the regulatory (receiver) class of proteins (rather than the sensory type). The *algR1* promoter has a σ^{54} (or a *P. aeruginosa* homologue) recognition sequence for RNA polymerase binding - σ^{54} is an alternative sigma factor which, like all other sigma factors, confers upon RNA polymerase the ability to recognise specific promoter sequences. Genes expressed by RNApol-σ^{54} control diverse cellular functions in *P. aeruginosa* such as expression of pilin and flagellin. AlgR1 has been purified and shown to bind to *algD* far upstream sites (Kato and Chakrabarty, 1991) at two high affinity positions (RB1 and RB2) centred at -468 and -391-bp, respectively from the

transcriptional start site. In addition, a low-affinity AlgR1 binding site, RB3, centred at -40, has been reported to participate in *algD* activation (Mohr *et al.*, 1991).

AlgR1 also binds to two regions of the *algC* promoter, upstream of the mRNA start site. These binding sites are similar to eukaryotic enhancer elements as they can be moved upstream, or downstream or even switched to the opposite orientation without disrupting the activation of the promoter (Zielinski *et al.*, 1992).

AlgB, also modulates expression of *algD* and like AlgR1, is thought to be a member of the NtrC subclass of two component prokaryotic regulators (Wozniak and Ohman, 1991). AlgB is not essential for alginate synthesis, but activates the *algD* promoter leading to a high level production of the polymer. AlgB, in common with NtrC, shows a conserved central domain which in NtrC interacts with RNA polymerase containing σ^{54} and a carboxy-terminal domain which contains a DNA binding motif and a conserved amino-terminal domain. AlgB has been purified but the isolated form does not appear to bind to a sequence upstream of the *algD* transcriptional start site (Goldberg and Dahnke, 1992; Wozniak and Ohman, 1991). Thus, there may be another gene (*algX*) in the circuit between *algB* and *algD*. In addition AlgB does not appear to be required for *algR1* expression or vice versa. Moreover, with the exception of a requirement for AlgU (see later), the AlgB and AlgR1 pathways appear to be operating independently of each other (Wozniak and Ohman, 1994).

Konyecsni and Deretic (1990) and Kato *et al.* (1989, 1990) characterised AlgR2 (AlgQ) and AlgR3 (AlgP). AlgR2 and AlgR3 do not appear to be regulated and the promoters of both have σ^{70} RNA polymerase recognition sequences and as such, are constitutively expressed at low levels in mucoid and non-mucoid *P. aeruginosa*. AlgR3 has an unusual structure and appears to be a histone-like protein. AlgR3 has been shown to be required for *algD* transcription (Kato *et al.*, 1990).

However more recent work by Deretic *et al.* (1993) revealed that AlgR3 is an intracellular protein with a wide target distribution suggestive of a more general role in transcription. AlgR2 is a kinase and undergoes autophosphorylation in the presence of ATP or GTP and transfers the acquired phosphate to AlgR1 (Roychoudhury *et al.*, 1992; Deretic *et al.*, 1992). Phosphorylation of response regulators of bacterial two component signal transduction systems is known to enhance their DNA binding ability. Indeed, Zielinski *et al.* (1992) showed that AlgR1-phosphate had the ability to bind its cognate binding site on the *algC* promoter with a higher affinity than AlgR1.

A requirement for integration host factor (IHF) from *P. aeruginosa* in *algD* expression was examined by Wozniak (1994) (IHF of *P. aeruginosa* is an analogue of the DNA-binding bending protein, IHF). Two consensus IHF binding sites were identified 75-bp upstream and 90-bp downstream from the *algD* transcriptional start site. Each bound with a different affinity, the second site having a 90-fold higher affinity for IHF than the first. An IHF binding site is also found in the *algB* promoter region, thus IHF controls transcription of *algB*, at least in part. Mutations in the IHF binding sites in *algD* reduced transcription by 3-4 fold. The role of IHF in *algD* expression may be to promote loop formation to allow activators bound far upstream or downstream to gain access to the promoter.

Binding of additional regulatory factors such as AlgB, may be accommodated by the physical organisation of the *algD* promoter. These additional elements and factors affecting the nucleoid structure may participate in the transduction of environmental signals, such as nitrogen limitation, dehydration, medium osmolarity, oxygen tension, or stress in general, into a concerted response of the *algD* promoter.

Conversion to mucoidy

Environmental activation of the *algC* and *algD* promoters does not lead to genotypically mucoid cells, although it may lead to transient alginate synthesis, allowing cells to attach to the lung surfaces and form a biofilm. However, continued proliferation in the CF-affected lung triggers an event that leads to constitutive production of alginate. Prolonged stress conditions (i.e. nutrient starvation, growth in the presence of antibiotics or ethanol, or mutagenesis of the *muc* locus) can also induce a genotypic switch to alginate production. A gene cluster mapping at 68 minutes on the *P. aeruginosa* chromosome, which consists of *algU* (also called *algT*), *mucB* (also called *algN*) and *mucA*, appears to regulate the conversion to constitutive alginate production, so that cells isolated from the CF affected lung are mucoid even when grown on laboratory culture media. AlgU affects the mucoid phenotype and *algD* transcription and shows sequence similarity with the alternative sigma factor σ^H (SigH or SpoOH) from *Bacillus subtilis* and *B. licheniformis* (Martin *et al.*, 1993) but more markedly to a global stress response factor, σ^E, previously putatively identified in *E. coli* and *S. typhimurium* (Erickson and Gross, 1989).

SigH of *B. subtilis* and *B. licheniformis* controls initial stages in the developmental processes of sporulation and competence as well as several post-exponential phase processes including stationary-phase-specific extracellular enzymes (Tatti *et al.*, 1989). In comparison, σ^E functions as a global stress response factor suggesting that the AlgU putative sigma factor is likely to contribute to a more comprehensive defence against environmental stress (possibly including heat shock and oxidative stress, Martin *et al.*, 1994) as well as a role in transcription of alginate biosynthetic genes. Unsurprisingly, inactivation of *algU* on the chromosome of a mucoid *P. aeruginosa* results in a loss of alginate production, and, more importantly abrogates transcription of the key alginate biosynthetic gene *algD* (Deretic *et al.*, 1993).

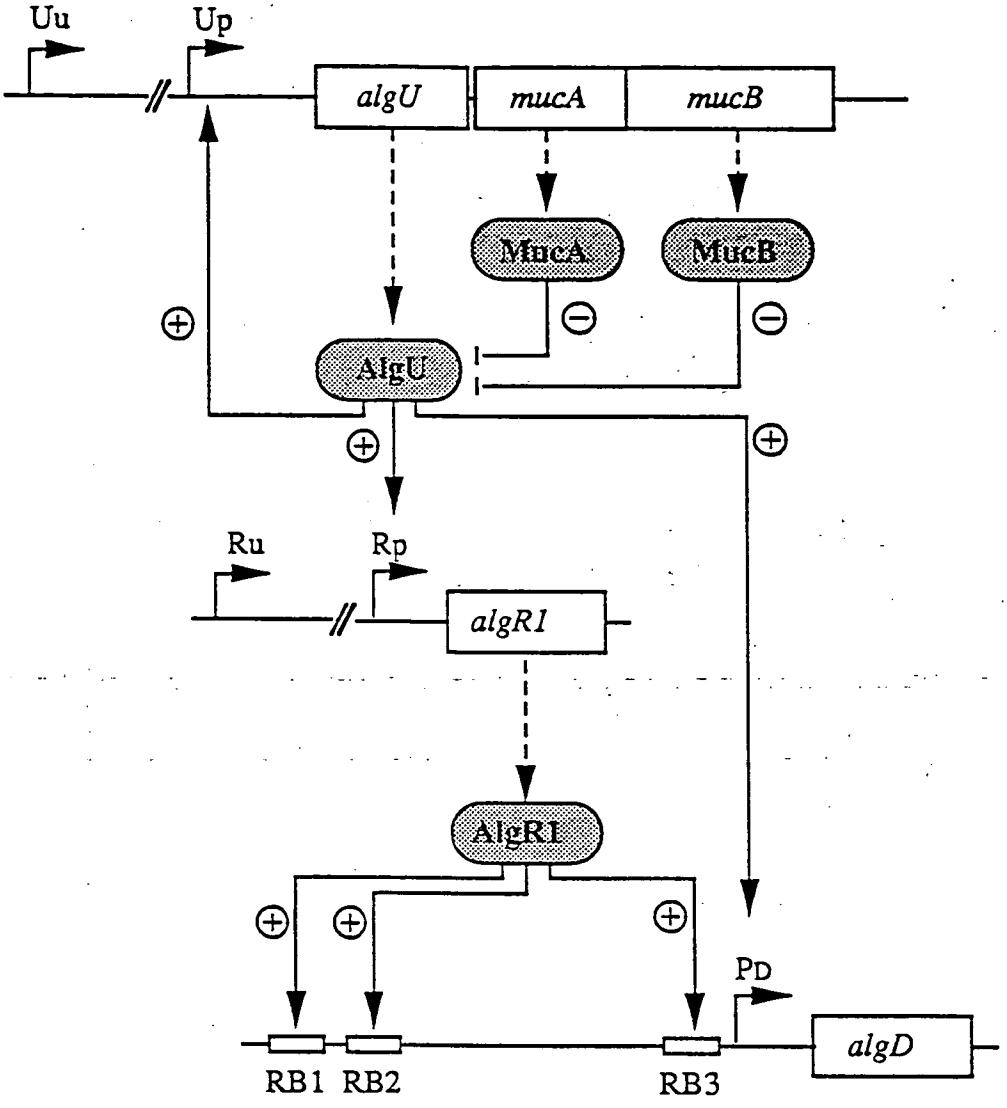
AlgU is required for the expression of *algB* and *algR1*, two genes that encode response regulators and both these genes are necessary for transcription of the *algD* promoter (Wozniak and Ohman, 1994). Transcription from the promoter proximal to the *algU* coding region was found to be dependent on AlgU. The -35 and -10 sequences of this promoter showed strong similarity to the promoters of *algD* and *algR1*. In wild type non-mucoid cells and in the absence of appropriate exogenous stimuli, transcription of *algU* and *algR1* from the AlgU independent upstream promoters most likely provides a baseline expression of these two critical regulatory factors. In mucoid mutants and presumably in the wild type cells under inducing conditions, transcription of *algU* and *algR1* is further stimulated as a result of the activation of their respective AlgU-dependent promoters. This, in turn, supplies adequate levels of the alternative sigma factor AlgU and the response regulated AlgR1 to activate the *algD* promoter and induce the synthesis of alginate.

The level of *algU* expression is additionally controlled by the accessory elements *mucA* and *mucB*, most likely at the post-translational level. Mutations in *mucA* or *mucB*, such as those described for mucoid *P. aeruginosa* isolated from CF patients, cause activation of the AlgU-dependent systems (DeVries *et al.*, 1994). Indeed, Martin *et al.* (1993) identified *mucA* as a major site for mutations causing mucoidy in *P. aeruginosa*. The *mucA* gene and a tightly linked downstream gene, *mucB*, are both required for suppression of mucoidy. A strong activation of *algD* transcription and conversion to mucoidy are observed when *mucA* or *mucB* are inactivated on the chromosome of previously non-mucoid strains, provided that the first gene of the cluster (*algU*) is intact.

Although the precise activities of *mucA* and *mucB* remain to be biochemically defined, the genetic evidence suggests that they suppress the function or expression of *algU* and therefore function as anti-sigma factors. This system along with signal transduction regulators and histone-like elements, is likely designed to control

development of biofilms in response to appropriate environmental cues and stresses. Superimposed on this regulatory network are mutations in *mucA* that lock the system in its constitutive state, (which is favourable and therefore selected) in the CF lung microenvironment because of the anti-phagocytic properties of the mucoid coating (Govan *et al.*, 1992). Mucoid *P. aeruginosa* isolated from CF patients were found to have frameshift mutations, deletions and also nonsense mutations (that generate premature stop codons) at different locations in *mucA*. However, the conversion of *P. aeruginosa* PAO to the mucoid phenotype has also been associated with a genetic rearrangement upstream of the endotoxin A gene. This genetic rearrangement was due to insertion elements (possibly of plasmid origin) and suggests a more complicated regulatory system may be involved (Sokol *et al.*, 1994). Interaction of the various regulatory elements involved in mucoidy of *P. aeruginosa* is shown in figure 1.11.

Figure 1.11. Regulatory circuitry controlling *algD* transcription and expression of mucoid phenotype. At the top of the regulatory cascade controlling *algD* expression is the *algU*, *mucA*, *mucB* gene cluster controlling the conversion to mucoidy in *P. aeruginosa*. At least two promoters, U_u (distal) and U_p (proximal), control the expression of *algU*. The transcription of *algR1* occurs from R_u (distal) and R_p (proximal) promoters. The *algD* gene appears to be transcribed from a single promoter, P_D. Expression of *algD* is absolutely dependent upon two factors: AlgU, the putative sigma factor initiating transcription at P_D, and AlgR1, a response regulator which binds to the three sites RB1, RB2, and RB3 upstream of P_D and further augments expression of *algD*. MucA and MucB act to suppress AlgU activity, and mutations in these genes can cause conversion to mucoidy by relieving AlgU from negative regulation. This is associated with increased transcription of alginate regulatory genes which up-regulate the expression from the critical *algD* promoter. Activation of *algD* is thought to affect the expression of the large cluster of alginate biosynthetic genes located downstream of *algD* (diagram adapted from Martin *et al.*, 1994).


Uu
Up
algU
mucA
mucB
MucA
MucB
AlgU
Ru
Rp
algR1
AlgR1
PD
algD
RB1
RB2
RB3

1.7. AIMS

The critical role played by alginate in the pathogenesis of *P. aeruginosa* within the CF lung has led to intensive studies to resolve the mechanism of its synthesis. Indeed, the majority of the biosynthetic, modification and regulatory proteins involved in *P. aeruginosa* alginate production have been characterised. However, as yet, the proteins involved in the integral steps of polymerisation of GDP-mannuronic acid to polymannuronate and cytoplasmic and periplasmic transport have failed to be elucidated. These processes are thought to be encoded within the uncharacterised regions of the 34 minute alginate biosynthetic operon. The general aim of this thesis is to characterise such unknown components of alginate synthesis.

Although undecaprenol-dependent polymerisation mechanisms have yet to be identified in the synthesis of bacterial alginate, it has been generally assumed to be the most probable scenario (Sutherland, 1982; May *et al.*, 1991). Thus, it is expected that polymannuronate (the alginate precursor) is formed by the sequential addition of mannuronic acid on undecaprenol lipids. Involvement of undecaprenol in such a polymerisation process provides an obligatory requirement for the cytoplasmic membrane (the subcellular location of undecaprenol lipids). Thus, the assembly of polymannuronate is likely to occur at the cytosolic face of the cytoplasmic membrane. Completion of polymerisation at such a location provides an intrinsic requirement for a post-polymerisation export mechanisms, channelling the polysaccharide through the inner membrane, periplasm and outer membrane. Previous surveys on many Gram-negative bacteria that utilise undecaprenol-dependent polymerisation mechanisms have revealed the inner membrane components required for polysaccharide translocation. All these bacteria appear to utilise inner-membrane transport systems homologous to members of the ABC

transporter superfamily. The initial aim of this thesis is to identify a similar homologue within the alginate biosynthetic cluster of *P. aeruginosa*. Such a homologue may well be provisionally identified and located by performing a DNA hybridisation survey using the conserved ATP-binding cassette (which typifies this superfamily) as a probe.

However, if this approach proves unsuccessful then an alternative strategy of directly sequencing one of the two uncharacterised regions of the alginate biosynthetic cluster can be employed. Sequencing of such a region should reveal potential open reading frames, the function of the putative gene products encoded may be deduced by amino acid or structural similarity to proteins previously characterised for other Gram-negative exopolysaccharide biosynthetic systems. Alternatively, polypeptide function maybe intimated through homologies observed through a more general database search or by a functional motif program. The putative open reading frame may be confirmed by overexpression of the gene-product and potential translational modification events can be monitored by pulse-labelling with [^{35}S]methionine. Subcellular location and membrane protein topography may be deduced by constructing gene fusions to a marker enzyme such as β-lactamase. Additionally, generation of a mutant in the chromosomally located gene, coupled with suitable complementation analysis may reveal whether the protein is essential for the production of the mucoid *P. aeruginosa* phenotype and may also help determine function of the protein.

Overall, this thesis aims to uncover further details in the production of the *P. aeruginosa* alginate exopolysaccharide as a virulence factor. It is hoped that this work will provide additional insights and information towards a fuller understanding of alginate synthesis within this microorganism.

Chapter 2

Materials and Methods.

2.1. MATERIALS

2.1.1. Enzymes, isotopes and chemicals

Stabilised solutions of [α-^{32}P] dCTP (3000 Ci/mmol) and [α-^{35}S] dATP (400 Ci/mmol) were purchased from Amersham International plc., Lincoln Place, Buckinghamshire. Stabilised [^{35}S]methionine (800 Ci/mmol) was purchased from NEN Du Pont Ltd, Wedgewood Way, Hartfordshire. Enzymes were purchased from Boehringer Mannheim UK, Bell Lane, East Sussex. Specialist chemicals: agarose was bought from Northumbria Biologicals Ltd., Nelson Industrial Estate, Cramlington, Northumberland; antibiotics were supplied by Sigma Chemical Co. Ltd., Fancy Rd., Poole. Dorset; Isopropyl-β-D-thiogalactoside (IPTG) and 5-bromo-4-chloro-3-indoyl-β-D-galactoside (X-Gal) were purchased from Boehringer Mannheim. Rabbit anti-β-lactamase antibodies were purchased from 5 Prime-3 Prime, Inc, 5603 Arapahoe Road, Boulder, Colorado, USA. Other chemicals, including goat anti-rabbit IgG, were supplied by Sigma Chemical Co., BDH Chemicals Ltd., Merck House, Poole, Dorset, and Fisons Scientific Equipment, Bishop Meadow Rd., Leicestershire. Hybond-N nylon membrane and Hybond-C super nitrocellulose membrane were purchased from Amersham International plc. and X-ray film (Curix RP1 and CRONEX 4 Pelicula, both 100 NIF) was supplied by AGFA (H.A. West, Edinburgh) and Du Pont UK, respectively.

Synthetic oligonucleotides were purchased from Oswel DNA Service, Mayfield Road, Edinburgh, UK. Oligonucleotides used are listed in Table 2.1.

2.1.2. Bacterial strains and plasmids

Bacterial strains and plasmids and their sources are shown in Table 2.2. and 2.3. respectively.

PRIMER	PRIMER SEQUENCE	COMMENTS AND USE
-40 Universal	5'-GTTTCCCAGTCACGAC-3'	Sequencing mp19 derived vectors (Yanisch-Perron *et al.*, 1985)
β-Lactamase (2294-2312)	5'-CTCGTGCACCCAACTGA-3'	Sequencing *algK-'blaM* fusion junctions (Broome-Smith and Spratt, 1986)
Primer 1 (225-251 of *algK*)	5'-GGTACCCGCATATGAAGATGCCCATC-3' *Nde*I	PCR from pAA1 for *algK* 5'-3' with additional *Nde*I site.
Primer 2 (379*-1642 of *algK*)	5'-GGGGATCCTCATAGGTTTCTGGCTCTTCTT-3' *Bam*HI	PCR from pAA1 for *algK* 3'-5' with additional *Bam*HI site.
Primer 3 (80-61 of pT7-7)	5'-CAACGATATCCCTCTAGAAA-3' *Eco*RV	PCR from pAA20 for *algK* and RBS 5'-3', with additional *Eco*RV site, (Tabor, 1990)
Primer 4 (381*-1642 of *algK*)	5'-CGGGATCCAGCTGTAGGCTTTCTGGCTCTTCTT-3' *Bam*HI *Pvu*II	PCR from pAA20 for *algK* 3'-5', with additional *Bam*HI and *Pvu*II sites.
Primer 5 (850-873 of *algL*)	5'-CCCTGCAGAGCGACGACTTCAAC-3' *Pst*I	PCR from PAB for *algL* fragment 5'-3' (Boyd *et al.*, 1993).

PRIMER	PRIMER SEQUENCE	COMMENTS AND USE
Primer 6 (1325-1302 of *algL*)	5'-GCCTGCAGGGCGCCGTGCCTTCTC-3' *Pst*I	PCR from PAB for *algL* fragment 3'-5', (Boyd *et al.*, 1993).
Primer 7 (1212-1233 of *algK*)	5'-CGCTGCAGGACCCGCGCAAGGC-3' *Pst*I	PCR from pAA22 for *algK* fragment 5'-3' with additional *Pst*I site.

Table 2.1. Oligonucleotides used for sequencing and for PCR where PAB is *P. aeruginosa* strain B and RBS is the equivalent of ribosome binding site. The numbers underneath each oligonucleotide refer to the base position either in figure 4.2 (for *algK*), in the reference source quoted or Chu *et al.* (1991)*.

Table 2.2. Bacterial strains. Table of all *E. coli* and *P. aeruginosa* strains mentioned in this thesis including genotype and source.

BACTERIAL STRAIN	GENOTYPE	SOURCE/REFERENCE
E. coli		
BL21(DE3)	F^-, *omp*$^+$, r_B^-, m_B^-, (DE3 is λ lysogen bearing *lac*I, *lac*UV5, *gene1*)	Studier *et al.* (1990).
HB101	F^-, *hsd*S20 (r_B^-, m_B^-), *sup*E44, *recA*13, *ara*14, *pro*A2, *rps*L20 (str^R), *xyl*-5, *mlt*-5, *sup*E44, λ-	Maniatis *et al.* (1983).
JM101	*thi*, Δ(*lac-pro*AB), [F'-*tra*D36, *pro*AB, *lac*I^q **Z**ΔM15]	Yanisch-Perron *et al.* (1985).
P. aeruginosa		
strain B (PAB)	Alg^+	Constitutive mucoid strain isolated and kindly provided by Professor A. Linker.
PA8873	*algA*, *his*-1	Darzins *et al.* (1984).
PABL	*algL*	Derived from PAB by insertional inactivation of the *algL* gene, this thesis.
PABK	*algK*	Derived from PAB by insertional inactivation of the *algK* gene, this thesis.
PA8873L	*algL*, *his*-1	Derived from PA8873 by insertional inactivation of the *algL* gene, this thesis.
PA8873K	*algL*, *his*-1	Derived from PA8873 by insertional inactivation of the *algK* gene, this thesis.

PLASMIDS	RELEVANT FEATURES	SOURCE/REFERENCE
pAA1	pBluescript derivative, containing region from *alg44* to *algE* on *Sma*I-*Eco*RI fragment.	See figure 4.1, this thesis.
pAA9	pJBS633 derivative, containing subcloned *Sma*I-*Bgl*II fragment of *algK*.	this thesis.
pAA13	pT7-3 derivative, containing subcloned region from *alg44* to *algE* on *Bam*HI-*Hin*dIII fragment from pAA1.	See figure 5.1, this thesis.
pAA15	pT7-5 derivative containing subcloned region from *alg44* to *algE* on *Bam*HI-*Hin*dIII fragment from pAA1.	See figure 5.1, this thesis.
pAA20	pT7-7 derivative containing *Nde*I-*Bam*HI fragment amplified from pAA1 by primers 1 and 2, encoding AlgK.	See figure 5.1, this thesis.
pAA22	pJBS633 derivative containing *Eco*RV-*Bam*HI fragment amplified from pAA20 by primers 3 and 2 encoding AlgK.	this thesis.
pAA23	pMMB66(EH) derivative containing *Eco*RV-*Bam*HI fragment amplified from pAA20 by primers 3 and 2 encoding AlgK.	this thesis.
pAA26	pJBS633 derivative subcloned region from *algE* to *alg44* on *Bam*HI-*Eco*RV fragment from pAA1.	this thesis.
pAA30	pBR325 derivative containing *Pst*I fragment of *algL* amplified from PAB chromosome using primers 5 and 6.	this thesis.

PLASMIDS	RELEVANT FEATURES	SOURCE/REFERENCE
pAA31	pBR325 derivative containing *Pst*I fragment of *algK* amplified from pAA1 using primers 7 and 3.	this thesis.
pAA32	pMMB66(EH) derivative with subcloned *Sma*I-*Hin*dIII fragment of pAD501 containing *algK* and *algE*.	this thesis.
pAA33	pAA32 derivative with subcloned 2-kb *Hin*dIII (*algA*) fragment of pAD4038.	this thesis.
pAA35	pAA22 derivative, expressing AlgK-'BlaM fusion protein.	this thesis (see chapter 6 for further details).
pAD4038	Amp^R, pMMB24 derivative with 2-kb fragment containing *algA*.	Sá-Correia *et al.* (1987).
pAD501	Tet^R, pCP13 derivative containing 24-kb fragment from 34 minute biosynthetic cluster.	Darzins *et al.* (1985).
pBluescript	ColE1, Amp^R, *lacZ*α	Short *et al.* (1988).
pBR325	Amp^R, Cml^R, Tet^R	Bolivar *et al.* (1978).
pCM1	Amp^R, pBR327 derivative with CAT cartridge on *Sal*I fragment.	Close and Rodriguez. (1982).
pJBS633	Tet^R, Kan^R, *'blaM*	Broome-Smith and Spratt, (1986).
pLysS	Cml^R, pACYC184 derivative containing T7 lysozyme gene.	Studier *et al.* (1990).
pMMB66(EH)	Amp^R, *lacI*^{q}, *tac*	Fürste *et al.* (1986).
pRK2013	ColE1, Tra^+-(RK2), Kan^R	Figurski and Helinski, (1979).

PLASMID	RELEVANT FEATURES	SOURCE/REFERENCE
pT7-3	Amp^R, ColE1, p_{T7}	Tabor and Richardson, (1985).
pT7-5	Amp^R, ColE1, p_{T7}	Tabor and Richardson, (1985).
pT7-7	Amp^R, ColE1, p_{T7} and RBS	Tabor and Richardson, (1985).

PHAGE	RELEVANT FEATURES	SOURCE/REFERENCE
mp19	*lacZ*α, *lac*I	Yanisch-Perron *et al.* (1985).
R408	*gene2*	Russel *et al.* (1986).

Table 2.3. Plasmids and phage. Table showing all the plasmids, derivatives and phage mentioned in this thesis. RBS is the equivalent of ribosome binding site and CAT is chloramphenicol acetyl transferase.

2.1.3. **Solutions**

TE buffer contained 10 mM Tris and 1 mM EDTA dissolved in dH_2O, adjusted to the required pH with hydrochloric acid. TE buffer used for DNA contained 50 mM Tris and 10 mM EDTA adjusted to pH 8.0 with HCl.

10 x TBE contained 0.9 M Tris, 0.9 M Boric acid and 0.02 M EDTA dissolved in dH_2O, pH 8.0.

20 x SSC contained 3 M NaCl and 0.3 M tri-sodium citrate dissolved in dH_2O, pH 7.0.

Antibiotic solutions were used as shown in Table 2.4. Aqueous solutions were filter-sterilised before use.

2.1.4. **Molecular weight standards**

The protein molecular weight standards (normal and pre-stained) used during SDS-PAGE analysis of proteins are listed in table 2.5 and table 2.6 below.

ANTIBIOTIC	SOLVENT	STOCK CONCENTRATION	FINAL CONCENTRATION *E. coli*	*P. aeruginosa*
Ampicillin	H_2O	10 mg/ml	50 μg/ml	200 μg/ml
Chloramphenicol	Ethanol	34 mg/ml	5 μg/ml	-
Kanamycin	H_2O	10 mg/ml	50 μg/ml	-
Tetracycline	Ethanol	5 mg/ml	10 μg/ml	200 μg/ml

Table 2.4. Antibiotics. The antibiotic solutions and concentrations used during this thesis.

PROTEIN STANDARD	APPROXIMATE MOLECULAR WEIGHT (Da)
α-Lactoalbumin	14,200
Trypsin inhibitor	20,100
Trypsinogen	24,000
Carbonic Anhydrase	29,000
Glyceraldehyde-Phosphate Dehydrogenase	36,000
Albumin (egg)	45,000
Albumin (bovine)	66,000

Table 2.5. Molecular weight markers (MW-SDS-70L, Sigma) used as a mixture of standards for SDS-PAGE.

PROTEIN STANDARD	NATIVE PROTEIN MOLECULAR WEIGHT (Da)	APPARENT MOLECULAR WEIGHT (Da)
α_2-Macroglobulin (human plasma)	180,000	191,000
β-Galactosidase (*E. coli*)	116,000	117,000
Fructose-6-phosphate (rabbit muscle)	84,000	91,800
Pyruvate kinase (chicken muscle)	58,000	72,700
Fumarase (porcine heart)	48,500	57,800
Lactic dehydrogenase (rabbit muscle)	36,500	40,800
Triosephosphate isomerase (rabbit muscle)	26,600	34,100

Table 2.6. Prestained protein molecular weight markers (MS-SDS-BLUE kit, Sigma). Table showing the native molecular weights and the apparent stained molecular weights of prestained markers as determined by the manufacturers on a 9.5% SDS-PAGE Laemmli gel (Laemmli, 1970).

2.2. MEDIA

Luria-Bertani (LB) medium and agar:

Difco bacto-tryptone (10 g), Difco yeast extract (5 g), NaCl (10 g), were dissolved in 1 litre of sterile dH_2O. The pH was adjusted as necessary to pH 7.2 with 5N NaOH and autoclaved at 20 lb/in^2 (190 kPa) at 120°C for 15 minutes. LB agar was formed by the addition of 15 g/l difco bacto-agar.

LB alginate agar:

LB agar as previous but supplemented with 0.1% Difco alginate and 1% agarose (Northumbria Biologicals limited).

MAP medium:

100 mM monosodium glutamate, 7.5 mM NaH_2PO_4, 16.8 mM K_2HPO_4, 10 mM $MgSO_4$ with 1 x M9 medium and glucose (0.4%). MAP agar was formed by the addition of 15 g/l Difco bacto-agar.

SOC medium:

Difco bacto-tryptone (20 g), Difco yeast extract (5g), NaCl (10 mM), KCl (2.5 mM); $MgCl_2$ (10 mM), $MgSO_4$ (10 mM), glucose (20 mM), were dissolved in 1 litre of sterile dH_2O. The medium was then autoclaved at 20 lb/in^2 (190 kPa) at 120°C for 15 minutes.

Minimal Agar:

Difco bacto-agar (15 g/l) was added to 300 ml distilled water and autoclaved. The following filter sterilised solutions were then added: 100 ml 4 x M9 salts (Miller, 1972); 4 ml 20% (w/v) glycerol; 1 ml 1 M $MgSO_4$ solution; 50 µl 1 M $CaCl_2$

solution. Minimal agar was supplemented with 0.4% glucose and 0.04% thiamine, for *E. coli* strains or alternatively with 0.4% glucose and 0.04% histidine for *P. aeruginosa* strains.

TY Broth:

Difco bacto-tryptone (5 g), Difco yeast extract (3 g) and $CaCl_2.2H_2O$ (1.03 g), dissolved in 1 litre dH_2O. The pH was adjusted to 7.0 and the broth autoclaved at 20 lb/in^2 (190 kPa) at 120°C for 15 minutes.

LC Top Agar:

7 g of Difco bacto-agar, Difco bacto-tryptone (10 g), Difco yeast extract (5 g) and NaCl (5 g) was added and dissolved in 950 ml of distilled deionised water (ddH_2O). The pH was adjusted to 7.0 and the volume was adjusted to 1 litre. The agar was sterilised by autoclaving at 20 lb/in^2 (190 kPa) at 120°C for 20 minutes.

2.3. METHODS

2.3.1. Manipulation of bacteria and phage

Growth of bacterial cultures

E. coli and *P. aeruginosa* cultures.

Liquid cultures were prepared by inoculating a single colony into 5 ml Luria-Bertani (LB) broth, followed by growth with shaking overnight at 37°C at 300 rpm. Where larger cultures were required, bacteria were grown as 5 ml pre-cultures as above and diluted 1/100 in fresh sterile LB in conical flasks, with a total capacity 5-10 times that of the final culture volume.

Storage of bacterial cultures

For long-term storage of bacterial cultures, cells were grown to stationary phase in a nutrient-rich broth, with the addition of antibiotic where required. 1 ml of the culture was removed, transferred to an Eppendorf tube (1.5 ml) and the cells were pelleted at 20,000*g*. The supernatant was removed and the pellet was resuspended in a further 1 ml of nutrient-rich broth, mixed with 70 µl DMSO (100%) and transferred to sterile glass bottles and maintained at -80°C. Bacteria were recovered by removing a small amount of the frozen cells with a sterile tooth-pick and placing the inoculum onto an appropriate agar plate, containing antibiotic where necessary. After overnight incubation at the correct temperature a single colony was picked and a fresh culture was propagated. Streak plates were stored at 4°C for 4-6 weeks.

2.3.2. **Nucleic acid manipulation and detection methods**

Agarose gel electrophoresis

DNA was resolved in electrophoresis grade agarose gels (0.7% w/v in 1 x TBE unless otherwise stated) containing 0.5 µg/ml ethidium bromide. DNA samples were mixed with 1/6th volumes of 6 x loading buffer (shown below). Generally, electrophoresis was carried out at 100 V for 1 hour in a 110 mm x 150 mm horizontal gel unless otherwise stated. Phage Lambda DNA was purchased, pre-digested with *Hin*dIII (from Boehringer Mannheim), and run on agarose gels as size markers. DNA was visualised under UV light (λ=313 nm) and photographed using a Mitsubishi video copy processor.

Loading buffer (6 x)

40% (w/v) sucrose

0.25% (w/v) Bromophenol blue

in dH_20

Preparation of dialysis tubing

The tubing was cut to convenient lengths then boiled for 10 minutes in a large volume of 2% (w/v) sodium bicarbonate solution, containing 1 mM EDTA (pH 8.0). The tubing was then rinsed thoroughly with ddH_2O and autoclaved in a loosely capped bottle containing fresh ddH_2O at 20 lb/in^2 at 120°C for 15 minutes. Thereafter the dialysis tubing was stored at 4°C.

Equilibration of phenol

Phenol was purchased from Rathburn Chemicals. To one volume of phenol, an equal volume of 0.5 M Tris/HCl pH 8.0 was added, with 8-hydroxyquinoline for a final concentration of 0.1% (w/v), followed by stirring for 15 minutes. After the two phases had separated, the upper aqueous phase was removed and replaced with an equal volume of 0.1 M Tris/HCl, pH 8.0, with stirring for a further 15 minutes. The upper phase was removed as before and extractions were repeated until the pH of the phenolic phase was greater than 7.8. The equilibrated phenol was then stored under 0.1 volumes of 0.1 M Tris/HCl pH 8.0 in a light-tight bottle at -20°C.

Phenol-chloroform extraction of DNA

One volume of equilibrated phenol was added to the DNA then mixed thoroughly by vortexing. The tube was then centrifuged at 20,000*g* for 2 minutes and the top aqueous layer removed to a fresh tube. One volume of chloroform was added to the decanted supernatant which was then vortexed and centrifuged at 20,000*g* for 2 minutes. The top layer was again transferred to a fresh tube and one volume of chloroform added, vortexed, centrifuged for 2 minutes at 20,000*g*. Finally, the top layer was removed leaving a trace amount of DNA extract to prevent carry-over of chloroform.

Ethanol-precipitation of DNA

One-tenth volume of 3M sodium acetate (pH 4.8) was added per volume of DNA followed by 2.5 (total) volumes of absolute ethanol. After thorough mixing the tube was placed at -80°C for 20 minutes, followed by centrifugation for 10-12 minutes at 20,000*g*. The pellet was then washed with 1 volume 70% (v/v) ethanol, dried and resuspended in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0).

Restriction digests of DNA

In all cases enzymes and buffers were supplied by Boehringer Mannheim. 0.1-20 μg of DNA was digested in a volume of 20-200 μl containing 1/10th volume of 10 x enzyme buffer with the volumes adjusted using dH_2O. Digestion was allowed to proceed for 1-2 hours at 37°C and chromosomal DNA was incubated overnight at 37°C. Where digestion with two enzymes was required, the DNA was digested first in the lower salt buffer, then the second enzyme was added in the appropriate buffer.

Ligation of DNA

All ligation reactions were performed using T4 DNA ligase and 10 x ligase buffer (supplied by Boehringer Mannheim). 50-100 ng of vector DNA (suitably digested) was incubated with a 5-10 times excess of DNA to be inserted in 1 x ligation buffer with 10 units of T4 DNA ligase. The reaction was made up to a total volume of 10-20 μl with dH_2O and incubated overnight at 12°C.

Dephosphorylation of vector DNA

In order to increase the efficiency of ligation, by preventing vector DNA religation (after digestion with restriction enzymes), vector DNA was dephosphorylated using calf intestinal phosphatase (Boehringer Mannheim). The cut

DNA was incubated at 37°C for 1 hour in 1 x ligation mix with 1 unit of calf intestinal phosphatase (CIP). The CIP was deactivated by incubation at 65°C for 10 minutes in the presence in 1/10th volume ethylene glycol-bis(β-aminoethyl ether) N, N, N', N',-tetraacetic acid (200 mmol/l) and the DNA was purified by phenol and chloroform extraction followed by ethanol precipitation as previously described. The precipitated DNA was resuspended in TE and stored at -20°C.

Recovery of DNA from agarose gels

The band of interest was excised from an agarose gel and the DNA was extracted using 'Geneclean II ^{TM}' (BIO101). 0.5 volumes of TBE modifier was added with 4.5 volumes of sodium iodide (this assumed that the weight of gel was 1 g/ml). The gel slice was then melted at 45-55°C for 5 minutes. 5 µl of "GLASSMILK" was added to the tube, and the sample was mixed and left on ice for 5 minutes with further mixing every 1-2 minutes to ensure the "GLASSMILK" stayed in suspension. The silica matrix with bound DNA was then pelleted by spinning in a microfuge at 20,000*g* for 5 secs. The supernatant was removed and the pellet was resuspended in 500 µl "NEW-WASH". This procedure was carried out three times before resuspending the pellet in 5 µl TE. The tube was incubated at 45°C for 2-3 minutes to elute the DNA form the "GLASSMILK". The tube was then spun at 20,000*g* for 30 seconds to pellet the "GLASSMILK", and the supernatant containing the eluted DNA was transferred to an Eppendorf tube.

Polymerase Chain Reaction

Polymerase chain reactions (PCR) were carried out in a Techne "Gene E" thermal cycler dry-block. Primers were designed and constructed as required (see table 2.1). Reaction mixes contained approximately 0.1 mg target DNA, 10 µl 10 x reaction buffer (supplied with the Taq DNA polymerase), 8 µl dNTP mix (containing

2.5 mM of each dNTP), 1 µl of each primer (0.1 nmol/µl) and 1 µl Taq DNA polymerase (1000 units/µl) in a volume of 100 µl made up with dH_2O. Sterile mineral oil (60 µl) was added to the surface of the reaction mixture to prevent evaporation during thermal cycling.

Denaturation was carried out at 94°C for 1 minute, followed by annealing at 55°C for 1 minute and elongation at 72°C for 2 minutes. This pattern was repeated for 30 cycles. An additional cycle of 72°C for 10 minutes was finally included to complete the process. PCR products were identified by agarose gel electrophoresis.

Small-scale preparation of *E. coli* plasmid DNA

Plasmid DNA was prepared by the method of He *et al.,* (1990, 1991). A single colony of the plasmid carrying strain was inoculated into 5 ml LB broth with the required antibiotic and grown with shaking overnight at 37°C. The culture was centrifuged at 20,000*g* in Eppendorf tubes and resuspended in 800 µl TELT solution (50 mM Tris/HCl pH 7.5; 62.5 mM EDTA, pH 7.5; 2.5 M LiCl; 0.4% [v/v] Triton X-100). 40 µl of a 40 mg/ml solution of lysozyme in TELT was added, mixed, and the tubes were allowed to stand for 2 minutes before boiling in a water bath for 90-120 seconds. The tubes were cooled on ice for 10 minutes and the cell debris was pelleted by centrifugation at 20,000*g* for 15 minutes. The supernatant was transferred to fresh tubes, 0.6 volumes of isopropanol was added and the tubes were placed on ice. 2 minute incubation on ice, was followed by centrifugation at 20,000*g* for 30 minutes. The pellets were washed in 70% ethanol, desiccated and then resuspended in 100 µl TE.

An identical approach was utilised for the extraction of the replicative (RF) form of M13 mp19 DNA and its derivatives.

Large-scale preparation of *E. coli* plasmid DNA

This protocol is based on the methods described by Birnboim and Doly (1979) and Ish-Horowicz and Burke (1981).

500 ml cultures (containing the appropriate antibiotic where necessary) of bacteria were grown overnight with shaking at 37°C in nutrient rich broth. Cells were pelleted by centrifugation at 5,000*g* for 5 minutes then resuspended in 18 ml of solution I (containing 50 mM glucose, 25 mM Tris/HCl pH 8.0, 10 mM EDTA) followed by 40 ml of solution II (0.2 M NaOH solution containing 1% SDS [w/v]) and placed on ice for 5 minutes. Then 20 ml of cold 5 M potassium acetate, pH 4.8, (solution III) was added with thorough mixing and the samples were left on ice for 15 minutes. The cell lysate was then centrifuged at 6,000*g* for 5 minutes. The supernatant was filtered through sterile non-absorbent cotton wool. The debris-free supernatant was then transferred to a fresh tube and 0.6 volumes of isopropanol was added, mixed and the samples were left at room temperature for 2 minutes. The DNA was pelleted by centrifugation at 6,000*g* for a further 5 minutes, rinsed with 70% (v/v) ethanol then dried under a vacuum. The pellet was then resuspended in 9 ml TE. To the DNA solution 10 g of caesium chloride was then added with 0.25 ml of ethidium bromide (10 mg/ml). Unwanted precipitates were spun from solution at 4,000*g*. The supernatant was transferred to Sorvall TV865 quickseal centrifuge tubes and centrifuged at 200,000*g* for 18 hours at room temperature. The plasmid band was visualised on the CsCl gradient under U.V. light and removed with a sterile needle and syringe. The ethidium bromide-containing DNA was then extracted with an equal volume of TE saturated butan-1-ol until the pink colour had completely disappeared from both phases. The aqueous phase was then dialysed in TE overnight at 4°C with several changes of the buffer, to remove the caesium chloride. The DNA solution was then run on an agarose gel for quantification purposes. If required the dialysate volume was reduced by extracting with butan-2-ol.

Small-scale preparation of plasmid DNA from *P. aeruginosa*

Essentially the method for small scale preparation of plasmid DNA from *P. aeruginosa* follows that of large scale preparation of plasmid DNA from *E. coli* but with smaller volumes.

1 ml of overnight culture was transferred into an Eppendorf tube and the cells were pelleted at 20,000*g* for 30 seconds. The supernatant was removed and the cell pellet was resuspended in 100 µl of ice-cold solution I, 200 µl of solution II was added and mixed thoroughly, before the tube was incubated on ice for 5 minutes. 150 µl of cold solution III was added and mixed, before a further 15 minute incubation on ice. The tube was then spun at 20,000*g* for 1 minute to pellet the cell debris. The supernatant was transferred to a second Eppendorf and mixed with 0.6 volumes of isopropanol before further incubation at room temperature for 2 minutes. The DNA was then pelleted by centrifugation at 20,000*g* for a further minute. The resulting DNA pellet was washed in 70% (v/v) ethanol, dried under vacuum and resuspended in 100 µl TE.

Preparation of chromosomal DNA from bacterial cells

The method used was as described by Silhavy *et al.* (1984). 100 ml of LB broth was inoculated with a single colony of the test strain and grown overnight with shaking at 37°C. Cells were then collected by centrifugation at 4,000*g* for 20 minutes. The pellet was resuspended in 5 ml 50 mM Tris/HCl pH 8.0, 50 mM EDTA and the cell suspension was frozen at -20°C. 0.5 ml of lysozyme solution (10 mg/ml in 0.25 M Tris/HCl, pH 8.0) was added to the frozen cells and the cell suspension was allowed to thaw in a water-bath at room temperature. When the cells had just thawed the tube was then placed on ice for 45 minutes. 1 ml STEP solution (0.5% [w/v] SDS; 50 mM Tris/HCl, pH 7.5; 0.4M EDTA; 1 mg/ml proteinase K) was added with thorough mixing, followed by incubation at 50°C for 60 minutes

with thorough mixing, followed by incubation at 50°C for 60 minutes with occasional gentle mixing. 6 ml of equilibrated phenol was then added to the lysate with gentle but thorough mixing for 5 minutes. The tube was centrifuged at 1,000*g* for 15 minutes and the top aqueous layer was removed to a fresh tube, taking care to exclude the material from the interface. 0.1 volume 3M sodium acetate was then added with gentle mixing, followed by 2 volumes of ethanol and mixed by inversion. The DNA/RNA precipitated as a glob and was spooled out with a glass micropipette. (Excess alcohol was removed by gently rotating the precipitate against the side of the tube). The precipitate was transferred to a fresh tube containing 5 ml 50 mM Tris/HCl pH 7.5 solution containing 1 mM EDTA and 200 μg/ml RNAse A. The precipitate was dissolved with gentle rocking overnight at 4°C. When the DNA had completely dissolved, an equal volume of chloroform was added and the solution was mixed gently followed by centrifugation at 1,000*g* for 15 minutes. The aqueous top layer was transferred to a fresh tube and 0.1 volumes 3M sodium acetate was added along with 2 (total) volumes of ethanol and the samples were treated as previously. The DNA precipitate was removed and dissolved in 1 ml TE and stored at 4°C.

Small-scale preparation of M13 single-stranded DNA

A single plaque bearing JM101, infected with M13 mp19, or a derivative was inoculated into 3 ml of TY broth together with a sample of uninfected JM101 and incubated with vigorous aeration overnight at 37°C. After overnight growth, the culture (1.5 ml) was centrifuged (20,000*g*) in Eppendorf tubes. The supernatant was transferred to fresh tubes, 150 μl of a solution containing 2.5 M NaCl and 20% (w/v) polyethylene glycol (PEG) grade 6000 was added, and the tubes were left at room temperature for 15 minutes. The PEG precipitate was then centrifuged for 5 minutes at 20,000*g* and the pellet was dried under a vacuum and resuspended in 100 μl of TE (10 mM Tris, 1 mM EDTA, pH 8.0). 50 μl of equilibrated phenol was added, mixed,

then spun at 20,000*g* in a microfuge for 2 minutes. The top aqueous layer was transferred to a fresh tube and extracted twice with chloroform. The final aqueous layer was then treated with 0.1 volumes of 3M sodium acetate buffer (pH 4.8) and 2 (total) volumes of ethanol and placed at -70°C for 15 minutes. The DNA was precipitated by centrifugation at 20,000*g* for 10 minutes and the supernatant was removed. The pellet was dried and then resuspended in 20 µl TE.

Small-scale single-stranded DNA preparation from pJBS633-derived vectors

A single colony of *E. coli* strain JM101, carrying plasmid pJBS633 (tetracycline/kanamycin resistant pBR322 derivative), was inoculated into 3 ml TY broth containing 25 mg/ml kanamycin together with 10^7 pfu/ml of R408 helper phage (a modification of the methods described by Broome-Smith and Spratt, (1986). The procedure then followed that of M13 mp19 single stranded DNA extraction.

One-step transformation of *E. coli* (using TSS)

This method is essentially the same as that described by Chung *et al.* (1989). An overnight culture of the strain to be transformed was grown with shaking in 5 ml sterile LB broth at 37°C. 5 ml of fresh LB broth was inoculated with 50 µl of the overnight culture and replaced at 37°C with shaking for approximately 2 hours (till the OD reached 0.4-0.6). Cells were then pelleted in sterile Eppendorf tubes at 20,000*g* for 2 minutes and the pellet was resuspended in 100 µl Transformation and Storage Solution (1 ml TSS contains: 0.5 ml LB broth, 0.5 ml 20% PEG 6000, 50 µl DMSO and 45 µl 1 M $MgSO_4$). The DNA to be transformed was then added to the cells and the tube was placed on ice for 1 hour. The cells were then heat-shocked at 37°C for 2-3 minutes in a water bath, then placed at 37°C and grown in LB with shaking for approximately 45 minutes. Thereafter, the cells were plated onto selective medium.

Electroporation of *E. coli* and *P. aeruginosa* cells with plasmid DNA.

Cells for electroporation were prepared as follows. A single colony of the strain to be transformed was inoculated into 5 ml LB broth and grown with shaking overnight at 37°C. 2.5 ml of this overnight culture was inoculated into 500 ml fresh LB broth and grown with shaking (300 rpm) to an optical density (λ=600 nm) of 0.5 to 0.7. The cells were then chilled on ice for 10 to 15 minutes then transferred to pre-chilled centrifuge pots and centrifuged at 2,000*g* for 20 minutes. The supernatant was poured off and the cells initially were resuspended in 5 ml of an ice-cold sterile solution of 1 mM HEPES ([N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulphonic acid]), pH 7.0, with a further 500 ml of solution added subsequently. The cells were centrifuged as before and the supernatant was poured off, resuspending the cells in the residual liquid. A further 500 ml of ice-cold water was added and the cells re-centrifuged under the same conditions. The supernatant was removed and the pellet was resuspended in the remaining liquid. 40 ml of an ice-cold sterile 10% (v/v) glycerol solution was added and the cells were centrifuged for 10 minutes at 2,000*g* at 4°C. The pellet volume was then estimated visually and the cells resuspended in an equal volume of ice-cold sterile 10% (v/v) glycerol solution with thorough mixing. Cells were aliquoted in 100 µl portions at a density of approx. 2 x 10^{11} cells/ml and stored at -80°C.

A Bio-Rad "Gene Pulser" apparatus was used for electroporation, with the voltage set at 2.5 kV and the Pulse controller set to 200 Ohms. 5 pg to 0.5 µg of DNA in 1 µl dH_2O was added to the tube containing freshly thawed cells on ice. The contents of the tube were mixed gently and transferred to an electroporation cuvette, pre-chilled on ice for 5 minutes. Ensuring even distribution of the cells across the bottom of the cuvette, the cuvette was then placed in the sample chamber. The pulse was then applied and the cuvette was removed and 1 ml SOC medium was added

immediately. The contents of the cuvette were then transferred to a sterile microfuge tube and incubated in LB with moderate shaking at 37°C for 30-60 minutes. Cells were then diluted in sterile LB broth where necessary and plated onto selective medium.

Pseudomonas conjugation

Triparental filter matings were carried out as described by Sá-Correia *et al.* (1987). 1 ml from overnight cultures of each of an *E. coli* donor strain, an *E. coli* strain bearing the pRK2013 helper plasmid and a recipient *P. aeruginosa* strain were spun at 20,000*g* for 30 seconds. The pelleted cells were resuspended in a further 1 ml of LB broth, mixed and filtered onto a sterile 0.45 μm Whatman filter (25 mm diameter). The filter was placed onto a LB agar plate and incubated for 5-8 hours at 37°C, the cells were then washed off with sterile 0.9% NaCl (saline) and plated as serial dilutions onto minimal agar plates supplemented with histidine, glucose and the appropriate selective antibiotics.

Propagation of M13 mp19 and derived phage

mp19 DNA (either derived using the TELT miniprep method or from ligation mixes, see 'ligation of DNA' above) was transfected into JM101 cells using the TSS method described. After incubating on ice for 45 minutes the contents of the Eppendorf tubes were added and mixed with 2 ml of melted LC top agar (at 46°C) containing 200 μl of an overnight of JM101, 40 μl of 2% X-gal (w/v dissolved in dimethylformamide) and 40 μl of 2% (w/v) IPTG (dissolved in H_2O). This mix was then poured onto on LB agar plate, allowed to set and incubated overnight at 37°C. Under these conditions, mp19 infected JM101 cells appear as blue plaques on a JM101 lawn, whilst mp19 with a DNA insert give clear plaques.

Blue/white screen for pBluescript derived phagemids

Transformation of JM101 cells with phagemid by either TSS or electroporation methods was described previously. Visualisation of recombinant plasmids is essentially as described by Short *et al*, (1988) and is by selection on LB agar containing 50 μg/ml ampicillin, combined with the blue/white screen provided by the addition of X-Gal and IPTG at the concentration described above. JM101 colonies harbouring pBluescript appear as blue colonies, whilst JM101 colonies bearing recombinant pBluescript appear as white or pale blue colonies.

Southern* Blotting and hybridisation

The gel (110 mm by 150 mm) to be blotted was electrophoresed for 12-15 hours at 1 V/cm. The gel was then soaked in 0.25 M HCl for 30 minutes with gentle agitation. Having removed the acid, a solution of 0.5 M NaOH containing 1.5 M NaCl was added and gentle shaking was continued for 40 minutes. The gel was then rinsed with sterile distilled water and soaked in 0.5 M Tris.HCl (pH 7.4), containing 1.5 M NaCl, for a further 40 minutes. Then a sheet of blotting paper, soaked in 20 x SSC (3 M NaCl, 0.3 M sodium citrate dissolved in H_2O), was placed on an upturned gel casting tray with its ends resting in a reservoir of 20 x SSC. The gel was then placed on top of this tray (well-side down) ensuring no air bubbles were trapped underneath. Strips of clingfilm were placed around the gel to prevent buffer passing directly from the wick to the blotting paper which eventually was to be placed on top of the gel, as this would short-circuit the system. A piece of nylon membrane (Hybond™-N, Amersham), cut to the same size as the gel, was soaked in 2 x SSC and then put directly on top of the gel, again expelling any air bubbles. A stack of blotting paper (0.5 cm), cut to approximately the same size as the gel, was placed on top of the nylon membrane, followed by a stack of paper towels (8 cm) to draw the buffer. Finally, a weight (approximately 500 g) was placed on top of the stack to

ensure good contact and the apparatus was left overnight at room temperature. The next day the nylon membrane was removed and rinsed in 2 x SSC, then dried in a vacuum oven at 80°C for 2 hours. The membrane was U.V. cross-linked for 2 minutes in a UV Stratalinker™1800 (Stratagene) prior to further treatment.

* Southern, 1975.

Random primed end-labelling of DNA

DNA was labelled using the Boehringer Mannheim Random Primed DNA Labelling kit.

50-100 ng of the DNA (0.2-kb upwards) to be labelled was denatured by heating for 10 minutes at 100°C with immediate cooling on ice thereafter. The denatured DNA was added to a sterile Eppendorf tube containing 1 µl each of dATP, dGTP, and dTTP, each at 0.5 mmol/l in unspecified Tris buffer. 2 µl of 10 x Reaction buffer (hexanucleotide mix in concentrated reaction buffer) was added and the volume was adjusted to 19 µl following the addition of 5 µl [α^{32}P]dCTP (3000Ci/mmol, 10 µCi/µl). 1 µl Klenow enzyme was added and the reaction mix was incubated at 37°C for 1 hour. The reaction was stopped by the addition of 2 µl of 200 mM EDTA, pH 8.0 and the volume was adjusted to 200 µl with TE. Non-incorporated radio-isotope was removed by passing the reaction mix through a 1 ml TE-equilibrated, pre-spun Sephadex G-50 column in a centrifuge at 2,000*g* for 4 minutes. Labelled DNA was eluted and unincorporated radio-isotope was retained on the column.

DNA hybridisation

Hybridisations were performed in a 'Techne Hybridiser HB-1D' oven and were carried out at 65°C (high stringency) or at 43°C with formamide (low stringency).

High stringency protocol:

65°C pre-hybridisation solution:

1 mM EDTA

0.5 M NaH_2PO_4, pH 7.2

7% SDS (w/v)

The membrane was placed in a roller bottle, covered by a pre-hybridisation solution (volume of pre-hybridisation used is 1/10th the surface area of the membrane for sufficient blocking of the membrane) and incubated in a Techne oven for 5 minutes at 65°C. Meanwhile, the labelled DNA (probe) was denatured by boiling for 10 minutes and then was added to the membrane with fresh buffer (same volume as above). Hybridisation was allowed to proceed overnight at 65°C. The membrane was then washed twice in a solution containing 1 mM EDTA, 40 mM NaH_2PO_4, pH 7.2 and 5% SDS (w/v) followed by two further washes containing 1 mM EDTA, 40 mM NaH_2PO_4, pH 7.2 and 1% SDS (w/v) for 30 minutes, at 65°C. The membrane was then wrapped in clingfilm and autoradiographed at -80°C.

Low stringency protocol:

Formamide pre-hybridisation solution:

50% formamide (v/v)

0.25 M NaH_2PO_4, pH 7.2

0.25 M NaCl

7% SDS (w/v)

1 mM EDTA

The hybridisation procedure follows that of the high stringency protocol except incubation was at 43°C. Overnight hybridisation was followed by three successive 15 minute washes in the solutions described below:

i) 2 x SSC/0.1% SDS

ii) 0.5 x SSC/0.1% SDS

iii) 0.1 x SSC/0.1% SDS

Sequencing of single stranded DNA by the dideoxynucleotide method

Sequencing reactions were performed following the Sequenase Version 2.0 (United States Biochemical) protocol. Sequencing gels were run using a 48 well (38 cm x 50 cm) Sequi-Gen Nucleic Acid Sequencing Cell apparatus (Bio-Rad).

Single stranded DNA was prepared as previously described. To 7 µl of single stranded DNA (containing approximately 1-2 µg DNA dissolved in dH_2O), 2 µl of 5 x sequenase reaction buffer and a further 1 pmol of primer DNA (1 µl) were added. The annealing reaction was performed at 65°C for 2 minutes and then allowed to cool to room temperature over 30 minutes before placing on ice. Extension was carried out by the addition of 1 µl of 0.1 M DTT, 2 µl of a 1:4 dilution of 7-deaza dGTP labelling mix (containing 7.5 µM 7-deaza dGTP, as well as dCTP and dTTP), 1 µl of [α-^{35}S] dATP (400Ci/mmol, 10 µCi/µl) and 2 µl of Sequenase enzyme diluted 1:8 in sequenase dilution buffer. The extension mix was left at room temperature for 0.5 minutes then 3.5 µl of the mix was added to four tubes, pre-incubated at 37°C, containing 2.5 µl of the following:

7-deaza ddGTP mix (80 µM dNTPs*; 8 µM ddGTP),

7-deaza ddATP mix (" " " ddATP),

7-deaza ddTTP mix (" " " ddTTP),

7-deaza ddCTP mix (" " " ddCTP),

dNTPs*; 80 µM each 7-deaza-dGTP, dATP, dTTP and dCTP.

Termination was performed after incubation at 37°C for 3 minutes with the addition of 4 µl of stop solution containing 95% (v/v) formamide, 20 mM EDTA,

0.05% (w/v) Bromophenol blue and 0.05% (w/v) Xylene cyanol FF. Samples were then immediately placed on ice.

The resultant products were separated on a 6% denaturing polyacrylamide gel of the following composition: 71.4 g urea, 17 ml 10 x TBE, 40 ml 30%(w/v) acrylamide: bisacrylamide (19:1) mixture, with the volume was adjusted to 170 ml with dH_2O. Initially 40 ml was removed and used in the casting tray in order to seal the gel apparatus, the solution was polymerised by the addition of 200 µl of fresh 20% (w/v) ammonium persulphate and 200 µl of TEMED (N, N, N', N',-Tetramethylethlenediamine). The sequencing gel was polymerised using 120 µl (w/v) ammonium persulphate solution and 120 µl TEMED. Samples were boiled for 2 minutes immediately prior to loading on the gel and electrophoresis was carried out at 60 mA in 1 x TBE buffer for 2.5-7.5 hours. The gel was then dried under vacuum for two hours at 80°C, prior to autoradiography overnight at room temperature.

2.3.3. **Protein expression**

Overexpression using pT7 vectors in *E. coli*

This method is based on that described by Sambrook *et al.* (1989) with precautions recommended by Studier and Moffat (1986) for protein expression in pT7-vectors.

5 ml of LB was inoculated with a single colony of the *E. coli* BL21 (DE3) pLysS strain (bearing the required plasmid to be analysed) and grown overnight in the correct selective antibiotics. 50 µl of the overnight culture was then inoculated into a further 5 ml LB medium containing 100 µg/ml ampicillin (maintaining selection for the pT7 plasmid) and 5 µg/ml chloramphenicol (maintaining the selection for the pLysS plasmid) and incubated for a further 2 hours at 37°C with shaking. 1 ml of uninduced culture was transferred into an Eppendorf tube and centrifuged at 20,000*g* for 1 minute at room temperature. The supernatant was

discarded and the cells resuspended in 100 µl of 1 x loading sample buffer. The remaining culture was induced by the addition of isopropylthio-β-D-galactoside (IPTG) to a final concentration of 1 mM. 1 ml aliquots of induced culture were removed at hourly intervals and transferred to Eppendorf tubes and processed as described previously for the uninduced sample. The aliquots were denatured at 100°C for 3 minutes and stored at 0°C before use. Prior to loading the samples were spun at 20,000*g* for 1 minute and 10 µl of each suspension was loaded onto a 10% SDS-polyacrylamide gel.

Overexpression using pMMB66(EH) derived vectors in *E. coli* and PA8873

Essentially a method identical to the one above for the pT7 vectors was employed. However the *E. coli* strain used in expression of pMMB66 derived vectors was HB101 and the *P. aeruginosa* strain was PA8873, neither of which requires the addition of chloramphenicol.

[^{35}S]Methionine labelling of proteins using pT7 vectors in *E. coli*

This procedure essentially followed that of Franklin *et al.* (1994). 5 ml cultures of BL21(DE3)pLysS bearing the pT7-7 vector were grown to mid-log phase at 37°C in LB and a 1 ml aliquot was transferred to an Eppendorf tube where cells were collected by centrifugation at 20,000*g*. The cells were washed twice in M9 medium and resuspended in 1 ml of M9 medium supplemented with 0.4% glucose, 0.4% thiamine, the common 18 amino acids (except methionine and cysteine) at 0.005% v/v final concentration and appropriate antibiotics. Cells were incubated for a further hour (to an OD_{600} of 0.6-0.8) and then IPTG was added to a final concentration of 1 mM to permit transcription from the T7 RNA polymerase gene. The incubation was either continued for a further 45 minutes or alternatively, after 15 minutes rifampicin was added to a final concentration of 200 µg/ml and the

incubation was continued for another 30 minutes at 37°C. The [^{35}S]methionine (10 μCi, 800 Ci/mmol) was added and the cells were incubated at 37°C for a further 5, 10 or 20 minutes before harvesting. These samples were centrifuged at 20,000*g* for 30 seconds and the supernatant was discarded. The cells were resuspended in 100 μl LSB.

[^{35}S]Methionine pulse-chase labelling of proteins in *P. aeruginosa*

This procedure essentially follows that of Franklin *et al.* (1994). Bacteria were grown overnight in LB broth at 37°C. Cells from 0.75 ml of culture were collected by centrifugation (5000*g*, 10 minutes), washed with M9 medium, resuspended in 3 ml of M9 medium supplemented with glucose (0.4%) and incubated at 37°C for 2 hours (to an OD_{600} of 0.6-0.8). IPTG was added (to a final concentration of a 1 mM) and incubation continued for a further 30 minutes. [^{35}S]Methionine (10 μCi, 800 Ci/mmol) was added and cells were incubated at 37°C for 2 minutes. Unlabelled L-methionine (0.1% final concentration) was added, and then samples (0.75 ml) were withdrawn at 0, 3, 15 and 30 minutes after the chase began. Samples were centrifuged at 20,000*g* for 30 seconds and the supernatant was discarded, the cells were resuspended in 100 μl LSB.

Preparation of SDS-polyacrylamide gels

Proteins were resolved with a (20 cm x 20 cm) PROTEAN II xi Slab Cell (Biorad) by one dimensional electrophoresis on a 10% polyacrylamide resolving gel preceded by a 4% polyacrylamide stacking gel prepared as follows:

Resolving gel (10%)

30% (w/v) acrylamide/bisacrylamide (29:1)	33.3 ml
distilled water	40.2 ml

1.5 M Tris/HCl pH 8.8	25.0 ml
10% (w/v) SDS	1.0 ml
10% ammonium persulphate (fresh)	500 µl
TEMED	50 µl

Stacking gel (4%)

30% (w/v) acrylamide/bisacrylamide (29:1)	6.5 ml
distilled water	30.5 ml
0.5 M Tris/HCl pH 6.8	12.5 ml
10% (w/v) SDS	500 µl
10% ammonium persulphate (fresh)	250 µl
TEMED	50 µl

Laemmli Sample Buffer (10 x) for SDS-polyacrylamide gels

25 mM Tris/HCl pH 6.8

2% (w/v) SDS

10% (v/v) glycerol

0.2 M β–mercaptoethanol

0.002% (w/v) Bromophenol Blue

5 x Running buffer for SDS-polyacrylamide gels

144 g glycine, 30 g Tris-base, 5 g SDS in 1 litre ddH_2O.

The resolving gel was poured between two glass plates and overlaid with H_2O during polymerisation. Once the gel had set the H_2O was poured off and the stacking gel (approximately 15 ml) was poured onto the resolving gel. Samples (10-20 µl) in

loading sample buffer (LSB) were loaded into individual wells and the gel was run for 18 hours at 100 V.

Once the gel had been removed from the electrophoresis apparatus it was stained with Coomassie Brilliant Blue R250 (Biorad) staining solution (0.5g of Coomassie Brilliant Blue R250 dissolved in 1 litre solution of ddH_2O: methanol: glacial acetic acid, (45: 45: 10), previously filtered through a Whatman No.1 filter. Staining was carried out on a shaking platform for 30-60 minutes. The stain was removed and the gel was destained in several changes of ddH_2O: methanol: glacial acetic solution (45: 45: 10) over a 5-8 hour period.

2.3.4. **Construction of fusion proteins**

Unidirectional digestion of DNA using exonuclease III and S1 nuclease

This method was described by Henikoff (1987) and modified for use with the vector pJBS633 as described by Broome-Smith and Spratt (1986).

10 μg of plasmid DNA was completely digested with *Sph*I and *Bam*HI to allow unidirectional digestion of DNA by exonuclease III. The DNA was purified by phenol and chloroform extraction followed by ethanol precipitation. The resultant pellet of DNA was redissolved in 40 μl of 1 x exonuclease III buffer and stored on ice.

7.5 μl of S1 reaction mix was added to 16 Eppendorfs tubes and stored on ice. The plasmid (in exonuclease III buffer) was incubated at 37°C for 5 minutes before 2.5 μl was removed and transferred to the first of the 16 S1 containing Eppendorf tubes. 300-400 units of exonuclease III was added and mixed with the remaining 37.5 μl of the DNA solution and returned to the 37°C water bath. At 30 second intervals, 2.5 μl aliquots of the DNA/exonuclease III solution were transferred to successive Eppendorf tubes containing the S1 nuclease reaction mix (the acidic

nature of the S1 buffer inactivates exonuclease III). After all the samples were taken the tubes were removed from the ice and incubated at 30°C for 30 minutes to allow S1 nuclease activity. After this period 1 µl of S1 stop mix was added to each tube and all were incubated for a further 10 minutes at 70°C to deactivate the S1 nuclease. 1 µl of Klenow mix was added to each 10 µl pooled sample and the tubes were incubated for 5 minute at 37°C. Incubation in the absence of dNTP's results in the 3'-5' exonuclease activity of the Klenow enzyme removing any protruding 3' termini. 1 µl of a dNTP solution (containing each dNTP at a concentration of 5 mM) was added and the incubation was continued at room temperature for a further 15 minutes.

The samples were then phenol and chloroform extracted, ethanol precipitated, resuspended in TE and digested with *Sma*I. The DNA was then religated overnight using T4 DNA ligase (see 'ligation of DNA')

The solutions used in this method are listed below.

10 x Exonuclease buffer mixture

0.66M Tris/HCl pH 8.0, containing 66 mM $MgCl_2$

10 x S1 buffer

5 M NaCl	5.0 ml
3M potassium acetate (pH 4.5)	1.1 ml
glycerol	5.0 ml
1 M $ZnSO_4$	20 µl

S1 reaction mix contains:

dH_2O	172 µl
10 x S1 buffer	27 µl
S1 nuclease	60 units

S1 stop mix contains:

0.3 M Tris base, 50 mM EDTA (pH 8.0).

Klenow mix contains:

dH_2O	20 µl
1 M $MgCl_2$	6 µl
0.1 M Tris/HCl (pH 7.6)	3 µl
Klenow enzyme	3 units

Preparation of cells for Western blotting

1 ml of an overnight culture of JM101 bearing the required plasmid was transferred to a microfuge tube and the cells were pelleted at 20,000*g* for 1 minute. The supernatant was removed by aspiration and the pellet was resuspended in 100 µl LSB. 10 µl of each suspension was loaded into wells and the gels were subjected to SDS-PAGE for 18 hours at 100 V as described above. The samples were then transferred from the gel to a nitrocellulose membrane (Hybond™-C super; Amersham) and probed with an appropriate antibody as described below.

i) **Electrophoretic transfer of protein to membranes;**

The method used for transferring proteins to nitrocellulose membranes from acrylamide gels was as suggested by the manufacturers of the apparatus utilised and essentially followed that of Towbin *et al.* (1979). Electrophoretic transfer of protein to the membrane was performed using a Biorad 'TRANS-BLOT™ CELL'.

After electrophoresis the polyacrylamide gel was placed on a piece of 3MM Whatman paper which was cut to size and pre-soaked in 1 x transfer buffer (10 x Transfer Buffer contains 250 mM Tris, pH 8.3, with 1.5 M glycine and 20%

methanol) on a transfer buffer soaked filter pad. A pre-wetted transfer sheet of nitrocellulose Hybond™-C super (cut to the appropriate size) was placed on top of the gel, taking care to avoid any air bubbles. The sandwich was completed by the addition of a second layer of soaked and trimmed filter paper, to the top of the membrane, followed by a further soaked filter pad. The entire sandwich was placed in the gel holder prior to being immersed in transfer buffer within the gel tank. Transfer from gel to membrane was achieved by overnight electrophoresis at 200 mA at 4°C. After transfer the apparatus was dismantled and the membrane was used immediately for immunodetection.

ii) **Immunodetection of β-lactamase fusions;**

The detection of nitrocellulose bound antisera was by alkaline phosphatase staining and was carried out essentially as described by Knecht and Dimond (1984). After electrophoresis the membrane was placed in 100 ml of 20% w/v milk powder solution (milk powder dissolved in ddH_2O) and shaken at room temperature for 1 hour to block the unbound sites on the membrane. This solution was removed and replaced with 10 ml of Tween wash containing a 1:1000 dilution of rabbit anti-β-lactamase antisera (5 prime to 3 prime INC.). After a further hour of shaking at room temperature this solution was discarded and the membrane was washed six times in 100 ml of Tween wash (pH 7.4) - each wash involved 10 minutes shaking at room temperature. Subsequent to washing, the membrane was then placed in 10 ml Tween wash containing 10 ml (1:1000 dilution) of alkaline phosphatase conjugated, goat anti-rabbit IgG (whole molecule, Sigma). After a further hours agitation, this solution was removed and unbound antisera was removed by washing as previously but with Tween wash at pH 8.8. Bound antisera was visualised by the addition of 10 ml of developing solution. Once bands were visible the phosphatase reaction was halted by rinsing the membrane in copious quantities of dH_2O.

Tween wash: 9 g NaCl, 0.5 ml of polyoxyethylene sorbitan monolaurate (Tween 20) dissolved in 10 ml of 1 M Tris/HCl at pH 7.4 or 8.8 all finally diluted to 1 l with dH_2O.

Developing solution: 1 ml of 0.1% (w/v) nitroblue tetrazolium in 10 mM Tris/HCl pH 8.8; 40 µl 1 M $MgCl_2$, 0.1 ml of 5-bromo-4-chloro-3-indolyl phosphate (p-toluidine salt) in dimethylformamide (5 mg/ml) and 9 ml of 0.5 M Tris/HCl pH 8.8.

2.4. Computer Analysis of Sequence Data

Sequence data was analysed using the University of Wisconsin Genetics Computer Group programme (Devereux *et al.*, 1984; GCG, 1991). The MAP programme was used to produce the amino acid sequences of all the reading frames and restriction enzyme sites in the sequence. Protein and nucleotide sequences were matched using GAP (algorithm of Needleman and Winch, 1970) and aligned using the BESTFIT programme (Smith and Waterman, 1981). Protein comparison searches were performed in the SwissProtein database using the FASTA programme in GCG7 (Pearson and Lipman, 1988).

Chapter 3

Hybridisation survey of the *P. aeruginosa* alginate biosynthetic cluster for a homologue of the ABC transporter superfamily.

3.1. INTRODUCTION

Although no undecaprenol-linked intermediates have been positively identified in the synthesis of bacterial alginate by *A. vinelandii* or *P. aeruginosa* (Sutherland, 1982; May *et al.*, 1991), it is thought that assembly of polymannuronate (the basic alginate structure) requires the sequential addition of mannuronic acid from its activated precursor (GDP-mannuronic acid) on undecaprenol lipids. The proposed involvement of a lipid acceptor in the assembly of polymannuronate provides an obligatory requirement for the cytoplasmic membrane. The polymerisation event for LPS, peptidoglycan and EPS in all Gram-negative bacteria is thought to be at the inner face of the cytoplasmic membrane. The polymerisation of GDP-mannuronic acid to polymannuronic acid in both *A. vinelandii* and *P. aeruginosa* is thought to be analogous to these systems and thus postulated to occur at the cytosolic face of the cytoplasmic membrane. Polymerisation of alginate within the cytoplasm provides an intrinsic requirement for a mechanism to transport the polysaccharide across the inner membrane.

For all sufficiently characterised Gram-negative bacteria, transport of polysaccharide (including LPS, Vazquez *et al.*, 1993) across the inner membrane is performed by putative ATP-dependent export mechanisms. All inner membrane polysaccharide transport systems so far analysed appear to belong to the ATP-Binding Cassette (ABC) transporter superfamily. The ABC superfamily includes both prokaryotic and eukaryotic proteins involved with a variety of ATP-dependant export and import processes. Within this family are the periplasmic binding protein-dependent permeases of enteric bacteria such as Mal, His and Opp which mediate maltose, histidine and oligopeptide transport, respectively P glycoprotein (MDR) responsible for multidrug resistance in mammalian tumour cells and CFTR, which is

defective in CF patients, responsible for cAMP mediated Cl^- transport (for a review see Higgins *et al.*, 1986, 1990). The members of the ABC-transporter superfamily share a common organisational pattern consisting of one or more hydrophobic membrane component and a hydrophilic, ATP-binding component (Higgins *et al.*, 1990). The general organisation of a typical ABC transporter is illustrated in figure 1.5 using the oligopeptide permease of *S. typhimurium* as an example. Two of these domains are highly hydrophobic, integral membrane proteins, each consisting of 5-6 membrane-spanning α-helices, while the other two domains are peripherally located on the cytoplasmic face of the membrane and couple the hydrolysis of ATP to substrate transport across the inner membrane. In the majority of bacterial transporters (such as Opp) these four domains are present as separate polypeptide chains. However, domains are frequently fused into larger multifunctional polypeptides, particularly in eukaryotes.

3.1.1. Polysaccharide ABC exporters

The polysaccharide ABC exporters discovered to date tend to consist of either two polypeptides encoded by two separate genes e.g. *kpsT/M* (Smith *et al.*, 1990, Pavelka *et al.*, 1990), *ctrC/D* (Frosch *et al.*, 1991), *bexA/B* (Kroll *et al.*, 1990), *nodI/J* (Vazquez *et al.*, 1993) or a single polypeptide encoded by a single gene *e.g. ndvA* (Stanfield *et al.*, 1988) and *chvA* (Cangelosi *et al.*, 1989). In the two polypeptide system, one of the polypeptides encodes an highly hydrophobic, inner membrane domain with six α helices whilst the other encodes the cytoplasmic, ATP binding domain, thus each polypeptide subunit is thought to exist as a homodimer. In the case of NdvA and ChvA only one polypeptide is produced, which contains both the ATP binding domain and a highly hydrophobic membrane domain, in this system the polypeptide is again thought to dimerise to give the four domain transporter.

Hydrophobic domains

The archetypal ABC transporter system requires two similar polypeptides or exists as part of a larger multi-domain protein. Comparison of the hydrophobic polypeptides between systems, including those involved in polysaccharide secretion, reveals little overall sequence similarity. However, all are highly hydrophobic and seem to be structurally related. Each protein consists of a core structure of six potential membrane-spanning α-helices separated by short stretches of hydrophilic sequence (Higgins *et al.*, 1990). Hydrophobicity comparisons of the inner membrane polysaccharide exporters KpsM, CtrC, BexC reveal the common secondary structural profile as well as the hydrophobic putative membrane-spanning stretches (see figure 3.1). The complete transporter, thus contains 12 membrane-spanning helices which are thought to be responsible for mediating the translocation of substrate across the lipid bilayer.

ATP-binding domains

These proteins bind ATP and couple ATP hydrolysis to the transport process (Mimmack *et al.*, 1989). For all polysaccharide export systems two domains are required, each possessing an ATP binding site (Higgins *et al.*, 1990). The concept of two ATP binding domains as part of each transport system is also consistent with evidence suggesting that two ATP molecules may be hydrolysed per transport event (Mimmack *et al.*, 1989). In general the ATP binding proteins are hydrophilic and contain no potential membrane-spanning helices, they are assumed to be peripherally associated with the cytoplasmic face of the membrane, compatible with their proposed role in coupling ATP hydrolysis to transport. A cytoplasmic location for the ATP binding proteins has been predicted for the polysaccharide export systems of *E. coli*, *N. meningitidis* and *H. influenzae*. However, NdvA and ChvA exhibit both putative membrane-spanning and ATP binding domains within the same polypeptide.

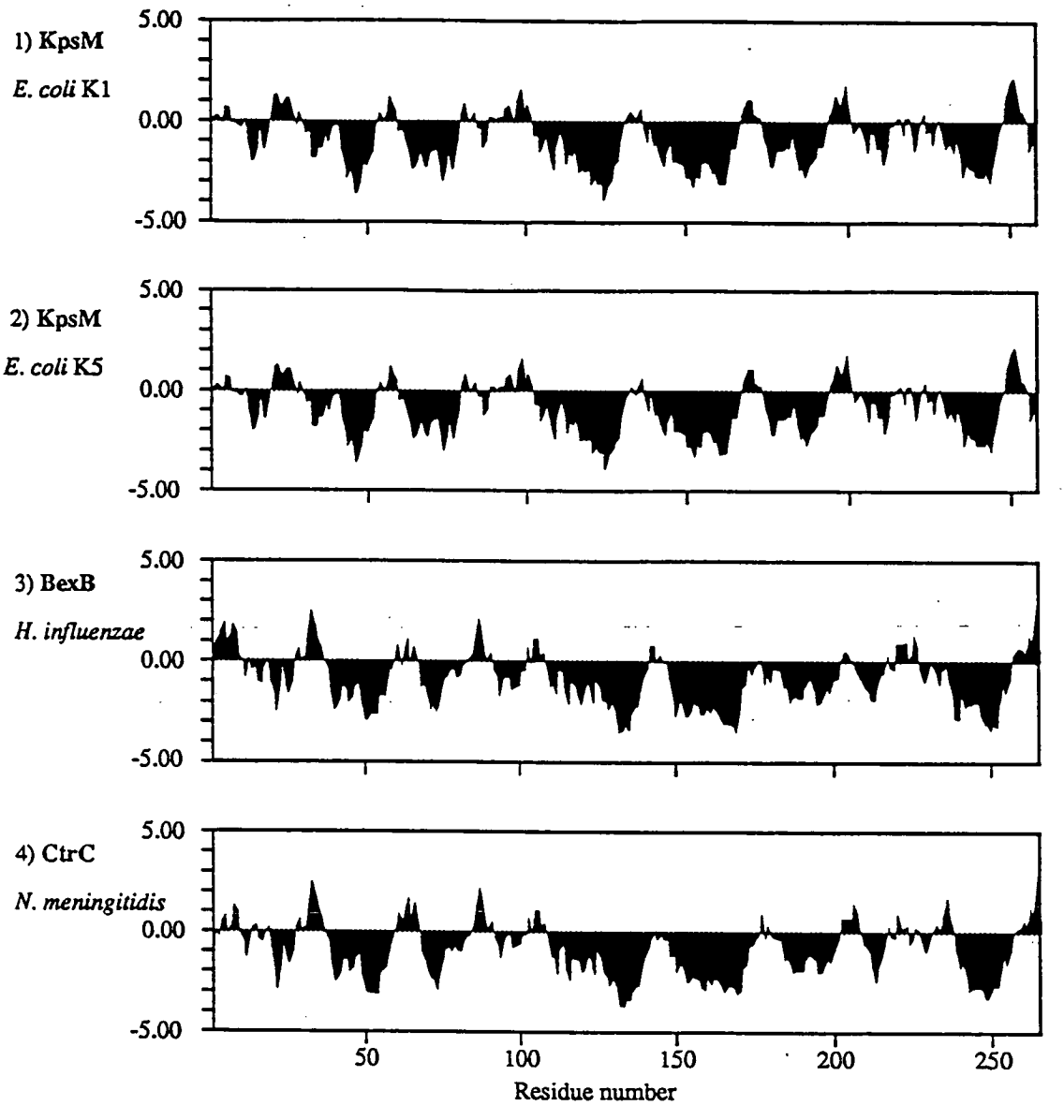


Figure 3.1. Hydropathy profiles of proposed membrane proteins involved in polysaccharide transport systems of Gram-negative bacteria. The profiles were performed according to Kyte and Doolittle (1982) with a window of 11. The amino acid sequence of KpsM (*E. coli* K1), KpsM (*E. coli* K5), BexB, and CtrC were derived from Pavelka *et al.* (1991), Smith *et al.* (1990), Kroll *et al.* (1990) and Frosch *et al.* (1991), respectively.

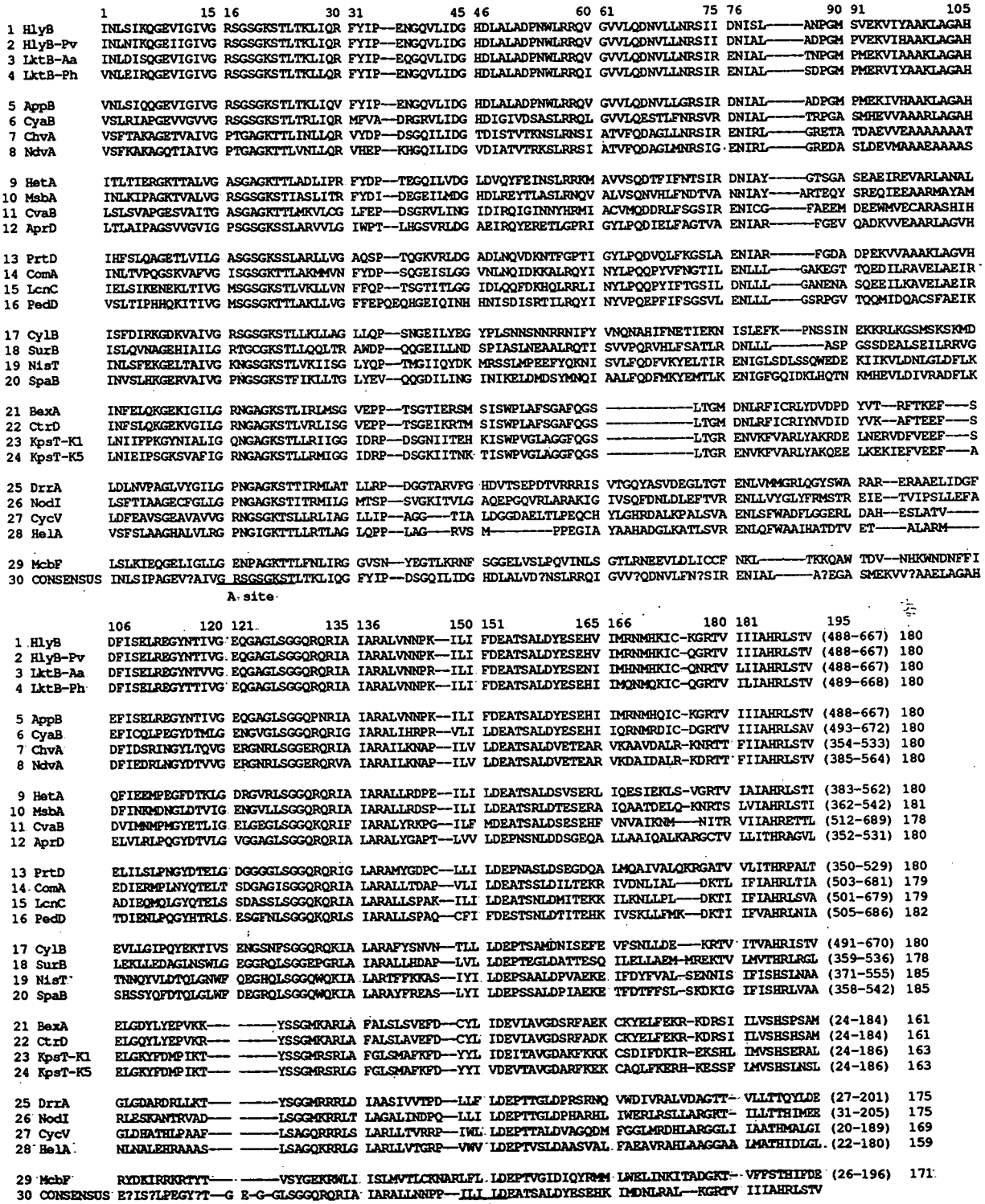
It is assumed however that the ATP binding portion is cytoplasmically orientated (a similar arrangement for the ATP binding domain of the histidine ABC transporter is postulated for the integral membrane protein, HisP (Ames *et al.*, 1986).

Unlike the hydrophobic membrane components, the hydrophilic ATP binding proteins share extensive amino acid sequence similarity. ATP binding proteins share approximately 30% sequence identity over their entire length, regardless of the system with which they are associated (Higgins *et al.*, 1990). Within the polysaccharide inner membrane exporters the similarity is even more marked. For instance KpsT (from *E. coli* K1) shows 84.9% similarity and 72.9% identity to KpsT from *E. coli* K5, 70.7% similarity and 46.7% identity with BexA (*H. influenzae*) and 58.7% similarity and 39.7% identity with that of CtrD (*N. meningitidis*) (Pavelka *et al.*, 1991; Vazquez *et al.*, 1993). This sequence similarity encompasses approximately 190 amino acids, and this region is known as the ATP binding cassette and contains two highly conserved ATP-binding motifs, A and B (see figure 3.2) which form an ATP binding pocket (for a review, see Fath and Kolter., 1993). The ATP binding site occurs at the end of an α helix, and the residues GXXGXGKST form a turn, bringing the lysine residue into close proximity with the phosphates of Mg^{2+}-ATP. The aspartic acid residue in the B site is in close proximity, in the three dimensional structure, to the A site and its negative charge may interact with the Mg^{2+} molecules (Walker *et al.*, 1982).

Location of genes encoding polysaccharide ABC transporters

Biosynthetic genes for EPS are arranged in clusters in most bacteria. Based on the similar arrangement of gene clusters involved in the synthesis of EPS in *E. coli*, *N. meningitidis* B and *H. influenzae*, Frosch *et al.* (1991) predicted a common evolutionary origin existed for the mechanism of capsule production in Gram-negative bacteria which express group II capsular polysaccharides. These gene

Figure 3.2. Sequence alignment and consensus sequence of the approximately 190 amino acid domain containing the ATP binding cassette. The alignment was adapted from Fath and Kolter (1993). The ABC's are involved in the transport of the following 1) alpha haemolysin (HlyB, *E. coli*). 2) alpha haemolysin (HlyB, *Proteus vulgaris*). 3 leukotoxin (LktB, *Actinobacillus actinomycetemcomitans*). 4) leukotoxin (LktB, *Pasturella haemolytica*). 5) haemolysin (AppB, *Actinobacillus pleuropneumoniae*). 6) cyclolysin (CyaB, *Bordetella pertussis*). 7) β-1,2-glucan (ChvA, *A. tumefaciens*). 8) β-1,2-glucan (NdvA, *R. meliloti*). 9) unknown (HetA, *Anaebaena* sp.). 10) unidentified (MsbA, *E. coli*). 11) coliconV (CvaB, *E. coli*). 12) alkaline protease (AprB, *P. aeruginosa*). 13) proteases A, B, C (PrtD, *Erwinia chrysanthemi*). 14) competence factor (ComA, *Streptococcus pneumoniae*). 15) lactococcin A (LcnC, *Lactococcis lactis*). 16) pediocin PA-1 (PedD, *Pediococcus acidilactici*). 17) haemolysin/bacteriocin (CylB, *Enterococcus faecalis*). 18) unknown (SurB, *E. coli*). 19) nisin (NisT, *Lactococcus lactis*). 20) subtilin (SpaB, *Bacillus subtilis*). 21) capsular polysaccharide (BexA, *H. influenzae*). 22) capsular polysaccharide (CtrD, *N. meningitidis*). 23) capsular polysaccharide (KpsT, *E. coli* K1). 24) capsular polysaccharide (KpsT, *E. coli* K5). 25) daunorubicin/doxorubicin (DrrA, *Streptomyces peucetius*). 26) lipopolysaccharide (NodI, *R. leguminosarum*). 27) heme for cytochrome *c* (CycV, *Bradyrhizobium japonicum*). 28) heme for cytochrome *c* (HelA, *Rhodobacter capsulatus*). 29) microcin B17 (McbF, *E. coli*). The consensus (30) shows the amino acid that occurs in each column with most frequency. The ATP binding motifs A and B, are underlined. Amino acids are shown as single letter code and **?** represents an amino acid position where there is no obvious overall consensus.



A site

B site

clusters contain three functional segments, one encoding genes involved in the translocation of the completed polysaccharide across the periplasm and outer membrane, the second contains genes responsible for biosynthesis of the polysaccharide and a third containing the genes involved in transport across the inner membrane. Similar genetic organisation occurs in the exopolysaccharide biosynthetic systems of *R. meliloti* and *A. tumefaciens* whereby the genes responsible for transport of β1-2 glucan are located as an operon within an EPS biosynthetic cluster.

On the basis of the conserved arrangement of the genes for EPS biosynthesis of other Gram-negative bacteria, it appeared likely that the 34 minute biosynthetic gene cluster of *P. aeruginosa* would contain one or more genes for export of polymannuronate across the cytoplasmic membrane which would encode a member of the ABC transport superfamily. This hypothesis was explored further in this chapter.

3.2. SOUTHERN BLOT ANALYSIS OF THE *P. AERUGINOSA* ALGINATE BIOSYNTHETIC CLUSTER.

The ABC transport superfamily is typified by an approximately 190 amino acid, conserved region surrounding the ATP-binding motifs, A and B (see figure 3.2). In order to determine whether such a system is encoded in the 34 minute gene cluster, the ability of a gene encoding one of the ATP-binding subunits of an ABC transporter to hybridise to plasmids bearing fragments from the biosynthetic cluster was examined. The *opp* locus which contains the polycistronic operon *oppABCDF* is known to encode such a transporter *oppD* and *oppF* encode separate ATP binding subunits and so, one of these (*oppD*) was chosen for use as a nucleic acid probe. Plasmid pMPG27 (provided by Dr. M. P. Gallagher) contains *oppD* as a 1410-bp

*Eco*RI/*Eco*RV fragment within vector pCH41 (Gallagher *et al.*, 1989) which itself is a derivative of the high copy number vector pLC2833 (Remaut *et al.*, 1983). A *Hin*dIII/*Alu*I 224-bp region surrounding the ATP-binding motif B was excised (see figure 3.3), purified and random prime labelled with [α-^{32}P]dCTP (see 2.3.2 for details).

All the Gram-negative bacteria which produce EPS and have been characterised to date have their biosynthetic genes located within a gene cluster. In *P. aeruginosa*, the alginate biosynthetic genes cluster, located at 34 minutes, is thought to have an operonic structure (Chitnis and Ohman., 1993) and appears to contain all the genes required for the synthesis of alginate except *algC* which encodes PMM/PGM, a dual function enzyme involved in both alginate and LPS synthesis. It seems likely then, that genes responsible for the transport of the polysaccharide (polymannuronate) across the inner membrane should reside within the biosynthetic cluster. Furthermore, if the *P. aeruginosa* EPS biosynthetic system was functionally related to that of most other EPS producing Gram-negative bacteria, it would also be probable that the inner membrane transporter would be a member of the ABC transport superfamily.

The 34 minute biosynthetic region was therefore probed with the *Hin*dIII/*Alu*I fragment of *oppD* which carries a segment of an ATP binding subunit of oligopeptide permease. The genes encoding the 34 minute biosynthetic region from *P. aeruginosa* was provided for hybridisation on two different constructs, pAD501 and pRM812. pAD501 is derived from the cosmid vector pCP13 and contains the entire 34 minute biosynthetic cluster (except *algA*) on a 24-kb *Bam*HI fragment (Darzins *et al.*, 1985) and pRM812, which contains a 9-kb *Hin*dIII fragment (*algD* to *algE*) cloned into the phagemid pUC119 (constructed by Dr. R. Maharaj, personal communication). These regions are shown in figure 3.4. The smaller construct was incorporated into the survey in order to help localise any hybridisation within the biosynthetic cluster.

OppD probe

```
                                                        aagct 4115
                                                          A

tTTGAAGAGTCCGTCAGAATGCTGGATGCCGTAAAAATGCCGGAAGCGCGCAAACGGATG 4175
F  E  E  S  V  R  M  L  D  A  V  K  M  P  E  A  R  K  R  M
A  E  L  A  G  A  H  E  ?  I  S  ?  L  P  E  G  Y  -  T  -

AAAATGTACCCGCACGAATTTTCCGGCGGAATGCGTCAACGTGTGATGATTGCGATGGCG 4235
K  M  Y  P  H  E  F  S  G  G  M  R  Q  R  V  M  I  A  M  A
G  E  -  G  -  G  L  S  G  G  Q  R  Q  R  I  A  I  A  R  A

CTGTTGTGCAGGCCGAAACTGCTTATTGCCGATGAGCCTACCACCGCGCTTGATGTTACC 4295
L  L  C  R  P  K  L  L  I  A  D  E  P  T  T  A  L  D  V  T
L  L  N  N  P  P  I  L  I  L  D  E  A  T  S  A  L  D  Y  E

GTGCAGGCGCAGATTATGACCTTGCTTAATGagct 4330
V  Q  A  Q  I  M  T  L  L  N  E
S  E  H  K  I  M  D  N  L  R  A
```

Figure 3.3. *Hin*dIII/*Alu*I region of *oppD* from *S. typhimurium* (Hiles *et al.*, 1987) selected to use as probe. The nucleic acid sequence of this region is shown, *Hin*dIII and *Alu*I sites are shown in lower case. The predicted protein sequence of this region of OppD is shown directly below the nucleic acid sequence, the consensus state (derived from the alignment in figure 3.2) is shown in **bold** type directly beneath the amino acid sequence. The ATP binding motif, B, is underlined. Amino acids are shown as single letter code and **?** represents an amino acid position where there is no obvious overall consensus.

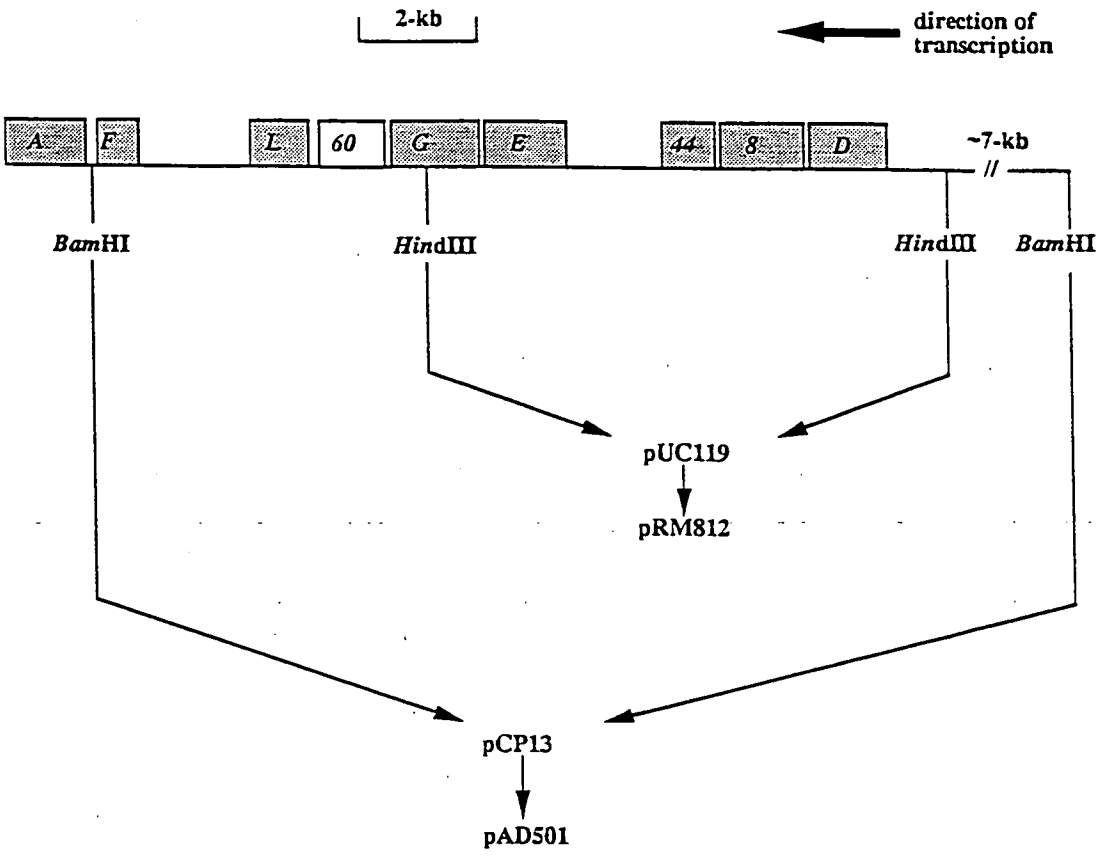


Figure 3.4. Section of the alginate biosynthetic cluster of *P. aeruginosa* cloned into vectors pAD501 and pRM812. *algD* encodes GDP-mannose dehydrogenase, *algE* encodes the outer membrane pore, *algG* encodes a C-5 epimerase, *algF* encodes an alginate acetylase and *algL* encodes the alginate lyase. The functions of the gene products of *alg8*, *alg44* and *alg60* have not been determined.

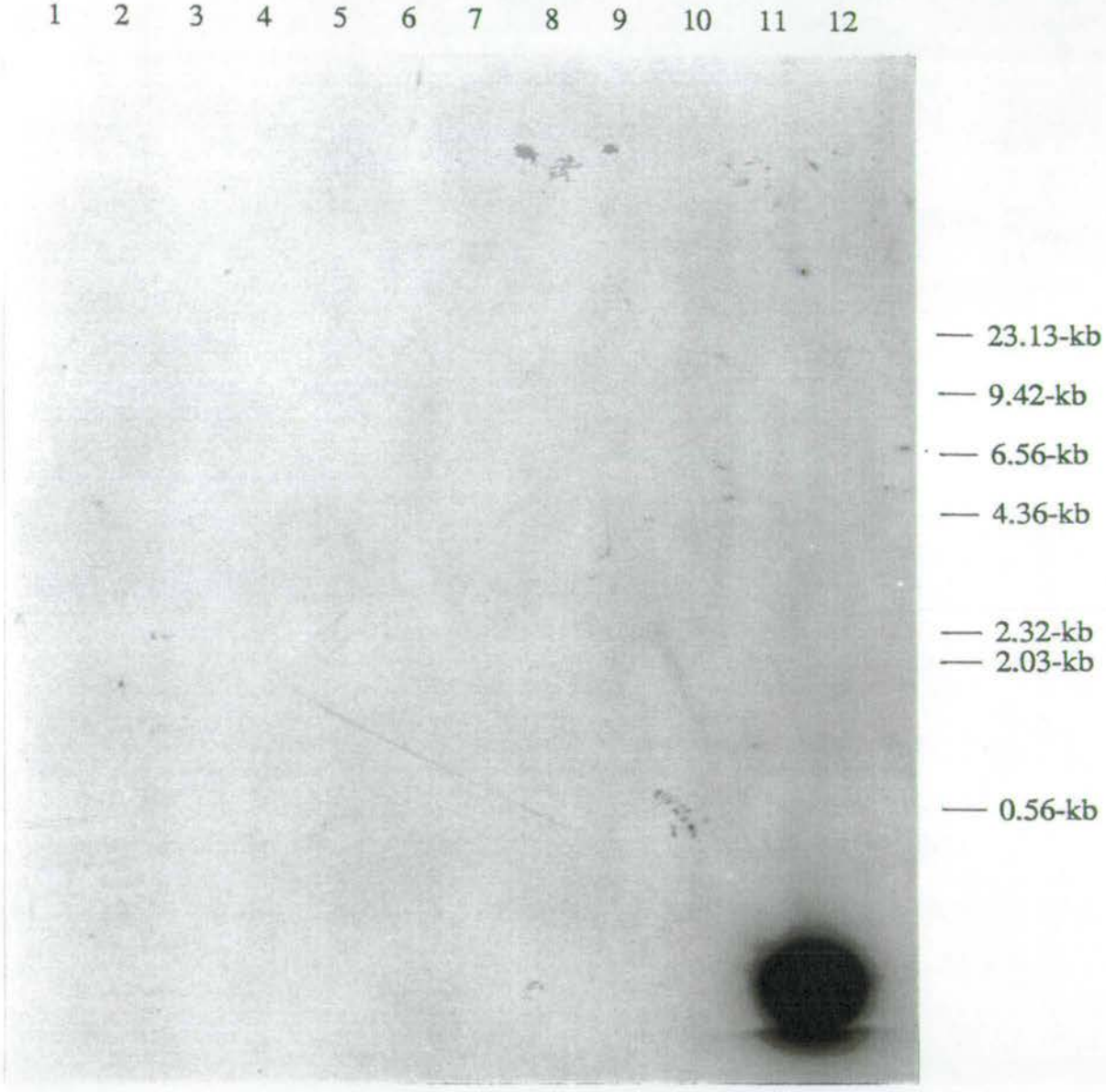
DNA from these vectors were maxiprepared using the CsCl method (as outlined by Sambrook *et al.*, 1989) and then 0.5 μg of each was digested with a variety of restriction endonucleases. Similar digests of 5 μg chromosomal DNA from the constitutively mucoid *P. aeruginosa* strain B (prepared as Silhavy *et al.*, 1984) were also carried out. The samples were electophoresed on an agarose gel before transfer to a nylon membrane. Hybridisation of the blots with the probe was then performed at 43°C in a formamide prehybridisation solution (see 2.3.2 for details). This relaxed criteria allows non-identical, but similar nucleic acid sequences to hybridise to the probe. Following hybridisation the blot was autoradiographed (see figure 3.5). Unfortunately, the autoradiogram showed only a single band and this corresponds to the control lane which contains the *Hin*dIII/*Alu*I fragment of *oppD* which was used as the probe in this hybridisation survey. Repeated digestion and hybridisation analysis failed to yield any further information (data not shown).

3.3. DISCUSSION

The lack of any cross hybridisation of the *oppD* probe, containing the highly conserved region of the ATP binding cassette, to either the 34 minute biosynthetic cluster or *P. aeruginosa* strain B chromosomal DNA can occur for a variety of reasons. The most likely explanation would seem to be due to the differences in the G:C content of *S. typhimurium* and *P. aeruginosa* DNA. *S. typhimurium* has a typical G:C content of 50% (Stanier *et al.*, 1981) and the probe itself had a G:C content of 50.4% and this is significantly different to the base ratio reported for *P. aeruginosa* which has a typical G:C content of ~70% (Darzins *et al.*, 1984; Schiller *et al.*, 1993; Chu *et al.*, 1991). This difference in G:C content may well be enough to prevent hybridisation of *S. typhimurium* DNA to *P. aeruginosa* DNA even under conditions of low stringency. The lack of hybridisation of the *oppD* probe to *P.*

Figure 3.5. Southern blot analysis of vector DNA containing sections of the 34 minute biosynthetic cluster of *P. aeruginosa* and chromosomal DNA from the mucoid *P. aeruginosa* strain B.

Lanes are as follows 1) undigested pAD501, 2) *Bam*HI digested pAD501, 3) *Eco*RI digested pAD501, 4) *Hin*dIII digested pAD501, 5) undigested pRM812, 6) *Eco*RI digested pRM812, 7) *Hin*dIII digested pRM812, 8) undigested chromosomal DNA, 9) *Bam*HI digested chromosomal DNA, 10) *Eco*RI digested chromosomal DNA, 11) *Hin*dIII digested chromosomal DNA, 12) *Hin*dIII/*Alu*I *oppD* fragment (used as probe). The DNA was separated by agarose gel (0.7% w/v) electrophoresis in TBE buffer. DNA was transferred to a Hybond™-N nylon membrane, prior to hybridisation with random primed *Hin*dIII/*Alu*I fragment of *oppD* from *S. typhimurium*. Hybridisation conditions were as stated in chapter 2 (2.3.2) for the low stringency protocol. Autoradiography was carried out for 16 hours. The positions of the lambda *Hin*dIII molecular weight markers are indicated on the right-hand side.


1 2 3 4 5 6 7 8 9 10 11 12
23.13-kb
9.42-kb
6.56-kb
4.36-kb
2.32-kb
2.03-kb
0.56-kb

aeruginosa chromosomal DNA provides further evidence for this hypothesis, as putative members of the ABC transport superfamily have already been characterised for other *P. aeruginosa* cellular processes (such as the alkaline protease exporter, Duong *et al.*, 1992). Further, the report that oligopeptide permease may play a role in peptidoglycan recycling suggests that homologues of oligopeptide permease may be widespread in bacteria (Goodell and Higgins, 1987). This suggests that even with an average amino acid identity of approximately 40% for members of the ABC transporter family over the region of the probe, the significance of the variation in G:C content between *E. coli* and *P. aeruginosa* and also the effect of codon degeneracy in codon usage between these two organisms is enough to eliminate hybridisation to any ABC transporters, using an *E. coli* probe of equivalent similarity.

From these experiments alone, it cannot be concluded whether one or more genes encoding an ABC transporter reside within the 34 minute biosynthetic region. A potential solution to the problem of codon bias would be to select a similar ABC probe, as utilised in this analysis, but derived from a *P. aeruginosa* rather than an *E. coli* host gene.

Chapter 4

Sequence analysis of the region between *alg44* and *algE* in the alginate biosynthetic cluster of *P. aeruginosa*, identification of a new gene, *algK*.

4.1. INTRODUCTION

Although the alginate biosynthetic pathway of *P. aeruginosa* has been very well characterised there are still a variety of poorly defined stages. Indeed, the genes involved in polymerisation of GDP mannuronic acid to polymannuronate, its subsequent transfer across the cytoplasmic membrane and periplasm are main areas which are yet to be identified. It also remains unclear whether the acetylation of *P. aeruginosa* polymannuronate which occurs at positions *O*-2 and *O*-3 requires two separate proteins and not just AlgF (alginate acetylase). All the genes involved in alginate biosynthesis are clustered at 34 minutes on the *P. aeruginosa* chromosome except the bifunctional *algC* gene and it is likely that genes encoding the undefined steps of polymerisation, export and possibly acetylation are located in this region (May *et al.*, 1991). At the outset of this project only the nucleic acid sequence for the *algA* (Darzins *et al.*, 1985), *algD* (Deretic *et al.*, 1987) and *algE* (Chu *et al.*, 1991) alginate biosynthetic cluster genes had been reported and a function postulated for each. The genes *alg8*, *alg44* and *alg60* had been sequenced but not published (cited in May *et al.*, 1991) and no function had been predicted. Also a mutant of the *algG* locus had been isolated which provided evidence for a C-5 epimerase (for a review, see May *et al.*, 1991). Only two regions of the alginate biosynthetic region were left uncharacterised, the first was a ~2-kb region between *alg44* and *algE* and the second a 4-kb fragment between *alg60* and *algA*. It would seem probable that the genes encoding the final uncharacterised steps of alginate synthesis in *P. aeruginosa* might reside in the unsequenced areas of the 34 minute biosynthetic cluster.

Initial studies were undertaken to discover the location of gene(s) predicted to be involved in the transport of polymannuronate across the inner membrane. However, hybridisation surveys with a potential transport homologue (see chapter 3)

had failed to elucidate any further information on location. In the *bex* and *ctr* loci involved in the capsule (EPS) biosynthetic clusters of *H. influenzae* and *N. meningitidis* the inner membrane transport genes lie together in an operon with genes encoding the outer membrane pore (Kroll *et al.*, 1991; and Frosch *et al.*, 1991 respectively). If the genetic arrangement for *P. aeruginosa* was similar to those of *H. influenzae* and *N. meningitidis*, it would appear probable that the gene(s) encoding the inner membrane transporter would reside in close proximity to *algE* (which encodes the outer membrane alginate permeable pore, Chu *et al.*, 1991). The only uncharacterised region in such a locality is that which lies between *algE* and *alg44* and therefore this area was selected for further, more detailed analysis.

4.2. NUCLEIC ACID SEQUENCE ANALYSIS OF REGION BETWEEN *alg44* AND *algE*

The region upstream of *algE* (which encodes the alginate permeable outer membrane pore), of *P. aeruginosa*, was selected for further DNA sequence analysis. Initial studies were confined to an (approximately 2-kb) *Eco*R1/*Sma*I fragment isolated from pRM812 (see figure 3.4) which contains the entire uncharacterised segment from *alg44* to *algE*, together with some flanking DNA. This segment was subcloned from pRM812 into the high copy number phagemid, pBluescript (Short *et al.*, 1988) to form plasmid pAA1. Sequencing of this region was carried out using variations on the dideoxy method of Sanger *et al.* (1977). In order to eliminate the sequencing problems imposed by the G:C rich nature of *P. aeruginosa* DNA, initial sequencing of this region was performed on DNA fragments subcloned from pAA1 into the M13 vector, mp19 (Messing *et al.*, 1983), -see table 4.1 for details of constructs. Resolution of G:C compressions, common in pseudomonal DNA sequencing (Deretic *et al.*, 1987; Darzins *et al.*, 1985; Zielinski *et al.*, 1991 etc), was

mp19 SUBCLONE	FRAGMENT DESCRIPTION	NUCLEOTIDE POSITION
1	*Kpn*I-*Pst*I 5'-3'	226-671
2	*Pst*I-*Sal*I 5'-3'	666-968
3	*Sal*I-*Sal*I 5'-3'	963-1346
4*	*Sal*I-*Sal*I 3'-5'	1346-963
5*	*Pst*I-*Kpn*I 3'-5'	671-226
PLASMID AND DELETION NUMBER	**DIRECTION OF SEQUENCE**	**NUCLEOTIDE POSITION OF FUSION JUNCTION**
pAA26 Δ1	5'-3'	515
pAA22 Δ1	3'-5'	1109
Δ2	3'-5'	1089
Δ3	3'-5'	715
Δ4	3'-5'	617
Δ5	3'-5'	386
Δ6	3'-5'	286

Table 4.1. Table showing the *algK* fragments sub-cloned and sequenced in M13 mp19 (Yanisch-Perron *et al.*, 1985) and description of exonuclease III deletions of pAA26 and pAA22. The asterisk denotes that the fragments were previously cloned into other vectors in order to reverse the fragment orientation when finally sub-cloned into mp19. The nucleotide position is determined from the *Sma*I site (see figure 4.2). Previously sequenced regions run from nucleotide positions 1-231 (Maharaj *et al.*, 1993) and 1295-*algE* (Chu *et al.*, 1991).

achieved by replacing dGTP with its 7 deaza dGTP analogue (as described in 2.3.2). Single stranded DNA was isolated and purified (as described by Sambrook *et al.*, 1989) and the 'Universal' -40 primer (USB) 5'-GTTTCCCAGTCACGAC-3' was used for all mp19 sequencing reactions.

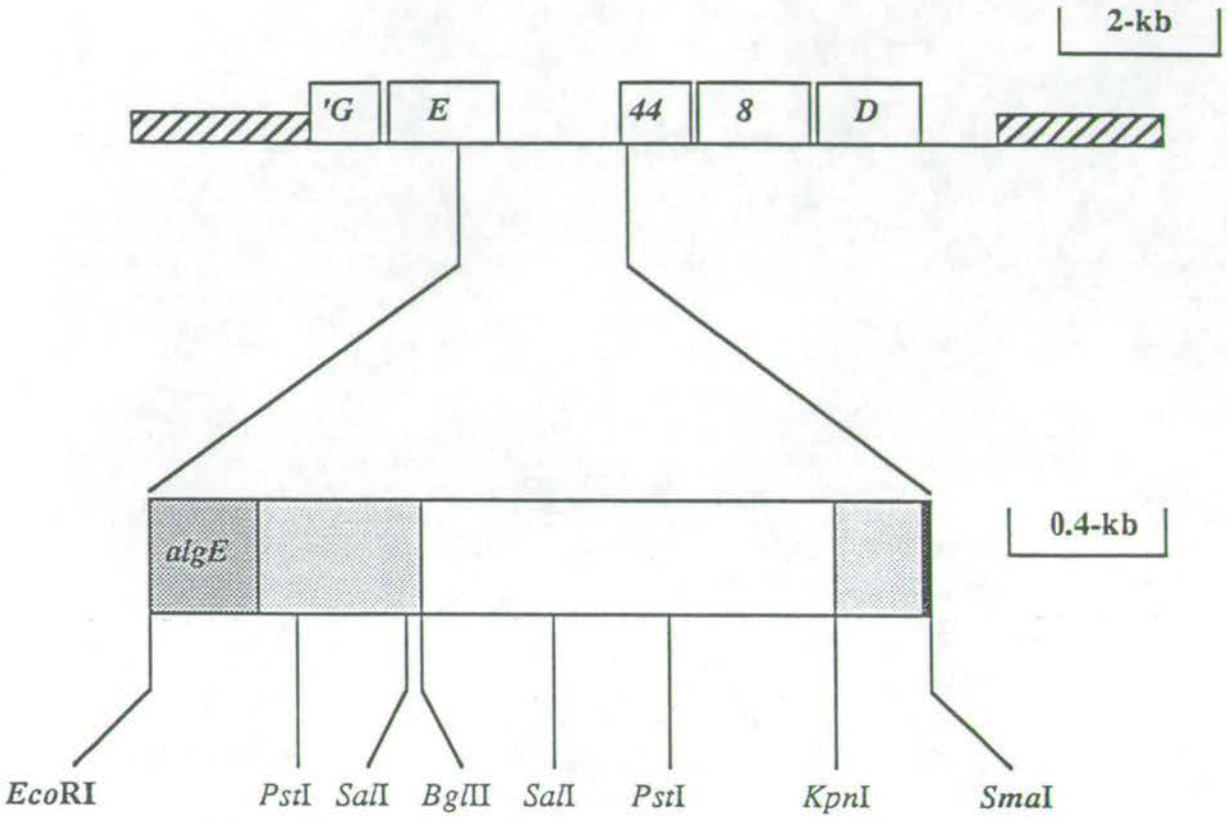
Gaps in the double strand sequence data were completed by exonuclease III deletion (as described in 2.3.2) of the *Eco*RI/*Sma*I fragment cloned in both orientations into pJBS633 (Broome-Smith and Spratt., 1986). The first orientation was achieved as a *Bam*HI/*Eco*RV subclone from pAA1 forming plasmid pAA26 and the reverse orientation was achieved when preparing the β-lactamase fusion plasmid pAA22 (see chapter 6). Both fragments were cloned into the *Bam*HI and *Eco*RV sites of pJBS633. Single stranded DNA was generated using the helper phage, R408, instead of IR1 suggested by Broome-Smith and Spratt (1986), and the DNA was isolated and purified as previously. The β-lactamase primer 5'-CTCGTGCACCCAACTGA-3' was utilised to sequence into the cloned segment (as generally described by Broome-Smith and Spratt., 1986).

The sequencing strategy, showing the extent and direction of the sequencing reactions from individual constructs is shown in figure 4.1.

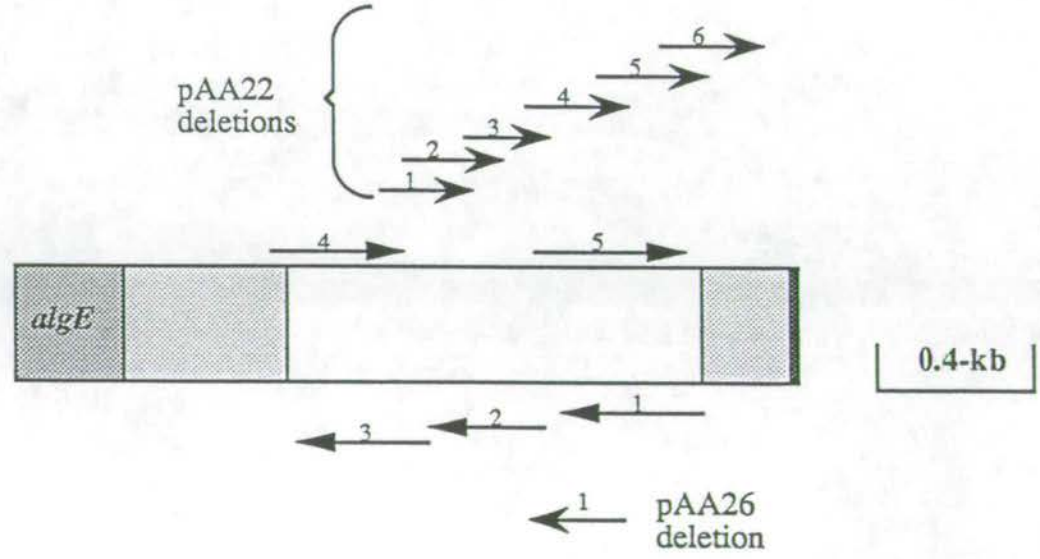
Within the nucleotide sequence, two long open reading frames (ORF's) were predicted, but only one contains potential ATG translational start sequences. However, the second open reading frame does have a potential valine (GTG) start site (548 base pairs downstream of the *alg44* stop codon) the ORF however then runs through *algE*. Chu *et al.* (1991) determined both the *algE* ATG (methionine) translational start and the *algE* transcriptional start point within the 34 minute biosynthetic cluster of *P. aeruginosa*. These *algE* start points are approximately a kilobase further downstream from the valine position, and thus it appears extremely unlikely that this valine is a further translational start point. Both of these ORF's are transcribed in the same direction as the other genes characterised in the 34 minute

Figure 4.1. Panel 1 shows the construction of vector pAA1 from pRM812 (see figure 3.1), and the respective internal restriction sites present in this region. Panel 2 shows the sequencing strategy applied to the unsequenced region between *alg44* and *algE*. The arrows indicate the extent and direction of the sequence obtained from specific M13 mp19 clones (⟶) or from exonuclease III deletions of plasmids pAA22 and pAA26 (⟶), see table 4.1 for further details. The regions of DNA previously sequenced and published (the upstream *algE* region by Chu *et al*, 1991 and the downstream *alg44* region by Maharaj *et al.*, 1993) are shown as shaded boxes (▭). The *alg44* coding region within pAA1 is indicated by ▬ and the *algE* coding region within pAA1 is indicated by ▭. pUC119 vector DNA of pRM812 is indicated by ▨.

PANEL 1



PANEL 2



biosynthetic region, that is to say, from *algD* towards *algA*. Furthermore, only one of these ORFs is predicted to be correct when the highly specific codon usage of *P. aeruginosa* is considered. Minton *et al.* (1984) found that the G:C rich nature of *P. aeruginosa* is confined to the third base (wobble base) which has a typical 90+% G:C content. This situation has been confirmed in all characterised genes of *P. aeruginosa* and all of *P. aeruginosa* genes sequenced within the alginate biosynthetic region (Darzins *et al.*, 1985; Deretic *et al.*, 1987; Franklin *et al.*, 1993., Chu *et al.*, 1991; Shinabarger *et al.*, 1993; Schiller *et al.*, 1993). The majority of the A:T content of the third base position is almost entirely due to the codon GAA which encodes glutamic acid (E), a codon bias typical of *P. aeruginosa* genes (see table 4.2).

Using this information an ORF of 1425-bp is predicted, encoding a protein of 475 amino acids with a predicted molecular mass of 52,470 Da. This gene was called *algK*. The complete sequence of *algK* is shown in figure 4.2. The ORF is shown initiating at an ATG codon 230 base pairs downstream from the stop codon of *alg44* (Maharaj *et al.*, 1993), and terminates at a TGA stop codon, which overlaps with the ATG translational start codon of *algE* (Chu *et al.*, 1991). Overlapping start and stop codons are not unusual in *P. aeruginosa* operons and typical examples are observed in the case of the *xcp* operon, which encodes nine proteins involved in the *P. aeruginosa* protein secretory apparatus (Bally *et al.*, 1989).

On analysis of the predicted protein sequence, it appears that 46.11% of the amino acids are non-polar, 31.36% are hydrophilic, 11.37% are acidic and 11.16% are basic and overall the protein has a pI of 5.64. A hydropathy plot of the protein was performed according to the method of Kyte and Doolittle (1982), and the pattern obtained is shown in figure 4.3. Integral membrane proteins are typified by repeated membrane-spanning α-helices consisting of 16-23 hydrophobic amino

Figure 4.2. Nucleic acid sequence and predicted protein sequence of the *algK* open reading frame. The sequence is shown from the *Sma*I site (lower case and underlined) of the C-terminal end of *alg44* to the translation initiation codon of *algE*. Encoded amino acids are specified by the standard one letter abbreviation and are shown beneath the nucleotide sequence. Termination codons are designated by asterisks. The putative signal sequence is underlined and the cysteine (C) residue required for lipid attachment is indicated in **bold** type.

SmaI

cccgggtgaacttccaggtcgccggcgacggctcaaccgcagcggccgcatcgtcaacac 60

(*alg44*)G *

cgcgccggtggacggcgacctgtcctcggaaatccgcgtgcagatccagcccgaccagcc 120

gctcgacgcccagtacgccggccgcccggccgaggtcagcatcggcggcctgcccggccg 180

caccctgctgaacaaggccgtgaccctggccaccgctcgctgaccggtacccgaccatga 240

M K 2

agatgcccatcctccctcccctgcccctggcgtcccggcacctgctgctggcctcggcca 300

M P I L P P L P L A S R H L L L A S A I 22

tcgccctggccgccggctgcgccggactgcccgaccagcgcctggcccaggaggccctgg 360

A L A A G C A G L P D Q R L A Q E A L E 42

agcgcggcgacctcgccaccgcccagagcaactaccaggccctggccgcgatgggttacg 420

R G D L A T A Q S N Y Q A L A A M G Y A 62

ccgacgcccaggtcggcctcgccgacatgcaggtcgccagcggcgacagtgcccagcagg 480

D A Q V G L A D M Q V A S G D S A Q Q A 82

ccaaggccgaaaagctctaccgcgaggccgcgcagacctcgccgcgggcccgcgcgcgcc 540

K A E K L Y R E A A Q T S P R A R A R L 102

ttggcaagtggctggcggccaagcccggcgccagcgacgccgagcaccgcgaggccgagc 600

G K W L A A K P G A S D A E H R E A E R 122

gcctgctcagccaggccttcgaacagggcgaggacagtgccctggtgccgctcatcgtgc 660

L L S Q A F E Q G E D S A L V P L I V L 142

tctacctgcagtacccgcagtcctggccggaaatcgacccgcagcagcgcatcgaccagt 720

Y L Q Y P Q S W P E I D P Q Q R I D Q W 162

ggcgcgcgcgcggactgccgcaggccgacctggcgcagatcatcctctaccgcacccagg 780

R A R G L P Q A D L A Q I I L Y R T Q G 182

gcacctacgcccagcacctgggcgagatcgagcaggtctgccagcgctggctgcggcgca 840

T Y A Q H L G E I E Q V C Q R W L R R M 202

tggatgtgtgctggtacgaactggccacggtctaccagatgcagggcaacgcggaaaagc 900

D V C W Y E L A T V Y Q M Q G N A E K Q 222

agaaggtcctgctcgaacagctgcgcgccgcctacaaggccggccgggtgcccggcgagc 960

K V L L E Q L R A A Y K A G R V P G E R 242

gggtcgactcggtggccggggtgctcgccgacggcgaactcggccagcccgacccgcaga 1020

V D S V A G V L A D G E L G Q P D P Q T 262

ccgcccaggcgctgctggaagagatcgcgccgagctaccccgcggcctgggtcagcctgg 1080

A Q A L L E E I A P S Y P A A W V S L A 282

ccaagctgctctacgactatcccgaccagggcgacctggaaaagatgctcggctacctga 1140

K L L Y D Y P D Q G D L E K M L G Y L K 302

agaacgcccaggacgccgcccagccacgcgccgaactgctcctcgggcgcctctactacg 1200

N A Q D A A Q P R A E L L L G R L Y Y D 322

acggcaagtgggcgccgcaggacccgcgcaaggccgagcggcacctgctcaaggccgccg 1260

G K W A P Q D P R K A E R H L L K A A A 342

ccagcgaaccccaggccaactactacctggggcagatctaccgccgcggcttcctcggca 1320

S E P Q A N Y Y L G Q I Y R R G F L G K 362

aggtctatccgcagaaggccgtcgaccacctgatcctcgccgcccgcgccggccaggcca 1380

V Y P Q K A V D H L I L A A R A G Q A S 382

gcgccgacatggccctggcccagctctggtcgcagggccgcggcatccagccgaaccggg 1440

A D M A L A Q L W S Q G R G I Q P N R V 402

tcaacgcctacgtcttcggccagctcgccgtgcagcagcaggtaccgcaggccagcgacc 1500

N A Y V F G Q L A V Q Q Q V P Q A S D L 422

tgctcgggcagatcgaagcgcaactgccgccggccgaacgcagccaggcgcagcaattgc 1560

L G Q I E A Q L P P A E R S Q A Q Q L L 442

tcaagcgcgaacaacagagccgcggcaacaactggcaggccaccgtcagcctgctgcaga 1620

K R E Q Q S R G N N W Q A T V S L L Q S 462

gccaggactcgccgatcaacgaagaagagccagaaagcctatga

Q D S P I N E E E P E S L *

M (*algE*)

Amino acid	Codon	N°	Amino acid	Codon	N°	Amino acid	Codon	N°	Amino acid	Codon	N°
Gly	GGG	4	Arg	AGG	0	Trp	TGG	9	Arg	CGG	7
Gly	GGA	2	Arg	AGA	0	END	TGA	1	Arg	CGA	0
Gly	GGT	1	Ser	AGT	2	Cys	TGT	0	Arg	CGT	0
Gly	GGC	24	Ser	AGC	14	Cys	TGC	3	Arg	CGC	24
Glu	GAG	13	Lys	AAG	17	END	TAG	0	Gln	CAG	52
Glu	**GAA**	**17**	Lys	AAA	0	END	TAA	0	Gln	CAA	3
Asp	GAT	1	Asn	AAT	0	Tyr	TAT	2	His	CAT	0
Asp	GAC	23	Asn	AAC	9	Tyr	TAC	19	His	CAC	5
Val	GTG	7	Met	ATG	8	Leu	TTG	1	Leu	CTG	36
Val	GTA	1	Ile	ATA	0	Leu	TTA	0	Leu	CTA	1
Val	GTT	0	Ile	ATT	0	Phe	TTT	0	Leu	CTT	1
Val	GTC	12	Ile	ATC	14	Phe	TTC	3	Leu	CTC	23
Ala	GCG	14	Thr	ACA	1	Ser	TCG	5	Pro	CCG	16
Ala	GCA	0	Thr	ACT	0	Ser	TCA	0	Pro	CCA	2
Ala	GCT	0	Thr	ACA	0	Ser	TCT	0	Pro	CCT	1
Ala	GCC	60	Thr	ACC	6	Ser	TCC	2	Pro	CCC	10

Table 4.2. Codon usage within AlgK, amino acids are shown in standard three letter code, and the number of times each amino acid is encoded by individual codons within the polypeptide is shown alongside. The GAA codon bias is highlighted in **bold** type.

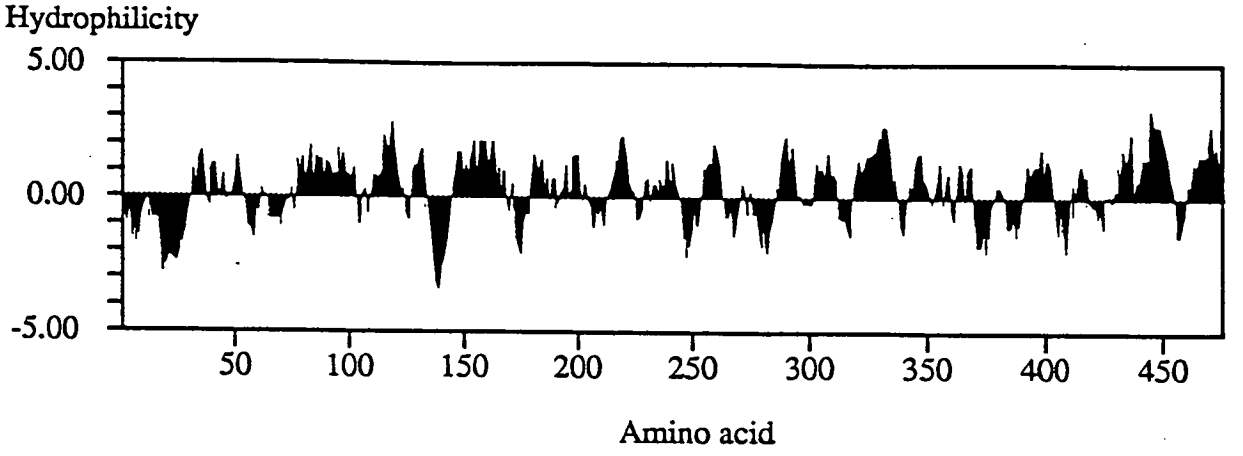


Figure 4.3. Kyte and Doolittle (1982) plot of relative hydropathy of the AlgK protein. Positive values represent hydrophilic regions of the protein (Y-axis). Amino acids are numbered from the N-terminus of the protein (X-axis). A window size of seven was used for analysis.

acids, separated by shorter regions of hydrophilic amino acids (thought to reside in the aqueous environment). The lack of any long stretches of hydrophobicity (as predicted by the Kyte and Doolittle plot) suggests that this protein is not an integral membrane protein and as such predicts that it is not the integral membrane component of a putative ABC transporter. However, the hydropathy plot does reveal, a single long hydrophobic (indicated by negative values) region, with a pI of 11.65, spanning the first 27 amino acids of the putative protein. Such an arrangement is indicative of a signal sequence, required for targeting of a protein beyond the cytoplasm, and similar sequences have been reported recently for AlgL (Schiller *et al.*, 1993, Boyd *et al.*, 1993), AlgF (Shinabarger *et al.*, 1993) and AlgG (Franklin *et al.*, 1994). Using the computer program 'motifs' (part of the UWGCG package, Devereux *et al.*, 1991) it was also predicted that the AlgK was a membrane lipoprotein and contained a prokaryotic membrane lipoprotein signal sequence with a lipid attachment site at the putative cleavage point.

In prokaryotes, membrane lipoproteins are synthesised with a precursor signal peptide, which is cleaved by a specific lipoprotein signal peptidase (signal peptidase II). The peptidase recognises a conserved amino acid sequence and cuts upstream of a cysteine residue to which a glyceride fatty acid lipid is subsequently attached. From the precursor sequences of all such known proteins, a consensus pattern and a set of rules to identify this type of post-translational modification was deduced and incorporated into the UWGCG 'motifs' program in order to predict further members of the family. The consensus pattern of the cleavage site for lipoprotein signal peptides is as follows:

~{DERK}(6)-[LIVMFSTAG](2)[IVMSTAGQ]-[AGS]-C.

that is to say the first 6 amino acids of the signal sequence must not be aspartic acid (D), glutamic acid (E), arginine (R) or lysine (K). This is followed by 2 amino acids which can be leucine (L), isoleucine (I), valine (V), methionine (M), serine (S), threonine (T), alanine (A) or glycine (G). Followed by a single amino acid which may be I, V, M, S, T, A, G, or glutamine (Q). Directly upstream of the cysteine lipid attachment site, must be a single A, G or S residue (for a review see Pugsley, 1993). This consensus is obeyed by the putative prokaryotic lipid attachment signal sequence of AlgK which is as follows for residues 18 to 28:

(LASAIA)-(LA)(A)-(G)-C

The putative signal sequence of AlgK also obeys the additional 'motif' rules of ;
1) the cysteine must always be between positions 15 to 35 of the sequence in consideration (it is actually in position 28 in AlgK),
2) there must be at least one lysine or arginine in the first seven positions of the sequence (in AlgK there is one lysine residue, at positions 2).

The position of the consensus pattern of the cleavage site for AlgK is indicated in figure 4.2.

Comparison of AlgK with other proteins was also undertaken using the UWGCG package 'FASTA' (see 2.3.5 for description). This revealed very little amino acid identity with other proteins in the database including those of the ABC transporter superfamily. A specific survey for the ATP/GTP binding site motif, A, (typical of ABC transporters) was also performed using the UWGCG program 'motifs'.

The ATP/GTP binding site motifs were generated from sequence comparisons and crystallographic data analysis. This analysis has shown that an appreciable proportion of proteins that bind ATP or GTP share a number of highly conserved sequence motifs in a glycine-rich region which probably forms a flexible loop between a β strand and an α helix. These groups interact with one of the phosphate groups of the nucleotide and are highly conserved amongst proteins that bind ATP or GTP. The consensus pattern (as defined in the UWGCG program) is [AG]-X(4)-G-K [ST], or alanine or glycine followed by any four amino acids, a glycine and lysine residue and finally serine or threonine. This consensus is used in the 'motifs' program to isolate proteins that possibly bind ATP (and thus includes all ABC transporters) or GTP. However no similarity was observed between this consensus and AlgK predicting that it is not an ATP-binding subunit of the ABC transporter superfamily.

Computer analysis of *algK* 5' region for possible promoter sequences

Chu *et al.* (1991) identified a putative *algE* internal promoter by sequence alignment with the promoter of *algD* (as defined by Deretic *et al.*, 1987). Since a non-coding sequence of 230-bp was found between the *alg44* stop codon and the initiation codon of *algK* computer analysis of the upstream region for a potential promoter region was undertaken, using the UWGCG program, 'Bestfit'. 'Bestfit' makes an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximise the number of matches using the local homology algorithm of Smith and Waterman (1981). The upstream region between the TGA stop of *alg44* and the *algK* ATG translational start was used as the first sequence in this alignment. The second sequence was the *algD* promoter, as defined by Deretic *et al.* (1987). The 'Bestfit' program showed that the region of greatest similarity between the upstream region of *algK* and that of the

algD promoter occurred between -151 and -111-bp before the ATG start, the sequence similarity is shown figure 4.4. (panel A).

Previously (Zielinski *et al.*, 1991; May *et al.*, 1991) alignment had been performed on the putative promoter regions of *algD*, *algR1* and *algC*, this revealed identity over a GG-N_{10}-GC consensus typified by promoters under the regulatory control of RpoN (σ^{54}) as identified by Kusto *et al.* (1989). Alignment of the upstream region of *algK* to these regions (see panel B figure 4.4) revealed complete identity over the RpoN (σ^{54}) consensus region and may suggest that *algK* has an internal promoter controlled by a similar sigma factor. However, more recent studies have shown that the conversion to mucoidy within *P. aeruginosa* and the regulation of both *algR1* and *algD* is actually controlled by a different sigma factor, σ^{E}, which is thought to recognise a different consensus sequence (Martin *et al.*, 1994). An alignment of the known σ^{E} regulated promoter sequences from both *P. aeruginosa* and *E. coli* is shown in panel C, figure 4.4, and reveals that there is very little identity between the *algK* upstream region and the σ^{E} consensus predicting that the putative *algK* promoter is not regulated by this sigma factor.

Incidentally, *algR2* and *algR3* share a TTGACA-17($^{+}$/-1-bp)-TATAAT canonical -35 and -10 sequences recognised by σ^{70} containing RNApol (Deretic and Konyecsni, 1990; Konyecsni and Deretic, 1990 respectively) and as such both promoters are constitutively transcribed *in-vivo*. However no obvious identity was observed by alignment of the upstream region of *algK* to these promoters when using the same 'Bestfit' program.

No obvious inverted repeats which might function as a *rho* independent transcription terminator (a hairpin structure is observed for transcription termination for *algR2*, Konyecsni and Deretic, 1990) were discovered downstream of the *algK* TGA stop. This is in keeping with the hypothesis that all of the genes are transcribed as a single transcriptional unit (Chitnis and Ohman., 1993).

Figure 4.4. Alignment of the upstream region of *algK* with promoter regions of genes with putative RpoN and σ^E consensus sequences.

Panel A shows the optimal alignment between the upstream region of *algK* and the promoter region of *algD* (identified by Deretic *et al.*, 1987). The '-' defines the *algD* transcriptional start point (Deretic *et al.*, 1987) and identity is marked with a vertical bar.

Panel B shows the alignment of the 5' upstream region of *algK* to the promoter regions of *algD*, *algR1*, and *algC*. Identical bases between all four 'promoters' are marked with a vertical bar. The transcriptional start points ('-') and promoter sequences for *algD*, *algR1*, and *algC* are derived from Deretic *et al*, (1987), Deretic and Konyecsni, (1989) and Zielinski *et al*, (1991) respectively. The RpoN (σ^{54}) consensus GG-N_{10}-GC (Kusto *et al.*, 1989) is marked in **bold.**

Panel C shows the alignment of the putative promoter region of *algK* with the σ^E consensus of *algU*, *algD* and *algR1* of *P. aeruginosa* and the *E. coli* genes, *htrA* and *rpoH*. The promoter region sequences and transcriptional start points (marked as previously) for *algU*, *algD*, *algR1*, *htrA* and *rpoH* were derived from Martins *et al.* (1994), Deretic *et al*, (1987), Deretic and Konyecsni, (1989), Lipinska *et al*, (1988) and Wang and Kiguni, (1989). The σ^E -35 and -10 consensus sequences (Erickson *et al.*, 1987) is marked in **bold**.

PANEL A

P. aeruginosa algK CGGAAATCCGCGTGCAGATCCAGCCCGACCAGCCGC

P. aeruginosa algD CGGAACTTCCCTCGCAGAGAAAACATCCTATCACCG

PANEL B

P. aeruginosa algK C**GG**AAATCCGCGT**GC**AGATCCAGCCCGACCAGCCGC

P. aeruginosa algD C**GG**AACTTCCCTC**GC**AGAGAAAACATCCTATCACCG

P. aeruginosa algR1 G**GG**CACTTTTCGG**GC**CTAAAGCGAGTCTCAGCGTCG

P. aeruginosa algC A**GG**AACTCGGCGG**GC**AACGCACTGCCAAACCCCCTG

PANEL C

P. aeruginosa algK CG**GAAAT**CCGCGTGCAGATCCAGCCC**GA**CCAGCCGC

P. aeruginosa algU GA**GAACTT**TTGCAAGAAGCCCGAG**TCT**ATCTTGGCA

P. aeruginosa algD CG**GAACTT**CCCTCGCAGAGAAAACA**TC**CT**A**TCACCG

P. aeruginosa algR1 GG**GCACTT**TTCGGGCCTAAAGCGAG**TCTCA**GCGTCG

E. coli htrA CG**GAACTT**CAGGCTATAAAACGAA**TCTGA**AGAACA

E. coli rpoH TT**GAACTT**GTGGATAAAATCACGG**TCTGA**TAAAACA

σ^E consensus **GAACTT** **TCTGA**

-35 -10

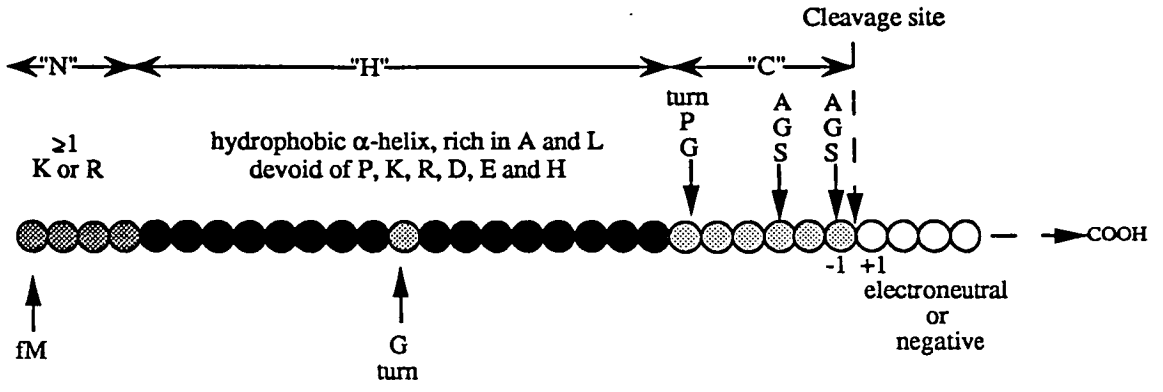
4.3. DISCUSSION

During this study a novel gene, '*algK*', was located in the 34 minute alginate biosynthetic cluster of *P. aeruginosa*. Sequencing of the region between *alg44* and *algE* revealed a single long open reading frame, transcribed in the same direction as the rest of the genes characterised in the operon. This open reading frame encoded a putative 475 amino acid protein. Comparison of AlgK with other proteins in the databases available failed to discover any proteins with obvious and significant/extensive similarities. AlgK was also subjected to a specific survey for the 'A' motif, which is common to the ATP binding subunits of all ABC transporters, but again there was no obvious similarity, suggesting that AlgK is not such a component of the ABC transporter superfamily.

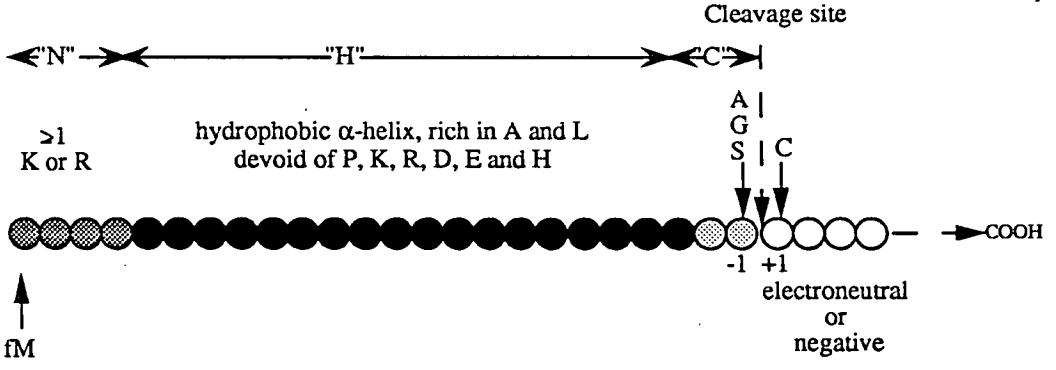
Hydrophobicity profiling of the AlgK polypeptide predicted that the protein was relatively hydrophilic suggesting that it was not an integral membrane component of an ABC or other transporter. The hydrophobicity plot predicted only a single long hydrophobic segment and this was located at the N-terminus of the AlgK protein and typified a signal peptide. A similar hydrophobic N-terminal arrangement has also been shown to occur in the proteins AlgE (Chu *et al.*, 1991), AlgG (Franklin *et al.*, 1993), AlgF (Shinabarger *et al.*, 1993) and AlgL (Boyd *et al.*, 1993) which are located in the periplasm or the outer membrane. Like all standard signal peptides, those of AlgE, AlgG, AlgF and AlgL share a consensus, three domain, 'N', 'H' and 'C' structure (see figure 4.5). The N-terminus ('N') domain is 5-6 residues long and has on average, two positively charged amino acids; the hydrophobic ('H' domain) core is approximately 12 residues long, contains highly hydrophobic residues and is uncharged (the high leucine and alanine content causes the signal peptides to adopt an α-helical configuration); the C-terminus ('C' domain) is 6 residues long, has neutral small side-chain amino acids at positions -1 and -3, and often includes a

Figure 4.5. Schematic representation of the standard and lipoprotein signal peptides found in precursors of bacterial secretory proteins. Each amino acid within the functional elements of the signal peptides is indicated by a solid or shaded circle. Solid circles represent mainly hydrophobic or apolar residues whereas shaded circles indicate polar or charged residues. Open circles represent residues elsewhere in the polypeptide. Standard and lipoprotein signal peptides have three domains, the N-terminal (N), hydrophobic (H) and C terminal (C). The N-domain is typically a polar domain of 2-15 residues and carries a net positive charge. The H-domain (>8 residues) is composed of predominantly hydrophobic residues and alanine and lacks strongly polar or charged residues. The C-domain is less hydrophobic and contains the signals that are recognised by signal peptidase I or II. Amino acids are indicated in the single letter code and the fMet (fM) translational start amino acid is also marked (modified from Pugsley, 1993).

Standard signal peptide



Lipoprotein signal peptide



proline or glycine turn residue at position -6. The 'C' domain also contains the cleavage signal recognised by signal peptidase I (Izard *et al.*, 1994). However, unlike the signal peptides of the other characterised proteins of the alginate biosynthetic cluster and pathway, computer analysis predicts that the signal peptide contained a prokaryotic lipoprotein lipid attachment site.

The signal peptides of prokaryotic membrane lipoproteins differ from the standard, signal peptidase I-cleaved, signal sequence (as described above) in numerous ways. Firstly, the hydrophobic domain is shorter, with more hydrophobic residues. Secondly, there is no obvious turn residue upstream of the cleavage site, and most importantly, there is a strict requirement for cysteine at the cleavage site. This amino-terminal cysteine residue is converted to glyceryl cysteine before cleavage occurs and is usually modified further by the addition of two fatty acids to the glyceryl residue and a fatty acid to the free amino group, after signal peptide cleavage by signal peptidase II. Lipoproteins within Gram-negative bacteria tend to be either anchored to the outer leaflet of the cytoplasmic membrane or in the inner or outer leaflets of the outer membrane, as shown with NlpA, Lpp and pullulanase respectively (review Pugsley *et al.*, 1993). Removal of the 27 amino acid signal sequence (pI 11.65) would generate a mature hydrophilic protein with an overall pI of 5.20 and a putative molecular weight of 49,730 Da. Attempts to overexpress AlgK should confirm both the molecular weight and thus reading frame of the *algK* gene product and may reveal evidence for any post translational modification of the protein product due to signal sequence cleavage.

Computer analysis using the UWGCG 'Bestfit' program revealed a possible internal promoter region upstream of the *algK* translation initiation codon, by alignment of this sequence with the *algD* promoter (Deretic *et al.*, 1987). The overall sequence identity of this region was 40% but was greater than 70% when considering

the putative RpoN (σ^{54}) consensus region alone. Bacterial sigma factors are a family of proteins that confer upon RNA polymerase the ability to recognise specific promoter sequences. Among σ factors, σ^{54} (encoded by RpoN) is the only member with no obvious amino acid similarities to the others (Morett and Buck, 1989). Genes expressed by RNA polymerase-σ^{54} control diverse cellular functions. For example in *P. aeruginosa* RpoN has been shown to control expression of genes involved in assimilation and degradation of nitrogen compounds as well as motility and adherence (Totten *et al.*, 1990) and their promoters are characterised by the consensus sequence 5'-GG-N_{10}-GC-3'.

Initiation of transcription by σ^{54} interaction requires an activator protein which bind sites located at least 80-bp upstream of the promoters which they regulate (a typical example is the two component regulatory protein, NtrC, for a review of RpoN function see Morett and Buck, 1989). Such activator proteins catalyse the isomerisation of closed complexes between σ^{54}-holoenzyme and the promoter to transcription in response to a distinct physiological signal such as oxygen tension or availability of nutrients, (Morett and Buck, 1989). In line with this, Kimbara and Chakrabarty (1989) demonstrated a massive decrease in *algD* and *algR1* transcription in a RpoN-deficient, non-mucoid, *P. aeruginosa* mutant when under conditions of high osmolarity. Homology between the promoters of *algR1*, *algD*, *algC* and *algK* would suggest that *algK* is directly controlled by σ^{54} and predicts that *algK* activity might be modulated by certain environmental conditions.

In conflict with this view however, Totten *et al.* (1990) generated a RpoN mutant (by insertional inactivation) in a mucoid strain of *P. aeruginosa* which had no effect on mucoidy and implied a less significant role for RpoN in alginate production. Indeed, recent work by Martin *et al.* (1994) suggested that conversion to mucoidy was due predominantly to control by AlgU, a putative sigma factor with a high degree of similarity to σ^{E} which participates in resistance to high temperatures and

oxidative stress in *E. coli* (Erickson and Gross, 1989). Conservation of the -35 and -10 sequences of AlgU-dependent and σ^E-dependent promoters revealed an entirely different consensus than that for the putative σ^{54}. However, alignment of the σ^E consensus with the putative promoter region of *algK* revealed very little similarity. This result predicts that *algK* does not have an AlgU/σ^E dependent promoter. Indeed the operonic structure of the 34 minute biosynthetic cluster (Chitnis and Ohman., 1993) would predict that the *algD* promoter is the major promoter. The relevance if any of σ^{54} consensus sites has been largely ignored since the emergence of clear evidence for the involvement of AlgU as the sigma factor responsible for mucoidy in *P. aeruginosa.* However, the experiments of Kimbara and Chakrabarty (1989) and Totten *et al.* (1990) might suggest a role for hierarchy of sigma factors involved in alginate synthesis, with RpoN being involved in minor stress responses and AlgU being involved in complete mucoid conversion.

Low level, weak internal promoters have been postulated for other genes in the synthetic cluster, such as *algL* (Schiller *et al.*, 1993), *algF* and *algA* (Shinabarger *et al.*, 1993) and *algE* (Chu *et al.*, 1991). *algL* and *algF*, when borne on vectors in the reverse orientation to the vector promoters were found to produce alginate lyase and acetylase respectively and *algA* complemented *P. aeruginosa* mutants even when in a promoterless vector. However, low expression may be caused by run off transcription from another promoter on these plasmids. The putative promoter region of *algE* was studied more closely. Chu *et al.* (1991) transcriptionally fused the upstream region from *algE* to a promoterless streptomycin gene which when mobilised into *P. aeruginosa* resulted in a 30 fold increase in streptomycin resistance in non-mucoid cells. The postulated similarity of the region upstream of the transcriptional start point from *algE* and *algD* led Chu *et al.* (1991) to infer that the *algE* promoter functions under the control of σ^{54} and as such is only activated under

specific growth conditions. No promoter sequences or regulating sigma factors have been postulated for *algA*, *algF* (Shinabarger et al., 1993) or *algL* (Schiller *et al.*, 1993). Whether these promoters are archaic or functional *in-situ* rather than just in plasmid constructs has yet to be elucidated. Interestingly, the CAP gene operon of *Staphylococcus aureus* involved in exopolysaccharide production have also been shown to have internal promoters (Lin *et al.*, 1994). Indeed, Lin *et al.* (1994) postulated that the more distal a gene is from the major promoter, the less efficiently it is transcribed. Thus, the internal promoters might play a role in ensuring that the downstream genes are produced in adequate amounts.

Ideally, analysis of the activity of the putative *algK* promoter under different environmental stresses, would be analysed experimentally. The systems typically used to detect the activation of *P. aeruginosa* promoters, under different environmental conditions, involve the generation of transcriptional fusions. Typically the promoter region of interest is fused (in a broad host range plasmid vector) to the reporter gene encoding β-galactosidase (*lacZ*) or catechol 2,3-dioxygenase (*xylE*). Zielinski *et al.* (1991) utilised the *lacZ* system for following the activation of the *algC* promoter and DeVault *et al.* (1990), amongst others (for a review see May and Chakrabarty., 1994) utilised the *xylE* system to study the activation of *algD* (Kimbara and Chakrabarty (1989) also used the *xylE* system to study the activation of *algR1*). Both these methods allow monitoring of the expressed gene fusion spectrophometrically; catechol 2,3-dioxygenase catalyses the formation of an intensely yellow product (2-hydroxy muconic semialdehyde) from catechol, and β-galactosidase hydrolyses ONPG (o-nitrophenyl-β-D-galactopyranoside) to release nitrophenol. The intensity of the yellow colouration (nitrophenol) is proportional to the amounts of reporter gene synthesised and thus the effects of various environmental stresses on reporter gene expression can be examined.

Cloning of the immediate upstream region of *algK* to form a fusion with either the *lacZ* or *xylE* transcriptional reporter systems would define any internal promoter activity. The system would also allow monitoring of the effect, if any, of environmental stresses which might normally result in the induction of *algK* transcription. This form of control could be monitored in both mucoid and non-mucoid *P. aeruginosa* strains, in the presence or absence of an *rpoN* mutation.

Chapter 5

Expression of AlgK in *E. coli* and *P. aeruginosa*.

5.1. INTRODUCTION

Sequence analysis of the region between *alg44* and *alg8* revealed a single long open reading frame of some 1425-bp which was termed *algK*. The *algK* gene encoded a putative polypeptide of 475 amino acids with a predicted (unprocessed) molecular weight of 52,470-Da. Hydrophobicity profiling suggested that AlgK did not possess the hydrophobicity characteristics required for an integral membrane protein. However, a hydrophobic stretch of approximately 30 amino acids at the N-terminus of AlgK predicted the presence of a signal sequence. Computer analysis, using a standard consensus sequence revealed that the N-terminal of the putative AlgK polypeptide was potentially a lipoprotein signal sequence for lipid attachment. In such systems, the cysteine residue at the C-terminal end of the signal peptide is converted into glycerylcysteine and the signal sequence immediately upstream of the cysteine residue is then cleaved off by signal peptidase II. The resulting N-terminal, glyceryl cysteine, is then further modified by the addition of two fatty acids to the glyceryl residue and a fatty acid to the free amino group.

A series of AlgK protein expression studies were conducted in both *E. coli* and *P. aeruginosa* in order to verify the information deduced from the nucleotide sequence. Initial studies were aimed at confirmation of the presence of the open reading frame as defined by the sequence obtained from *algK* gene, using the pT7 series of vectors within *E. coli* (Tabor and Richardson, 1985). This system allows the expression of genes under the control of the bacteriophage T7 RNA polymerase. Such an approach has a number of advantages over vector systems utilising the *E. coli* RNA polymerase. First, T7 RNA polymerase is a very active enzyme and synthesises RNA at approximately 5 times the rate of *E. coli* RNA polymerase. In addition, it terminates transcription less frequently because efficient termination

sequences for T7 polymerase are found only rarely outwith T7 DNA (Tabor and Richardson, 1985). As a further point, T7 RNA polymerase is highly selective for initiation at its own promoter sequences and it does not initiate transcription from *E. coli* promoter sequences with any significant efficiency (Studier and Moffatt, 1986). Finally, T7 RNA polymerase is resistant to antibiotics such as rifampicin that inhibit *E. coli* RNA polymerase and consequently, the addition of rifampicin to cells that are producing T7 RNA polymerase results in the exclusive expression of genes under the control of a T7 RNA polymerase promoter (or pT7).

In most T7 expression systems, the RNA polymerase of bacteriophage T7 is not encoded for on the plasmid vector. Instead, a specific *E. coli* bacterial host lysogen is required, BL21(DE3). BL21 is an ideal host for expression of foreign gene products as it is deficient in both a protease of the heat shock system (encoded by the *lon* gene) and the OmpT outer membrane protease that can degrade proteins during their purification. Bacteriophage DE3 is a λ derivative that carries a DNA fragment containing the *lac*I gene, the *lac*UV5 promoter and the T7 phage *gene1* which encodes T7 RNA polymerase. Thus, the DE3 lysogen results in the *lac*UV5 promoter (which is derepressible with IPTG) being the sole promoter directing transcription of the T7 RNA polymerase gene (Studier *et al.*, 1990).

The BL21(DE3) strain used in the present study also contains the plasmid, pLysS, which carries the T7 lysozyme gene. pLysS causes a low level accumulation of T7 lysozyme which binds to T7 RNA polymerase and prevents transcriptional activity from traces of this enzyme which occur in the non-induced situation as a result of basal gene expression. This reduces the problems of low level transcription of toxic target genes in the uninduced cell.

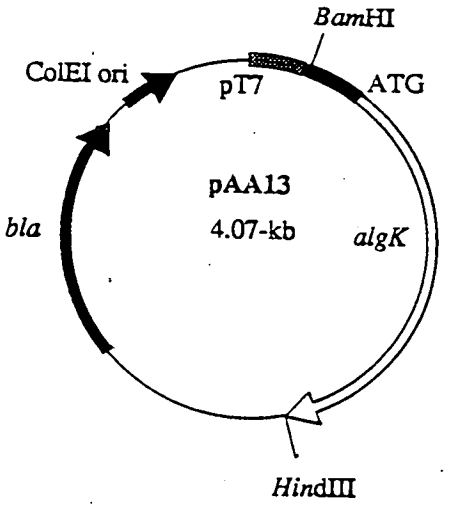
A second expression system was also examined as a means of AlgK production and this used the broad host range vector pMMB66(EH), which is

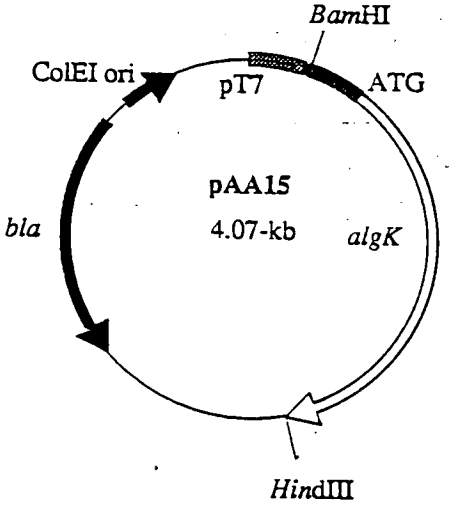
suitable for gene expression in most Gram-negative bacteria (Fürste *et al.*, 1986). The plasmid contains a polylinker cloning site downstream of the hybrid *tac* (*trp/lac*) promoter (Russell and Bennett, 1982). Expression of cloned sequences from p*tac* is negatively regulated by the *lac* repressor (which is synthesised from the *lacI*^{q} gene present on the vector) which binds to the *lac* operator in order to mediate repression: IPTG acts as a gratuitous inducer. The plasmid also contains the vegetative and transfer origins of replication from the broad host range plasmid R300B, enabling it to be mobilised for conjugal transfer by incompatibility group P1(IncP) or I (IncI) helper plasmids. In this case pRK2013, an IncP plasmid (Figurski and Helinski., 1977) was used in order to mobilise the plasmid into *P. aeruginosa* strains from *E. coli* cells.

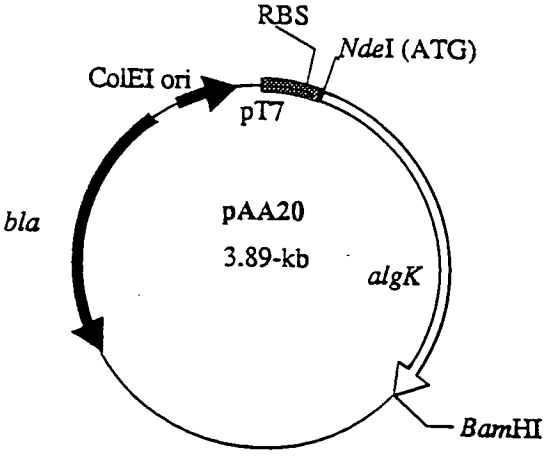
5.2.1. Subcloning and overexpression of AlgK in pT7 vectors

The initial expression survey of AlgK was performed utilising the vector pT7-5. A *Bam*HI/*Hin*dIII DNA fragment from pAA1 (containing the entire *Eco*RI/*Sma*I fragment from pRM812, see previous chapter) was subcloned into pT7-5 to form plasmid pAA15 (see figure 5.1). This plasmid was then transformed into the *E. coli* strain BL21(DE3) pLysS. T7 polymerase was induced with IPTG as described by Sambrook *et al.* (1989) and is described in detail in 2.3.3. Essentially, the appropriate transformed *E. coli* strain was grown in nutrient rich broth until an OD_{600} of 0.4-0.6 was reached, expression from the vector was induced by the addition of IPTG (1 mM final concentration). Samples were harvested at the point of induction and at hourly intervals. These samples were then denatured in Laemmli sample buffer (LSB) prior to separation by SDS-PAGE (polyacrylamide gel electrophoresis). However, Coomassie Brilliant Blue staining of the resultant gel failed to reveal any novel protein bands (see lane 4, figure 5.2).

Figure 5.1. Diagrammatic representation of AlgK expression vectors pAA13, pAA15 and pAA20 derived from the vectors pT7-3, pT7-5 and pT7-7 (derivatives of pT7-1 described by Tabor and Richardson, 1985) respectively. All three vectors contain a T7 RNA polymerase promoter, the *bla* gene encoding resistance to ampicillin (β-lactamase), the ColE1 origin of replication and *algK*. pAA13 has the ampicillin resistance gene in the same transcriptional orientation as the pT7 promoter and therefore will be overexpressed upon induction of T7 RNA polymerase, whilst pAA15 carries the *bla* gene in the opposite orientation. pAA20 differs from pAA13 and pAA15 in that it has a strong ribosome binding site (RBS) directly upstream of the *algK* start codon (ATG). The non-coding upstream region of *algK* is represented by ▒▒▒▒.


BamHI
ColEI ori
pT7
ATG
pAA13
4.07-kb
bla
algK
HindIII


BamHI
ColEI ori
pT7
ATG
pAA15
4.07-kb
bla
algK
HindIII


RBS
NdeI (ATG)
ColEI ori
pT7
pAA20
3.89-kb
bla
algK
BamHI

In order to deduce the possible reasons for lack of obvious AlgK expression, the same *Bam*HI/*Hin*dIII fragment was subcloned into pT7-3 to form plasmid pAA13 (see figure 5.1). pT7-3 differs from pT7-5 in that the gene (*bla*) for ampicillin resistance (β-lactamase) is in the opposite orientation and this results in the *bla* gene also being placed under the control of the pT7 promoter in the correct orientation for expression. Induction of T7 polymerase with IPTG should therefore result in the overexpression of both the cloned sequence when inserted in the polylinker between the *p*T7 and *bla* (see figure 5.1) and also, β-lactamase. Induction of BL21 (DE3) pLysS strains, carrying such constructs, was carried out as before and the samples were again analysed by SDS-PAGE followed by Coomassie Brilliant Blue staining. The gel revealed overexpressed β-lactamase but no other novel protein bands were detected (see lane 8, figure 5.2). This result predicted that T7 RNA polymerase was functioning properly, and thus the problem must either lie with inefficiency at the translational stage, instability of the overexpressed protein product or an error in the predicted open reading frame of the *algK* sequence.

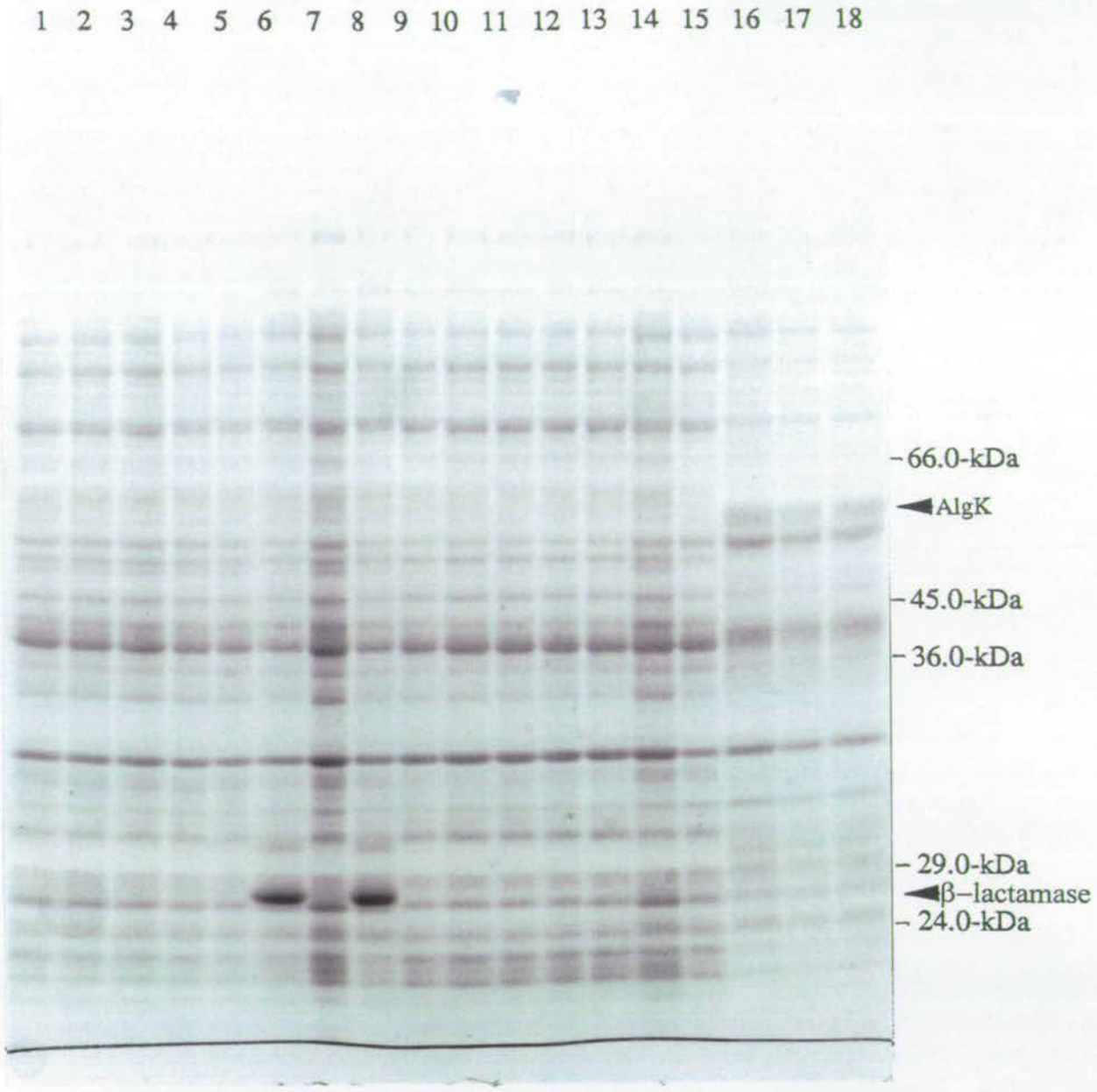
Previous overexpression studies with alginate biosynthetic genes by Sá-Correia *et al.* (1987), Deretic *et al.* (1987), Darzins *et al.* (1984) used maxicell analysis which required [^{35}S]methionine labelling to detect over-produced protein. Zielinski *et al.* (1991) also utilised the pT7-5 vector for expression of AlgC (phosphomannomutase) but over-production was only visualised by [^{35}S]methionine labelling. This earlier work suggested that the latter proteins are expressed at low levels and may reflect on inefficient translation. Since this was a formal possibility for *algK*, the experiment was repeated using the pT7-7 vector in order to increase the efficiency of translation. pT7-7 differs from pT7-3 and pT7-5 by having a strong ribosome binding site (RBS) and translational start site upstream of the polylinker sequence. To insert the *algK* gene into pT7-7 it was amplified using the polymerase chain reaction, with pAA1 as the DNA template. Synthetic oligonucleotide primers

were designed, such that the upstream primer (primer 1) contained an *Nde*I site encompassing the natural ATG site of *algK* (thus generating a complete ATG start site at the correct position downstream of the RBS when subcloned into pT7-7). The downstream primer (primer 2) contained a *Bam*HI site adjacent to the complementary sequence of the last 22 bases of the *algK* ORF (including the TGA stop codon). The resulting PCR product (*algK* with an N-terminal *Nde*I site and C-terminal flanking *Bam*HI site) was then digested with *Nde*I/*Bam*HI and ligated into the respective sites of pT7-7 to form plasmid pAA20 (see figure 5.1). This directly replaced the initiating pT7-7 ATG with the ATG translational start codon of *algK* followed by the complete *algK* ORF in close proximity to the strong RBS of the pT7-7 vector. pAA20 was transformed into BL21(DE3) pLysS and induced with IPTG, as described previously. SDS-PAGE analysis of the fractions was followed by Coomassie Brilliant Blue staining and revealed a novel protein band only in the induced pAA20 lanes. This polypeptide ran as a broad band at approximately 50 to 53-kDa (as shown in figure 5.2) which is in agreement with the molecular weight predicted for the AlgK pre-protein, as deduced from the *algK* nucleotide sequence (52.5-kDa).

5.2.2. **Overexpression of AlgK in *P. aeruginosa***

In order to deduce whether AlgK was expressed in *P. aeruginosa* as it is in *E. coli* the *algK* gene was cloned into the broad host, p*tac* expression vector pMMB66(EH). Initial experiments mimicked those of the pT7 system with a *Bam*HI/*Hin*dIII fragment from pAA1 being subcloned into the vector. The same expression conditions used for the pT7 systems were employed with the pMMB66(EH) construct but as before no novel protein bands were visualised (data not shown). Since the previous experiments predicted that inefficient translation may be occurring from the natural AlgK RBS in the pT7 vectors, it seemed likely that a

Figure 5.2. SDS-PAGE analysis of the expression of AlgK from pAA15, pAA13 and pAA20 in *E. coli* BL21(DE3) pLysS cells. BL21(DE3)pLysS cells carrying pT7-5/3 or pAA15/13 were either uninduced (lanes 1, 3, 5 and 7 respectively) or were treated with IPTG, to induce the expression of T7 RNA polymerase from the *lac*UV5 promoter for 4 hours (lanes 2, 4, 6 and 8 respectively). pT7-7 or pAA20 were either uninduced (lanes 9 and 14 respectively) or were treated with IPTG for 1-4 hours. Samples are shown for cells carrying BL21(DE3) [pLysS, pT7-7] induced for 1, 2, 3 or 4 hours (lanes 10-13 respectively), or for BL21(DE3) carrying [pLysS, pAA20] induced for the same time periods (lanes 15-18 respectively). Proteins from lysed cells were denatured in Laemmli sample buffer (Laemmli, 1970), separated by SDS-PAGE, and stained with Coomassie Brilliant Blue. Numbers on the right indicate the locations of the molecular size markers, the position of the protein expressed from *algK* and overexpressed β-lactamase are also shown on the right.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
66.0-kDa
AlgK
45.0-kDa
36.0-kDa
29.0-kDa
β–lactamase
24.0-kDa

similar phenomenon was being observed in the pMMB66 construct. In order to circumvent this problem it was decided to incorporate the RBS of pT7-7 into the pMMB66(EH) expression vector. Thus, another synthetic oligonucleotide was designed which would result in the amplification of the powerful RBS from pT7-7 along with *algK* when pAA20 was used as the DNA source. Primer 3 primes upstream of both the RBS and the ATG start of *algK* in pAA20 and also contains an *Eco*RV restriction site towards the 5' end of the primer. Thus, PCR with primer 3, combined with (the C-terminal) primer 2 results in the amplification of the *algK* gene, containing the strong RBS from pT7-7, with two unique flanking restriction sites: *Eco*RV at the 5' end and *Bam*HI at the 3' terminus.

Digestion of DNA with *Eco*RV results in the production of a 'blunt' end and as such the PCR fragment (once digested with *Eco*RV and *Bam*HI) could be cloned into the *Sma*I/*Bam*HI sites of pMMB66(EH) (as digestion of DNA with *Sma*I also results in 'blunt' ended DNA) in the correct orientation with the N-terminus end immediately downstream of the p*tac* promoter. Cloning of the PCR product into pMMB66(EH) generated plasmid pAA23 (see figure 5.1).

Initially, pAA23 was introduced into *E. coli* strain HB101, for triparental mating into *P. aeruginosa*. pRK2013 (also within HB101) was used as a helper plasmid to mobilise the recombinant pAA23 into *P. aeruginosa*. pRK2013 contains the RK2 *tra* functions and the kanamycin resistance gene ligated into a ColE1 replicon. *P. aeruginosa*, non-mucoid strain 8873 (Darzins and Chakrabarty, 1984) was used as the recipient strain and triparental filter matings were performed as described by Sá-Correia *et al.* (1987). After matings, exconjugants were selected by ampicillin resistance on a minimal agar, thus eliminating HB101 which is a multiple auxotroph.

Overexpression from the p*tac* promoter of pAA23 was carried out in both *E. coli* (HB101) and PA8873 and the induction protocol used was identical to that used

for the pT7 systems. Denatured samples were separated by SDS-PAGE and stained with Coomassie Brilliant Blue (see figure 5.3). The AlgK polypeptide migrated as in the pT7 system as a broad band at approximately 50 to 53-kDa protein. This size is consistent with the predicted molecular weight of the unprocessed protein encoded by the *algK* gene, however the size of the band may be sufficient to obscure any processing (cleavage of the putative signal sequence would result in a 3-kDa size decrease).

5.2.3. [^{35}S]Methionine labelling of AlgK under the T7 and *tac* promoters

In order to visualise any post-translational processing of AlgK a series of [^{35}S]methionine polypeptide labelling experiments were performed. Initially, [^{35}S]methionine labelling was performed using pAA20 (pT7-7 with *algK* insert) in BL21 (DE3) [pLysS] following the procedures of Franklin *et al.* 1994 (as described in 2.3.3). Cells were grown in minimal medium (supplemented with an 18 amino acid mixture) until an OD_{600} of 0.6-0.8 is reached, expression from the T7 promoter was then induced by the addition of IPTG (1 mM final concentration). Proteins were then labelled with [^{35}S]methionine for 1, 5, and 20 minutes and were then harvested, denatured and separated by SDS-PAGE (as outlined in Franklin *et al.*, 1994). Alternatively, rifampicin (200 μg/ml, final concentration) was added during IPTG induction stage and then [^{35}S]methionine labelling continued as previously. Examination of [^{35}S]methionine labelled proteins after autoradiography revealed two unique bands at approximately 53 and 50-kDa (figure 5.4, lanes 2, 3, and 4, or the rifampicin treated lanes 6, 7, and 8), that were not seen in the control strain BL21 (DE3) carrying [pLysS and pT7-7] (figure 5.4 lanes 1 and 5). This suggests that processing of the signal peptide is occurring.

The [^{35}S]methionine labelling experiments were subsequently extended to AlgK synthesis in *P. aeruginosa*, to see if the same modification processes were

Figure 5.3. SDS-PAGE analysis of the expression of AlgK from pAA23 in *P. aeruginosa* strain 8873, and *E. coli* strain HB101. Lanes 1 and 2 show HB101 (pMMB66 EH) treated with IPTG to induce expression from the *tac* promoter for 4 hours; lanes 3-7 contain HB101 (pAA23) treated with IPTG for 1, 2, 3 and 4 hours, respectively. Lanes 8 and 9, contain PA8873 and PA8873 (pMMB66 EH) treated with IPTG for 4 hours whilst lanes 10-14 PA8873 (pAA23) treated with IPTG for 1, 2, 3 and 4 hours, respectively. Cells were lysed and denatured in LSB (Laemmli, 1970), separated by SDS-PAGE, and stained with Coomassie Brilliant Blue. Numbers on the right indicate molecular size markers, the protein expressed from *algK* is also indicated.

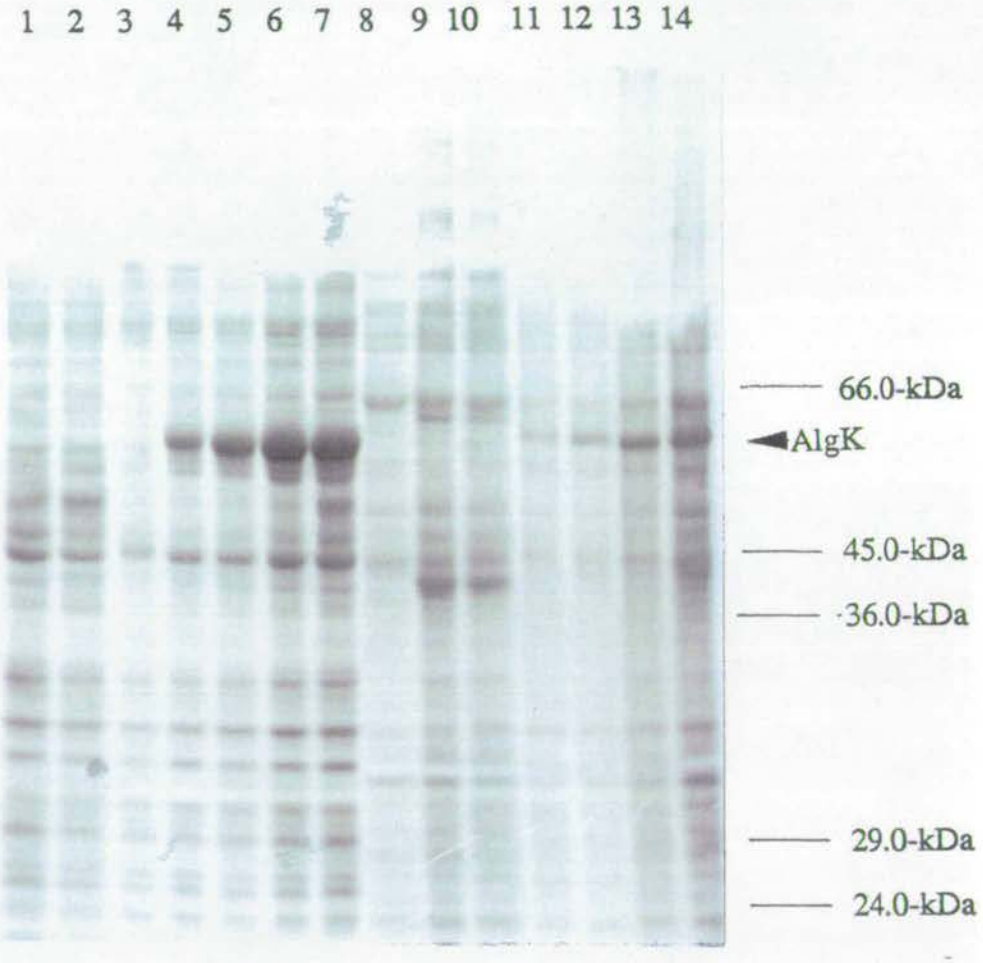
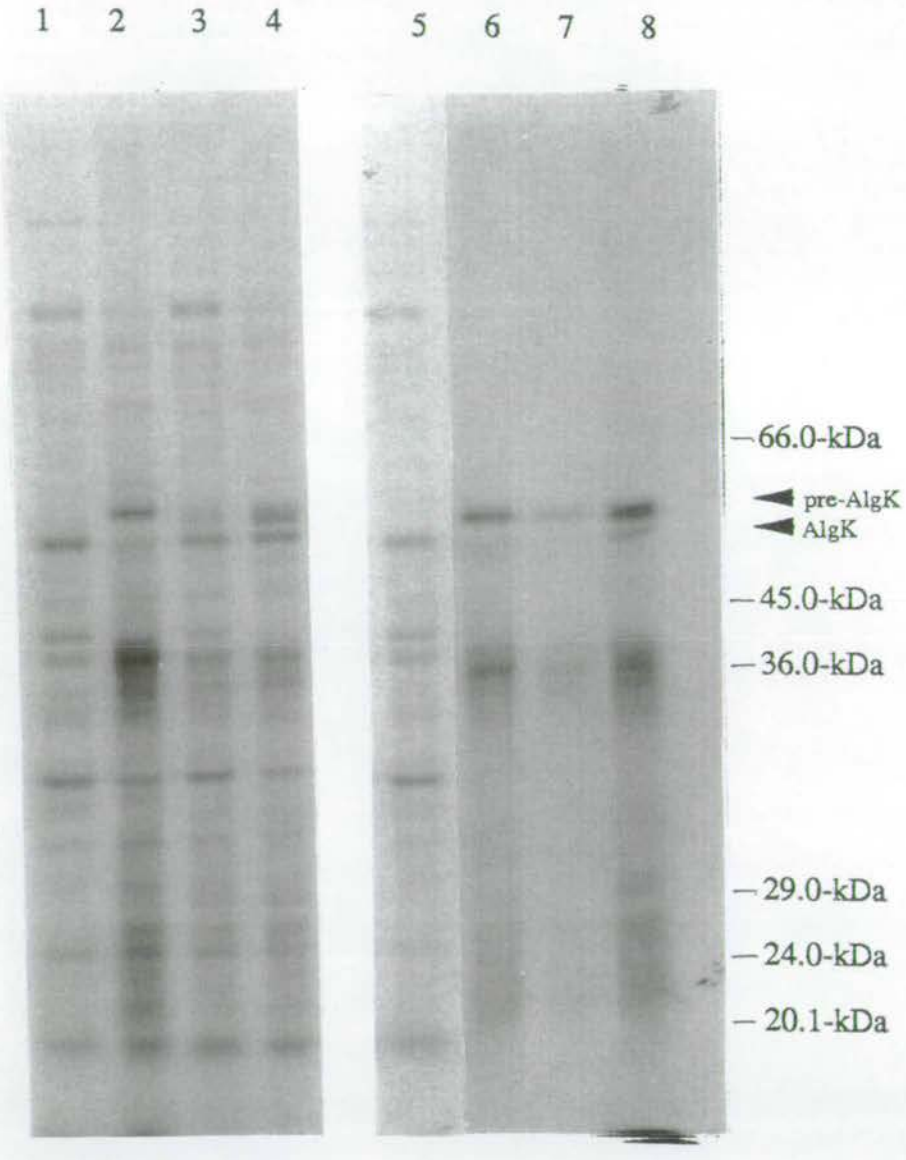

1 2 3 4 5 6 7 8 9 10 11 12 13 14
66.0-kDa
AlgK
45.0-kDa
36.0-kDa
29.0-kDa
24.0-kDa

Figure 5.4. Autoradiogram of proteins expressed from the T7 promoter in *E. coli* strains BL21(DE3)pLysS. Samples were labelled with [^{35}S]methionine for various time intervals, denatured in LSB (Laemmli, 1970) and separated by SDS-PAGE. Samples were as follows: Lane 1, BL21(DE3) [pLysS, pT7-7] treated with IPTG and labelled with [^{35}S]methionine for 5 minutes; Lanes 2-4, BL21(DE3)[pLysS, pAA20] treated with IPTG and labelled with [^{35}S]methionine for 1, 5 and 20 minutes respectively. Lane 5 BL21(DE3)[pLysS, pT7-7] treated with IPTG and rifampicin and labelled with [^{35}S]methionine for 5 minutes; Lanes 6-8 BL21(DE3)[pLysS, pAA20] treated with IPTG and rifampicin and labelled with [^{35}S]methionine for 1, 5, and 20 minutes, respectively. Putative AlgK preprotein and mature protein is marked. The size and positions of molecular weight markers are indicated.

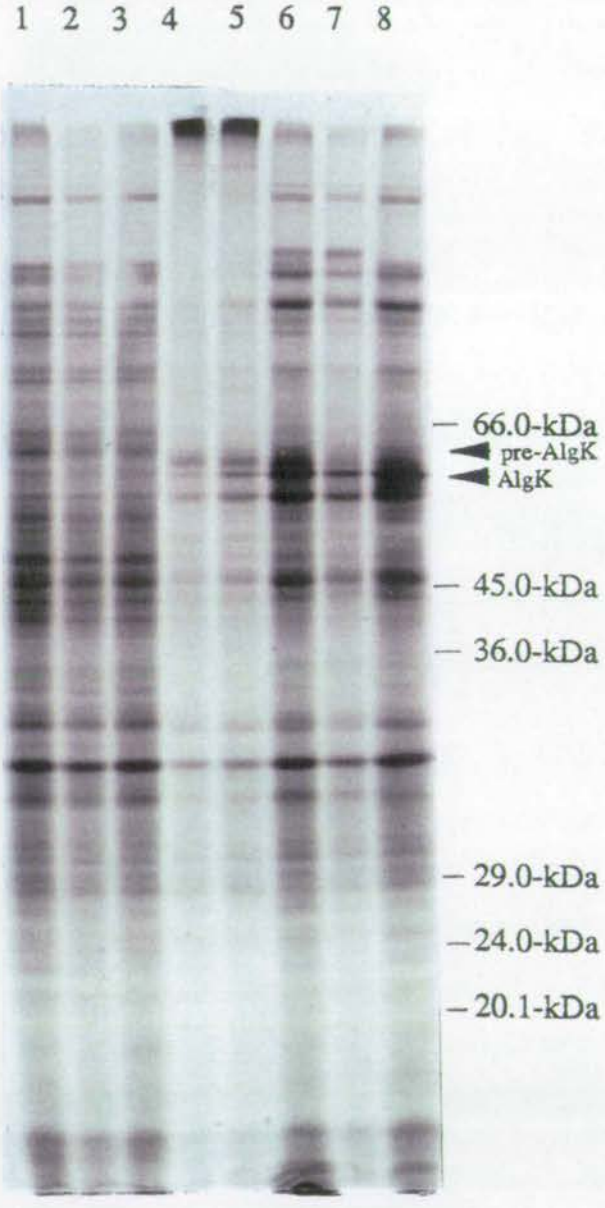

1
2
3
4
5
6
7
8
66.0-kDa
pre-AlgK
AlgK
45.0-kDa
36.0-kDa
29.0-kDa
24.0-kDa
20.1-kDa

exhibited. Pulse-chase labelling was performed on PA8873(pAA23) using the methods of Franklin *et al.* (1994) as described in 2.3.3. Cells were grown in minimal media (to an OD_{600} of 0.6-0.8) prior to being treated with IPTG to induce expression from the *tac* promoter. A 2 minute pulse of [^{35}S]methionine was followed by chasing with 'cold' methionine for 0, 3, 5, 15, and 30 minutes, proteins were denatured, separated by SDS-PAGE and then autoradiographed. Analysis of ^{35}S-labelled proteins was not as clear as with the pT7 system but nevertheless AlgK did appear to be modified with time, from an original higher band at approximately 53-kDa to a lower band approximately 2-3-kDa smaller (see lanes 4-9 of figure 5.5).

5.3. DISCUSSION

Initial overexpression studies of AlgK, utilising vectors containing pT7 and p*tac* promoters in *E. coli* and *P. aeruginosa* respectively, failed to produce any positive results with Coomassie Brilliant Blue staining after SDS-PAGE separation. This problem was attributed to inefficient translation as T7 RNA polymerase was found to transcribe and translate β-lactamase efficiently (when in the correct orientation). In order to circumvent poor translational efficiency, studies with the vector pT7-7 were undertaken. This vector also utilises a T7 promoter but has a strong ribosome binding site incorporated upstream of an ATG translational start codon. Cloning of *algK* into this vector allows for transcription directed from the pT7 promoter and translation directed from the strong RBS. Induction of *algK* transcription in this vector with IPTG resulted in observable overexpression of a polypeptide which migrated as a 53-kDa band. This strongly correlates with the protein size of 52,470-Da predicted from the nucleotide sequence of the open reading frame of *algK*.

Figure 5.5. Autoradiogram of pulse-labelled proteins in *P. aeruginosa* cells expressing *algK*. Cells were labelled with [^{35}S]methionine for intervals between 0 and 30 minutes. Lanes 1, 2, and 3 show PA8873 (pMMB66 EH) labelled with [^{35}S]methionine for 2 minutes and chased for 0, 5 and 30 minutes respectively, with unlabelled methionine. Lanes 4, 5, 6, 7 and 8 show PA8873 (pAA23) labelled with [^{35}S]methionine for 2 minutes and then chased with 'cold' methionine for 0, 3, 5, 15 and 30 minutes, respectively. Numbers on the right indicate the positions of molecular size markers, Pre-AlgK and AlgK indicate the precursor and mature proteins expressed by *algK*, respectively. Cells were lysed and proteins were denatured by boiling in LSB (Laemmli, 1970) and samples were separated by SDS-PAGE.


1 2 3 4 5 6 7 8
66.0-kDa
pre-AlgK
AlgK
45.0-kDa
36.0-kDa
29.0-kDa
24.0-kDa
20.1-kDa

Visible overexpression of AlgK within *P. aeruginosa* (PA8873) also required the incorporation of a strong ribosome binding site and this was derived from the pT7-7 vector. Expression of *algK* within PA8873, under IPTG inducing conditions, resulted in the production of an equivalent protein band as observed for the pT7-7 system migrating at about 50 to 53-kDa. The close similarity between the predicted protein size of AlgK and that found in the expression studies suggests that the single long open reading frame predicted from the deoxy nucleotide sequence is correct.

Computer analysis of the predicted amino acid sequence of AlgK revealed the potential presence of a signal sequence and lipid attachment site. Such signal sequences are cleaved by signal peptidase II (on the periplasmic face of the cytoplasmic membrane) generating an N-terminal cysteine residue which is modified by attachment of fatty acids, allowing the protein to be anchored to the inner or outer membrane. Synthesis of an AlgK lipoprotein would result in the removal of a 27 amino acid signal sequence, with the mature AlgK protein having a molecular weight 3-kDa less than the pre-AlgK protein. In the present study, overexpression followed by Coomassie Brilliant Blue staining was not a sensitive enough procedure to detect signal sequence cleavage under the conditions examined and only revealed a single diffuse band from approximately 50-53-kDa (see figures 5.2 and 5.3). Thus, the expression survey was extended to [^{35}S]methionine labelling of AlgK in both *E. coli* and *P. aeruginosa* in order to detect the occurrence of any post-translational modification event. Pulse labelling of AlgK, expressed in *E. coli* strain BL21(DE3) pLysS, revealed a definite change in migration pattern over the 20 minute pulse, with the upper (~53-kDa) band being converted to a faster migrating (~50-kDa) band. [^{35}S]methionine labelling of AlgK in PA8873 produced a similar, but, less clear result (see figure 5.5). Taken in conjunction with the labelling in *E. coli*, however, it can be concluded that AlgK does appear to be undergoing a post-translational

modification event, in *P. aeruginosa*, as would be predicted from the N-terminal amino-acid sequence.

The cleavage of signal sequences in general, is commonly confirmed by N-terminal sequencing of the mature protein. Experiments on *P. aeruginosa* biosynthetic proteins AlgE (Chu *et al.*, 1991), AlgF (Shinabarger *et al.*, 1993) and AlgG (Franklin *et al.*, 1994) have revealed the N-terminal amino acid, of the mature protein, following cleavage of their respective signal sequences from the pre-protein. Identification of a terminal glycerylcysteine residue (which has a slightly different profile to cysteic acid, Perumal and Minkley, 1984) in the mature AlgK protein by N-terminal sequencing would be definitive for a lipoprotein, as only signal peptidase II, the lipoprotein peptidase, recognises this signal sequence, cleaves at a cysteine residue and generates an N-terminal glyceryl cysteine residue (Hayashi *et al.*, 1984).

Confirmatory experiments might involve globomycin and radioactive palmitate. The activity of lipoprotein signal peptidase (signal peptidase II) is specifically and characteristically inhibited by the action of globomycin. Incorporation of globomycin into protein expression experiments results in accumulation of non-modified, high molecular weight, pre-proteins (as typically exhibited in experiments by Yamaguchi *et al.*, 1988) and as such predicts that the proteins are processed lipoprotein. [^{3}H]Palmitic acid is incorporated into lipoproteins at the signal peptidase cleavage stage, and thus labels lipoproteins specifically. [^{3}H]Palmitic acid is commonly used in conjunction with normal protein expression systems for prediction of the subcellular localisation of such lipoproteins, once membrane fractions have been prepared (Yamaguchi *et al.*, 1988). N-terminal sequencing of the AlgK mature protein combined with globomycin and [^{3}H]palmitic acid would definitively prove that AlgK was indeed a lipoprotein. Radioactive labelling of the lipid moiety of an AlgK lipoprotein would allow identification of the

membrane to which it is usually anchored but would not reveal the subcellular location of other regions of the polypeptide.

Chapter 6

Analysis of the subcellular location and topology of AlgK using β-lactamase as a reporter gene.

6.1. INTRODUCTION

Analysis of the predicted amino acid sequence of AlgK has revealed a protein with a putative lipoprotein signal sequence and lipid attachment site. Post-translational modification of lipoproteins is a periplasmic event, signal peptidase II being a polytopic cytoplasmic membrane protein which cleaves the signal peptide as the protein emerges through the inner membrane into the periplasm. Once cleaved the lipoprotein is anchored at the N-terminus by fatty acids, to one of three possible subcellular locations: the outer leaflet of the inner membrane, the inner leaflet of the outer membrane or the outer leaflet of the outer membrane (See Pugsley, 1993 for review). However, Gennity *et al.* (1992) demonstrated that the amino acid nature of the rest of the polypeptide dictates how the lipoproteins are assembled at these locations. It is possible that lipoproteins may have polypeptide segments that span between the membranes or interact peripherally with one or other of the membranes in the cell envelope (Pugsley, 1993). Lipoproteins can even be anchored in one membrane by fatty acids and in another by a typical membrane anchor, a system that has been artificially generated by Gennity *et al.* (1992). Despite having obtained the nucleic acid sequence of *algK* and having examined its production the subcellular location of AlgK and the presence of any membrane-spanning topology still remained unclear. Thus, in order to clarify the undefined subcellular location and possible membrane topology, a series of products from *algK-'bla* gene fusions were constructed and immunological studies were undertaken.

To analyse the subcellular location and possible membrane interaction of AlgK, the plasmid pJBS633 (Broome-Smith and Spratt, 1986) was used. pJBS633 is a derivative of pBR322 and encodes proteins that confer both tetracycline and kanamycin resistance but its main research value is that it also contains the *'blaM* gene. *'blaM* encodes only the mature form of TEM β-lactamase, and lacks coding

information for the N-terminal signal peptide which is required for targeting the polypeptide across the cytoplasmic membrane to the periplasm (Kadonago *et al.*, 1984). This β-lactamase derivative also lacks a ribosome binding site and promoter and thus there is no expression of this β-lactamase from pJBS633 under normal conditions.

In its normal periplasmic location, β-lactamase protects *E. coli* from ampicillin-induced lysis by hydrolysing the antibiotic (β-lactam ring) before it can disrupt the correct formation of the periplasmic peptidoglycan layer. Ampicillin acts by covalent inactivation of the cytoplasmic membrane-bound penicillin binding proteins. Cytoplasmic forms of mature β-lactamase cannot intervene between the incoming antibiotic and its cytoplasmic membrane anchored, periplasmically oriented, enzyme target(s), leaving the host cell totally unprotected from lysis by ampicillin. Thus a single bacterium, expressing periplasmic TEM β-lactamase, if inoculated onto agar containing ampicillin, can inactivate the β-lactam and thus survive, multiply and form a colony. However, a bacterium synthesising cytoplasmic β-lactamase is as sensitive to the bactericidal action of ampicillin as a non-β-lactamase producing cell and, as such, is duly lysed.

Although individual cells producing cytoplasmic β-lactamase lyse and therefore fail to form colonies on agar containing ampicillin, such cells if inoculated at sufficiently high density, do grow. In this instance, ampicillin-induced lysis of a number of bacteria results in the release of the cytoplasmic β-lactamase into the medium, and the subsequent hydrolysis of the ampicillin, thus enabling the remaining cells to survive and grow (Kadanaga and Knowles, 1985). The vector pJBS633 exploits these features as a means for determining the subcellular location of polypeptide domains.

In order for the mature form of TEM-β-lactamase encoded in pJBS633 to be expressed, its N-terminus has to be fused in-frame, to an upstream gene or fragment

containing an efficient ribosome binding site so that a fusion protein is synthesised. In high density, cells expressing the in-frame fusion protein will grow on agar containing ampicillin even if the fusion protein is located in the cytoplasm or periplasm. If however, the fusion of the β-lactamase is out-of-frame with the upstream open reading frame then cells will show no significant resistance to ampicillin (<5μg/ml), even when plated at high concentration. In this way, by constructing an in-frame fusion between *algK* and *'blaM* (the gene encoding mature β-lactamase), the predicted reading frame of *algK* can be confirmed by showing that cells expressing this fusion protein are resistant to ampicillin when plated in suitable density.

A second use of pJBS633 is the analysis of the membrane topology of the target protein. Again, this involves gene fusion studies, but this time, β-lactamase is fused to progressively truncated forms of the target protein (by *in-situ* deletion of the target gene within pJBS633 from the 3' end, before fusion formation). Since polytopic membrane proteins span the cytoplasmic membrane a number of times, particular polypeptide domains will be located in either the cytoplasm or periplasm or embedded in the membrane. So, if β-lactamase is fused to part of a membrane protein that is normally periplasmic, individual cells will be protected and thus survive high concentrations of ampicillin (>800μg/ml). Alternatively, if the β-lactamase is fused to a region of the protein that is normally cytoplasmic then individual cells are unprotected from ampicillin action and are lysed when plated on agar containing 5μg/ml ampicillin. When β-lactamase is fused to a transmembrane region of the protein an intermediate level of resistance is expected. Nucleotide sequencing across the DNA fusion junction then allows localisation of the fusion point at the polypeptide level. Thus, with a sufficient number of fusions, the β-lactamase (pJBS633) system can be used to generate two dimensional topological maps for bacterial inner membrane proteins (Broome-Smith and Spratt, 1986). The

ability or inability of the fusion protein to protect single cells against lysis by ampicillin should indicate whether the β-lactamase moiety of the fusion protein has been translocated across the cytoplasmic membrane to the periplasm or is retained within the cytoplasm.

pJBS633 has been used successfully to predict cytoplasmic, periplasmic and inner membrane-spanning domains for a variety of bacterial proteins. These include the cytoplasmic, membrane-anchored, penicillin binding proteins 1B and 3 of *E. coli* (PBP3 is a predicted lipoprotein)(Edelman *et al.*, 1987; Bowler and Spratt, 1989, respectively), the HlyB and HlyD proteins involved in the translocation of haemolysin through the inner membrane and periplasm (Wang *et al.*, 1991), the inner membrane subunit of the anion translocating ATPase (ArsD) of *E. coli* (Wu *et al.*, 1992) and the anchor sub-unit of dimethyl sulfoxide reductase (DmsC) of *E. coli* (Weiner *et al.*, 1993).

6.2.1. **Construction of a translational fusion between AlgK and mature TEM β-lactamase**

Construction of pAA9

Studies of an *algK-'bla*M fusion vector, initially involved subcloning the *Bgl*II/*Sma*I fragment from vector pAA1 (see figure 4.1) directly into the *Bam*HI/*Eco*RV sites of pJBS633 to form pAA9. Generation of a fusion system from this construct was found to be inadequate, with little expression of β-lactamase even when fused in-frame. Poor expression was concluded to be due to the same inefficient translation problems exhibited with the expression systems (see previous chapter), thus a new construct was designed and assembled which would contain an efficient ribosome binding site.

Construction of pAA22

pAA22 was formed from the cloning of *algK* gene into pJBS633. The *algK* gene used for this construct was amplified by PCR using pAA20 as the DNA source. Primer 3 was used (see previous chapter) as the N-terminal primer to allow the amplification of an efficient ribosome binding site along with *algK*. The downstream C-terminal primer involved (primer 4) contains *Pvu*II and *Bam*HI restriction sites which prime from the complementary strand and immediately downstream from and overlapping the complementary sequence of the TGA stop codon. The resulting PCR product (*algK* with flanking *Eco*RV and *Bam*HI restriction sites were digested with *Eco*RV and *Bam*HI and ligated into the respective sites located within the tetracycline gene of pJBS633. These digests allow the retention of the tetracycline promoter which acts as the transcriptional promoter for *algK*'-'*blaM* fusions.

Fusion of *algK* to *'blaM* in pAA22

The gene encoding the mature form of β-lactamase (*'blaM*) within pJBS633 has been engineered to introduce a blunt end cloning site (either *Pvu*II or *Sma*I, a derivative available in the laboratory) at the original junction of the coding region and signal peptide. When amplified by primers 3 and 4 the *algK* gene contains a *Pvu*II restriction site at the C-terminus. To construct a directed translational fusion, pAA22 was digested with *Pvu*II and *Sma*I, deleting a 1.7-kb fragment of 'spacer' DNA from pJBS633, followed by religation. This resulted in plasmid pAA35. pAA35 was transformed into JM101 selecting for kanamycin resistance. Transformants were then 'patched' onto LB agar containing 200 μg/ml ampicillin. All patched transformants grew on 200 μg/ml ampicillin indicating that the proposed open reading frame for *algK* was correct. To confirm this result, a further, more reliable 'in-frame' spot-test was employed (Broome-Smith and Spratt, 1986). In this test, a 6 μl drop of an undiluted overnight culture is spotted onto an LB test plate

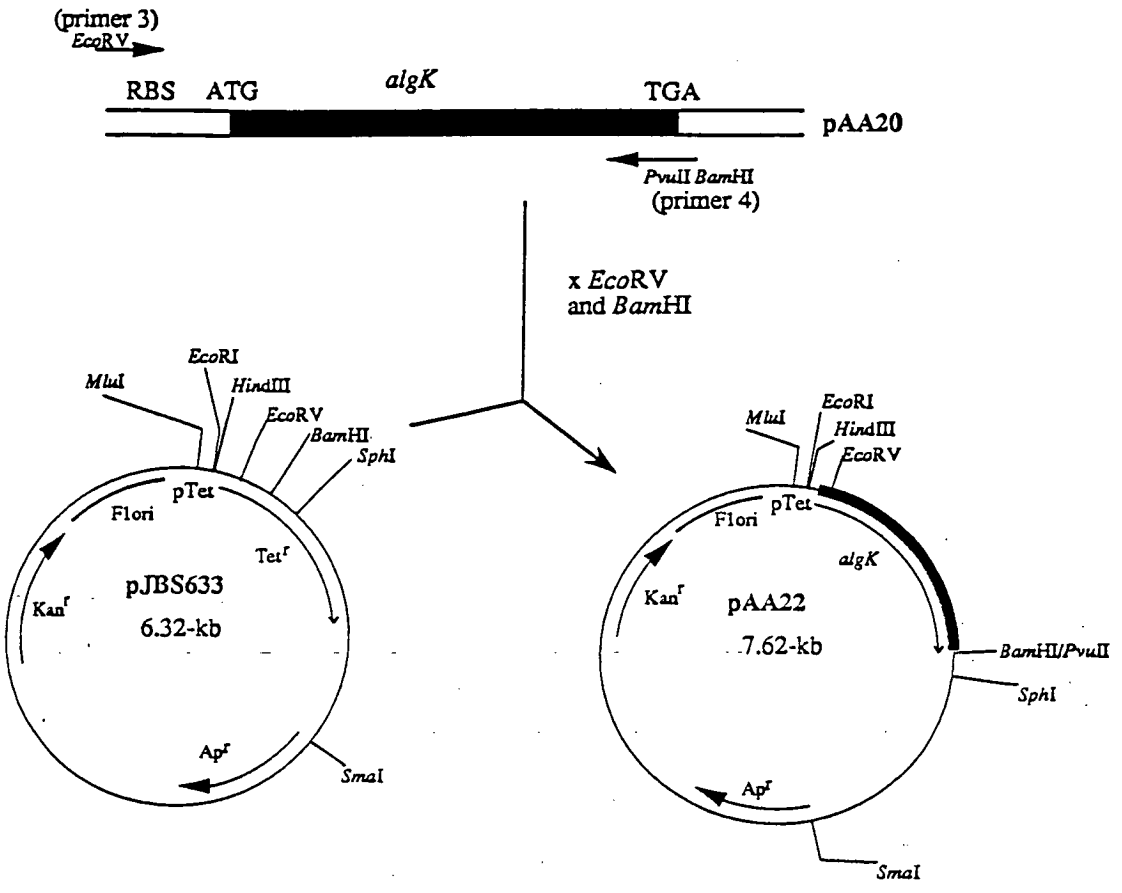


Figure 6.1a. Construction of pAA22. pAA22 is formed by the insertion of an *Eco*RV-*Bam*HI PCR amplified fragment from pAA20 (encoding *algK*) into pJBS633. pAA35 is formed from the deletion of the *Sma*I-*Pvu*II fragment from pAA22. A random series of fusions of the *algK* gene to the coding region for mature TEM β-lactamase were produced as described in text. The solid box represents the *algK* gene in pAA20 or pAA22. Abbreviations used are RBS, ribosome binding site; Ap^r, ampicillin resistance gene; Tet^r, tetracycline resistance gene; Kan^r, kanamycin resistance gene.

containing 200 μg/ml ampicillin. Only in-frame recombinants will grow under these conditions and as such JM101 bearing pAA35 was found to behave in this manner.

In order to predict the subcellular location of the AlgK-β-lactamase fusion protein, single cell minimum inhibitory concentration (MIC) experiments were performed (as described by Broome-Smith and Spratt, 1986). An overnight culture of JM101(pAA35) was diluted 10^5 fold to give a bacterial concentration of approximately 10 bacteria/μl and 4 μl aliquots were spotted onto LB plates, containing increasing levels of ampicillin. Single cell minimum inhibitory concentrations (MIC) are defined as the minimum concentration of antibiotic (in this case ampicillin) required to prevent colony formation under these culture conditions. The single cell MIC obtained for JM101(pAA35) cells was in excess of 800 μg/ml ampicillin; resistance to this degree is only exhibited if the β-lactamase fusion moiety has a periplasmic location. Thus, the C-terminus of AlgK is predicted to have a periplasmic location.

6.2.2. **Analysis of the membrane interaction of *algK* using mature TEM β-lactamase as a reporter gene**

pAA22 was also used to construct a series of in-frame fusions of the mature form of β-lactamase to random positions within AlgK. The subcellular location of the β-lactamase moiety of the fusion protein could then be deduced by the level of ampicillin resistance conferred on single cells of the host *E. coli* strain. Fusions directing the translocation of β-lactamase to the periplasm, will provide individual *E. coli* cells with a high level of resistance to ampicillin. However, fusions that result in the cytoplasmic location of the β-lactamase moiety will confer negligible ampicillin resistance to individual *E. coli* cells.

Construction of a series of *algK*-β-lactamase fusions

The subcellular localisation of domains of AlgK was analysed in *E. coli* by generating a range of fusions of *algK* to β-lactamase in pAA22. The initial stage in synthesising a spectrum of truncated *algK* fusions requires construction of a series of unidirectional deletions from the 3' terminus of *algK*. Deletions from 3' to 5' within *algK* were achieved using the unidirectional DNA digestion activity of exonuclease III. This enzyme catalyses the stepwise removal of 5' mononucleotides from the recessed or blunt 3'-hydroxyl terminus of double stranded DNA. However, protruding 3' termini are completely resistant to the activity of the enzyme (Rogers and Weiss, 1980) and this lack of activity at a protruding 3' terminus provides the basis for the generation of unidirectional DNA deletions within pJBS633 and derived vectors such as pAA22. To achieve this situation, the plasmid is digested with two restriction enzymes whose sites of cleavage lie between the target DNA (*algK*) and *'blaM* (the gene encoding mature β-lactamase). The enzyme that cleaves nearer the target sequence must generate a 5'-protrusion or a blunt end, whilst the second enzyme which digests further downstream of *algK* (nearer *'blaM*), generates a 3'-protrusion or overhang that protects the remainder of the vector from exonuclease III digestion.

pAA22 was completely digested with *Bam*HI, which results in a 5' DNA protrusion at the 3' terminus of *algK* (and is thus a substrate for exonuclease III 3'-5' digestion), and *Sph*I which generates a 3' overhang upstream of mature β-lactamase (thus protecting *'blaM* from exonuclease III activity). Complete digestion of pAA22 with *Bam*HI and *Sph*I, followed by the addition of exonuclease III, therefore results in appropriate termini for the upstream and unidirectional 3' to 5' deletion of *algK*.

Following digestion of pAA22 with *Bam*HI and *Sph*I, linearised pAA22 was digested with exonuclease III over a time course of 8 minutes with aliquots taken at 30 second intervals (exonuclease III removes approximately 200 nucleotides per

minute under the conditions used - described in 2.3.4) and then treated with S1 nuclease in the presence of zinc cations at acid pH, conditions which inhibits the action of exonuclease III. S1 nuclease removes the remaining undigested single stranded DNA. Klenow polymerase was then added to the aliquots, to create blunt DNA ends at the 3' end of the truncated *algK* remnants and *Sma*I was added to generate the unique blunt end at the start of the *'blaM* coding sequence. The DNA was ligated and then transformed into *E. coli* JM101 cells and kanamycin resistant transformants were selected.

In-frame fusions to β-lactamase were initially identified by patching onto LB ampicillin (200μg/ml) plates and surviving transformants were then retested using the more reliable spot-test (as described previously) to predict 'in-frame' fusions. Putative in-frame *algK'-'blaM* fusions were then sequenced in order to define the position of the junctions and to verify the continuity of the open reading frame.

Nucleotide sequencing of fusion junctions

In order to avoid the problems associated with sequencing double stranded (ds) DNA derived from *P. aeruginosa* (Deretic *et al.*, 1987 etc) all sequencing reactions were performed on single stranded (ss) DNA using the GTP analogue 7-deaza-dGTP (see chapter 4). Single stranded DNA was generated from pAA22 by superinfection with R408 helper phage (Russel *et al.*, 1986) instead of IR1 which is suggested by Broome-Smith and Spratt (1986). PEG-purified, ssDNA (as described in 2.3.2) was then used for sequencing reactions. The nucleotide sequence across the fusion junctions was obtained by using a specific β-lactamase primer that is complementary to codons 14-18 of the mature form of β-lactamase and primes the sequencing reactions across the fusion junction towards the cloned fragment. Sequencing was performed on both the predicted in-frame clones and some control, out-of-frame, clones to confirm that the system was functioning properly. Using this

method, nine in-frame fusions were isolated out of a total of 48 transformants and the precise locations of the fusion junctions were determined. The fusion junction of pAA35 was also sequenced using this protocol and was found to be in-frame with the *'blaM* gene.

Prediction of the subcellular location of the β-lactamase moieties of the AlgK-β-lactamase fusion proteins

Using the single cell MIC method previously described for pAA35, the subcellular locations of the β-lactamase moiety of the fusion proteins can be predicted. Fusions that result in the transfer of β-lactamase to the periplasm provide single cells with a high degree of resistance to ampicillin (>800 μg/ml) on solid media, whereas fusions which convey a cytoplasmic location on the β-lactamase domain provide no additional ampicillin resistance to the individual cell and thus the cell is lysed by <5 μg/ml ampicillin.

Using this method it was discovered that eight of the nine fusions provided single cells with the ability to survive >800 μg/ml of ampicillin, suggesting that in these cases, the β-lactamase moieties of the fusion proteins were periplasmic. Only one AlgK-β-lactamase fusion protein provided no resistance to ampicillin in single cells (which were lysed by an ampicillin level of 5 μg/ml). This predicts a cytoplasmic location for this β-lactamase moiety which was fused in-frame to the seventh amino acid of the AlgK pre-protein. No other putative cytoplasmic orientated β-lactamase moieties were identified and no intermediate single cell MIC's were observed suggesting that AlgK does not contain any other membrane-spanning regions. A summary of the position of the fusion junction, the proposed location of the β-lactamase domain and MIC of ampicillin for each fusion is displayed in table 6.1.

POSITION OF FUSION JUNCTION	MINIMUM INHIBITORY CONCENTRATION OF AMPICILLIN (μg/ml)	PROPOSED SUBCELLULAR LOCATION OF DOMAINS
Pro^{7}	<5	CYTOPLASMIC
Glu^{42}	>800	PERIPLASMIC
Gln^{50}	>800	PERIPLASMIC
Gln^{80}	>800	PERIPLASMIC
Lys^{86}	>800	PERIPLASMIC
Ser^{125}	>800	PERIPLASMIC
Leu^{284}	>800	PERIPLASMIC
Gln^{291}	>800	PERIPLASMIC
Ala^{388}	>800	PERIPLASMIC
$AlgK^{476}$	>800	PERIPLASMIC

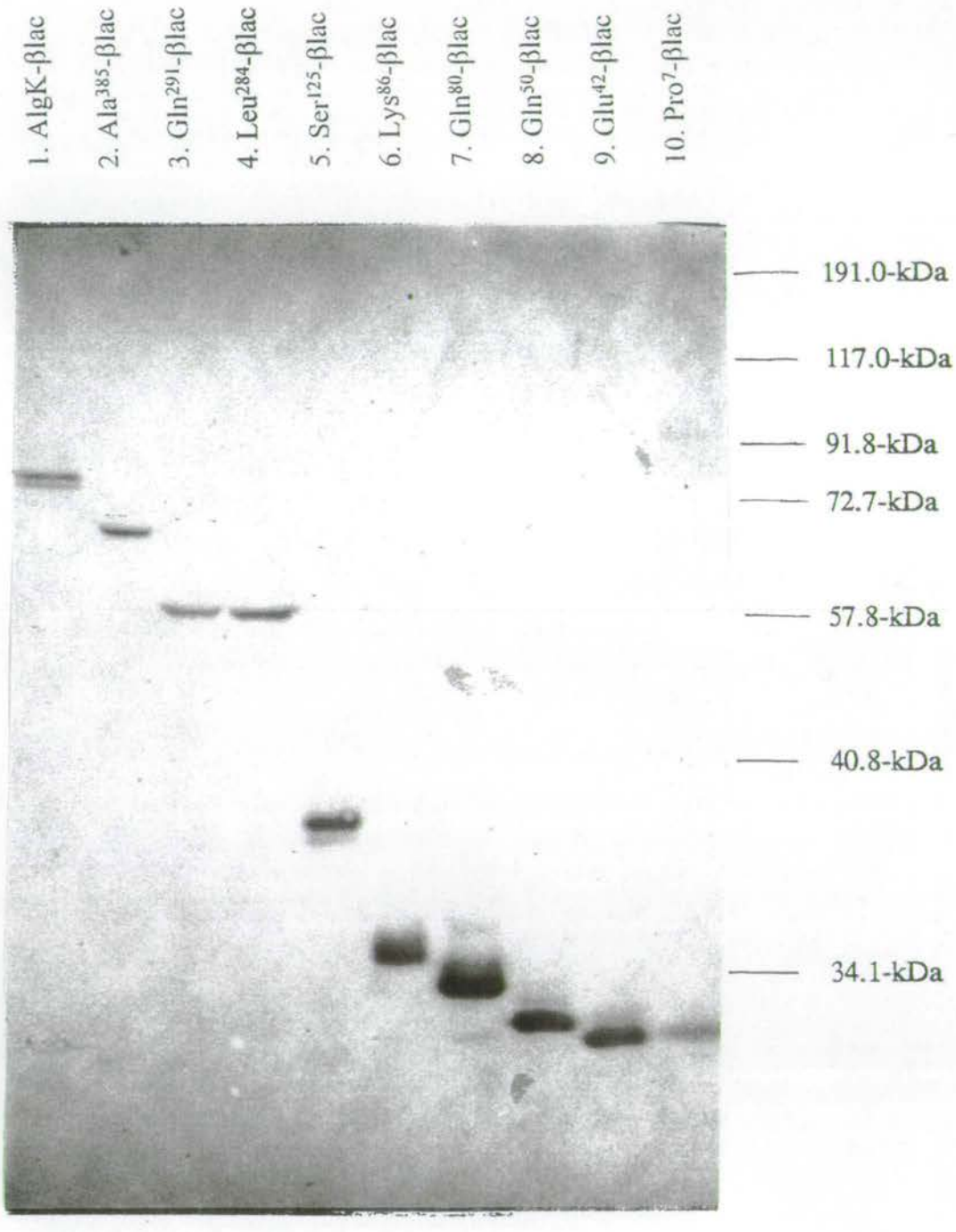
Table 6.1. Listing of the in-frame junctions of each *algK"blaM* fusion as determined by nucleotide sequencing. The minimum inhibitory concentration (MIC) of ampicillin required to lyse single cells of JM101 expressing each fusion protein is indicated and the position of the β-lactamase moiety of each fusion as inferred from the single cell MIC is shown alongside.

6.2.3. **Detection of AlgK-β-lactamase fusion proteins produced in *E. coli***

A Western Blot was performed in order to visualise the fusion proteins produced from pAA35 or pAA22 deletions. Whole cell fractions of overnight cultures of JM101 containing either pAA35 or truncated fusions derived from pAA22 were disrupted and denatured using Laemmli sample buffer, LSB, (Laemmli, 1970) and separated by SDS-PAGE (10%). Two gels were run in parallel and following electrophoresis, one was stained with Coomassie Brilliant Blue to check that protein loadings were even while the other was transferred to a nitrocellulose membrane (using a 'Transblot' apparatus, as described in 2.3.4) and then subjected to Western blot analysis. Rabbit anti-β-lactamase antisera (purchased from 5 Prime-3 Prime inc.) was used to bind to the β-lactamase moieties of the chimeric proteins. Detection and visualisation of bound rabbit anti-β-lactamase antibody was then undertaken using goat anti-rabbit IgG conjugated with the alkaline phosphatase (Sigma). Immuno-localised alkaline phophatase is then visualised by addition of the chromogenic substrates mixture 5-bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium (BCIP/NBT) which is converted by alkaline phosphatase action *in-situ* to a readily visible, dense blue compound. The resultant Western Blot is shown in figure 6.1.

The Western blot reveals the extent of truncation of AlgK, resulting in the stepwise reduction in band size from fusions containing the complete AlgK (lane 1) through to only seven amino acids of the AlgK N-terminus (lane 10). The blot also reveals evidence to substantiate the occurrence of a post-translational modification event to the AlgK portion of the fusion protein. The putative post-translational event is most strikingly observed when considering relative short fusions (lanes 9 and 10). Lane 10 contains a β-lactamase fusion protein with just seven N-terminal amino acids of AlgK which directs the β-lactamase moiety to the cytoplasm (see table 6.1).

Figure 6.1b. Detection of AlgK and truncated AlgK-β-lactamase fusion proteins. Strain JM101, containing pAA35 or pAA22 derivatives carrying *algK* partial deletions were grown overnight. Cells were harvested, denatured with Laemmli sample buffer (Laemmli, 1970) and equal loadings were run on a 10% SDS-polyacrylamide gel. The gel was immunoblotted using anti-β-lactamase antibodies, followed by detection with alkaline phosphatase-conjugated, goat anti-rabbit IgG. The position and size of the molecular weight markers are indicated.


1. AlgK-βlac
2. Ala385-βlac
3. Gln291-βlac
4. Leu284-βlac
5. Ser125-βlac
6. Lys86-βlac
7. Gln80-βlac
8. Gln50-βlac
9. Glu42-βlac
10. Pro7-βlac
191.0-kDa
117.0-kDa
91.8-kDa
72.7-kDa
57.8-kDa
40.8-kDa
34.1-kDa

However, this band comigrates to the same position (approximately 30-kDa) as a fusion protein some 35 amino acids larger (lane 9). In the latter case, the fusion occurs to amino acid Glu^{42} (glutamic acid) and the polypeptide would be expected to run at approximately 33.5-kDa (3.5-kDa larger in size). This is consistent with the cleavage and removal of a signal sequence from the longer polypeptide, indeed, all the AlgK-β-lactamase fusion bands also appear to run 3-4-kDa smaller than the predicted size for the combined fusion protein (as shown in table 6.2), except for Pro^{7}-β-lactamase. In lane 1, two full length AlgK-β-lactamase fusion proteins are clearly visible, the upper band is approximately 3-kDa larger than the lower protein band, this result may suggest that the post-translational modification event is occurring rather slowly and has not been fully completed.

6.3. DISCUSSION

In the present study, a gene fusion between the 3' terminus of the entire *algK* gene and *'blaM* was constructed. The resulting translational fusion was found to confer a high degree of resistance to ampicillin on individual *E. coli* cells, indicating that the open reading frame proposed for AlgK is correct and that the C-terminus of the AlgK protein is located in the periplasm.

Nine other gene fusions, encoding C-terminally truncated AlgK fused in-frame to mature β-lactamase have also been constructed and characterised to obtain further insight into the subcellular location of AlgK. These fusions were produced randomly by utilising the 3'-5' exonuclease activity of exonuclease III to delete coding sequence from the 3' terminus of *algK*. Sequencing of the fusion junctions of the in-frame hybrids revealed their location

POSITION OF FUSION JUNCTION	PREDICTED MOLECULAR WEIGHT OF AlgK DOMAIN (Da)	PREDICTED MOLECULAR WEIGHT OF FUSION PROTEIN (Da)	ACTUAL MOLECULAR WEIGHT OF FUSION PROTEIN (Da)
Pro^{7}	774	29674	30500
Glu^{42}	4642	33542	30500
Gln^{50}	5526	34426	31000
Gln^{80}	8842	37742	33000
Lys^{86}	9506	38406	35000
Ser^{125}	13816	42716	39000
Leu^{284}	31391	60291	57000
Glu^{291}	32164	61064	57500
Alu^{388}	42886	71786	68000
$AlgK^{476}$	52612	81522	79000 + 82000

Table 6.2. Listing of the in-frame *algK'-'blaM* fusions, the predicted molecular weight of the AlgK domain is indicated. The predicted combined molecular weight of the fusion protein is listed alongside the actual molecular weight of the fusion protein as deduced from the Western blot shown in figure 6.1.

within the *algK* gene and predicts their location within the AlgK protein. Comparison with the hydropathy plot (figure 4.3) found that fusion junctions were located in both hydrophilic and hydrophobic regions of the AlgK protein. However, the subcellular locations for these protein fusions, as predicted by their single cell MIC's, suggested that the location of AlgK was almost entirely periplasmic. No intermediate single cell MIC's were observed, indicating that none of the fusions adopted a trans-membrane location.

One AlgK-β-lactamase fusion protein, that of Pro^{7} (see table 6.1) did generate a cytoplasmically located β-lactamase fusion protein. The fusion junction of Pro^{7} lies within an N-terminal hydrophobic region of AlgK (see figure 4.2). The level of ampicillin resistance conveyed by this polypeptide, together with the signal sequence consensus and pulsing experiments supports the view that this region may constitute a membrane-spanning α-helix, anchoring the AlgK protein to the cytoplasmic membrane or alternatively may be part of a lipoprotein signal sequence which mediates the trans-membrane export of AlgK before cleavage. It seems probable that the seven amino acids of the N-terminus are not in themselves sufficiently hydrophobic to target the fusion protein across the cytoplasmic membrane and thus the fusion protein adopts a cytoplasmic location.

Close analysis of the Western blot obtained from the fusion proteins also predicts that the hydrophobic N-terminal sequence encodes a signal peptide and as such, is cleaved from the mature AlgK protein rather than being able to act as a membrane-spanning anchor. The most striking evidence for this event is revealed when considering the predicted sizes of the fusion proteins Pro^{7}-β-lactamase and Glu^{42}-β-lactamase. If no post-translational cleavage was occurring these proteins would be expected to migrate to positions of ~29.5-kDa and 33.5-kDa (the molecular weight of mature TEM β-lactamase is 28,900-Da as determined by Ambler and Scott, 1978). However the Pro^{7}-β-lactamase and Glu^{42}-β-lactamase fusion proteins

comigrate to a position of ~30-kDa. This result predicts a post-translational cleavage event with the removal of 3-4-kDa of the AlgK protein. The lipoprotein signal sequence predicted for AlgK is 27 amino acids long with a molecular weight of ~3-kDa and removal of this signal sequence would be consistent with the anomalous migration of the fusion protein (Pro^7-β-lactamase does not contain a complete signal sequence or have the required amino acid sequence for cleavage and thus is not processed). Indeed all the other AlgK-β-lactamase fusion proteins examined appeared to migrate to positions ~3-4-kDa smaller than that predicted for a complete, non-modified fusion protein (see table 6.2).

The construct pAA35 which results in the expression of the full length AlgK protein fused at its C-terminal to β-lactamase was found to produce two bands when visualised on the Western blot (figure 6.1). The difference in the molecular weights of these two bands as predicted from their migration is ~3-kDa. It therefore seems likely that the bands correspond to the pre-protein and mature forms of the AlgK fusion protein. Thus, unlike the smaller truncated fusion proteins the complete AlgK-β-lactamase fusion protein appears to require a longer time interval to fully process the signal sequence. There are several potential reasons why this cleavage event may be slower in the larger protein. Firstly the β-lactamase domain of the fusion protein may interfere with the interaction between the membrane lipids and the signal peptide which is required for the insertion of the hydrophobic region into the membrane prior to the 'flipping' of the protein across the cytoplasm. Alternatively the shear size of the AlgK-β-lactamase fusion protein may well inhibit the translocation across the cytoplasmic membrane via a putative translocase and thus, the blot may represent the cytosolic (preprotein) and periplasmic (mature protein) locations of the fusion protein. Alternatively, the AlgK-β-lactamase fusion may have completely traversed the cytoplasmic membrane, with the signal sequence acting as an uncleaved hydrophobic cytoplasmic membrane anchor for the

periplasmic protein. It may well be that the actual nature of the fusion protein, either due to its charge or conformation, may inhibit the action of the polytopic signal peptidase II and thus, allow the pre-protein form to persist. Interestingly, studies utilising pJBS633 with the putative lipoprotein, PBP3 (penicillin binding protein) revealed only periplasmic oriented, cytoplasmic membrane-anchored pre-protein. In the latter case, it was proposed that the N-terminal 'signal sequence' was acting as a cytoplasmic membrane spanning anchor and that conversion to the mature protein was somehow inhibited (Bowler and Spratt, 1989).

Another point raised by the Western blot (figure 6.1) is that in general, the quantity of protein fusion product observed appears to be inversely proportional to the molecular weight of the fusion product. This correlation is not unusual in protein fusion systems. Calamia and Manoil (1990), using pulse chase experiments on LacY-PhoA (lactose permease-alkaline phosphatase) hybrid proteins showed that in general, as the length of the LacY-PhoA hybrid proteins increased their stability (and therefore their abundance) decreased.

In the present study, the Gln^{80} and Lys^{86} fusion proteins were found to migrate as slight smears. This may reflect on the quantity of protein produced or possibly on heterogeneity in conformation or detergent binding, a phenomenon also observed by Calamia and Manoil (1990) with certain LacY-PhoA fusions.

In conclusion, analysis of AlgK-TEM-β-lactamase fusion proteins has successfully confirmed the open reading frame of *algK* and in addition, has provided evidence that AlgK is predominately periplasmic. Close analysis of the migration of the AlgK-β-lactamase fusion protein suggested that a post-translational processing event may be occurring and this is consistent with the cleavage of a signal sequence for cytoplasmic export. Previous computer analysis of the amino acid sequence predicted that AlgK contains a 27 amino acid lipoprotein signal sequence which

could be removed by signal peptidase II prior to monotopic anchoring with fatty acid via its terminal cysteine residue. The polypeptide could be anchored to either the outer leaflet of the cytoplasmic membrane or either the inner or outer leaflets of the outer membrane (review Pugsley, 1993). However, analysis using the β-lactamase reporter gene predicts that AlgK is periplasmic (as only periplasmic orientated β-lactamase moieties protect individual *E. coli* cells from killing by ampicillin, as shown by Bowler and Spratt, 1989 and Edelman et al., 1987) which predicts that AlgK is not anchored to the outer leaflet of the outer membrane. Additionally, analysis of the proteins produced by non-mucoid and mucoid *P. aeruginosa* strains (Grabert *et al.*, 1990) revealed only a single additional, mucoid associated, outer membrane protein which was subsequently identified as AlgE, the alginate outer membrane porin (Chu *et al.*, 1991). Thus, it appears most probable that AlgK is a periplasmic lipoprotein anchored to the outer leaflet of the cytoplasmic membrane. The actual presence and position of an AlgK membrane anchor may be revealed by determining the location of these β-lactamase fusions by Western blots of cytoplasmic and outer membrane cellular fractions, separated on sucrose gradients.

Chapter 7

Construction of an *algK* mutant in *P. aeruginosa*.

7.1. INTRODUCTION

In this report, a novel gene, designated *algK* has been identified and characterised. However, searches with the AlgK polypeptide through the available protein databases failed to reveal any obvious or significant similarities, but suggested that the protein was not a component of an ABC transporter. Analysis of the nucleic and putative amino acid sequence predicted that the *algK* gene encodes a probable lipoprotein with a cleavable signal sequence giving rise to a mature protein with a molecular weight of approximately 50-kDa (see chapter 4). Data from expression of AlgK (chapter 5) confirmed the open reading frame and gave further supporting evidence for a post-translational signal sequence cleavage event. A β-lactamase protein fusion survey revealed that AlgK was most likely a periplasmic orientated lipoprotein which is anchored to either the outer leaflet of the cytoplasmic membrane or the inner leaflet of the outer membrane. However, none of this data revealed whether AlgK was 'essential' or even definitely involved in alginate production (although its genetic location within the alginate biosynthetic operon implicates a role).

Previous surveys on the alginate biosynthetic cluster in order to determine whether particular loci, were 'essential' or 'non-essential' for the production of a mucoid phenotype, have required the construction of mutants within *P. aeruginosa.* Darzins and Chakrabarty (1984) generated a constitutive mucoid *P. aeruginosa* strain, 8830, using EMS (ethylmethanesulphonate) mutagenesis. Further EMS mutagenesis on this mutant by both Darzins and Chakrabarty (1984) and Darzins *et al*, (1985) generated a cluster of mutations within the 34 minute biosynthetic region that result in a loss of the mucoid phenotype. Complementation analysis by introduction of broad host range (BHR) vectors bearing DNA fragments of the

cluster defined regions that restored the mucoid phenotype on these *P. aeruginosa* strains. Thus, complementation analysis revealed the presence of *algA* which complemented the non-mucoid EMS mutant *P. aeruginosa* strain 8873 (Darzins *et al.*, 1986), *algD* which complemented strain 8835 (Deretic *et al.*, 1987), *algE* which complemented mutant 8897 (Chu *et al.*, 1991) *alg8*, *alg44* and *alg60* which complemented strains 8838, 8874 and 8887 respectively (Wang *et al.*, 1987). All of these genes (except *alg60*) have subsequently been sequenced and characterised more fully. EMS mutagenesis thus revealed the genes which were essential for the mucoid phenotype within this cluster. Not surprisingly, the mucoid to non-mucoid selection criteria utilised by Darzins and Chakrabarty (1984) and Deretic *et al.* (1985) failed to isolate any more subtle or less obvious phenotypes such as inactivation of genes involved in alginate modification genes which may not prevent a mucoid phenotype or potentially any genes whose lack of function would be lethal to *P. aeruginosa* or only reduce the efficiency of alginate production. An exception to this was the isolation of an EMS-generated 'non-essential' gene locus by Chitnis and Ohman (1990). They discovered a mutant whose alginate was resistant to guluronase (G'ase), and thus retained a mucoid phenotype on G'ase containing agar. The retention of mucoidy predicted that the alginate consisted entirely of polymannuronate, and thus complementation analysis and sequencing lead to the identification of *algG* encoding a C-5 mannuronan epimerase.

Regions of the alginate biosynthetic cluster not defined by EMS mutations that cause a loss in mucoidy, include the region between *alg44* and *algE* (Wang *et al.*, 1987) and the 4-kb region between *alg60* and *algA* (Darzins *et al.*, 1985). Generation of chromosomal 'knock-outs' in the second of these two regions has relied on insertional inactivation by single homologous recombination of suicide vectors bearing the required gene fragment. This method necessitates preliminary isolation, sequencing and characterisation of the individual genes present in order to select the

homologous fragment and as such, has only been performed on *algL* which encodes alginate lyase (Boyd *et al.*, 1993) and *algF* which encodes an alginate acetylase (Shinabarger *et al.*, 1993). Construction of *algL* or *algF* mutants in *P. aeruginosa* by single homologous recombination was achieved by cloning a small fragment of the gene into the narrow host range vector, pBR325 (Bolivar, 1978) and this vector was then mobilised into *P. aeruginosa* strains. pBR325 acts as a suicide vector as it cannot replicate within *P. aeruginosa* therefore the only way for the plasmid to be maintained in the mucoid *P. aeruginosa* host is by integration via a single homologous recombination event with the appropriate chromosomal gene. Application of this procedure by Boyd *et al.* (1993) and Shinabarger *et al.* (1993) generated insertionally inactivated chromosomal mutants of *algL* and *algF* and resulting in a loss of mucoidy.

It was decided that in order to help elucidate the possible role of AlgK in alginate EPS synthesis further, the chromosomal *algK* gene would be inactivated. As such, a method similar to that used for the chromosomal inactivation of *algL* (Boyd *et al.*, 1993) and *algF* (Shinabarger *et al.*, 1993) was employed.

7.2.1. Construction of *algK* insertion mutant

In order to generate an inactivate chromosomal *algK* gene, pBR325 was used as a suicide vector. pBR325 or derivatives can be mobilised from *E. coli* strains to *P. aeruginosa* by the use of a helper plasmid bearing *tra* functions (such as pRK2013) using a triparental mating procedure (as described in 2.3.2). However, pBR325 (or derivatives) cannot replicate within a *P. aeruginosa* host and thus the plasmid will only be maintained if a single homologous recombination event occurs, resulting in the insertion of the entire pBR325 plasmid into the chromosomal target gene. Mucoid *P. aeruginosa* strains are extremely resistant to a large number of antibiotics but, even at high density, can be eliminated *in-vitro* with high concentrations of

tetracycline. Thus, by cloning the *algK* fragment into the β-lactamase gene of pBR325 (resulting in loss of ampicillin resistance) selection for recombination in the final *P. aeruginosa* strain can be conveyed by the inheritance of tetracycline resistance.

Construction of suitable *algK* bearing vectors

The β-lactamase gene within pBR325 has only a single unique restriction site, *Pst*I. The initial cloning strategy employed involved the direct sub-cloning of a 1-kb *Pst*I *algK* fragment from pAA1 into this site, resulting in loss of ampicillin resistance when transformed into HB101. This procedure generated the plasmids pAA24 and pAA25 where the *Pst*I fragment of *algK* is in opposite orientations. Both pAA24 and pAA25 were introduced into mucoid *P. aeruginosa* strain B by triparental mating using pRK2013 as a helper plasmid (also within HB101).

However, repeated attempts using this protocol failed to isolate any tetracycline resistant *P. aeruginosa* colonies (frequency $<10^{-9}$). It was noted that in previous *algL/F* insertional inactivation systems the fragments cloned into pBR325 were much smaller, approximately 0.4-kb in length (Boyd *et al.*, 1993; Shinabarger *et al.*, 1993, respectively). Furthermore, a previous study by Quandt and Hynes (1993) on single and double recombination frequencies of plasmid with *R. leguminosarum* chromosome determined that as the length of the homologous fragment increased, the likelihood of single recombination diminished and inversely the probability of a double recombination event increased. Indeed, with the systems employed by Quandt and Hynes (1993) a homologous fragment of 0.7-kb was equally likely to recombine by either method whereas a 1.2-kb fragment was 50 times more likely to incorporate into the chromosome by double recombination event as opposed to a single recombination event. The frequency of generating a recombinant strain with a single crossover dropped in efficiency by over 20 fold with an increase in fragment

an increase in fragment length of 0.5-kb. A double recombination event when using pBR325 as a suicide vector would result in an insertion of the homologous fragment into the appropriate chromosomal loci but not the incorporation of the plasmid. Thus, recombinants would not be tetracycline resistant and therefore not selectable.

Thus it appeared possible that by reducing the size of the *algK* fragment cloned into pBR325 the efficiency of the final recombination event within mucoid *P. aeruginosa* may be improved. To test this hypothesis a 0.4-kb fragment was selected for cloning into the pBR325 *bla* gene. In order to amplify this region with the required *Pst*I sites a further synthetic oligonucleotide primer was designed, Primer 7. Primer 7 (as described in table 2.1) contains a *Pst*I site towards the 5' end and was used in conjunction with Primer 4 (see chapter 6) which primes from the complementary strand at the *algK* TGA stop codon but does not contain a *Pst*I site. PCR using pAA22 as the DNA source thus generated a product of approximately 460 bases. Digestion of the product with *Pst*I gives rise to a 5' *Pst*I site as coded for by the synthetic oligonucleotide and a 3' *Pst*I site as coded for within the *algK* nucleotide sequence at position 1614-1619 giving rise to a smaller *algK* fragment of approximately 410-bp. The digested PCR product was then ligated into the respective sites of pBR325 to form pAA31.

A control plasmid was also generated by repeating the experiments of Boyd *et al.* (1993). Boyd *et al.* (1993) constructed a pBR325 derivative used for the successful generation of an *algL* null mutant by direct sub-cloning of an internal *algL* 470-bp *Pst*I fragment into pBR325. Unfortunately the required donor vector was not available within our laboratory so PCR was utilised to generate the identical fragment with identical *Pst*I restriction sites. Two oligonucleotides were designed and synthesised. Primer 5, primed at the 5' *Pst*I site of *algL* (base position 853-858, from Boyd *et al.*, 1993) encompassing a region of 23-bp with the *Pst*I site towards the 5' end and Primer 6 spanned 24-bp at the *Pst*I site nearer the 3' terminus of *algL*

(at position 1318-1323, from Boyd *et al.*, 1993). The PCR product was digested with *Pst*I to give the 470-bp fragment which was then ligated into pBR325 to give plasmid pAA30.

HB101 bearing either pAA30 or pAA31 was then triparentally filter mated using pRK2013 as a helper plasmid into both the mucoid *P. aeruginosa* strain B (PAB) or PA8873 (Darzins and Chakrabarty, 1984). Single homologous recombinants were selected for on minimal plates containing 200 μg/ml tetracycline. Stable tetracycline colonies were obtained for both plasmids in both *P. aeruginosa* strains. These colonies were analysed by Southern hybridisation to confirm that plasmids pAA31 and pAA30 had inserted into chromosomally located *algK* and *algL* genes (see below).

Southern hybridisation of prospective *algK* and *algL* recombinants

The nucleic acid probe selected for identifying the insertion of pBR325 derived vectors into *P. aeruginosa* was the chloramphenicol resistance gene (chloramphenicol acetylase transferase) which is present on the pBR325 vector. The actual DNA used for the probe was isolated from vector pCM1 (Close and Rodriguez, 1982) which contains a promoterless chloramphenicol resistance cartridge flanked by *Sal*I restriction sites. Both of these chloramphenicol acetylase transferase (CAT) genes are derived from the same Tn*9* source. Thus pCM1 was digested with *Sal*I and the 780-bp CAT fragment isolated, prior to being random prime labelled with [α-^{32}P]dCTP (as described in 2.3.2).

Chromosomal DNA from appropriate tetracycline resistant *P. aeruginosa* strains was isolated (as Silhavy *et al.*, 1984) and 3μg of each digested with both *Eco*RI and *Hin*dIII restriction endonucleases. The samples were electrophoresed on an agarose gel before transfer to a nylon membrane (as described in 2.3.2).

Hybridisation of the Southern blot was performed using a high stringency (65°C) protocol which should only allow hybridisation to identical or extremely similar nucleic acid sequences. Following hybridisation the blot was autoradiographed for 16 hours prior to developing (see figure 7.1). The autoradiogram revealed hybridisation in lanes 3-10 predicting the presence of a CAT gene within these *P. aeruginosa* strains that is not present in the original *P. aeruginosa* strains PAB and PA8873 (lanes 1 and 2 respectively) predicting that the recombination event was successful. The size of the *Eco*RI fragments reveals the orientation of both pAA30 (pBR325 with *algL* fragment) and pAA31 (pBR325 with *algK* fragment) once recombined within the respective chromosomal genes. In both cases the CAT gene of pBR325 is immediately upstream of the subcloned fragment. The *algK* recombination event is represented schematically in figure 7.2.

7.2.2. **Complementation of the *algK* mutant**

All the tetracycline resistant *P. aeruginosa* strain B isolates were found to be visibly non-mucoid when plated on LB, minimal and MAP plates (MAP medium specifically promotes alginate production in *P. aeruginosa*, Franklin *et al.*, 1994). This, combined with the Southern hybridisation data, predicts that both *algK* and *algL* had been successfully inactivated.

The non-mucoid phenotype of the tetracycline, *algK* mutant (PABK) may have been caused by an absence of AlgK activity. To test this hypothesis, plasmid pAA23, which contains the *algK* gene under the control of p*tac* (see chapter 5) was mobilised (using pRK2013 helper plasmid) from *E. coli* strain HB101 into PABK (*algK*::pAA31 mutant of *P. aeruginosa* strain B). After filter-mating, exconjugants were selected by ampicillin resistance (200 μg/ml) on minimal agar supplemented with tetracycline (200 μg/ml), HB101 is eliminated as it is both tetracycline sensitive and auxotrophic (see table 2.2). The presence of vector pAA23 within PABK was

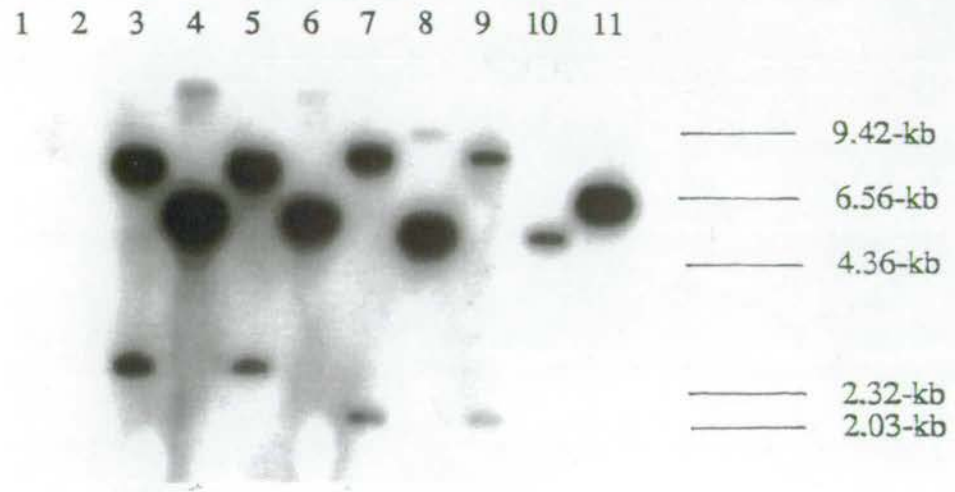
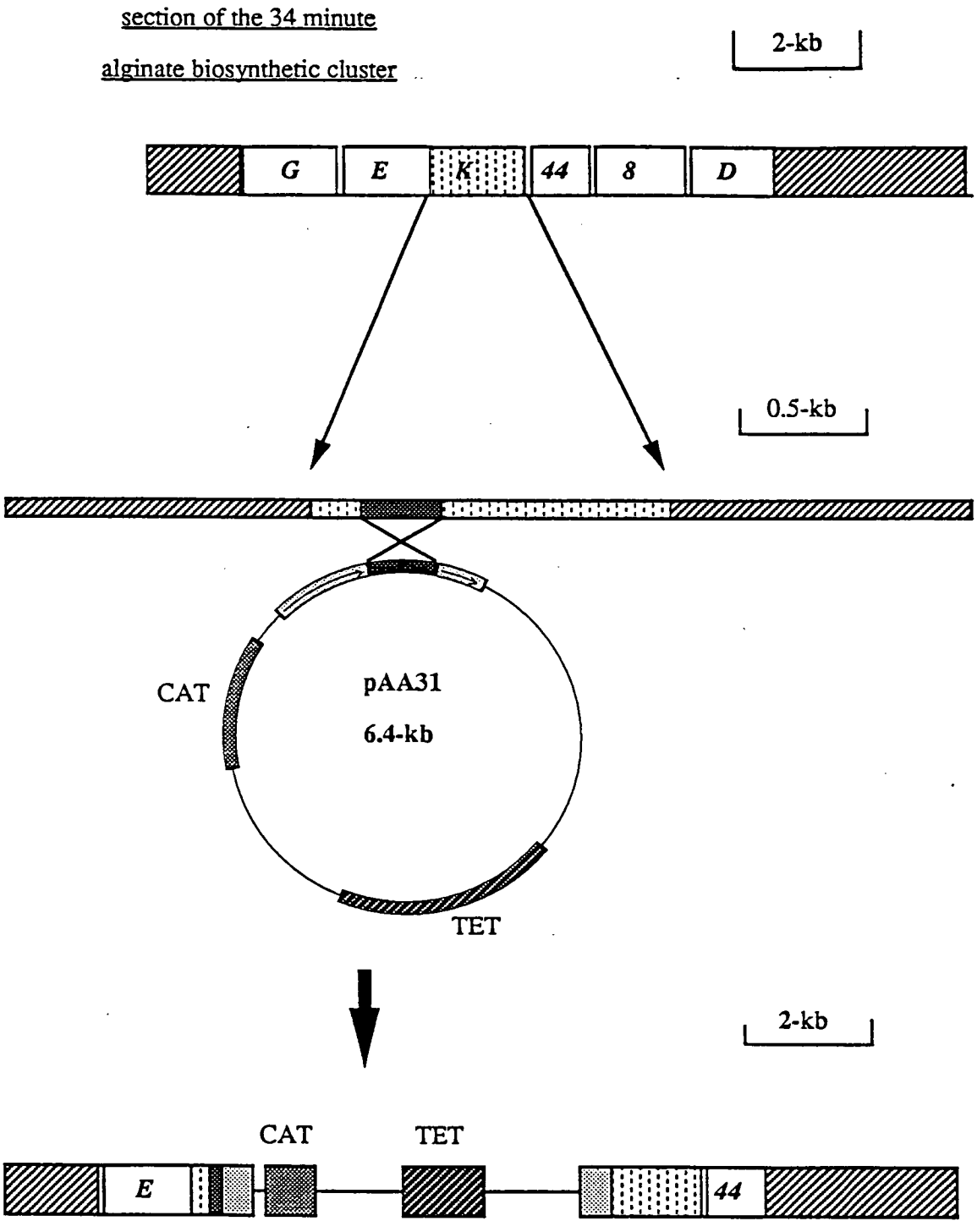


Figure 7.1. Southern blot analysis of chromosomal DNA derived from putative *algK* and *algL* insertion mutants of PAB and PA8873. Chromosomal DNA from tetracycline resistant isolates were analysed for hybridisation to the CAT gene present on pBR325.

Lanes are as follows. 1) *Eco*RI digested PAB, 2) *Eco*RI digested PA8873, 3) *Eco*RI digested PAB *algL* mutant, 4) HindIII digested PAB *algL* mutant, 5) *Eco*RI digested PA8873 *algL* mutant, 6) *Hin*dIII digested PA8873 *algL* mutant, 7) *Eco*RI digested PAB *algK* mutant, 8) *Hin*dIII digested PAB *algK* mutant, 9) *Eco*RI digested PA8873 *algK* mutant, 10) *Hin*dIII digested PA8873 *algK* mutant, 11) *Sal*I fragment of pCM1 (used as probe). The DNA was separated by agarose gel electrophoresis in TBE buffer. DNA was transferred to Hybond™-N nylon membrane, prior to hybridisation with random primed *Sal*I (CAT) fragment of pCM1. Hybridisation conditions were as stated in 2.3.2 for the high stringency protocol. Autoradiography was carried out for 16 hours. The position of the λ *Hin*dIII molecular weight markers is indicated on the right-hand-side.

Figure 7.2. Diagrammatic representation of the insertional inactivation of *P. aeruginosa* chromosomal *algK* gene by pAA31. The *bla* (ampicillin resistance) gene, tetracycline resistance gene and chloramphenicol (chloramphenicol acetyltransferase, CAT) resistance genes are represented by  ,  and  respectively. Chromosomal *algK* is represented by  and the region of homology is represented by  . The remaining *P. aeruginosa* chromosome is represented by  .


section of the 34 minute
alginate biosynthetic cluster
2-kb
G
E
K
44
8
D
0.5-kb
CAT
pAA31
6.4-kb
TET
2-kb
CAT
TET
E
44

confirmed by reisolating the plasmid from the *P. aeruginosa* strain using the miniprep procedure mentioned described in 2.3.2. Expression of AlgK from pAA23 was achieved by plating the transformants onto LB agar containing the respective selective antibiotics and 1 mM IPTG. Unfortunately expression of AlgK failed to complement PABK to a mucoid phenotype.

The lack of complementation of the *algK* mutant with a plasmid expressing AlgK is most probably due to the operonic nature of the 34 minute biosynthetic region (Chitnis and Ohman, 1993). Indeed, Chitnis and Ohman (1993) showed that any Tn*501* insertion (~8.2-kb DNA) in the cluster was polar on downstream alginate genes and thus prevented their transcription which resulted in a non-mucoid phenotype. Mucoidy was only restored when broad host range (BHR) plasmids were introduced that contained the remaining genes of the pathway (i.e. the genes downstream of the Tn*501* insert. It seems likely that insertional inactivation of a gene using a pBR325 (~6-kb) suicide vector would also have a polar effect on downstream genes and block the production of alginate. Inactivation of *algF* and *algL* using this suicide system also resulted in a loss of mucoidy, and the mucoid phenotype was only restored by the introduction of a BHR plasmid bearing the downstream gene *algA* alone (Shinabarger *et al.*, 1993). Thus, loss of mucoidy was entirely due to a polar effect on transcription of downstream genes and therefore the inactivated *algL/algF* genes were 'non-essential' in production of a visible mucoid phenotype.

To predict whether downstream genes were still active after the *algK* recombination event a plate-test was employed to identify any alginate lyase activity (as *algL* lies downstream of *algK* see figure 1.9). PABK $^{+}$/-pAA23 was plated onto LB agar containing commercial alginate, the respective selective antibiotics and $^{+}$/-IPTG (1mM) to induce expression from *algK*. Active lyase gives rise to large zones of clearing around individual colonies (after incubation at 37°C, 48 hours) which is

visualised by the addition of 10% cetyl pyridinium chloride to the surface of the plate (see Boyd *et al.*, 1993). PAB shows such a phenotype resulting in large, clear zones encircling each colony, but PABK and PABK harbouring induced pAA23 failed to show any obvious lyase activity (see table 7.1). This additional evidence along with the failure of pAA23 to obviously complement the *algK* mutant suggests that loss of mucoidy may be at least partially due, to a polar effect on expression of downstream genes of the operon which are essential for the production of the alginate EPS.

Construction of alternative vectors for complementation analysis

Downstream of *algK* only two genes, *algE* and *algA*, have been definitively shown to be essential in production of mucoidy (Chu *et al.*, 1991 and Darzins *et al.*, 1986, respectively). To circumvent the polar problems of insertion on downstream genes *algA*, *algE* and *algK* were introduced into the BHR vector, pMMB66(EH). Initially *algE* and *algK* were subcloned into pMMB66(EH) on a *Sma*I/*Hin*dIII fragment derived from pRM812 (see figure 3.4 for diagram of vector), which contains the entire coding sequence of both genes, to form pAA32. *algA* was sub-cloned from pAD4038 (pMMB24 containing *algA* on a *Hin*dIII fragment, see Sá-Correia *et al.*, 1987) into the *Hin*dIII site of pAA32. The resulting plasmid, pAA33 was then mobilised into PABK from HB101 using the helper plasmid pRK2013. Exconjugants were selected on minimal agar supplemented with tetracycline and ampicillin (both at 200 μg/ml).

The presence of plasmid pAA33 within PAB was confirmed by performing mini-preps (see 2.3.2 for details of procedure). Expression of the subcloned *algK*, *algE* and *algA* genes within pAA33 was induced from the p*tac* promoter by plating onto LB plates containing the appropriate antibiotics and 1 mM IPTG. Unfortunately, expression of AlgK, AlgE and AlgA failed to restore a mucoid phenotype to the PABK mutant. The probable reason for the failure of pAA33 to

P. aeruginosa STRAIN	IPTG INDUCED	PHENOTYPE	ALGINATE DIGESTION
PAB	NO	MUCOID	+
PABK	NO	NON-MUCOID	-
PABK(pAA23)	NO	NON-MUCOID	-
PABK(pAA23)	YES	NON-MUCOID	-

Table 7.1. Alginate digestion characteristics of different *P. aeruginosa* strains. The phenotype of *P. aeruginosa* was determined by 24 hour growth at 37°C on MAP plates. Alginate digestion was measured as a zone of clearing on LB alginate agar plates after cells were grown for 48 hours at 37°C and then the alginate was visualised by the addition of 10% cetyl pyridnium chloride (Boyd *et al.*, 1993). +, clear zone on opaque background of the plate; -, no clearing.

complement the *algK* mutant is due to a further uncharacterised 'essential' gene downstream of *algK*. This is most probably *alg60* which is thought to lie between *algG* and *algL* within the biosynthetic cluster. In agreement with this suggestion, EMS generated mutants in this region were found to be non-mucoid (Darzins *et al.*, 1985). The region between *algL* and *algA* is known to be free of any essential genes required for mucoidy. Indeed Boyd *et al*, (1993) showed that insertional inactivation of the *algL* gene prevented mucoidy by a polar effect on downstream genes, but that the mucoid phenotype was restored by the introduction of a broad host range vector expressing *algA* alone. Thus neither *algL* or the intervening genes between *algL* and *algA* are essential for mucoidy. However the lack of information on the nucleotide sequence of the *alg60* loci and the lack of suitable restriction sites on pAA33 prevented the inclusion of *alg60* into this vector and the extension of the complementation survey.

7.3. DISCUSSION

In this study, a mutation of chromosomal *algK* was generated in both mucoid and non-mucoid *P. aeruginosa* strains by the single homologous recombination of a pBR325 derivative into the *algK* gene. Inactivation of the *algK* gene using this procedure resulted in the conversion of the mucoid *P. aeruginosa* strain PAB to a non-mucoid phenotype (PABK). Initial complementation analysis utilising pAA23, a BHR vector expressing AlgK failed to restore the mucoid phenotype. This was assumed to be due to the proposed operonic nature of the 34 minute biosynthetic cluster (Chitnis and Ohman, 1993).

Previous successful complementation surveys had been achieved in EMS generated non-mucoid *P. aeruginosa* mutants (Darzins and Chakrabarty, 1984;

Darzins *et al.*, 1985) or on guluronase (G'ase) resistant mucoid mutants (*algG*, Chitnis and Ohman, 1990). Alkylating agents such as EMS tend to induce direct mispairing and EMS mutagenic action often results in G mispairing with T instead of C *in vivo* (Drake and Baltz, 1976). The effect of EMS mispairing is usually a frameshift mutation resulting in erroneous translation of the mutagenised gene. In an operon, such as that proposed for the 34 minute biosynthetic region, it is thought that a large polycistronic message is produced. An EMS mutagenic event within a single gene of the operon would only prevent correct translation of the mutagenised gene and allow the normal translation of the rest of the message. This is clearly demonstrated by the EMS generated *algG*4 (Chitnis and Ohman, 1990) a mutant of *P. aeruginosa* which is still mucoid but the polymer of which contains no guluronate (i.e. it consists solely of polymannuronate) due to an inactivated *algG* gene (which normally encodes C-5 epimerase). The production of a near perfect EPS suggests that the EMS-induced mutation in *algG* has no effect on any other 34 minute alginate biosynthetic cluster genes transcribed with it. The defect can also be complemented by the introduction into the *P. aeruginosa* strain of a BHR plasmid bearing an inducible *algG* gene. However, inactivation of a gene by a single homologous recombination event, resulting in the incorporation of the entire suicide vector, is likely to disrupt the transcription of the normal polycistronic message. Thus, genes downstream of the insert are not synthesised due to this block in transcription. Such a characteristic has been shown to occur by the inactivation of *algL* and *algF* with a pBR325 derived suicide vector (Boyd *et al.*, 1993 and Shinabarger *et al.*, 1993) as well as Tn*501* insertions (Chitnis and Ohman, 1993) within the cluster. Mucoidy was restored in each case by providing the chromosomally inactivated essential downstream genes on an inducible BHR vector.

In comparison therefore, it appeared probable that the non-mucoid PABK phenotype which failed to be complemented with plasmid-expressed AlgK was

caused by a polar defect in essential downstream genes, such as *algE* and *algA*, and not necessarily due to inactivation of *algK*. Indeed, it was shown that PABK had an alginate lyase negative phenotype unlike the PAB parent strain, indicating that the genes downstream of *algK* are not active.

The lack of expression of genes downstream of *algK* also suggests that the putative internal promoters of *algE*, *algL*, *algF* and *algA* (combined with plasmid expressed AlgK) was not adequate for the resumption of the mucoid phenotype. Intriguingly these putative promoters were thought to be functional in conditions akin to those examined in this system. This may suggest that for the above plasmid-borne alginate biosynthetic genes putative internal promoter activities are either enhanced by their plasmid location or that expression is actually caused by run through from cryptic or other plasmid promoters.

Complementation using a further broad host range construct bearing the known characterised essential genes downstream of *algK* (*algE* and *algA*) also failed to complement the *algK* mutation to the mucoid phenotype, predicting the presence of a further essential gene within the biosynthetic cluster. This inactive essential gene was most probably *alg60*, which is located between *algG* and *algL*. Previous surveys by Darzins *et al*, (1985) showed that EMS-generated mutations in the *alg60* locus resulted in a loss of the mucoid phenotype. However, since this early work no additional data on this region has been published and the sequence of this region is still unavailable. The scarcity of information concerning the *alg60* loci and the lack of any suitable restriction sites within pAA33, prevented the incorporation of *alg60* within this complementation vector. To include *alg60* into the complementation survey would necessitate either the cloning of the majority of the *algK* to *algA* (~12-kb) region into pMMB66 (or similar BHR vector) or alternatively would require expression of *alg60* on a further broad host range vector of a different incompatibility group. The latter method has the additional problem of transconjugant selection as *P.*

aeruginosa is very resistant to most antibiotics whilst PABK is resistant to both tetracycline and chloramphenicol (due to the insertion of pBR325) and additionally pMMB66 vectors if used to express *algK* separately confer ampicillin resistance. Time prevented any further development of either procedure.

Chapter 8

Characterisation of *algK*: an overview.

This thesis describes the identification and sequencing of a novel gene *algK* from the 34 minute alginate biosynthetic cluster of *P. aeruginosa* and the expression of its protein product. The *algK* gene is thought to encode a protein involved in the production of alginate exopolysaccharide of *P. aeruginosa*. *algK* was initially isolated on a 2-kb fragment containing a previously uncharacterised region of the 34 minute alginate biosynthetic cluster. Sequence analysis of this region revealed a single long open reading frame of approximately 1.4-kb which encoded a hydrophilic protein with a predicted molecular mass of 52,470-Da. Analysis of the amino acid sequence revealed a consensus lipoprotein cleavage site and predicted that the protein would undergo a post-translational modification event resulting in the removal of an N-terminal signal sequence, together with targeting of the protein outwith the cytoplasm. Overexpression of the AlgK protein was achieved in both *E. coli* and *P. aeruginosa* and the size of the protein product was consistent with that predicted from the amino acid sequence. [^{35}S]Methionine labelling of the protein product gave confirmatory evidence towards the occurrence of a post-translational cleavage event which was compatible with the removal of such a signal sequence. Construction of AlgK-β-lactamase fusion proteins and detection by immunoblotting, predicted a periplasmic orientation for the AlgK protein and also, provided further evidence for the removal of an N-terminal portion of the protein. However, insertional inactivation of the chromosomally located *algK* gene followed by preliminary complementation analysis, failed to reveal whether AlgK is essential for the production of the alginate phenotype.

Previous analysis of the EPS biosynthetic systems of other Gram-negative bacteria predicted that the export of the polymannuronate across the cytoplasmic membrane required an ABC transporter which would probably be encoded within the alginate biosynthetic cluster of *P. aeruginosa*. The initial thrust of this project was

directed towards the location of such an ABC homologue within the alginate biosynthetic cluster via a hybridisation survey. This Southern hybridisation analysis used a DNA probe isolated from a region surrounding an ATP binding motif which is particularly conserved amongst all ABC transporter genes. However, this protocol proved unsuccessful in locating any ABC-transporter homologues within either the biosynthetic cluster or the rest of the chromosome. Lack of cross-hybridisation was considered to be a manifestation of the considerable difference in G:C content between *P. aeruginosa* and the *S. typhimurium* DNA from which the probe had been isolated (70% compared to 50%, respectively). As such the presence or absence of any such ABC homologues involved in alginate EPS synthesis could not be confirmed.

Lack of success using this hybridisation survey led to a switch in protocols and it was decided to sequence a previously uncharacterised region within the 34 minute alginate biosynthetic cluster of *P. aeruginosa*. This 2-kb region between *alg44* and *alg8* was selected as the locus appeared to be a particularly good candidate for containing gene(s) responsible for transport across the inner membrane. Fragments of this region were subcloned into M13 mp19 and the nucleic acid sequence determined.

Initial analysis of the nucleic acid sequence between the *alg44* and *alg8* genes revealed a single open reading frame (ORF) of 1425-bp. This ORF began with an ATG translational start codon 230-bp downstream of the *alg44* TGA stop codon and ended with a TGA stop codon overlapping the ATG translational start codon of *algE* - these are the only genes identified within the 34 minute biosynthetic cluster that overlap and may suggest that they are translationally coupled. The 1425-bp ORF encoded a putative polypeptide of 475 amino acids with a predicted molecular weight of 52,470-Da with a pI of 5.64. A hydropathy plot of the predicted AlgK protein did not reveal any large degree of hydrophobicity (as anticipated by the acidic pI of the

polypeptide). This suggested that AlgK was not an integral cytoplasmic membrane protein and as such excluded the possibility that it was an integral membrane component of an ABC transporter. A computer search for the typical ATP binding motif, found in all ABC transporters, also proved unsuccessful, providing further confirmatory evidence that this protein was not the ATP binding component of an ABC transporter. Indeed comparison of the AlgK polypeptide with all the other proteins available in the current databases failed to reveal any significant amino acid identity with other proteins.

The only further data provided from the predicted amino acid sequence of AlgK was the possibility of an N-terminal signal sequence. The hydropathy plot of AlgK revealed a single long hydrophobic stretch of 27 amino acids at the beginning of the polypeptide sequence. Such an N-terminal stretch of hydrophobic and basic amino acids (pI of region is 11.65) is typical of either a signal sequence or a hydrophobic, cytoplasmic membrane anchor. Possession of a signal-sequence suggests that AlgK undergoes additional post-translational modification resulting in the cleavage/removal of the signal sequence and a non-cytoplasmic location for the mature protein. Computer analysis of the putative signal peptide revealed that the AlgK signal sequence contained a consensus cleavage site, typical of membrane lipoproteins.

The relatively large 'non-coding' region between the C-terminus of *alg44* and the ATG translational start of *algK* was also scrutinised further. Computer-aided comparison with the defined promoters of other genes involved in alginate biosynthesis or regulation, revealed strong homology with promoters under the regulatory control of RpoN (σ^{54}). Indeed it was shown that this upstream region of AlgK contained the GG-N_{10}-GC consensus typified by all the promoters under RpoN control as well as *algD*, *algR1* and *algC* involved in *P. aeruginosa* alginate synthesis. Transcription of *algD* and *algC* under the control of their RpoN promoters has been

shown to be induced by environmental factors typically exhibited in the CF lung. These factors include high osmolarity, dehydration, nitrogen and phosphate limitation. Thus, the homology between these promoter regions would predict that the *algK* gene was also under similar control mechanism and thus transcription may be induced or up-regulated by specific environmental factors.

However, more recent data has shown that the σ factor that controls the conversion to the constitutive mucoid phenotype in *P. aeruginosa* that is exhibited in the CF lung, is actually σ^E (encoded by *algU*). This σ factor is thought to recognise an entirely different consensus which is not exhibited in the *algK* upstream region and predicts that *algK* transcription is not directly regulated by the σ^E sigma factor at this putative internal promoter. In agreement with this, Totten *et al.* (1990) showed that inactivation of the *rpoN* gene within a mucoid *P. aeruginosa* strain had no effect on the mucoid phenotype. Thus, *algK* 'promoter' activity is unlikely to be of much significance when considering the constitutive mucoid state exhibited by *P. aeruginosa* in CF where transcriptional control of the biosynthetic cluster is predominately via the *algD* promoter. Nevertheless, it may well be possible that for some alginate genes σ^{54} is involved in the maintenance or control of minor stress responses that require alginate synthesis.

Initial results from overexpression surveys of AlgK within both *E. coli* and *P. aeruginosa* gave further credence to the observation that proteins of the biosynthetic pathway are produced in very low quantities. Even when *algK* was under the transcriptional control of the extremely strong pT7 and p*tac* promoters, no observable protein was produced. Indeed visible overexpression of AlgK required the addition of a powerful ribosome binding site from vector pT7-7. Once *algK* was under the control of both a strong promoter and RBS, the AlgK product was clearly observable in both *E. coli* and *P. aeruginosa* hosts. Overexpression from *algK* resulted in the production of a large diffuse protein band at approximately 50 to 53-kDa. The close

correlation between the predicted molecular weight of the AlgK and that of the overexpressed product indicated that the single long ORF predicted from the nucleic acid sequence is likely to be correct.

Overexpression and detection of AlgK by Coomassie Brilliant Blue staining, in itself failed to reveal any evidence of post-translational modification, the protein band probably being too diffuse to exhibit any alterations in molecular weight. However, [^{35}S]Methionine labelling of AlgK proteins provided corroboratory evidence that a post-translational modification event was occurring. This event appeared to result in a loss of molecular weight of approximately 3-kDa, and as such was entirely consistent with the cleavage of the 27 amino acid lipoprotein signal sequence (which has a predicted molecular weight of approximately 3-kDa).

The initial results from the AlgK-β-lactamase fusion survey was to give a strong indication of the subcellular location of the bulk of the AlgK protein. Concentrations of ampicillin required to lyse *E. coli* cells (single cell minimum inhibitory concentrations) expressing each truncated versions of 'mature' AlgK fused to β-lactamase were over 800 μg/ml of ampicillin. This predicted that the β-lactamase domain of the fusion must be in the required periplasmic position to block the inhibitory action of the β-lactamase and as this occurred for all fusions generated throughout the mature protein it appears likely that AlgK is an entirely periplasmic oriented protein. Only one AlgK-β-lactamase fusion protein did not direct the β-lactamase moiety to the periplasm and this contained the N-terminal seven amino acids of the putative preprotein. These seven amino acids are hydrophobic but probably not sufficiently so for targeting of the fusion protein into or across the cytoplasmic membrane and as such the fusion protein retains a cytoplasmic location.

Visualisation of the fusions by immunoblotting (Western blots) gave the clearest indication of a post-translational N-terminal signal sequence cleavage event. A post-translational modification event was predicted by the fact that all the putative

periplasmically oriented fusion proteins migrated approximately 3-4-kDa smaller than that predicted for the combined fusion protein (truncated AlgK plus β-lactamase) molecular weight. The fact that the cytoplasmically oriented fusion protein with only seven amino acids fused to β-lactamase migrated as predicted but comigrated with the next smallest truncated protein with 42 amino acids from AlgK fused to β-lactamase suggested that a post-translational cleavage event must be occurring between positions 7 and 42 of the N-terminus. Thus the N-terminus sequence before amino acid residue 42 of AlgK must both target the protein from the cytoplasm to the periplasm (as predicted by the single cell MIC's) whilst also containing a signal that results in the removal of the majority of this region *in vivo* (as the two truncated fusions co-migrate). This situation is highly analogous to the cleavage and removal of a signal sequence. Cleavage of the N-terminus sequence of AlgK can only be by signal peptidase II, as signal peptidase I which cleaves standard signal peptides does not recognise signal peptidase II sequences and a signal peptidase I consensus sequence is not present at the N-terminus of AlgK. Thus it is highly likely that AlgK is a periplasmically oriented lipoprotein, anchored to either of the periplasmically exposed leaflets of the inner or outer membrane. As a cleavage event by signal peptidase II has been shown, the ATG translational start of AlgK is also predicted to be correct as this start point alone contains all the signals required for signal peptidase II cleavage.

Although the AlgK-β-lactamase fusion survey predicted that mature AlgK was an entirely periplasmically orientated membrane lipoprotein it gave no clue as to whether the protein was anchored to the outer face of the inner membrane or the inner face of the outer membrane. However, gathering information about the proposed membrane anchor position of AlgK may prove to be of little relevance when considering the work previously published by Grabert *et al.* (1990). Grabert *et al.* (1990) performed an extensive survey on mucoid and non-mucoid revertants of *P.*

aeruginosa strains revealed only a single novel protein band in the non-mucoid outer membrane. This 54-kDa protein was detected in gels independently of the method applied to isolate the outer membranes including both differential detergent solubilisation and also, sucrose density gradient centrifugation. This protein was subsequently isolated and its N-terminal amino acids sequenced. Chu *et al.* (1991) upon nucleic acid sequencing of *algE* determined that the 54-kDa outer membrane protein preliminary identified by Grabert *et al.* (1990) corresponded to the mature AlgE protein. The work by Grabert *et al.* predicts two main points, firstly that it is unlikely that AlgK is an outer membrane-associated protein as it would be obvious from the membrane protein data presented by Grabert *et al.* (1990). Secondly, if AlgK was not isolated because of the procedures used then new methods of isolating membranes would have to be developed as the subcellular isolation procedures used by Grabert *et al.* (1990) were essentially the same as those commonly used to isolate membrane-associated lipoproteins (i.e. differential detergent solubilisation with lauryl sarcosinate and discontinuous sucrose gradients of sonicated cells, Gennity and Inouye, 1991). Thus it would appear likely that AlgK is anchored to the outer leaflet of the inner membrane. Nevertheless, it may prove useful to definitively show the position of AlgK anchorage.

Insertional inactivation of chromosomal *algK* within a mucoid *P. aeruginosa* strain due to the single homologous recombination of a suicide vector bearing a suitable fragment of *algK*, was also achieved in the present study. Inactivation of chromosomal *algK* resulted in the loss of the mucoid phenotype. However, initial analysis using pAA23 which overexpressed AlgK within the mutant (PABK), failed to restore the mucoid phenotype. This result predicted that the conversion from a mucoid to non-mucoid phenotype was at least partially due to a polar effect on transcription of genes downstream of AlgK that are essential in the production of the mucoid phenotype. This result in itself provided additional evidence towards the

theory that the 34 minute alginate biosynthetic cluster exists as an operon (Chitnis and Ohman, 1993). In this scenario a single polycistronic message is transcribed from the beginning of the cluster (*algD*) and insertion of the suicide vector within the chromosomally located *algK* blocks the transcription of the genes downstream of the site of insertion. In order to circumvent this problem a series of broad host range vectors were constructed bearing the previously characterised 'essential' genes downstream of *algK*. However, pAA33 which harbours *algK*, *algE* and *algA* genes under the control of the p*tac* promoter also failed to complement the *algK* mutation and mucoidy was not restored. This suggests that an uncharacterised region downstream of *algK* must also be essential for the mucoid phenotype - this is most likely to be the *alg60* loci (as Boyd *et al*, 1993 showed that the genes between *alg60* and *algA* are not essential). Unfortunately, broad host range vectors containing all the required essential genes downstream of *algK* were not available and it proved impossible under the time constraints and with the vectors available in the laboratory, to construct a suitable complementation vector to overcome this problem.

Thus, in summary the region between *alg44* and *algE* contains a novel gene *algK*. This gene encodes AlgK, a probable membrane lipoprotein with a mature molecular weight of approximately 50-kDa. This lipoprotein is orientated in the periplasm and is most likely anchored to the outer leaflet of the cytoplasmic membrane. AlgK appears to have no homologues within any of the protein databases and shares little functional or structural amino acid identity/similarity with any other characterised proteins. Although incomplete, this data nevertheless narrows the putative function of AlgK within alginate EPS production.

P. aeruginosa alginate EPS exists as a linear co-polymer of β(1-4)-linked D-mannuronate (2, 3 di-*O*-acetylated) and its C-5 epimer L-guluronate. Information on the genes involved in alginate synthesis has revealed the vast majority of the proteins required for the steps from the activated glucose 6-phosphate precursor to the

completed exopolysaccharide. The only steps thought to be required and as yet unclarified which may be encoded within the 34 minute biosynthetic cluster, are those of polymerisation of GDP-mannuronic acid to polymannuronic acid, export through the cytoplasmic membrane and periplasm, a second acetylation enzyme and an acyl carrier protein. If the synthesis of alginate exopolysaccharide is similar to all characterised EPS producing Gram-negative bacteria then it would appear extremely probable that the polymerisation event would occur on the cytoplasmic face of the inner membrane. The subcellular orientation of such an event entirely excludes the periplasmic AlgK protein from such a role. AlgK can also be excluded from a central role in cytoplasmic membrane polysaccharide transport as this invariably requires a hydrophobic, basic, polytopic integral cytoplasmic membrane spanning protein, a role which this highly hydrophilic and acidic AlgK protein is unlikely to fulfil.

Thus, it would appear likely that AlgK must either act in a polymannuronate/alginate modification capacity or perhaps have a role in post-cytoplasmic membrane transport. However, a possible alginate modification function is not obvious as May and Chakrabarty (1994) revealed that previous unpublished work had identified an acyl-carrier protein encoded upstream of *algF* (the gene for an alginate acetylase). Also, Shinabarger *et al.* (1993) demonstrated that *P. aeruginosa* with an insertionally deactivated *algF* gene could be complimented back to a mucoid phenotype upon expression of *algA* in *trans*. However, this EPS was devoid of acetylated mannuronate residues. Upon introduction of a BHR vector expressing both *algF* and *algA* an acetylated EPS was observed. The levels of acetylation in this EPS was equivalent to that of EPS in the constitutive mucoid strain. Complete acetylation upon expression of AlgF predicts that either AlgF was solely required for acetylation of polymannuronate or that the protein is required in a more complex acetylation mechanism involving more than

one protein, this process only being functional in the presence of AlgF. It appears most likely that AlgF is indeed the exclusive acetylase which functions in the modification of polymannuronate residues. No other obvious alginate modification enzymes appear to be required for the production of the completed exopolysaccharide.

Possibly the most likely candidate for AlgK function is that of a periplasmic alginate transport protein. As mentioned previously, periplasmically oriented proteins putatively involved in EPS transport have been identified for *E. coli*, *R. meliloti*, *H. influenzae*, *N. meningitidis* and *A. xylinum*. The best characterised of these proteins are the periplasmically orientated proteins KpsD and KpsE involved in CPS synthesis in *E. coli*. Initial studies on *kpsE kpsD* deletion mutants by Bronner *et al.* (1993) revealed a loss of surface expression of the polysialic acid (PSA) capsule. Electron micrograph (EM) surveys indicated that the polysialic acid accumulates in the periplasmic space of mutant cells. Analysis by Wunder *et al.* (1994) on a *kpsD* mutant revealed similar accumulation. Both KpsD and KpsE were found to have a periplasmic orientation. KpsD has an N-terminal signal sequence which was cleaved and resulted in a periplasmic protein (Wunder *et al.*, 1994) and KpsE was anchored to the cytoplasmic membrane by N and C-terminal membrane spanning α-helices. From EM and protein orientation data both groups postulated that KpsD and/or KpsE were involved in polysaccharide export after its translocation across the cytoplasmic membrane (by KpsM/T). It was postulated that these proteins may be components of a large biosynthetic-translocation complex that spans both inner and outer membranes. This complex would generate a pore or channel through which the polymer passes to the surface of the bacterial cell, by-passing the periplasm. Alternatively, as the *E. coli* PSA EPS has previously been reported to exit the cell where cytoplasmic and outer membranes are in close apposition (Krönke *et al.*, 1990) the KpsD and/or KpsE may participate in a transient membrane association.

The periplasmic ExoF protein involved in the production of succinoglycan EPS of *R. meliloti* was (by similarity to KpsD) thought to have a similar periplasmic transport function (review in Leigh and Walker, 1994). Also the primary structure of KpsE is similar to that of BexC of *H. influenzae* and CtrB of *N. meningitidis*, components of the respective EPS biosynthetic pathways. Indeed the KpsE sequence is approximately 30% identical and 70% similar to that of BexC and CtrB, respectively (Cieslewicz *et al.*, 1993). Both BexC (Kroll *et al*, 1990) and CtrB (Frosch *et al*, 1991) are predicted to have a periplasmic orientation. The location of each of the genes (encoding these periplasmic proteins) within an operon containing genes for the inner membrane ABC transporter and the outer membrane exopolysaccharide porin of *H. influenzae* and *N. meningitidis* suggested that BexB and CtrC are the periplasmic components of their respective CPS transporters (Kroll *et al.*, 1990 and Frosch *et al.*, 1991, respectively).

A similar genetic arrangement is observed for cellulose synthesis in *A. xylinum* (Saxena *et al.*, 1994). In this bacterium the genes encoding the polymerase (*acsAB*) are in an operon with the cellulose outer membrane porin gene *acsC* and *acsD*, a gene encoding a putative periplasmic protein. AcsD was thought either to be involved with the outer membrane pore structure or the organisation of the pores within the terminal synthesising complexes (intramembranous particles associated with cellulose synthesis and extrusion). In the absence of the AcsD polypeptide, the AcsC pore structure is present and capable of cellulose ribbon synthesis. However, the pore appears to be more sensitive to the exiting cellulose product which clogs the pore, restricting further polymerisation. Thus, AcsD appears to be required for the efficient transfer of cellulose from synthesis at the inner membrane by cellulose synthase (AcsAB) to the outer membrane pore.

Very recently, Saxena *et al.* (1995) demonstrated (using hydrophobic cluster analysis) that Alg8 showed similarity in overall structure and thus possibly function

to the putative polymerisation region of AcsAB. This similarity predicted that Alg8 belonged to a family of β-glycosyl transferases which putatively polymerise their substrates by a double addition mechanism whereby growth of the polysaccharide at its reducing end, using the nucleotide diphospho sugar as the substrate, is by the sequential addition of further nucleotide diphospho sugars in groups of two. Interestingly, polymerisation by the double addition mechanism does not require a primer, such as undecaprenol. No inner membrane transport protein for polysaccharide synthesis has been identified but the model proposed by Ross *et al*, (1990) suggested that the cellulose synthase protein may be associated with both polymerisation and inner membrane transport functions. Indeed the other proposed members of this family, chitin synthase of *Saccharomyces cerevisiae*, (Bulawa *et al.*, 1986), hyaluronic acid synthase of Group A Streptococci (Dougherty and Rijn, 1994) and NodC of *Azorhizobium caulinodans* (Geremia *et al.*, 1994) appear to both polymerise and transport across the membrane that they are located within. The membrane position of both Alg8 and Alg44 had previously lead both Maharaj *et al.* (1993) and May and Chakrabarty (1994) to speculate that these proteins may fulfil similar functions. The location of genes encoding the periplasmic oriented AlgK protein and the outer membrane alginate permeable pore, AlgE directly downstream of genes encoding a mannuronate polymerase/inner membrane transporter is not dissimilar to the genetic arrangements of *bcs* (Wong *et al.*, 1990) and *acs* (Saxena *et al.*, 1993), the cellulose synthesis and transport operon of *A. xylinum* (*bcs* and *acs* are similar operons isolated from different *A. xylinum* strains), the *bex* locus of *H. influenzae* (Kroll *et al.*, 1990) and the *ctr* locus of *N. meningitidis* (Frosch *et al.*, 1991), suggesting that AlgK may have a role in the periplasmic transport of polymannuronate. Thus, it is conceivable that AlgK (as similarly postulated for KpsD) is a periplasmic component of a large biosynthetic-translocation complex. Such a complex may help explain why the proteins involved in alginate production in

P. aeruginosa are expressed in such low quantities even in highly mucoid strains (Sá-Correia *et al.*, 1987).

Functional and structural analysis of the AlgK protein can now be started as the protein can be overexpressed easily in both *E. coli* and *P. aeruginosa* hosts. Thus purification of sufficient denatured and stable protein may allow subsequent analysis of its three dimensional structure by X-ray diffraction analysis. Purification of AlgK may also allow the production of anti-AlgK antibodies which may prove useful in confirming the subcellular location of AlgK as well as providing a means for assessing AlgK/protein or AlgK/polymannuronate interactions in any biosynthetic or transport complex.

Analysis of the structure and more importantly, the function of AlgK is required for understanding of all the steps in alginate EPS production, from activated precursor to fully formed polymer and virulence factor. Understanding of the method of alginate production in *P. aeruginosa* is vitally important in the development of methods of colonisation control of mucoid *P. aeruginosa* within the CF patient. Prevention of the mucoid phenotype within CF patients should lead to the effective elimination of the infection either by antibiotic treatment or by enhancing the effectiveness of the immune response of the patient (which is usually blocked by the alginate coating). Blocking of the mucoid phenotype may also prevent the immune hyper-responsive reactions which are elicited by the mucoid microcolonies and which causes much of the lung damage.

Any prospective anti-*P. aeruginosa* or anti-mucoid agents and possibly vaccines can potentially be researched using mouse models. Dorin *et al.* (1992), Shouwaert *et al.* (1992) and Ratcliff *et al.* (1993) have all managed to generate transgenic mice with *cftr* mutations that very closely resemble the electrophysiological defect in human CF. These transgenic mice maintain a CF electrophysiological phenotype, but also exhibit long-term survival. This, coupled

with the ability to control the exposure of these mice to infection with specific CF associated pathogens may allow the close study of this aspect of the disease. Indeed, surveys on CF mice have recently progressed to examine the effect of pulmonary infections within the animal model. Initial experiments by Davidson *et al.* (1995) have followed pulmonary colonisation of CF mice by *S. aureus* and *B. cepacia*, which appears to closely resemble the infection exhibited in humans. The natural advancement of these studies is to survey the effect of *P. aeruginosa* infection and colonisation. After this work it should be possible to evaluate the potential of any anti-*P. aeruginosa* drug(s) or alginate blocking agents as well as possible vaccines in the mouse model prior to human testing.

References.

Ambler, R. P. and Scott, G. K. (1978). Partial amino acid sequence of penicillinase coded by *Escherichia coli* plasmid R6K. *Proceedings of the National Academy of Science USA* **75**, 3732-3736.

Ames, G. F. (1986). Bacterial periplasmic transport systems: structure mechanism and evolution. *Annual Review in Biochemistry* **55**, 397-425

Anderson, M. P., Gregory, R. J., Thompson, S., Souza, D. W., Paul, S., Mulligan, R. S., Smith, A. E. and Welsh, M. J. (1991). Demonstration that CFTR is a chloride channel by alteration of its anion selectivity. *Science* **253**, 202-205.

Anderson, R. G., Hussey, H. and Baddiley, J. (1972). The mechanism of wall synthesis in bacteria: the organisation of enzymes and isoprenoid phosphates in the membrane. *Biochemical Journal* **127**, 11-25.

Annunziato, P. W, Wright, L. F, Varin, W. F and Silver, R. P. (1995). Nucleotide sequence and genetic analysis of the *neuD* and *neuB* genes in Region 2 of the polysialic acid gene cluster of *Escherichia coli* K1. *Journal of Bacteriology* **177** 312-319.

Bagdasarian. M.M., Amann, E., Lurz, R., Ruckert, B. and Bagdasarian, M. (1983). Activity of the *trp-lac* promoter of *Escherichia coli* in *Pseudomonas putida*. Construction of broad host-range controlled expression vectors. *Gene* **26**, 273-282.

Bally, M., Filloux, A., Akrim, M., Ball, G., Lazdunski, A. and Tommassen, J. (1992). Protein secretion in *Pseudomonas aeruginosa*: characterisation of seven *xcp* genes and processing of secretory apparatus components by prepilin peptidase. *Molecular Microbiology* **6**, 1121-1131.

Banerjee, P.C., Vanags, R. I., Chakrabarty, A. M. and Maitra, P. K. (1985). Fructose-1,6-bisphosphate aldolase activity is essential for synthesis of alginate from glucose. *Journal of Bacteriology* **161**, 458-460.

Bayer, A. S., Eftekhar, F., Tu, J., Nast, C. C. and Speert, D. P. (1990). Oxygen dependent up regulation of mucoid exopolysaccharide (alginate) production in *Pseudomonas aeruginosa*. *Infection and Immunity* **58**, 1344-1349.

Bear, C. E., Li, C., Katner, N., Bridges, R. J., Jenson, T. J. Ramjeesingh, M. and Riordan, J. R. (1992). Purification and functional reconstitution of the cystic fibrosis transmembrane conductance regulator (CFTR). *Cell* **68**, 809-818.

Becker, A., Kleickmann, A., Arnold, W. and Puhler, A. (1993a). Analysis of the *Rhizobium meliloti exoH/exoK/exoL* fragment: ExoK shows homology to excreted endo β-1-3, 1-4 glucanases and ExoH resembles membrane proteins. *Molecular and General Genetics* **238**, 145-154.

Becker, A., Kleickmann, A., Arnold, W. and Puhler, A. (1993b). Identification and analysis of the *Rhizobium meliloti exoAMONP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. *Molecular and General Genetics* **241**, 367-371.

Berry, A., DeVault, J. D. and Chakrabarty, A. M. (1989). High osmolarity is a signal for enhanced *algD* transcription in mucoid and non-mucoid *Pseudomonas aeruginosa* strains. *Journal of Bacteriology* **171**, 2312-2317.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7, 1513-1523.

Bolivar, F. (1978). Construction and characterisation of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique *Eco*RI sites for selection of *Eco*RI generated recombinant DNA molecules. *Gene* **4**, 121-136.

Boulnois, G. J. and Jann, K. (1989). Bacterial polysaccharide capsule synthesis export and evolution of structural diversity. *Molecular Microbiology* **3**, 1819-1823.

Bowler, L. and Spratt, B. G. (1989). Membrane topology of penicillin binding protein 3 of *Escherichia coli*. *Molecular Microbiology* **3**, 1277-1286.

Boyd, A., Ghosh, M., May, T. B., Shinabarger, D., Keogh, R. and Chakrabarty, A. M. (1993). Sequence of the *algL* gene of *Pseudomonas aeruginosa* and purification of its alginate lyase product. *Gene* **131**, 1-8.

Bronner, D., Clarke, B. R. and Whitfield, C. (1994). Identification of an ATP-binding cassette transport system required for translocation of lipopolysaccharide O-antigen side-chains across the cytoplasmic membrane of *Klebsiella pneumoniae* serotype O1. *Molecular Microbiology* **14**, 505-519.

Bronner, D., Sieberth, V., Pazzani, C., Roberts, I. S., Boulnois, G. J., Jann, B. and Jann, K. (1993). Expression of the capsular K5 polysaccharide of *Escherichia coli*; biochemical and electron microscopic analysis of mutants with defects in region I of the K5 gene cluster. *Journal of Bacteriology* **175**, 5984-5992.

Broome-Smith, J. K. and Spratt, B. G. (1986). A vector for the construction of translational fusions to TEM β-lactamase and the analysis of protein export signals and membrane topology. *Gene* **49**, 341-349.

Bulawa, C. E., Slater, M., Cabib, E., An-Young, J., Sburlati, A., Adair, W. L. jr. and Robbins, P. W. (1986). The *Saccharomyces cerevisiae* structural gene for chitin synthase is not required for chitin synthesis *in vivo*. *Cell* **46**, 213-225.

Bye, M. R., Ewig, J. M. and Quittell, L. M. (1994). Cystic fibrosis. *Lung* **172**, 251-270.

Calamia, J and Manoil, C. (1990). *lac* permease of *Escherichia coli*: topology and sequence elements promoting membrane insertion. *Proceedings of the National Academy of Science USA* **87**, 4937-4941.

Cangelosi, G. A., Martinetti, G., Leigh, J. A., Lee, C-C., Theines, C. Nester, E. W. (1989). Role of *Agrobacterium tumefaciens* ChvA protein in export of β–1-2-glucan. *Journal of Bacteriology* **171**, 1609-1615.

Cheng, S. H., Rich, D. P. and Marshall, J. (1991). Phosphorylation of the R-domain cAMP-dependant protein kinase regulates the CFTR chloride channel. *Cell* **66**, 1027-1036.

Chitnis, C. E., and Ohman, D. E. (1990). Cloning of *Pseudomonas aeruginosa algG*, which controls alginate structure. *Journal of Bacteriology* **172**, 2894-2900.

Chu, L., May, T. B., Chakrabarty, A. M. and Misra, T. K. (1991). Nucleotide Sequence and expression of the *algE* gene involved in alginate biosynthesis by *Pseudomonas aeruginosa*. *Gene* **107**, 1-10.

Chung, C. T, Siemala, S. L. and Miller, R. H. (1989). One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Science USA* **86**, 2172-2175.

Cieslewicz, M. J., Steenbergen, S. M. and Vimr, E. R. (1993). Cloning and sequencing, expression and complementation analysis of the *Escherichia coli* K1 *kps* region 1 gene, *kpsE*, and identification of an upstream open reading frame encoding a protein with homology to GutQ. *Journal of Bacteriology* **175**, 8018-8023.

Clark, P. H. (1990). *Pseudomonas aeruginosa*, an opportunistic pathogen. In Pseudomonas aeruginosa *infection and alginates: biochemistry, genetics and pathology*, (Gacesa, P. and Russell, N. J. editors) pp1-12,Chapman and Hall, London.

Close, T. J. and Rodriguez, R. L. (1982). Construction and characterisation of the chloramphenicol resistance cartridge: A new approach to the transcriptional mapping of extra-chromosomal elements. *Gene* **20**, 305-316.

Coyne, M. J., Russel, K. S., Coyle, C. L. and Goldberg, J. B. (1994). The *Pseudomonas aeruginosa algC* gene encodes phosphoglucomutase, required synthesis of a complete lipopolysaccharide core. *Journal of Bacteriology* **176**, 3500-3507.

Darzins, A. and Chakrabarty, A. M. (1984). Cloning of genes controlling alginate biosynthesis from a mucoid cystic fibrosis isolate of *Pseudomonas aeruginosa*. *Journal of Bacteriology* **166**, 9-18.

Darzins, A., Frantz, B., Vanags, R. I. and Chakrabarty, A. M. (1986). Nucleotide sequence analysis of the phosphomannose isomerase gene (*pmi*) of *Pseudomonas aeruginosa* and comparison with the corresponding *Escherichia coli* gene *manA*. *Gene* **42**, 293-302.

Darzins, A., Wang, S-K, Vanags, R. I. and Chakrabarty A, M. (1985). Clustering of mutations affecting alginic acid biosynthesis in mucoid *Pseudomonas aeruginosa*. *Journal of Bacteriology* **167**, 516-524.

Datta, N. and Kontomichalou, P. (1965). Penicillinase synthesis controlled by infectious R factors in Enterobacteriaciae. *Nature* **318**, 618-624.

Davidson, D. J., Dorin, J. R., McLachlan, G., Ranaldi, Y., Lamb, D., Doherty, C., Govan, J. and Porteous, D. J. (1995). Lung disease in the cystic fibrosis mouse exposed to bacterial pathogens. *Nature Genetics* **9**, 351-357.

Davies, D. G., Chakrabarty, A. M. and Geesey, G. G. (1993). Exopolysaccharide production in biofilm-substratum. Activation of alginate gene-expression by *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology* **59**, 1181-1186.

Deretic, V., Dikshit, R., Konyecsni, W. M and Chakrabarty, A. M. (1989). The *algR* gene, which regulates mucoidy in *Pseudomonas aeruginosa*, belongs to a class of environmentally responsive genes. *Journal of Bacteriology* **171**, 1278-1283.

Deretic, V., Gill, G. F and Chakrabarty, A. M. (1987). *Pseudomonas aeruginosa* infection and cystic fibrosis: nucleotide sequence and transcriptional regulation of the *algD* gene. *Nucleic Acids Research* **15**, 4567-4581.

Deretic, V., Govan, J. R. W., Konyecsni, W. M. and Martin, D. W. (1990). Mucoid *Pseudomonas aeruginosa* in cystic fibrosis: mutations in the *muc* loci affect transcription of the *algR* and *algD* genes in response to environmental stimuli. *Molecular Microbiology* **4**(2) 189-196.

Deretic, V. and Konyecsni, W. M. (1990). A prokaryotic regulatory factor with a histone H1-like carboxy-terminal domain: clonal variation of repeats within *algP*, a gene involved in regulation of mucoidy in *Pseudomonas aeruginosa*. *Journal of Bacteriology* **172**, 5540-5554.

Deretic, V., Leveau, J. H. J., Mohr, C. D. and Hibler, N. S. (1992). *In vitro* phosphorylation of AlgR, a regulator of mucoidy in *Pseudomonas aeruginosa*, by a

histidine protein kinase and effects of small phospho-donor molecules. *Molecular Microbiology* 19, 2761-2767.

Deretic, V., Martin, D. W., Schurr, M. J., Mudd, M. H., Hibler, N. S., Curcic, R. and Boucher, J. C. (1993). Conversion to mucoidy in *Pseudomonas aeruginosa*. Bio/Technology 11, 1133-1136.

Deretic. V and Konyecsni, W. M. (1989). Control of mucoidy in *Pseudomonas aeruginosa*: transcriptional regulation of *algR* and identification of the second regulatory gene, *algQ*. *Journal of Bacteriology* **171**, 3680-3688.

DeVault, J. D., Kimbara, K. and Chakrabarty, A. M. (1990). Pulmonary dehydration and infection in cystic fibrosis: evidence that ethanol activates alginate gene expression and induction of mucoidy in *Pseudomonas aeruginosa*. *Molecular Microbiology* 4, 737-745.

DeVault, J. D., Berry, A., Misra, T. K., Darzins, A. and Chakrabarty, A. M. (1989). Environmental sensory signals and microbial pathogenesis: *Pseudomonas aeruginosa* infection in cystic fibrosis. *Bio/technology* 7, 352-357.

Devereux. J, Haebek, P. and Smithes, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387-395.

DeVries, C. A. and Ohman, D. E. (1994). Mucoid to non-mucoid conversion in alginate-producing *Pseudomonas aeruginosa* often results from spontaneous mutations in *algT*, encoding a putative alternate sigma factor and shows evidence for autoregulation. *Journal of Bacteriology* **176**, 6677-6687.

Doggett, R. G. and Harrison, G. M. (1972). *Pseudomonas aeruginosa*: Immunostatus in patients with cystic fibrosis. *Infection and Immunity* **6**, 628-635.

Dorin, J. R, Dickinson, P., Alten, E. W. F. W., Smith S. N., Geddes, D. M., Stevenson, B. J., Kimber, W. L., Fleming, S., Clarke, A. R., Hooper, M. L., Aderson, L., Beddington, R. S. P. and Porteous, D. J. (1992). Cystic fibrosis in the mouse by targeted insertional mutagenesis. *Nature* **359**, 211-215.

Dougherty, B. A. and van de Rijn, I. (1994). Molecular characterisation of *hasA* from an operon required for hyaluronic acid synthesis in Group A Streptococci. *Journal of Biological Chemistry* **269**, 169-175.

Drake, J. W. and Baltz, R. H. (1976). The biochemistry of mutagenesis. *Annual Review of Biochemistry* **45**, 11-37.

Duong, F., Lazdinski, A., Cami, B. and Murgier, M. (1992). Nucleotide sequence of a cluster of genes involved in synthesis and secretion of alkaline protease in

Pseudomonas aeruginosa: relationships to other secretion pathways. *Gene* **121**, 47-54.

Edelman, A., Bowler, L., Broome-Smith, J. K. and Spratt, B. G. (1987). Use of the β–lactamase fusion vector to investigate the organisation of penicillin-binding protein 1B in the cytoplasmic membrane of *Escherichia coli*. *Molecular Microbiology* **1**, 101-106.

Erickson, J. W. and Gross, C. A. (1989). Identification of the subunit of *Escherichia coli* RNA polymerase: a second alternate factor involved in high-temperature gene expression. *Genes and Development* **3**, 1462-1471.

Erickson, J. W., Vaugn, W., Walter, W. A, Niedhardt, F. C. and Gross, C. A. (1987). Regulation of the promoters and transcripts of *rpoH*, the *Escherichia coli* heat shock regulatory gene. *Genes and Development* **1**, 419-432.

Ertesvåg, H., Doseth, B., Larsen, B., Skjåk-Bræk, G. and Valla, S. (1994). Cloning and expression of an *Azotobacter vinelandii* mannuronan C-5-epimerase gene. *Journal of Bacteriology* **176**, 2846-2853.

Evans, L. R. and Linker, A. (1973). Production and characterisation of the slime polysaccharide of *Pseudomonas aeruginosa*. *Journal of Bacteriology* **116**, 915-924.

Fath, M. J. and Kolter, R. (1993). ABC transporters: Bacterial exporters. *Microbiological Reviews* **57**, 995-1017.

Fett, W.F., Osman, S. F., Fishman, M. L. and Siebles, T. S. III. (1986). Alginate production by plant pathogenic pseudomonads. *Applied and Environmental Microbiology* 52, 466-473.

Fett, W. F., Wijey, C. and Lifson, E. R. (1992). Occurrence of alginate gene sequences among members of the pseudomonad rRNA homology groups I-IV. *FEMS Microbiology Letters* **99**, 151-158.

Figurski, D. H. and Helinski, D. R. (1979). Replication of an origin-containing derivative of plasmid pRK2 dependent on a plasmid function provided in *trans*. *Proceedings of the National Academy of Science USA* **76**, 1648-1652.

Finke, A., Bronner, D., Nikolaev, A. V., Jann, B. and Jann, K. (1991). Biosynthesis of the *Escherichia coli* K5 polysaccharides: polymerisation *in vitro* and characterisation of the product. *Journal of Bacteriology* **173**, 4088-4094.

Flynn, J. L. and Ohman, D. E. (1988). Cloning of genes from mucoid *Pseudomonas aeruginosa* which control spontaneous conversion to the alginate production phenotype. *Journal of Bacteriology* **170**, 1452-1460.

Franklin, M. J., Chitnis, C. E, Gacesa, P., Sonesson, A., White, D. C. and Ohman, D. E. (1994). *Pseudomonas aeruginosa* AlgG is a polymer level alginate C5-mannuron epimerase. *Journal of Bacteriology* **176**, 1821-1830.

Frosch, M., Edwards, V., Bousset, K., Kraube, B. and Weisgerber, C. (1991). Evidence for a common molecular origin of the capsule gene loci in Gram-negative bacteria expressing Group II capsular polysaccharides. *Molecular Microbiology* **5**, 1251-1263.

Frosch, M. and Müller, A. (1993). Phospholipid substitution of capsular polysaccharides and mechanisms of capsule formation in *Neisseria meningitdis*. *Molecular Microbiology* **8**, 483-493.

Frosch, M., Müller, D., Bousset, K. and Müller, A. (1992). Conserved outer membrane protein of *Neisseria meningitidis* involved in capsule expression. *Infection and Immunity* **60**, 798-803.

Frosch, M., Weisgerber, C. and Meyer, T. F. (1989). Molecular characterisation and expression in *Escherichia coli* of the gene encoding the polysaccharide capsule of *Neisseria meningitidis* Group B. *Proceedings of the National Academy of Sciences. USA.* **86**, 1669-1673.

Fürste, J. P., Pansegrau, W., Frank, R., Blöcker, H., Sholz, P., Bagdasarian, M. and Lanka, E. (1986). Molecular cloning of the plasmid RP4 primase region in multi-host-range *tacP* expression vector. *Gene* **48**, 119-131.

Gacesa, P. (1988). Alginates. *Carbohydrate Polymers* **8**, 2-21.

Gallagher, M. P., Pearce, S. R. and Higgins, C. F. (1989). Identification and localisation of the membrane associated, ATP binding subunit of the oligopeptide permease of *Salmonella typhimurium*. *European Journal of Biochemistry* **180**, 133-141.

Gennity, J. M. and Inouye, M. (1992). Structural determinants in addition to the amino-terminal sorting sequence influence membrane localisation of *Escherichia coli* lipoproteins. *Journal of Bacteriology* **174**, 2095-2101.

Geremia, R. A., Mergaert, P., Geelan, D., Van Montagu, M. and Holsters, M. (1994). The NodC protein of *Azorhizobium caulinodans* is an N-acetylglucosaminyl transferase. *Proceedings of the National Academy of Science USA* **91**, 2669-2673.

Glucksmann, M. A., Reuber, T. L. and Walker, G. C. (1993). Family of glycosyl transferases needed for the synthesis of succinoglycan by *Rhizobium meliloti*. *Journal of Bacteriology* **175**, 7033-7044.

Glucksmann, M. A., Reuber, T. L. and Walker, G. C. (1993) Genes needed for the modification, polymerisation, export and processing of succinoglycan by *Rhizobium meliloti*: a model for succinoglycan biosynthesis. *Journal of Bacteriology* **175**, 7045-7055.

Goldberg, J. B. and Dahnke, T. (1992). *Pseudomonas aeruginosa algB*, which modulates the expression of alginate, is a member of the *ntrC* subclass of prokaryotic regulators. *Molecular Microbiology* **6**, 59-66.

Goodell, E. W. and Higgins, C. F. (1987). Uptake of Cell Wall peptide by *Salmonella typhimurium* and *Escherichia coli*. *Journal of Bacteriology* **169**, 3861-3865.

Gorin, P. A. T. and Spencer, J. F. T. (1966). Exocellular alginic acid from *Azotobacter vinelandii*. *Canadian Journal of Chemistry* **44**, 993-998.

Govan, J. R. W. (1988). Alginate biosynthesis and other unusual characteristics associated with the pathogenesis of *Pseudomonas aeruginosa* in cystic fibrosis p67-96. In Griffiths, E., Donachie, W and Stephen, J (ed.). *Bacterial infection in respiratory and gastrointestinal mucosae*. IRL Press, Oxford.

Govan, J. R. W. and Fyfe, J. A. M. (1978). Mucoid *Pseudomonas aeruginosa* and cystic fibrosis: Resistance of mucoid form to carbenicillin, flucloxacillin and tobramycin and isolation of mucoid variants *in vitro*. *Journal of Antimicrobial Chemotherapy* **4**, 233-240.

Govan, J. R. W., Fyfe, J. A. M. and Jarman, T. R. (1981). Isolation of alginate producing mutants of *Pseudomonas fluorescens*, *Pseudomonas mendocina* and *Pseudomonas putida*. *Journal of General Microbiology* **125**, 217-220.

Govan, J. R. W. and Glass, S. (1990). The microbiology and therapy of cystic fibrosis lung infections. *Clinical Microbiological Reviews* **1**, 19-28.

Govan, J. R. W. and Harris, G. S. (1986). *Pseudomonas aeruginosa* and cystic fibrosis, unusual bacterial adaptations and pathogenesis. *Microbiological Sciences* **3**, 302-308.

Govan, J. R. W., Martin, D. W. and Deretic, V. P. (1992). Mucoid *Pseudomonas aeruginosa* and cystic fibrosis: the role of mutations in *muc* loci. *FEMS Microbiological Letters* **100**, 323-330.

Govan, J. R. W. and Nelson, J. W. (1992). Microbiology of lung infection in cystic fibrosis. *British Medical Bulletin* **48**, 912-931.

Grabert, E., Wingender, J. and Winkler, V. K. (1990). An outer membrane protein characteristic of mucoid strains of *Pseudomonas aeruginosa. FEMS Microbiology Letters* **68**, 83-88.

Graham, L. L and Beveridge, T. J. and Nanninga, N. (1991). Periplasmic space and the concept of the periplasm. *Trends in Biochemical Sciences* **16**, 328-329.

Harris, A. (1992). Cystic fibrosis gene. *British Medical Bulletin* 48, 737-754.

Hayashi, S., Chang, S-Y., Chang, S. and Wu, H. C. (1984). Modification and processing of *Bacillus licheniformis* prepenicillinase in *Escherichia coli.* Fate of mutant penicillinase lacking a lipoprotein modification site. *Journal of Biological Chemistry* **259**, 10448-10454.

He, M., Kaderhai, M. A., Adcock, I. and Austen, B. M. (1991). An improved and rapid procedure for isolating RNA-free *Escherichia coli* plasmid DNA. *Genetical Analysis: Techniques and Applications* **8**, 107-110.

He, M., Wilde, A and Kadershai, M. A. (1990). A simple single-step procedure for small-scale preparation of *Escherichia coli* plasmids. *Nucleic Acids Research* **18**, 1660.

Henikoff, S. (1987). Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods in Enzymology* **155**, 156-165.

Higgins, C. F. (1992). Cystic fibrosis transmembrane conductance regulator (CFTR). (1992). *British Medical Bulletin* **48**, 754-766.

Higgins, C. F., Haag, P. D., Nikaido, K., Ardeshir, F., Garcia, G. and Ames, G. F. (1986). Complete nucleotide sequence and identification of membrane components of the histidine transport operon of *Salmonella typhimurium. Nature* **298**, 723-727.

Higgins, C. F., Hyde, S. C., Mimmack, M. M., Gileadi, U., Gill, D. R. and Gallagher. M. P. (1990). Binding protein-dependent transport systems. *Journal of Bioenergetics and Biomembranes* **22**, 571-592.

Hiles, I. D., Gallagher, M. P., Jamieson, D. J. and Higgins, C. F. (1987). Molecular characterisation of the oligopeptide permease of *Salmonella typhimurium. Journal of Molecular Biology* **195**, 125-142.

Hodson, M. E. and Warner, J. O. (1992). Respiratory problems and their treatment. *British Medical Bulletin* **48**, 931-949.

Hoyle, B. D., Williams, L. J. and Costerton, J. W. (1993). Production of mucoid exopolysaccharide during development of *Pseudomonas aeruginosa* biofilms. *Infection and Immunity* **61**, 777-780.

Hyde, S. C., Emsley, P., Hartshorm, M. J., Mimmack, M. M., Gileadi, V., Pearce, S. R., Gallagher, M. P., Gill, D. R., Hubbard, R. E. and Higgins, C. F. (1990). Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* **346**, 362-365.

Ielpi, L., Couso, R. D. and Dankert, M. A. (1993). Sequential assembly and polymerisation of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. *Journal of Bacteriology* **175**, 2490-2500.

Ish-Horowicz, D. and Burke, J. F. (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Research* **9**, 2989-2998.

Izard, J. W. and Kendall, D. A. (1994). Signal peptides-exquisitely designed transport promoters. *Molecular Microbiology* **12**, 765-773.

Jann, K. and Jann, B. (1990). Structure and biosynthesis of the capsular antigens of *Escherichia coli*. *Current Topics in Microbiology and Immunology* **150**,18-40.

Kadonago, J. T., Gautier, A. E., Strauss, D. R., Charles, A. D., Edge, M. D. and Knowles, J. R. (1984). The role of the β-lactamase signal sequence in the secretion of proteins by *Escherichia coli*. *Journal of Biological Chemistry* **259**, 2149-2154.

Kadonago, J. T. and Knowles, J. R. (1985). A simple and efficient method for chemical mutagenesis of DNA. *Nucleic Acids Research* **13**, 1733-1745.

Kato, J. and Chakrabarty, A. M. (1991). Purification of the regulatory protein AlgR1 and its binding in the far upstream region of the *algD* promoter in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Science USA* **88**, 1760-1764.

Kato , J., Chu, L., Kitano, K., DeVault, J. D., Kimbara., K., Chakrabarty, A. M. and Misra, T. K. (1989). Nucleotide sequence of a regulatory region controlling alginate synthesis in *Pseudomonas aeruginosa*: characterisation of the *algR2* gene. *Gene* **84**, 31-38.

Kato, J., Misra, T. K. and Chakrabarty, A. M. (1990). AlgR3, a protein resembling eukaryotic histone H1, regulates alginate synthesis in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Science USA* **87**, 2887-2891.

Kimbara, K. and Chakrabarty, A.M. (1989). Control of alginate synthesis in *Pseudomonas aeruginosa*, regulation of the *algR1* gene. *Biochemical and Biophysical Research Communications* **164**, 601-608.

Knecht, D. A. and Dimond, R. L. (1984). Visualisation of antigenic proteins on Western Blots. *Analytical Biochemistry* **136**, 180-184..

Konyecsni, W. M. and Deretic, V. (1990). DNA sequence and expression analysis of *algP* and *algQ*, components of the multigene system transcriptionally regulating mucoidy in *Pseudomonas aeruginosa*: *algP* contains multiple direct repeats. *Journal of Bacteriology* **172**, 2511-2520.

Köplin, R., Arnold, W., Hötte, B., Simon, R., Wang, G. and Pühler, A. Genetics of xanthan production in *Xanthomonas campestris*: the *xanA* and *xanB* genes are involved in UDP-glucose and GDP-mannose biosynthesis. *Journal of Bacteriology* **174**, 191-199.

Kroll, J. S., Loynds, B., Brophy, L. N. and Moxon, E. R. (1990). The *bex* locus in encapsulated *Haemophilus influenzae*: a chromosomal region involved in capsule polysaccharide export. *Molecular Microbiology* **4**,1853-1862.

Kroll, J. S., Zamze, S., Loynds, B. and Moxon, E. R. (1989). Common organisation of chromosomal loci for production of different capsular polysaccharides in *Haemophilus influenzae*. *Journal of Bacteriology* **171**, 3343-3347.

Krönke, K. D., Boulnois, G., Roberts, I., Bitter-Suermann, D., Golecki, J. R., Jann, B. and Jann, K. (1990). Expression of the *Escherichia coli* K5 capsular antigen: immunoelectron microscopic and biochemical studies with recombinant *E. coli*. *Journal of Bacteriology* **172**, 3469-3472.

Kustu, S., Santero, E., Keener, J., Poham, D and Weiss, D. (1989). Expression of sigma-54 (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiological Reviews* **53**, 367-376.

Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying hydropathic character of a protein. *Journal of Molecular Biology* **157**, 195-120.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* **227**, 680-685.

Lam, J., Chan, R. and Costerton, J. W. (1980). Production of mucoid microcolonies by *Pseudomonas aeruginosa* within infected lungs in cystic fibrosis. *Infection and Immunity* **28**, 546-556.

Lee, C. J. (1987). Bacterial capsular polysaccharides, biochemistry, immunity and vaccine. *Molecular Immunology* **24**, 1005-1019.

Leigh, J. A. and Walker, G. C. (1994). Exopolysaccharides of Rhizobium: synthesis, regulation and symbiotic function. *Trends in Genetics* **10**, 63-67.

Lin, W. S., Cuneen, T. and Lee, C. Y. (1994). Sequence analysis and molecular characterisation of genes required for the biosynthesis of type 1 capsular polysaccharides in *Staphylococcus aureus*. *Journal of Bacteriology* **176**, 7005-7016.

Lin, T. Y. and Hassid, W. Z. (1966). Pathway of alginic acid synthesis in the marine brown algae, *Fucus gardneri silva*. *Journal of Biological Chemistry* **241**, 5284-5297.

Linker, A. and Evans, L. R. (1984). Isolation and characterisation of an alginase from mucoid strains of *Pseudomonas aeruginosa*. *Journal of Bacteriology* **159**, 958-964.

Linker, A. and Jones, R. S. (1966). A new polysaccharide resembling alginic acid isolated from Pseudomonads. *Journal of Biological Chemistry* **241**, 3845-3851.

Lipinska, B., Sharma, S. and Georgopaulus, C. (1988). Sequence analysis and regulation of the *htrA* gene of *Escherichia coli*: a σ^{32}-independent mechanisms of heat inducible transcription. *Nucleic Acids Research* **16**, 10053-10067.

Maharaj, R., May, T. B., Wang, S-K. and Chakrabarty, A. M. (1993). Sequence of *alg8* and *alg44* genes involved in the synthesis of alginate by *Pseudomonas aeruginosa*. *Gene* **136**, 267-269.

Mai, G. T., McCormack, J. G., Seow, W. K., Pier, G. B., Jackson, L. A. and Thong , Y. H. (1993). Inhibition of adherence of mucoid *Pseudomonas aeruginosa* by alginase specific monoclonal antibodies and antibiotics. *Infection and Immunity* **61**, 4338-4343.

Maniatis, T., Fritsch, E. and Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbour, New York; Cold Spring Harbour Laboratory Press.

Martin, D. W, Schurr, M. J., Mudd, M. H. and Deretic, V. (1993). Differentiation of *Pseudomonas aeruginosa* into the alginate producing form: inactivation of *mucB* causes conversion to mucoidy. *Molecular Microbiology* **9**, 495-506.

Martin, D. W, Schurr, M. J., Yu, H. and Deretic, V. (1994). Analysis of promoters controlled by the putative sigma factor AlgU regulating conversion to

mucoidy in *Pseudomonas aeruginosa*: relationship to σ^E and stress response. *Journal of Bacteriology* **176**, 6688-6696.

Matthews, C. K. (1993). The cell-bag of enzymes or network of channels. *Journal of Bacteriology* **175**, 6377-6381.

May, T. B. and Chakrabarty, A. M. (1994) *Pseudomonas aeruginosa*: genes and enzymes of alginate synthesis. *Trends in Microbiology* **2**, 151-156.

May, T. B., Shinabarger, D., Boyd, A. and Chakrabarty, A. M. Identification of amino acid residues involved in the activity of phosphomannose isomerase-guanosine 5'-diphospho-D-mannose pyrophosphorylase. *Journal of Biological Chemistry* **269**, 4872-4877.

May, T. B., Shinabarger, D., Maharaj, R., Kato, J., Chu, L., DeVault, J. D., Roychoudhury, S., Zielinski, N. A., Berry, A., Rothmel, R. K., Misra, T. K. and Chakrabarty, A. M. (1991). Alginate synthesis by *Pseudomonas aeruginosa*: a key pathogenic factor in chronic pulmonary infections of cystic fibrosis patients. *Clinical Microbiology Reviews* **4**, 191-206.

Messing, J. (1983). New M13 vectors for cloning. *Methods in Enzymology* **101**, 20-78.

Miller, J.H. (1972). Experiments in molecular genetics. Cold Spring Harbor, New York; Cold Spring Harbor Laboratory Press.

Mimmack, M. L., Gallagher, M. P., Hyde, S. C., Pearce, S. R., Booth, I. R. and Higgins, C. F. (1989). Energy coupling to periplasmic binding-protein dependent transport systems. *Proceedings of the National Academy of Science USA* **86**, 8257-8261.

Minton, N. P., Atkinson, T., Brunton, C.J., Sherwood, R. F. (1984). The complete nucleotide sequence of the Pseudomonas gene coding for caboxypeptidase-G2. *Gene* **31**, 31-38.

Mohr, C. D., Hibler, N. S. and Deretic, V. (1991). AlgR a response regulator controlling mucoidy in *Pseudomonas aeruginosa*, binds to the FUS sites of the *algD* promoter located unusually Far Upstream from the mRNA start site. *Journal of Bacteriology* **173**, 5136-5143.

Mohr, C.D., Martin, D. W, Konyecsni, W. M., Govan, J. R. W., Lory, S. and Deretic, V. (1990). Role of the far-upstream sites of the *algD* promoter and the *algR* and *rpoN* genes in environmental modulation of mucoidy in *Pseudomonas aeruginosa*. *Journal of Bacteriology* **172**, 6576-6580.

Morett, E and Buck, M. (1989). *In vivo* studies on the interaction of RNA polymerase σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti nifH* promoters. *Journal of Molecular Biology* **210**, 65-77.

Moxon, E. R. and Kroll, J. S. (1990). The role of polysaccharide capsules as virulence factors. *Current trends in Microbiology and Immunology* **150**, 65-85.

Needleman, S. B. and Winsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48**, 443-453.

Nelson, J. W., Tredgett, M. W., Sheehan, J. K., Thornton, D. J., Notman, D. and Govan. J. R. W. (1990). Mucinophilic and chemotactic properties of *Pseudomonas aeruginosa* in relation to pulmonary colonisation in cystic fibrosis. *Infection and Immunity* **58**, 1489-1495.

Nikaido, H. and Vaara, M. (1985). Molecular basis of bacterial outer membrane permeability. *Microbiological Reviews* **49**, 1-32.

Pavelka, M. S., Wright, L. F. and. Silver, R. P. (1991). Identification of two genes, *kpsM* and *kpsT*, in region 3 of the polysialic acid gene cluster of *Escherichia coli* K1. *Journal of Bacteriology* **173**, 4603-4610.

Pazzani, C., Rosenow, C., Boulnois, G. J, Bronner, D, Jann, J. and Roberts, I. S. (1993). Molecular analysis of region 1 of the *Escherichia coli* K5 antigen gene cluster: a region encoding proteins involved in cell surface expression of capsular polysaccharide. *Journal of Bacteriology* **175**, 5978-5983.

Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Science USA* **85**, 2444-2448.

Perumal, N. B. and Minkley, E. G. jr. (1984). The product of the F sex factor *traT* surface exclusion gene is a lipoprotein. *Journal of Biological Chemistry* **259**, 5357-5360.

Pier, B. P., Grout, M. and Desjardins, D. (1991). Complement deposition by antibodies to *Pseudomonas aeruginosa* mucoid exopolysaccharide (MEP) and by non-MEP specific opsonins. *Journal of Immunology* **147**, 1869-1876.

Pindar, D. F. and Bucke, C. (1975). The biosynthesis of alginic acid by *Azotobacter vinelandii*. *Biochemical Journal* **152**, 617-622.

Pugashetti, B. K., Vadas, L., Priheir, H. S. and Feingold, D. S. (1983). GDP mannose dehydrogenase and biosynthesis of alginate like polysaccharide in a mucoid strain of *Pseudomonas aeruginosa*. *Journal of Bacteriology* **153**, 1107-1110.

Pugsley, A. P. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiological Reviews* **57**, 50-108.

Quandt, J. and Hynes, M. F. (1993). Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. *Gene* **127**, 15-21.

Ratcliff, R., Evans, M. J., Cuthbert, A. W., Macvinnish, L. J., Foster, D., Anderson, J. R. and Colledge, W. H. (1993). Production of a severe cystic fibrosis mutation on mice by gene targeting. *Nature Genetics* **4**, 35-41.

Regue, M. and Wu, H. C. (1988). Synthesis and export of lipoproteins in bacteria. In *Protein transfer and organelle biogenesis*, Das, R. C. and Robbins, P. W (ed.). pp 587-606. Academic Press Inc. San Diego.

Rehm, B. H. A, Boheim, G., Tommassen, J. and Winkler, U. K. (1994a). Overexpression of *algE* in *Escherichia coli*: subcellular localisation, purification and ion channel properties. *Journal of Bacteriology* **176**, 5639-5647.

Rehm, B. A., Grabert, E., Hein, J. and Winkler, U. K. (1994b). Antibody response of rabbits and cystic fibrosis patients to an alginate-specific outer membrane protein of a mucoid strain of *Pseudomonas aeruginosa*. *Microbial pathogenesis* **16**, 43-51.

Remaut, E., Tsao, H. Fliers, W. (1983). Improved plasmid vectors with thermoinducible expression and temperature-regulated runaway replication. *Gene* **22**, 103-113.

Reuhs, B. L., Carlson, R. W. and Kim, J. S. (1993) *Rhizobium fredii* and *Rhizobium meliloti* produce 3-deoxy-D-manno-2-octulosonic acid-containing polysaccharides that are structurally analogous to group II K antigens (capsular polysaccharides) found in *Escherichia coli*. *Journal of Bacteriology* **175**, 3570-3580.

Rhamphal, R. and Pier, G. B. (1985). Role of *Pseudomonas aeruginosa* mucoid exopolysaccharide in adherence to tracheal cells. *Infection and Immunity* **47**, 1-4.

Rogers, S. G. and Weiss, B. (1980). Exonuclease III of *Escherichia coli* K-12. An AP endonuclease. *Methods in Enzymology* **65**, 201-211.

Rosenow, C., Esumeh, F, Roberts, I. S and Jann, K. (1995) Characterisation and localisation of the KpsE protein of *Escherichia coli* K5, which is involved in polysaccharide export. *Journal of Bacteriology* **177**, 1137-1143.

Ross, P., Mayer, R. and Benziman, M. (1991). Cellulose biosynthesis and function in bacteria. *Microbiological Reviews* **55**, 35-38.

Roychoudhury, S., Chakrabarty, K., Ho, Y. K. and Chakrabarty, A. M. (1992). Characterisation of guanosine diphospho D-mannose dehydrogenase from *Pseudomonas aeruginosa*. Structural analysis by limited proteolysis. *Journal of Biological Chemistry* **267**, 990-996.

Roychoudhury, S., May, T. B., Gill, J. F., Singh, S. K., Fengold, P. S. and Chakrabarty, A. M. (1989). Purification and characterisation of guanosine diphospho D-mannose dehydrogenase-A key enzyme in the biosynthesis of alginate by *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* **264**, 9380-9385.

Russel, M., Kidel, S. and Kelley, M. R. (1986). An improved filamentous helper phage for sequencing single stranded plasmid DNA. *Gene* **45**, 333-338.

Russell, D. R. and Bennett, G. N. (1982). Cloning of small DNA fragments containing the *Escherichia coli* tryptophan operon promoter and operator. *Gene* **17**, 9-18.

Sá-Correia, I., Darzins, A., Wang, S., Berry, A. and Chakrabarty, A. M. (1987). Alginate biosynthetic enzymes in mucoid and non-mucoid *Pseudomonas aeruginosa*: overproduction of phosphomannose isomerase, phosphomannomutase, and GDP-mannose pyrophosphorylase by overexpression of the phosphomannose isomerase (*pmi*) gene. *Journal of Bacteriology* **169**, 3224-3231.

Saiman, L., Cacalano, G., Gruenert, D. and Prince, A. (1992). Comparison of adherence of *Pseudomonas aeruginosa* to respiratory epithelial cells from cystic fibrosis patients and healthy subjects. *Infection and Immunity* **60**, 2808-2814.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular cloning, a laboratory manual. Cold Spring Harbour Laboratory, New York; Cold Spring Harbour Press.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitions. *Proceedings of the National Academy of Science USA* **74**, 5463-5467.

Saxena, I. M., Brown, M. B., Fevre, M., Geremia, R. A. and Henrissat, B. (1995). Multidomain architecture of β-glycosyl transferases: implications for mechanisms of action. *Journal of Bacteriology* **177**, 1419-1424.

Saxena, I. M., Kudlicka, K., Okuda, K. and Brown, R. M. (1994). Characterisation of genes in the cellulose synthesising operon (*acs* operon) of *Acetobacter xylinum*: implications for cellulose crystallisation. *Journal of Bacteriology* **176**, 5735-5752.

Schiller, N. L., Monday, S. R., Boyd, C. M., Keen, N. T. and Ohman, D. E. (1993). Characterisation of the *Pseudomonas aeruginosa* alginate lyase gene (*algL*):

cloning, sequencing and expression in *Escherichia coli*. *Journal of Bacteriology* **175**, 4780-4789.

Shinabarger, D., Berry, A., May, T. B., Rothmel, R., Fialho, A. and Chakrabarty, A. M. (1991). Purification and characterisation of phosphomannose isomerase guanosine diphospho-D-mannose pyrophosphorylase. *Journal of Biological Chemistry* **266**, 2080-2088.

Shinabarger, D., May, T. B., Boyd, A., Ghosh, M. and Chakrabarty, A. M. (1993). Nucleotide sequence and expression of the *Pseudomonas aeruginosa algF* gene controlling acetylation of alginate. *Molecular Microbiology* **9**, 1027-1935.

Short, J. M., Fernandez, J. M., Sorge, J. A. and Huse, W. D. (1988). λ ZAP: a bacteriophage λ expression vector with *in-vivo* excision properties. *Nucleic Acids Research* **16**, 7583-7600.

Silhavy, T. J., Berman, M. and Enquist, L. W. (1984). DNA extraction from bacterial cells pp137-140. In *Experiments with Gene Fusions*. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press.

Skjåk-Bræck, G., Grasdalen, H. and Larsen, B. (1986). Monomer sequence and acetylation patterns in some bacterial alginates. *Carbohydrate Research* **154**, 239-250.

Smith, A. N, Boulnois, G. J. and Roberts, I. S. (1990). Molecular analysis of the *Escherichia coli* K5 *kps* locus: identification and characterisation of an inner-membrane capsular polysaccharide transport system. *Molecular Microbiology* **4**, 1863-1869.

Smith, T. F. and Waterman, M. S. (1981). Identification of common molecular subsequences. *Journal of Molecular Biology* **147**, 195-197.

Snoewaert, J. N., Brigman, K. K., Latour, A. M., Malouf, N. N., Boucher, R. C., Smithies, O and Koller, B. H. (1992). An animal model for cystic fibrosis made by gene targeting. *Science* **257**, 1083-1088.

Sokol, P. A., Luan, M-Z., Storey, D. G. and Thirukkumaran, P. Genetic rearrangement associated with *in vivo* mucoid conversion of *Pseudomonas aeruginosa* PAO is due to insertion elements. *Journal of Bacteriology* **176**, 553-562.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503-517.

Stanfield, S. W., Ielpi, L., O' Brochta, D, Helinski, D. R. and Ditta, G. S. (1988). The *ndvA* gene product of *Rhizobium meliloti* is required for β(1-2) glucan

production has homology to the ATP-binding export protein HlyB. *Journal of Bacteriology* **170**, 3523-3530.

Stanier, R. Y, Adelberg, E. A. and Ingraham, J. L. (1981). The enteric group and related organisms. In *General Microbiology* (fourth edition), Chapter 20, pp 612-629. The Macmillan Press Ltd, Hong Kong.

Studier, F. W. and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective, high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorf, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology* **185**, 60-89.

Sutherland, I. W. (1982). Biosynthesis of microbial exopolysaccharides. *Advances in Microbial Physiology* **23**, 79-150.

Sutherland, I. W. (1989). Bacterial exopolysaccharides-their nature and production. *Antibiotics and Chemotherapy* **42**, 50-55.

Sutherland, I. W. (1990). Biotechnology of microbial exopolysaccharides. *Cambridge Studies in Biotechnology* **9**. Cambridge University Press.

Tabor, S. (1990). Expression using the T7 RNA polymerase/promoter system. In *Current Protocols in Molecular Biology.* Ausobel, F. M., Brent, R., Kingston, R. E., Moore, D. M., Seidman, J. G., Smith, J. A. and Strahl, K. (ed.). Greene Publishing Associates, New York.

Tabor, S. and Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Science USA* **82**, 1074-1078.

Tatti, K. M., Carter, H. L., Moir, A. and Moran, C. P. jr. (1989). Sigma-H directed transcription of *citG* in *Bacillus subtilis*. *Journal of Bacteriology* **171**, 5228-5932.

Terry, J., Pina, S. E. and Mattingly, S. J. (1991). Environmental conditions which influence mucoid conversion in *Pseudomonas aeruginosa*. *Infection and Immunity* **59**, 471-477.

Totten, P. A., Lara, J. C. and Lory, S. (1990). The *rpoN* gene product of *Pseudomonas aeruginosa* is required for expression of diverse genes, including the flagellin gene. *Journal of Bacteriology* **172**, 389-396.

Towbin, H., Staehalin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Science USA* **76**, 4350-4354.

Troy, F. A., Frerman, F. E. and Heath E. C. (1971). The biosynthesis of capsular polysaccharide in *Aerobacter aerogenes*. *Journal of Biological Chemistry* **246**, 118-125.

Vasil, M. L., Pritchard, A. E. and Ostrof, R. M. (1990). Molecular biology of Exotoxin A and Phospholipase C of *Pseudomonas aeruginosa*. *Pseudomonas: Biotransformations, Pathogenesis and Evolving Biotechnology,* edited by Silver, S., Chakrabarty, A. M., Iglewski, B. & Kaplan, S. **Chapter 1**, 3-14.

Vazquez, M., Santana, O. and Quinto, C. (1993). The NodI and NodJ proteins from rhizobium and bradyrhizobium strains are similar to capsular polysaccharide secretion proteins from Gram-negative bacteria. *Molecular Microbiology* **8**, 369-377.

Vimr, E. R., Bergstrom, R., Steenbergen, S. M., Boulnois, G. and Roberts, I. (1992). Homology among *Escherichia coli* K1 and K92 polysialyl transferases. *Journal of Bacteriology* **174**, 5127-5131.

Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982). Distantly related sequences in the a and b subunits of ATP synthase, myosin, kinases and other ATP requiring enzymes and a common nucleotide binding fold. *EMBO* **1**, 945-951.

Wang, Q. and Kaguni, J. M. (1989). A novel sigma factor is involved in expression of the *rpoH* gene of *Escherichia coli*. *Journal of Bacteriology* **171**, 4248-4253.

Wang, R., Seror, S. J., Blight, M., Pratt, J. M., Broome-Smith, J. K. and Holland, I. B. (1991). Analysis of the membrane organisation of an *Escherichia coli* protein translocator, HlyB a member of a large family of prokaryote and eukaryote surface transport proteins. *Journal of Molecular Biology* **217**, 441-454.

Wang, S. K., Sá-Correia, I. Darzins, A. and Chakrabarty, A. M. (1987). Characterisation of the *Pseudomonas aeruginosa* alginate (*alg*) gene region II. *Journal of General Microbiology* **133**, 2303-2314.

Warner, J. O. (1992). Immunology of cystic fibrosis. *British Medical Bulletin* 48, 893-912.

Weiner, J. H., Shaw, G., Turner, R. J. and Treiber, C. A. (1992). The topology of anchor subunit of dimethyl-sulfoxide reductase of *Escherichia coli*. *Journal of Biological Chemistry* **268**, 3238-3244.

Whitfield, C. and Valvano, M. A. (1993). Biosynthesis and expression of cell-surface polysaccharides in Gram-negative bacteria. *Advances in Microbial Physiology* **35**, 135-246.

Wong, H. C., Fear, A. L., Calhoon, R. D., Eichinger, G. H., Mayer,R., Amikam, D., Benziman, M., Tal, R., Gelfound, D. H., Meade, J. H., Emerick, A. W., Bruner, R. and Ben-Bassiot, A. (1990). Genetic organisation of the cellulose synthase operon in *Acetobacter xylinum*. *Proceedings of the National Academy of Science USA* **87**, 8130-8134.

Wozniak, D. J. (1994). Integration host factor and sequences downstream of *Pseudomonas aeruginosa algD* transcription start site are required for expression. *Journal of Bacteriology* **176**, 5068-5076.

Wozniak, D. J. and Ohman, D. E. (1991). *Pseudomonas aeruginosa*, AlgB, a two-component response regulator of the NtrC family, is required for *algD* transcription. *Journal of Bacteriology* **173**, 1406-1413.

Wozniak, D. J. and Ohman, D. E. (1994). Transcriptional analysis of the *Pseudomonas aeruginosa* genes *algR*, *algB* and *algD* reveals a hierarchy of alginate gene expression which is modulated by *algT*. *Journal of Bacteriology* **176**, 6007-6014.

Wu, J., Tisa, L. S. and Rosen, B. P. (1992). Membrane topology of the ArsB protein, the membrane subunit of an anion-translocating ATPase. *Journal of Biological Chemistry* 267, 12570-12576.

Wunder, D. E., Aaronson, W., Hayes, S. F., Bliss, J. M. and Silver, R. P. (1994). Nucleotide sequence and mutational analysis of the gene encoding KpsD, a periplasmic protein involved in transport of polysialic acid in *Escherichia coli* K1. *Journal of Bacteriology* **176**, 4025-4033.

Yamaguchi, K., Yu, F. and Inouye, M. (1988). A single amino acid determinant of the membrane localisation of lipoproteins in *Escherichia coli*. *Cell* **53**, 423-432.

Yanisch-Peron, C., Vieira, J. and Messing. J. (1985). Improved M13 phage cloning vectors and host strains; nucleotide sequences of the M13 mp18 and pUC19 vectors. *Gene* **33**, 103-109.

Zielinski, N. A, Chakrabarty, A. M. and Berry, A. (1991). Characterisation and regulation of the *Pseudomonas aeruginosa algC* gene encoding phosphomannomutase. *Journal of Biological Chemistry* **266**, 9754-9761.

Zielinski, N. A., DeVault, J. D., Roychoudhury, S., May, T. B., Kimbara, K., Kato, J., Shinabarger, D., Kitano, K., Berry, A., Misra, T. K. and Chakrabarty, A. M. (1990). Molecular genetics of alginate biosynthesis in *Pseudomonas*

aeruginosa. Chapter 2 In *Pseudomonas: biotransformations, pathogenesis and evolving biotechnology*. Silver, S., Chakrabarty, A. M., Iglewski, B. and Kaplan, S (editors). American Society for Microbiology. Washington, D. C.

Zielinski, N. A., Maharaj, R., Roychoudhury, S., Danganan, C. E., Hendrickson, W. and Chakrabarty, A. M. (1992). Alginate synthesis in *Pseudomonas aeruginosa*: environmental regulation of the *algC* promoter. *Journal of Bacteriology* **174**, 7680-7688.