The Unidirectional Flagellum of *R.sphaeroides*: Cloning and Analysis of Genes encoding Regulatory, Structural and Motor Components

by

Ian Gordon Goodfellow, B.Sc.(Hons)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy.

November 1996

Acknowledgements

I would like to thank my supervisor, Liz Sockett, for her constant help and encouragement throughout the course of this project. All the members of the *Rhodobacter* research group, past and present, are also thanked for their friendship and help over the past three years.

I would also like to thank my friends in the Dept of Life Science, Nottingham for making my three years there very memorable, especially : Charles Pollitt, Deepan Shah, Shaun Morroll, Karen Searle, Matt Edge, Liz Sockett, Marion Wood, Glen Thorlby, Caroline Ramm and those whose names I can't spell.

A special thanks goes to Gary 'Gaz' Grimsey, Phil 'Giffer' Heath, Mark 'Big' Perrins, Richard 'Pilch' Hastings, Raj Nair, Steve Hyde and the members of the Nottingham University and Trent University Jitsu clubs for keeping me socially "active".

I would also like to thank my family for their support during the course of my academic career. Finally, special thanks goes to Kerry Utton for her love, friendship and support over the past 4 years.

This work was supported by a BBSRC studentship to R.E.S. for I.G.G.

Abbreviations

SI units were used with the standard abbreviations suggested by the Biochemical Society. All additional abbreviations and units are listed below:

ATP	:	Adenosine triphosphate
BCIP	:	5-Bromo-4-chloro-3-indolyl phosphate
bp	:	Base pairs
BSA	:	Bovine serum albumin
deaza	:	Deazapurine
DTT	:	Dithiothreitol
EtBr	:	Ethidium bromide
GST	:	Glutathione S-tranferase
h.	:	Hour(s)
IPTG	:	Isopropyl β-galactopyranoside
Kb	:	Kilobases
KDa	:	Kilodaltons
o/n	:	Overnight
OD	:	Optical density
min.	:	Minute(s)
NBT	:	Nitro Blue tetrazolium-5-bromo-4-chloro-3-indolyl phosphate
PBS	:	Phosphate buffered saline
psi.	:	Pounds per square inch
rpm	:	Revolutions per minute
SDS	:	Sodium dodecyl sulphate
SDW	:	Sterile distilled water

Figures

(cont...)

Tables

Contents

Chapter 2: Methods

ABSTRACT

The Unidirectional Flagellum of *R.sphaeroides*: Cloning and Analysis of Genes Encoding Regulatory, Structural and Motor Components

Ian Gordon Goodfellow, B.Sc.(Hons)

Submitted for the degree of Doctor of Philosophy

In this study several components responsible for the formation and function of the unidirectional flagellum of *R.sphaeroides* WS8 were identified via the characterisation of motility impaired Tn*phoA* mutants. The role of the alternative sigma factor sigma 54 in flagellar gene regulation was also examined.

Mutant M18 was defective in a *flil* homologue, characterisation of this mutant revealed that FliI is not essential for flagellar formation in *R.sphaeroides*. This differs from that reported in the literature for *S.typhimurium* and so highlights the importance of studying *R.sphaeroides* as a model for flagellar motility.

Analysis of another mutant Nm7 revealed that it was defective in FliF, a rotor component around which other flagellar components assemble. Overexpression of a FliF fusion protein allowed the production of anti FliF antiserum.

DNA sequencing upstream and downstream of the *fliF* gene, revealed several other genes encoding flagellar components and a potential flagellar gene regulator (Torf).

fliE, encoding a component of the basal body of unknown function, was identified upstream of *fliF*, an interposon mutant was created and was unable to be complemented by the wild type gene in *trans* suggesting a dominant effect. This is the first dominant mutation to be isolated in any *fliE*.

The gene encoding the motor component FliG was also identified downstream of *fliF* and its C-terminal motility domain was found to contain regions that are conserved between FliG proteins from unidirectional and bidirectional motors, these may play a role in motor rotation and not switching. An overexpressed poly histidine FliG fusion protein was found to form a complex with the FliF-GST fusion protein *in vitro*.

The *torf* gene encodes a protein with homology to sigma 54 enhancer binding proteins. The Torf protein lacks any obvious DNA binding motif and may represent a novel member of the sigma 54 enhancer binding protein family.

Chapter 1

Introduction

1.1 Bacterial Motility - an introduction

In the early seventeenth century Leeuwenhoek wrote;

' ...when we see little living animals (protozoa) and see their legs and must judge the same to be ten thousand times thinner than a hair of my beard, and when I see animals living that are more than a hundred times smaller and am unable to observe any legs at all, I still conclude from their structure and the movements of their bodies that they do have legs' (Leeuwnehoek, reported in Piire, 1964). Well over 300 hundred years later, we still don't fully understand how the 'legs' reported by Leeuwenhoek function, however, we have gained a greater understanding of how microorganisms respond to changes in their environment by movement. Two modes of motility are most common; gliding motility and swimming motility. Gliding bacteria (e.g. bacteria from the *Myxococcus* genus) move by a mechanism thought to involve motility organelles in the cell wall with a layer of extracellular slime (Burchard, 1981; Lapidus & Berg, 1982; Wolkin & Pate, 1984; Wolkin & Pate, 1986; Godchaux *et al.*, 1990). Bacterial flagellar motility is the most widely studied system and over the past 300 years many insights into this area of motility have been gained.

The bacterial flagellum is a rotary organelle powered by a motor at its base (Berg, 1974) using a transmembrane gradient of protons (Larsen *et al.*, 1974; Manson *et al.*, 1977) or sodium ions (Hirota & Imae, 1983) as the energy source for torque generation. The torque generated at the base of the flagellum leads to rotation of a semi-rigid helical filament (Macnab & DeRosier, 1988) by a mechanism which is not understood (Caplan & Kara-Ivanov, 1993) but which leads to the propulsion of the bacterial cell.

1.1.1 Why be motile?

Due to the large number of proteins involved, the formation and functioning of the bacterial flagellum is a costly process in terms of energy required by the cell (Macnab, 1992), therefore the bacterial cell must gain some advantage by possessing a flagellum/flagella. In low nutrient environments, bacteria that can sense and move towards high concentrations of nutrients have an obvious advantage over those that cannot and bacteria use flagella to respond to environmental stimuli such as light (Harayama & Iino, 1976; Armitage *et al.*, 1985), oxygen (Taylor, 1983), temperature (Imae, 1985) etc. *Rhodobacter sphaeroides*, the organism used in this study, is a freshwater photosynthetic bacterium and is known to respond to light in the form of phototaxis as well as numerous other environmental stimuli (Armitage *et al.*, 1995). In its freshwater environment, *R.sphaeroides* will undoubtedly encounter many periods of nutrient deprivation. Its ability to respond to stimuli in the form of taxis via its single sub-polar flagellum (Armitage & Macnab, 1987) allows it to thrive in such environments. Flagella are also important in determining virulence of pathogens, and non-motile strains of various pathogenic organisms are avirulent (Pierce *et al.*, 1988).

1.2 The bacterial flagellum - structure

The bacterial flagellum, as seen under the light microscope consists of a filament which acts like a ship's propeller (Fig 1.1). The number of filaments present on the cell, the position of each and the method by which they are rotated, differs markedly throughout bacterial species. Of the bacteria studied as models for flagellar formation/function/rotation namely *Escherichia coli, Salmonella typhimurium, Caulobacter crescentus, Bacillus subtilis, Rhodobacter sphaeroides*, only two possess one flagellum per cell (*R.sphaeroides* and *C.crescentus*). The others possess approximately 5 flagella per cell. The photosynthetic bacterium *R.sphaeroides* is the only organism with a flagellum that rotates in only one direction (Armitage & Macnab, 1987),

1987), whereas the other bacterial flagella have the ability to switch the direction of rotation. It is the modulation of the switching of flagellum rotation, or stopping of the flagellum in the case of *R.sphaeroides,* which governs the response to environmental stimuli i.e. the movement of the cell towards stimuli (positive taxis) or the movement away from stimuli (negative taxis). It is the characteristics of the unidirectional motor which this project has concentrated on.

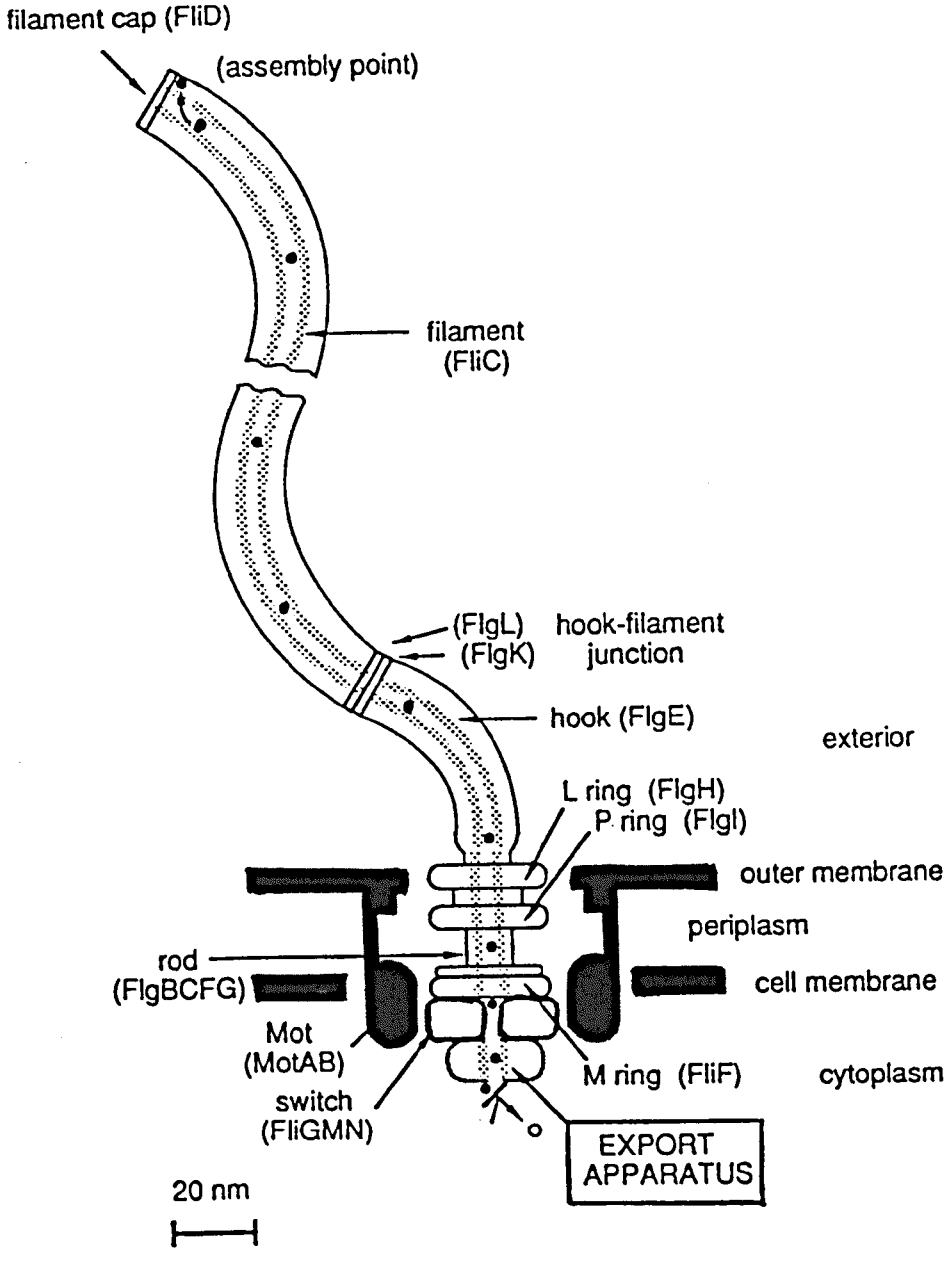
Gram negative bacterial flagella have a similar structure (Fig 1.1) and can be divided into four main regions, the filament, the hook, the basal body complex and the export apparatus. The basal body complex, which consists of a series of rings in the membrane connected by a rod structure, is responsible for the rotation of the flagellum and many models have been proposed as to how this occurs, some of which will be discussed later. During this project, components of the export apparatus and basal body complex were examined, as well as the regulation of flagellar synthesis. These features will all be discussed below to give the reader the background required for an understanding of this project. Before addressing any of these features, the fundamental mechanism of flagellum rotation will be discussed so as to give the reader an idea of what causes the flagellum to rotate and what follows is a brief introduction to the many theoretical models that have been proposed

1.3 Flagellar rotation - hypothetical models

Many models have been proposed for flagellum rotation;

The elucidation of the method by which the transmembrane proton gradient is converted into torque generation and rotation of the flagellum is the ultimate goal of most of the studies in flagellar function and the aim of this project was to shed some light on this by analysing motor components from a unidirectional motor. Due to the intriguing nature of this miniature 'engine', many theoretical models for rotation have been proposed. The large number of

Fig 1.1 The structure of the Gram negative bacterial flagellum. After Vogler *et. al.* (1991).



models and their high degree of complexity precludes me from giving a highly complex discussion of all the models, instead the reader is referred to the recent review by Caplan and Kara-Ivanov which gives an excellent coverage of all models predicted to date (Caplan & Kara-Ivanov, 1993). What follows is a summary of many of the models, with specific reference to those that in the light of recent molecular evidence may be correct.

Before considering the models of flagellum rotation, it is important to understand what features of the motor have been demonstrated by experimental evidence.

The motor is made up of complexes of MotA and MotB;

The motor is made up of force generators situated around the periphery of the MS-ring complex, thought to consist of 10-16 complexes of MotA/B (Khan *et al.*, 1988) (Fig 1.1). This is in close agreement with the finding that there is eight independent force generating units and that each unit is capable of rotating the rotor in either ccw or cw directions at approximately equal torque (Blair & Berg, 1988). There is evidence to suggest that MotA is the proton conducting component of the motor with MotB functioning simply to join MotA to the cell wall (Blair & Berg, 1988; Blair & Berg, 1990). This evidence has been debated (see below) and it may in fact be that the proton channel is made up of a complex of MotA and MotB (Sharp *et al.*, 1995).

The motor has several modes of rotation;

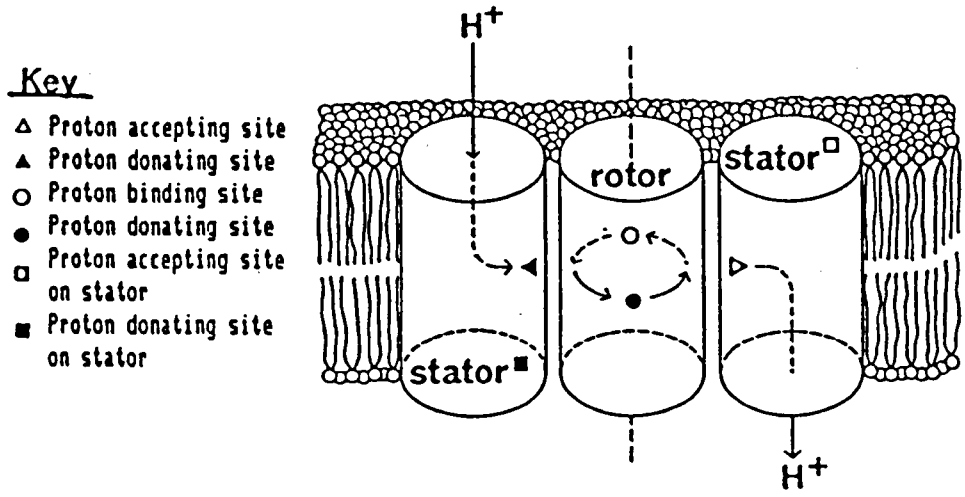
The biophysical characteristics of the motor from *E.coli* and *S.typhimurium* have been the centre of many studies (Kami-Ike *et al.*, 1991; Iwazawa *et al.*, 1993; Berry *et al.*, 1995; Fung & Berg, 1995), with the result being that much is known about the speed, torque etc. of the motor. These studies revealed that the bidirectional motors of *E.coli* and *S.typhimurium* have three modes of rotation; ccw rotation, which is the default 'natural' state of the motor, pausing where the motor does not rotate and cw rotation (Eisenbach, 1990). The key for pausing and cw rotation appears to be the phosphorylated

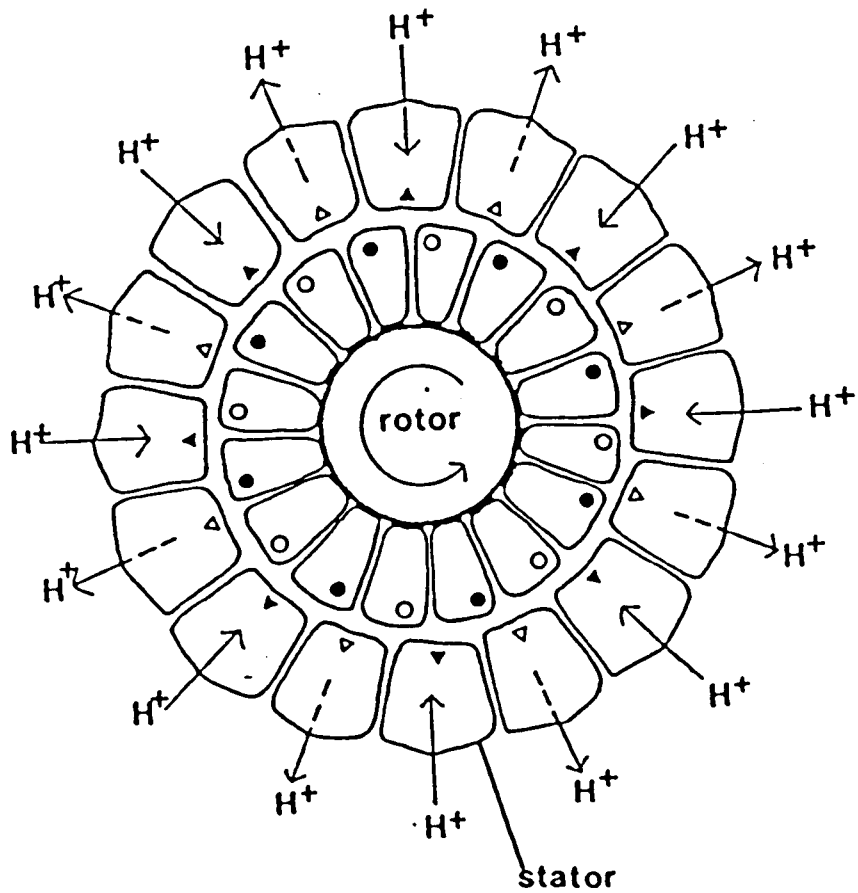
CheY protein. In the unidirectional motor of *R.sphaeroides* there are only two states; cw rotation and pausing (Armitage & Macnab, 1987). The pausing state of the motor may well be similar in both systems but the duration is significantly different i.e. seconds in *R.sphaeroides* (Poole *et al.*, 1988) and fractions of a second in *E.coli* and *S.typhimurium* (Eisenbach, 1990). The torque generated in the motors from *E.coli* and *Streptococcus* is proportional to the proton motive force (Manson *et al.*, 1980; Conley & Berg, 1984; Fung & Berg, 1995) with the proton flux coupled to flagellum rotation equating to between 200 - 1000 protons per revolution (Meister *et al.*, 1989).

Many models are based on a similar theme;

As stated above, many models for flagellar rotation have been proposed (e.g. Oosawa & Hayashi, 1986; Wagenknecht, 1986; Lauger, 1988; Fuhr & Hagedorn, 1989; Murata *et al.*, 1989). The majority of models rely on a mechanism of protons passing through the proton channel into the membrane whereby they bind to a proton binding site on the rotor (usually presumed to be the M-ring of the MS-ring complex), whereupon the rotor rotates to allow the proton to pass down the remainder of the channel into the cytoplasm (see Caplan and Kara-Ivanov for a more detailed discussion). The simplest model for this has been presented by Mitchell (Mitchell, 1984) (see Fig 1.2) whereby the ion enters a 'half channel' and charges a site within the membrane on a stator protein, possibly within the MotA/B complex. Then, by after transfer of the proton to the rotor site (possibly on the MS-ring) and mechanism of electrostatic repulsion/attraction, the rotor rotates so as to allow the passage of the proton to another 'half channel' on a different stator protein and subsequently to the cytoplasm. The spatial rearrangement of either the stator binding site or rotor binding sites would result in a bias towards either cw or ccw rotation. This would presumably be the role of phospho-CheY binding (the switching signal molecule). The rotor binding sites in Mitchell's model were presumed to be on the M-ring, however it is possible that they are actually

Fig 1.2 Mitchell's "well and turnstile" model for flagellar rotation. Taken from Mitchell, 1984. Top diagram represents a simplistic view of the model. Bottom is a representation of the arrangement of proton accepting and donating sites around the rotor.



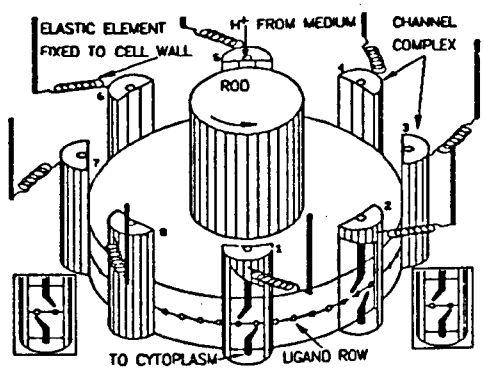


within FliG part of the switch complex (see below) and the second stator binding site is not required. A more detailed discussion of this will be presented in chapter 6.

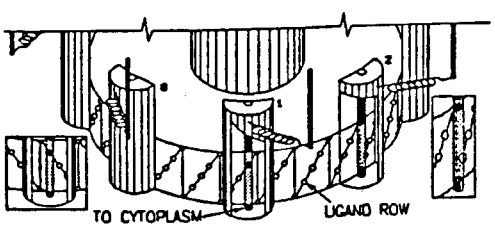
Several models are based on more direct molecular evidence

Berg and Khan, and Lauger have proposed models referred to as gated channel models (Khan & Berg, 1983; Lauger, 1988) whereby the role of MotB is to act as an elastic linkage of the proton channel (MotA) to the cell wall. These force generators are distributed around the periphery of the rotor (M-ring or the C-ring made up of the switch protein complex) (see Fig 1.3c) and that torque is generated when the channel complex (force generating unit) moves around the rotor (possibly the MS-ring or the switch complex - Fig 1.1), binds to a specific site and then moves in the direction opposite to what was required initially to bind the site, pulling the rotor with it. The two models differ mainly on the array of proton binding sites on the periphery of the rotor (see Figs 1.3 a & 1.3b). In the Berg and Khan model (Fig 1.3a) a single continuous row of proton binding sites runs around the periphery of the rotor and proton flux is allowed when one site in the channel is charged, the channel moves and allows the proton to pass through to another site on the same channel and subsequently into the cytoplasm. In contrast, in the Lauger model (Fig 1.3b) the proton binding sites on the rotor are arranged in a helical pattern and passage of the proton through the channel is only allowed by the interaction of the channel with a single proton binding site on this helix. It is then proposed that the proton alternates between the channel and the rotor via a series of proton binding sites allowing the passage of the proton to the cytoplasm. The mechanism of action of the switch proteins in both these models also differs; in the Berg and Khan model the switch acts to alter the arrangement of proton binding sites in the channel, whereas in the Lauger model, the switch acts to alter the arrangement of the proton binding sites on the rotor.

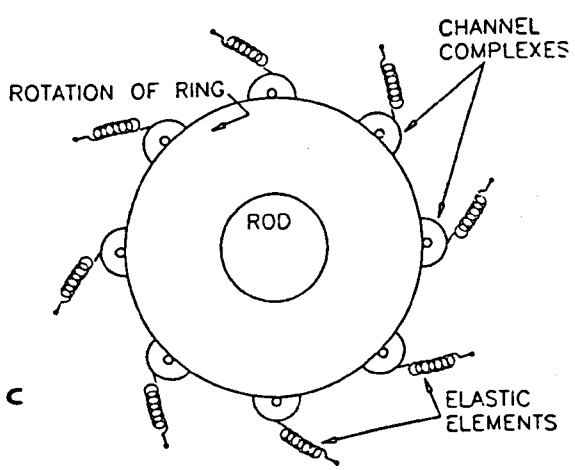
Fig 1.3 Berg-Khan and Lauger models of flagellar rotation. Taken from Caplan and Kara-Ivanov, 1993. Open circles represent empty proton-binding sites with filled circles representing bound protons. (A) Berg-Khan model. The right-hand inset shows the configuration after channel complex 1 has made a single incremental step in the cw direction allowing the proton access to the cytoplasm. The left-handed inset shows the configuration of the channel upon switching. (B) Lauger model. The right-hand inset shows the situation whereby a proton binds to sites at the top of the channel before the proton has left the bottom of the channel. The left-hand inset shows the modified array of proton binding sites upon switching. (C) simplistic view from above both models showing the elastic elements.



A BERG-KHAN



B LÄUGER



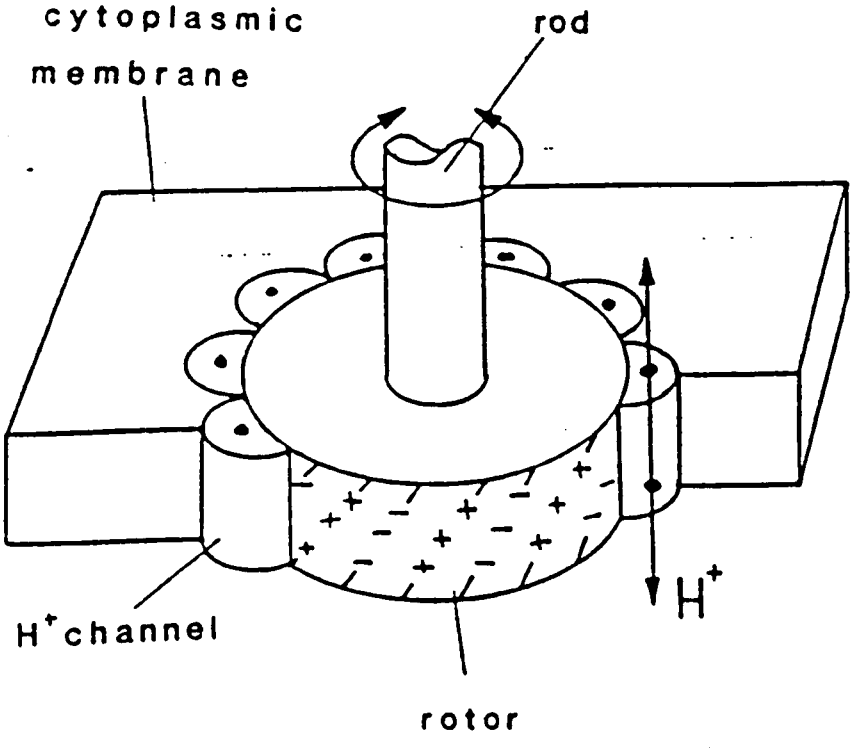
C

A model similar to that proposed by Lauger has been proposed by Berry (Berry, 1993) whereby a helical row of sites on the rotor interact with the proton binding sites in the channel to allow proton flux (see Fig 1.4). In Berry's model, the rotor has alternating lines of positive and negative charges close to its perimeter and torque is generated by the attraction of the negative sites for the positively charged proton bound to sites in the channel. Again, due to the helical arrangement of the charges on the rotor, the attraction results in rotation of the rotor. The switching in this model could occur either via a conformational change in the rotor resulting in the reversal of the tilt of the charged residues or by the action of phospho-CheY (the 'switching' signal) on the pK of the channel proton binding sites i.e. to increase the occupancy probability of the cytoplasmic proton binding sites and a reversal of rotation. The author correctly notes that the rotor component could in fact be FliG as it has been shown to has a large degree of clustered charges (see later for a greater discussion) (Irikura *et al.*, 1993).

Many other models of rotation have been proposed but have been deemed implausible, although intriguing, and therefore consequently not of relevance to this project (see Caplan and Kara-Ivanov for a discussion of these models).

Reviewing the literature, it is clear that a unidirectional motor, as is present in *R.sphaeroides* (Armitage & Macnab, 1987), would have many of the features of the bidirectional motor i.e. the method of torque generation would be predicted to be the same, but it would lack the mechanism of switching. The mechanism of this unidirectionality, may be in the form of a 'brake' or a 'clutch' that functions to uncouple rotation of the rotor from proton transport. Insights into the mechanism of flagellum rotation may be gained by comparing flagellar proteins from a unidirectional flagellum to those from bidirectional flagella which was the main aim of this project. In order to address this further

Fig 1.4 Berry's model for flagellum rotation. Taken from Berry, 1993. The rotor is shown with its tilted alternating lines of positive and negative charged residues thought to be important during rotation.



it is necessary for the reader to understand in more detail the structure and function of the flagellum.

1.4 Structure and function of the flagellum

1.4.1 The filament

The filament, as studied in *E.coli* and *S.typhimurium* is around 10-20μm long and 20nm in diameter (Macnab & DeRosier, 1988) (Fig 1.1), compared to 10-15μm and 16nm for *R.sphaeroides*, the bacterium used in this study (Sockett, 1986). Usually the filament consists of thousands of subunits of a single protein known as flagellin (Macnab, 1992).

In *R.sphaeroides* the filament coils up to form a large coil when it stops rotating, this is assumed to allow the cell to re-orientate via Brownian motion until the flagellum again starts rotating (Armitage & Macnab, 1987). This contrasts to what is seen in *E.coli* and *S.typhimurium*, whereby the filament adopts a right-handed helical conformation upon switching of the direction of flagellum rotation (ccw to cw) (Macnab & Ornston, 1977), rotation of which leads to re-orientation of the cell. As will be discussed later, the flagellin protein is thought to be exported through the centre of the filament (Namba *et al.*, 1989) and assembled onto the distal end of the flagellum (Iino, 1969; Emerson *et al.*, 1970).

1.4.2 The Hook

The hook, as the name suggests, is bent in most species, this is to allow the formation of a flagellar bundle in bacteria which are peritrichously flagellated (De Pamphilis & Adler, 1971). Consequently, uniflagellate organisms such as *R.sphaeroides* and *C.crescentus* possess hook structures that are less bent (Shapiro & Maizel, 1973; Sockett, 1986). The *S.typhimurium* hook, is composed of a single protein, FlgE (Kagawa *et al.*, 1976), and is thought to act as a universal joint between the rod and the filament (Macnab, 1992).

1.4.3 The basal body complex

The basal body is a complex of many proteins which interact to form this unique mechanoenzyme that causes the rotation of the filament. It is components of this part of the flagellum that this project has concentrated on as differences between unidirectional and bidirectional flagella may give some insights into the method of flagellum rotation. The complex itself consists of a series of rings (Fig 1.1) connected to each other via a rod (De Pamphilis & Adler, 1971). The basal body as isolated from *E.coli* and *S.typhimurium* lacks parts of the flagellum, namely the motor proteins and also the export apparatus (Khan *et al.*, 1992; Francis *et al.*, 1994). The existence of these additional structures has been demonstrated by genetic studies (see reviews Macnab, 1992; Blair, 1995) and they will be described along with the other structures shown to be present in the basal body.

1.4.3a The L,P and E rings

In Gram negative bacteria the L and P rings (Fig 1.1) interact with the outer membrane and peptidoglycan layers of the cell respectively and are thought to act as washers, allowing the free rotation of the flagellum in the outer membrane (Berg, 1974). Gram positive bacteria lack such structures as they have no outer membrane (De Pamphilis & Adler, 1971). The presence of the L and P rings in *R.sphaeroides* have been demonstrated by electron microscopy of isolated basal bodies (Sockett, 1986). There also appears to be an additional ring structure associated with the L and P rings, the E ring (Sockett, 1986). This ring structure is also present in the uniflagellate *C.crescentus* (Stallmeyer *et al.*, 1989).

1.4.3b The rod

The rod structure, known to join the L,P and E rings with the components in the inner membrane (Stallmeyer *et al.*, 1988), is thought to transmit the torque generated at the cytoplasmic membrane to the external components of the flagellum (Macnab & DeRosier, 1988). Five gene products constitute the rod, namely *flgB, flgC, flgF, flgG and fliE* (Homma *et al.*, 1990; Muller *et al.*, 1992), with FliE as the adapter between the MS-ring complex and the rest of the rod (Muller *et al.*, 1992). The *R.sphaeroides* rod structure looks similar to that from *E.coli* and *S.typhimurium* and is assumed to function in a similar way (Sockett, Pers. comm.).

1.4.3c The MS-ring complex

Given the close proximity of the MS-ring complex to the motor proteins, many models have implicated it in playing a major role in torque generation. Part of this project was to identify the gene encoding the protein that forms the MS-ring and compare it to that from a bi-directional motor with the aim to gaining insights into flagellum rotation. What follows is a brief description of what is currently known about the MS-ring protein FliF.

The FliF protein forms the MS-ring;

The gene encoding the protein that forms the MS ring complex (Fig 1.1) (*fliF*) has been identified in many organisms (Homma *et al.*, 1987; Matsumurra *et al.*, 1995; Zuberi *et al.*, 1991; Ramakrishnan, *et al.*, 1994; Arora *et al.*, 1996; Heinzerling *et al.*, 1995). The MS-ring complex is thought to consist of 20 to 26 subunits of FliF (Jones *et al.*, 1990; Sosinsky *et al.*, 1992; Ueno *et al.*, 1994). The M ring is integral to the cytoplasmic membrane with the S ring immediately distal and connected to it (Stallmeyer *et al.*, 1989). All the FliF proteins are predicted to have two membrane spanning regions (Homma *et al.*, 1987; Matsumurra *et al.*, 1995; Zuberi *et al.*, 1991; Ramakrishnan, *et al.*, 1994; Arora *et al.*, 1996) and Ueno and co-workers showed that the N-terminal membrane spanning region forms the central core of

the complex with the C terminus constituting the M ring region (Ueno *et al.*, 1994) (Fig 1.5).The MS-ring complex has been identified in *R.sphaeroides* basal body preparations, appears to have similar dimensions to that from other bacteria and is presumed to function in a similar way (Sockett, 1986).

The centre of the MS-ring complex is closed;

The MS ring complex was thought to be a complex of 2 gene products, although Homma and co-workers proposed that the S ring was a domain structure of FliF (Homma *et al.*, 1987). This was confirmed by Ueno and co-workers in 1992 by overexpressing the FliF protein in *E.coli* (Ueno *et al.*, 1992). They found that the over-expressed protein formed ring structures in the cytoplasmic membrane of the host cell, identical to those found in the hook-basal-body preparations. The finding that *E.coli* cells were alive when over 50% of their surface was occupied by the FliF complexes indicated that the central core of the complex is physiologically closed (Ueno *et al.*, 1992), as the cytoplasm would 'leak out' if it was open. Vogler and co-workers have proposed that the central core of the basal body is open in order for the export of flagellar components to occur and that proteins on the cytoplasmic side of the basal body are responsible for this export (Vogler *et al.*, 1991). It is obvious that the cytoplasmic flagellar specific export apparatus must be responsible for the opening of this central core. Recently, Macnab has suggested that the export apparatus may be inserted into a small portion of lipid bilayer in the centre of the MS ring complex which would be 'trapped' when the FliF monomers meet and complex in the cytoplasmic membrane (Macnab pers. comm.) although direct evidence of this is yet to be provided. This lipid bi-layer would presumably also be responsible for the 'plugging' of the central core of the FliF complexes seen by Ueno and co-workers when FliF was overexpressed in *E.coli* (Ueno *et al.*, 1992), and would explain the ability of the host cells to tolerate the 50% coverage of their surface by these complexes.

FliF may play a role in export;

Fig 1.5 Domain organisation of *S.typhimurium* FliF.

Taken from Ueno *et. al.* 1994.

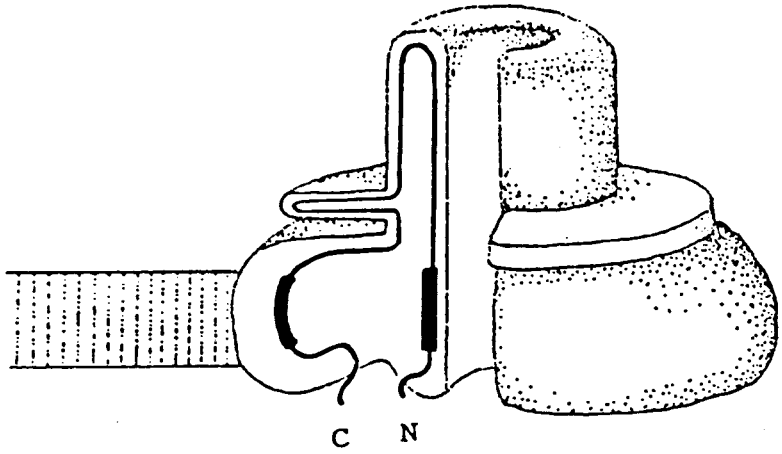
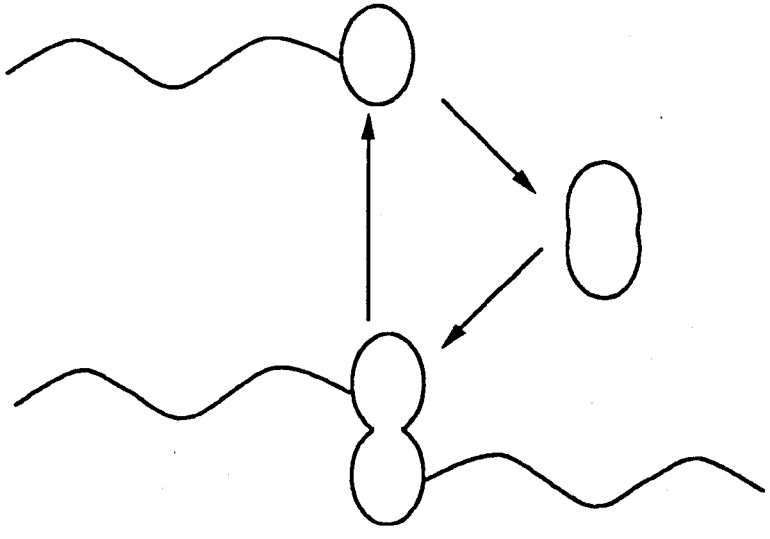


Fig 1.6 Flagellar targeting in *R.sphaeroides*



Recent evidence may further strengthen the idea that the centre of the MS-ring contains the export apparatus (Katayama *et al.*, 1996), although more direct proof is still required. The findings that part of FliF is homologous to various virulence-factor export components (Van Gijsegem *et al.*, 1995; Harshey & Toguchi, 1996) and that the homologous region is in fact the central core of the MS-ring complex (Ueno *et al.*, 1992) is consistent with the idea of Vogler and co-workers (Vogler *et al.*, 1991) that flagellar components are exported through the central core of the MS-ring complex by a flagellar specific export apparatus. The virulence-factor export components, all of the YscJ family (Michiels *et al.*, 1991), are thought to be outer-membrane proteins and include MxiJ (Abdelmounaaim *et al.*, 1992), NolT (Meinhardt *et al.*, 1993) and HrpB3 (Fenselau *et al.*, 1992). The role of these export factors in the export of virulence determinants is not completely understood, but it is thought that HrpB3 forms part of a pore complex in the outer membrane to allow export (Fenselau *et al.*, 1992), which may explain the homology with FliF. A more extensive analysis of the process of flagellum specific export will be covered in section 1.4.3k

The MS-ring was thought to play a role in torque generation;

As the M and S rings are proposed to be in close proximity to the motor proteins, they were originally central to many models of flagellum rotation e.g. Lauger, 1974; Glagalin, 1978; Khan & Berg, 1983, with the S ring being the stator and the M ring the presumed rotor (Berg, 1974). Ueno and co-workers later ruled out this possibility by showing that the S ring would rotate with the M ring (Ueno *et al.*, 1992). This has therefore led to the question:

Does the MS ring play an active or passive role in the flagellum rotation?

The finding that a *fliF* null mutant has no detectable flagella and that no motility mutants with paralysed or switch defects map to *fliF* (Yamaguchi *et al.*,

1986) led Homma and co-workers (Homma *et al.*, 1987) to conclude that either

1) The FliF protein is involved in energy transduction and switching, but the residues that are critical for these functions are also crucial for assembly, or
2) The FliF protein plays a structural role and is not involved in energy transduction or switching, merely serving as a passive structure onto which the switch complex is mounted, allowing its interaction with the Mot proteins to generate torque. In the light of more recent evidence for the role of the switch proteins (see 1.4.3d-i) the latter appears to be the more accepted view.

FliF may play a role in flagellum targeting;

In uniflagellate organisms, and to some extent in multiflagellated organisms, the process of flagellar targeting is an intriguing question i.e. **What determines exactly where on the cell's surface the flagellum is assembled?** The finding that the MS-ring complex is the first detectable substructure that is assembled during the process of flagellar biogenesis has led many workers to conclude that it is the target for the remaining flagellar substructures (Suzuki *et al.*, 1978; Jones & Macnab, 1990; Kubori *et al.*, 1992). Whether or not FliF monomers are targeted to a specific region of the inner membrane where they form the MS-ring complex, or if they simply insert randomly and fortuitously meet and form the MS-ring complex, is yet to be completely answered. The latter idea may well be the case for the peritrichously flagellated organisms as the position of the flagellum on the cells surface is random, but in the uniflagellate organisms e.g.. *R.sphaeroides* and *C.crescentus*, there appears to be a specific flagellum targeting process (Shapiro & Maizel, 1973; Foster, 1991). In *C.crescentus*, a bacterium with an asymmetric cell cycle where two morphologically and physiologically distinct progeny are produced (a motile swarmer cell and a sessile stalked cell), the flagellar assembly process is tightly coupled to the cell cycle resulting in the construction of a single flagellum at the pole opposite the stalk in the

predivisional cell.(Shapiro & Maizel, 1973). In the photosynthetic organism *R.sphaeroides*, there also appears to be a cell cycle specific effect on flagellar formation and targeting (Foster, 1991). It has been shown that during a shift from high light growth to low light growth *R.sphaeroides* sheds its flagellum. A decrease in motility was also observed during cell division (Foster, 1991). The finding that the newly synthesised flagellum of a daughter cell is positioned on the opposite wall to that of the mother cell, always on the longer side of the cell (Armitage & Macnab, 1987) (see Fig 1.6) and that the flagellum appears to be inserted into membrane devoid of the photosynthetic apparatus (Foster, 1991) suggests that a cell cycle mediated flagellum targeting/assembly process does control flagellar biosynthesis in *R.sphaeroides*. These cell cycle processes raise several intriguing questions; **How are the flagellar components regulated during the cell cycle? What determines where the flagellum is inserted and how are the motor components directed to that site? What causes the release of the flagellum and how does the cell cycle control this process? How is flagellar assembly controlled with respect to photosynthetic membrane differentiation?** The elegant work of Jenal and Shapiro in *C.crescentus* has shed some light on how these processes may occur (Jenal & Shapiro, 1996). They have shown that FliF is specifically targeted to the swarmer cell pole and that it is sequence in the C-terminus of FliF that is responsible not only for this targeting but also the cell cycle controlled proteolysis of FliF (Jenal & Shapiro, 1996). They have also shown that FliM and FliG, components of the flagellar motor switch complex, are also degraded at the same time as FliF and conclude that it is probably the proteolysis of these basal components of the flagellum that result in the cell cycle specific release of the flagellum in *C.crescentus* (Jenal *et al.*, 1994; Jenal & Shapiro, 1996). The comparison of the FliF, FliG and FliM proteins from *R.sphaeroides* and *C.crescentus* may show similar cell cycle proteolysis/targeting domains in these proteins and possibly give some clues as

to how the process of flagellum targeting and shedding occurs in *R.sphaeroides*. This shall be discussed in more detail in chapter 5.

Okino found that wild type *S.typhimurium* released hook-filament structures when allowed to grow for extended periods (Okino *et al.*, 1989). These findings may indicate that a similar cell cycle controlled release of flagella occurs in peritrichously flagellated organisms as well as the uniflagellates, however, no further reports has substantiated these findings.

FliF interacts with the switch complex;

One interesting set of mutations have been isolated in *fliF* (Francis *et al.*, 1992), these two spontaneous mutations resulted in the fusion of the MS-ring protein FliF to the switch protein FliG. This was the first direct evidence for the association of a switch protein with the basal body, and it further demonstrated that fusing FliF to FliG still permitted flagellum assembly and rotation. The two mutants, one which resulted in the 4 C-terminal residues of FliF being lost during the fusion to full length FliG, and the other that resulted in the removal of the C-terminal 56 residues of FliF and the 94 N-terminal residues of FliG being lost, both had rotational defects i.e. biased cw rotation and ccw rotation respectively. These data suggest that the interaction between FliF and FliG must be modifiable in order for correct switching to occur and that the fusion of FliF to FliG prevents this modification. The unidirectional flagellum of *R.sphaeroides* does not switch direction, this suggests that in this bacterium these proteins may be fused, as the modification that causes stopping in *R.sphaeroides* may not be as substantial as that which causes switching in other bacteria. As I will show in chapter 5, this appears not to be the situation as the genes are not fused.

Many workers have also demonstrated that FliF can interact with other components of the flagellar switch complex (Oosawa *et al.*, 1994; Jenal & Shapiro, 1996; Marykwas *et al.*, 1996) namely FliG and FliM. Oosawa and co-workers demonstrated an interaction using purified proteins and they also

showed that the C-terminal 100 residues of FliF were responsible for the interaction with FliG (Oosawa *et al.*, 1994). This was further verified by Jenal and co-workers on the uniflagellate *C.crescentus* (Jenal & Shapiro, 1996). The yeast-two hybrid system (Fields & Song, 1989; Chien *et al.*, 1991) has proved to be of great use in the study of flagellar protein interactions (Marykwas & Berg, 1996; Marykwas *et al.*, 1996) with one study again demonstrating an interaction between FliF, FliG and FliM (Marykwas *et al.*, 1996). They also showed that the N-terminal 61 residues of FliF are not required for an interaction between FliF and FliG/FliM. The analysis of the interaction between these switch and rotor proteins from *R.sphaeroides* may reveal interesting features that determine the unidirectionality of the flagellar motor.

1.4.3d The Switch complex

One of the major aims of this project was to determine the nature of the switch proteins in a unidirectional motor, comparing them to those from bidirectional motors and what follows is a description of what is known to date on the nature of the switch from other organisms. Most of the knowledge of switch proteins comes from the analysis of *E.coli* and *S.typhimurium*, both of which possess bidirectionally rotating flagella. The role of such a switch complex in the bacteria is, as the name suggests, switching of the direction of flagellum rotation from ccw to cw and it is the interaction of this switch complex with the chemosensory apparatus of the cell that leads to chemotaxis (covered in more detail below) and in *R.sphaeroides* phototaxis i.e. a response to light. One might hypothesise that such a switch complex might not exist in a unidirectional flagellum, or in fact play a different role i.e. stopping of the flagellum. As will become apparent from chapter 5, the latter is in fact the case.

The switch complex is a complex of three proteins (FliG, FliM and FliN) at the base of the flagellum;

The first evidence for the existence of a complex responsible for the switching of direction of flagellar rotation came from genetic studies in *S.typhimurium* (Yamaguchi *et al.*, 1986). Using mutants carrying mutations in the switch proteins FliG, FliM and FliN, they isolated suppressor mutants in several genes known to be involved in chemotaxis and also within themselves. They concluded that the proteins formed a complex which determines the direction of rotation and also participated in the conversion of proton energy into rotational energy. The presence of this complex at the base of the flagellum was subsequently confirmed using electron microscopy and antibodies directed against the proteins (Driks & DeRosier, 1991; Francis *et al.*, 1992; Khan *et al.*, 1992). The binding of FliG and FliM with the MS-ring complex has also been seen (Oosawa *et al.*, 1994) and the two-hybrid system has confirmed the interaction of FliG with FliF, FliG with FliM, FliM with itself and FliM with FliN (Marykwas & Berg, 1996; Marykwas *et al.*, 1996). Tang and co-workers have recently shown that FliG and FliM can bind to MotA (Tang *et al.*, 1996) as well as themselves. The proteins shall be covered separately with specific reference to the factors that govern switching/rotation.

1.4.3e FliG

As will be discussed in chapter 5, during the course of this project I identified the *fliG* gene from *R.sphaeroides* and what follows is a brief description of what is known on the role of the protein with particular emphasis on the features that govern switching/rotation.

FliG plays a role in switching and torque generation;

Extensive mutational analysis of FliG has led to the idea that it is involved both in flagellum assembly and rotation (Irikura *et al.*, 1993) and as stated above, there is evidence to suggest that FliG is the cytoplasmic rotor component of the torque generation apparatus (Garza *et al.*, 1995; Garza *et al.*, 1996). The actual interaction point of FliG with the motor proteins is suggested

to be a region within the C-terminus of FliG where the majority of rotation affecting mutations have been isolated (Irikura *et al.*, 1993), as shown in Fig 1.7. Residues within this region that are conserved between the FliG protein from *R.sphaeroides* and from bacteria with bidirectional flagella may be the residues important for determining rotation of the flagellum. This idea is addressed more thoroughly in chapter 5

FliG interacts with FliF, FliM FliN and MotA.

The interactions of FliG with other basal body components has been the focus of several recent studies (Oosawa *et al.*, 1994; Marykwas & Berg, 1996; Marykwas *et al.*, 1996; Tang *et al.*, 1996) and a more detailed discussion of this will be given in chapter 5.

The interaction of FliG with FliM may govern switching;

As will be shown below, it has been suggested that it is the modulation of FliM that controls switching of flagellum rotation, probably via its interactions with FliG (Sockett *et al.*, 1992; Welch *et al.*, 1993; Marykwas & Berg, 1996). It will be intriguing to analyse the interactions that take place between these proteins in *R.sphaeroides* as it could be predicted that as FliM is the major switching determinant (Sockett *et al.*, 1992; Tang & Blair, 1995; Lloyd *et al.*, 1996), these interactions in *R.sphaeroides* would result in flagellum stopping such as FliM acts as a 'brake' or 'clutch' for the motor. These features will be addressed more thoroughly in chapter 5.

FliG is present at high concentrations within the cell;

An interesting feature of FliG is that it is thought to be present at extremely high concentration within the cell, namely 3,700 copies (Roman *et al.*, 1993). This would equate to over 700 copies of FliG per flagellum. Mutants in which FliF is fused to FliG have been isolated which still allow flagellum rotation (Francis *et al.*, 1992). Assuming the stoichiometry of FliF to be up to 26 copies per flagellum as predicted by a variety of methods (Jones *et al.*, 1990; Sosinsky *et al.*, 1992; Ueno *et al.*, 1994), and given that the FliF-FliG fusion

still rotate, the predicted stoichiometry of FliG would also be up to 26 per flagellum. The question could then be asked: **Why produce such a large quantity of FliG?** MotA also appears to be expressed at high levels (Wilson & Macnab, 1988) and as shall be shown below, it has also been shown for the switch proteins FliM and FliN (Tang *et al.*, 1995; Tang & Blair, 1995). **Does the large excess of switch proteins and motor proteins reflect a transient interaction of proteins at the flagellar motor/switch?** There is no evidence to suggest whether or not this is the case, and this area requires more substantial study.

A potential role for FliG in controlling flagellar gene expression has also been noted (Marykwas *et al.*, 1996) as it has been found to bind to H-NS, an abundant histone-like protein. H-NS has been shown to be essential for the activation of class 1 and class 2 flagellar genes (Bertin *et al.*, 1994) (see below for a more in depth discussion of flagellar gene regulation). This finding may suggest that the binding of H-NS to FliG may be the key to activation of flagellar gene expression although whether or not this interaction takes place *in vivo* requires further investigations.

1.4.3f FliM

As for FliG, FliM has been shown to be present in most flagellar systems studied to date (Kihara *et al.*, 1989; Zuberi *et al.*, 1991) and is the site at which the chemosensory apparatus interacts with the flagellum (Sockett *et al.*, 1992; Welch *et al.*, 1993). The interaction with phosphorylated-CheY leads to flagellar reversal (ccw to cw rotation) which ultimately controls the response of the cell to tactic stimuli (for a review see Eisenbach, 1996). The chemotactic response will be covered in more detail below.

FliM binds to FliG, FliF, MotA and FliN;

The binding of FliM to the other components of the basal body has been

Fig 1.7 Domain organisation of *S.typhimurium* FliG, FliM and FliN. Taken from Irikura *et. al.*, 1993. Regions implicated in being important for ccw and cw states of the motor were isolated as suppressors of *cheY* and *cheZ* defects respectively.

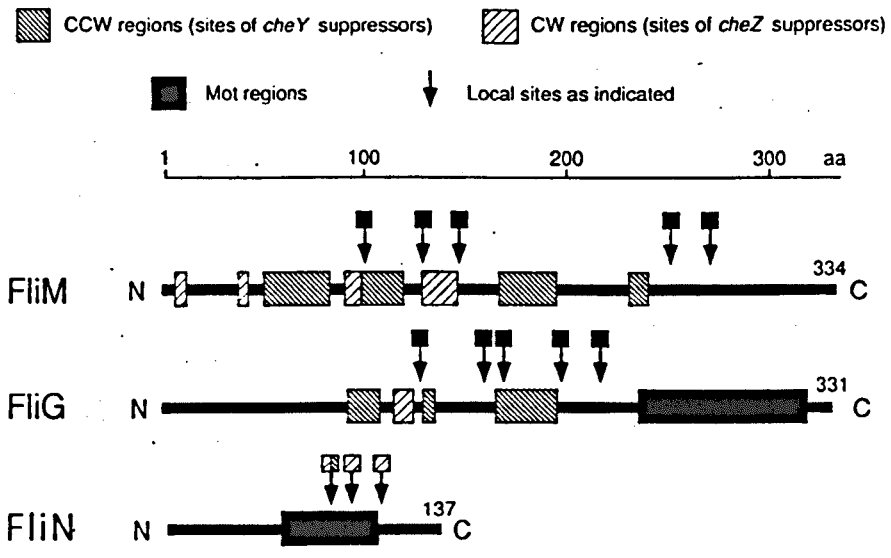
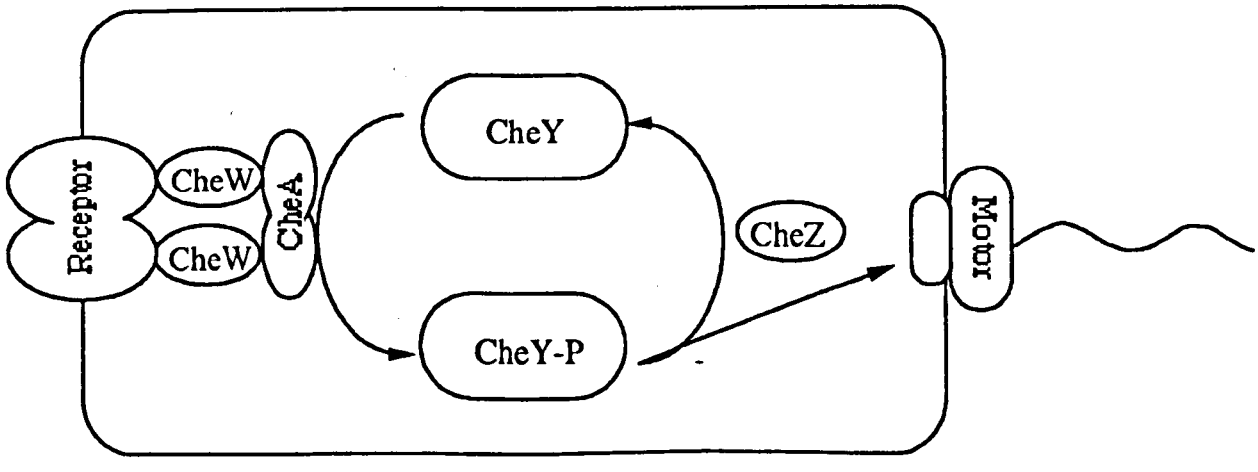


Fig 1.8 Simplified scheme of signal transduction in bacterial chemotaxis emphasising the central role of CheY. The receptor is shown as an MCP (see section 1.4.3g). The switch protein with which CheY interacts is FliM. For simplicity, the cytoplasmic proteins CheR and CheB, known to be involved in adaptation at the receptor level, are not shown. Redrawn from Eisenbach, 1996.



demonstrated (Oosawa *et al.*, 1994; Tang *et al.*, 1995; Marykwas & Berg, 1996; Marykwas *et al.*, 1996; Tang *et al.*, 1996), it has also found that there are approximately 1,400 copies of FliM per cell.(Tang *et al.*, 1996). The significance of this high stoichiometry is unknown, but as previously stated, it may reflect a poor affinity of these proteins for the basal body or a transient interaction. It will be interesting to determine the stiochiometries of all flagellar basal body components in the uniflagellate *R.sphaeroides* as this may give some idea as to why such a high stoichiometry exists in the multiflagellated organisms i.e. if due to poor affinity of the proteins for the basal body then *R.sphaeroides* would be predicted to also have a high stoichiometry, or due to the multiple flagella whereby *R.sphaeroides* would be predicted to possess low stoichiometries of the proteins..

FliM may hold FliG in the correct position to allow its interaction with the motor proteins;

Tang and co-workers found that cells containing approximately one third of the wild type levels of FliM showed rapid fluctuations in the speed of the motor and were poorly flagellated (Tang & Blair, 1995). This confirmed the findings of Sockett and co-workers that mutations in FliM can give rise to non-flagellate cells, motor defects and switch defects (Sockett *et al.*, 1992) (see Fig 1.7). The speed fluctuations observed may have been due to the fact that FliG was not in the correct conformation to allow maximum torque generation. This is consistent with the finding that mutations in MotB can be suppressed by FliG and FliM (Garza *et al.*, 1995; Garza *et al.*, 1996) and the model produce by Garza and co-workers whereby FliM holds FliG in the correct position to allow its interaction with the motor complex (Garza *et al.*, 1996). Tang and co-workers also showed that the overexpression of FliM results in reduced flagellation and reduced motility further confirming that FliM interacts with the basal body components FliG and FliN and that the effects seen were due the sequestration of FliG and FliN away from the basal body (Tang & Blair, 1995).

FliM is the point at which the chemosensory apparatus interacts with the flagellum

As FliM is the interaction point between the chemosensory apparatus and the flagellum (Sockett *et al.*, 1992; Welch *et al.*, 1993), it is appropriate to give a brief outline of how the chemosensory apparatus responds to stimuli in this section.

1.4.3g The chemosensory apparatus and chemotactic response

The ability of bacteria to swim towards favourable stimuli and away from unfavourable stimuli is the basis of taxis. This majority of what is known to date about taxis, and more specifically chemotaxis has stemmed from work on *E.coli* and *S.typhimurium*. The chemotactic response is a complex process and only a simple knowledge of it is required to understand the work carried out throughout this project. What follows will be a very brief summary and for a more in depth discussion of chemotaxis and other types of taxis the reader is referred to the following reference and citations therein: Bourret *et al.*, 1991; Armitage, 1992; Armitage *et al.*, 1995; Eisenbach, 1996.

Cells sense their environment via receptors and cause switching of rotation by a phosphorelay.

E.coli and *S.typhimurium* sense their environment mainly via receptors in the cytoplasmic membrane known as methyl-accepting proteins (MCP's). These MCP's, which can bind methyl groups, have a periplasmic domain that binds to the attractant/repellent molecules and via conformational change, signal to the cytoplasmic components of the chemosensory apparatus (see Armitage, 1992 for a review). There then follows a phosphotransfer relay that results in the production of high levels of the signal molecule CheY-phosphate (see Fig 1.8 for a brief description). It is the levels of CheY-phosphate that determine the direction of cell movement via its interaction with the switch protein FliM (Welch *et al.*, 1993). During stimulation by attractants, whereby the cell is moving up a gradient of attractant, there is prolonged swimming with less

switching (via less CheY-phosphate). When the cell starts to move down the gradient of attractant or up a gradient of repellent then the level of flagellum reversal increases (via the increase in CheY-phosphate) resulting in cellular tumbling and changing of direction of movement. This situation differs markedly in *R.sphaeroides,* as CheY-phosphate is presumed to cause flagellum stopping, whereupon the flagellum coils up, the cell reorientates by Brownian motion until the flagellum restarts and the cell swims off in another direction (Armitage & Macnab, 1987). It also appears that in *R.sphaeroides,* the basis of the chemotactic response is to sense a drop in stimuli or step down in the concentration of attractant instead of an increase as is seen in *E.coli* and *S.typhimurium* (Armitage *et al.*, 1995).

R.sphaeroides has two *cheY* genes - does it have two *fliM* genes?

Other significant differences between *R.sphaeroides* and other bacteria include the finding that *R.sphaeroides* possesses two *cheY* genes (Ward *et al.*, 1995) and that the MCP protein identified is predicted not to be membrane bound (Ward *et al.*, 1995). The significance of two CheY proteins is not known but it has been suggested that they are linked to two distinct phospho-donors i.e. one responding to the classical MCP dependant activation and the other to a metabolite donor (Armitage *et al.*, 1995). It may be that there are also two *fliM* genes encoding two distinct FliM proteins. The identification of *fliM* in *R.sphaeroides* (as will be discussed in chapter 5) should shed some light on the interactions that takes place between FliM and CheY to cause flagellum stopping.

1.4.3h FliN

The role of FliN in flagellum formation, switching and rotation has been debated as many motility mutations have been isolated within it that affect all three phenotypes (Irikura *et al.*, 1993) (see Fig. 1.7). Recently, a role for FliN in flagellum rotation has been discounted as being due to the poor insertion of mutant FliN proteins, containing mutations thought to affect rotation, into the basal complex (Tang *et al.*, 1995; Lloyd *et al.*, 1996).

FliN is also involved in holding FliG in the correct conformation for torque generation and FliM in the correct conformation for switching;

It has also been shown that the underexpression of FliN results in the same phenotype as that seen by the underexpression of FliM, namely motor speed fluctuations (Tang *et al.*, 1995). This may be due to the incorrect conformation of the switch complex the consequence of which is that FliG cannot interact correctly with the motor complex and FliM cannot cause switching.

FliN plays a major role in export;

FliN may play a major role in the export of flagellar components, possibly stabilising the interaction of the flagellum specific export apparatus with the basal body (Tang *et al.*, 1995). This is consistent with the finding that it is homologous to Spa33 (Tang *et al.*, 1995), a protein thought to be involved in the export of virulence determinants (Venkatesan *et al.*, 1992; Sasakawa *et al.*, 1993), and also that the region that is homologous is sufficient for the formation and rotation of the flagellum (Tang *et al.*, 1995). A more in depth description of the flagellar specific export apparatus can be found below

The presence of a FliN homologue in *R.sphaeroides* has been demonstrated (Sockett Pers. comm.) and the implications of this will be discussed in chapter 4.

1.4.3i The switch complex - a summary

Reviewing the literature, it is now apparent that the switch complex, consisting of three proteins in *E.coli* and *S.typhimurium* (FliG, FliM and FliN), is a protein complex in which all three proteins must be present at their correct stoichiometries and relative geometries in order for flagellum rotation, switching and assembly to occur efficiently. It appears that the presence of all three proteins allows the proteins involve rotation (FliG), switching (FliM) and export (FliN) to adopt the correct conformation to allow them to carry out their

function. The presence of the complex and the ability of the proteins to interact with each other and the MS-ring complex is indicative of the fact that their location at the base of the flagellum is essential for their correct function. The presence of such a complex in *R.sphaeroides* is probable and as will be shown in chapter 5 the presence of part of this complex has been demonstrated as a result of this project.

1.4.3j The MotA/B complex

The following section describes what is known to date about the motor proteins MotA and MotB which are thought to constitute the stator part of the motor. An understanding of how these proteins are thought to function and the evidence for this is essential for the reader to completely understand the rationale behind the rotation model predicted in Chapter 6.

The motor complex forms a proton pore;

As stated in section 1.3, the bacterial flagellum is powered by a motor at its base (Berg, 1974) using a transmembrane gradient of protons (Larsen *et al.*, 1974; Manson *et al.*, 1977) or sodium ions (Hirota & Imae, 1983). The MotA and MotB proteins are thought to be responsible for formation of the proton pore that is involved in the generation of torque (Blair & Berg, 1990; Wilson & Macnab, 1990) although the method by which they do this is not understood. It is the interaction of the motor proteins with the switch proteins that is thought to cause flagellum rotation and switching. The comparison of the motor proteins from the unidirectional flagellum of *R.sphaeroides* with those from bidirectional flagella may reveal critical domains important for rotation and not switching and in order to address this further what follows is a description of the current knowledge of the motor proteins.

The genes encoding homologues of MotA and MotB have been identified in many bacteria; *E.coli* (Dean *et al.*, 1984; Stader *et al.*, 1986); *B.subtillis* (Mirel *et al.*, 1992); *Vibrio parahaemolyticus* (McCarter, 1993)

and also from *R.sphaeroides* (Shah *et al.*, 1995; Shah & Sockett, 1995). Null mutations in these genes give rise to paralysed flagella that cannot rotate but are morphologically identical to those that do rotate.

MotA from *E.coli* is predicted to possess 4 membrane spanning helices with helix two and three being separated by a cytoplasmic domain that begins with a highly basic region and end with a highly acidic region (Dean *et al.*, 1984), the same features have also been noted in other MotA proteins (Mirel *et al.*, 1992; Shah & Sockett, 1995; McCarter, 1993).

MotB is thought to bind to peptidoglycan;

The MotB proteins predicted to possess only a single membrane spanning helix, with a large periplasmic domain (Shah *et. al.*, 1995; Stader *et al.*, 1986 ; Mirel *et al.*, 1992), which in the case of *E.coli* and *B.subtillis* contains residues thought to be involved in binding to peptidoglycan in the bacterial outer membrane (De Mot & Vanderleyden, 1994). The absence of these residues in the MotB component from the unidirectional flagellum of *R.sphaeroides* has led Shah and co-workers to conclude that if MotB requires some interaction with the outer membrane, that it does so using different residues or via the interaction with another peptidoglycan binding protein (Shah *et. al.*, 1995). The authors also suggest that MotB may act as the stopping determinant of the flagellar motor, possibly interacting with the MS-ring complex. Whether or not this process occurs has yet to be proved, although the evidence for the role of the switch proteins in switching the direction of rotation may suggest that they may also play some role in determining the unidirectionality. This will be discussed in chapter 5.

The motor proteins form complexes surrounding the base of the flagellum;

The only direct evidence for the presence of the motor proteins at the base of the flagellum comes from the use of freeze-fracture electron microscopy which has revealed the presence of 10 - 12 'studs' surrounding a doughnut-

shaped depression formed in the cytoplasmic membrane by the M-ring (Khan *et al.*, 1988). Such 'studs' are only formed in the presence of the genes encoding the motor proteins MotA and MotB and have therefore been assumed to be the result of the interaction of the motor proteins with the flagellar basal body (Khan *et al.*, 1988). The fact that both MotA and MotB must be present in order for these 'studs' to be formed, combined with the findings that MotB is incorporated into the membrane in a site limited manner (Stader *et al.*, 1986) whereas MotA is incorporated in a site unlimited manor (Wilson & Macnab, 1988) suggests that it is the localisation of MotB to the basal body and its interaction with MotA that results in the formation of the 'stud' particles seen.

The proton pore is made up of the membrane spanning helices of MotA and MotB;

The fundamental question to be answered from the study of the motility proteins MotA and MotB is: **What is the nature of the proton channel and how does it function?** Extensive mutational analysis of *motA* and *motB* from *E.coli* have revealed several domains thought to be important for the production of the proton channel (Blair *et al.*, 1991; Blair *et al.*, 1991; Stolz & Berg, 1991; Garza *et al.*, 1995; Sharp *et al.*, 1995; Garza *et al.*, 1996). Most of the work has centred on the generation of second site suppressor mutations and the analysis of dominant mutations that when transferred in *trans* into wild type cells, interfere with motility. The analysis of such mutants of MotA have revealed that almost all of these mutations fall within the membrane spanning helices as shown in Fig 1.9 (Blair *et al.*, 1991). These mutations led Blair and Berg to conclude that it is the membrane spanning helices that are the most important for the function of MotA. These results preceded the finding that MotA is in fact not a proton channel on its own as

Fig 1.9 Mutations isolated in *E.coli* MotA. Taken from Blair and Berg, 1991. Hexagons represent residues altered in slightly impaired mutants. Squares are residues that were altered in severely impaired, dominant alleles. Circles are residues that were altered in non-functional, dominant alleles.

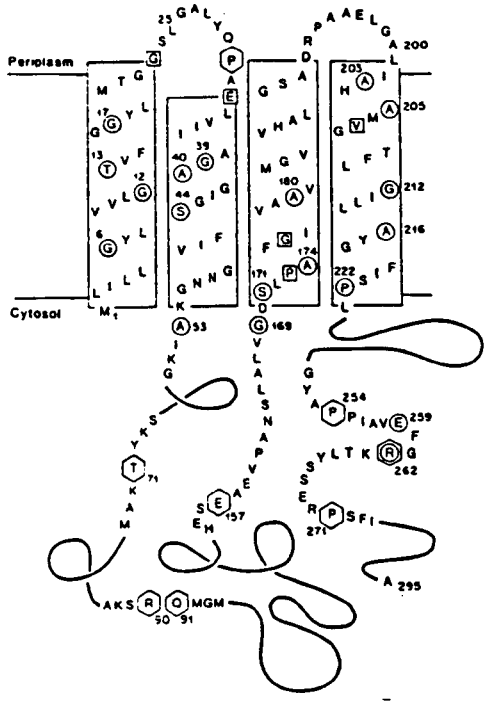
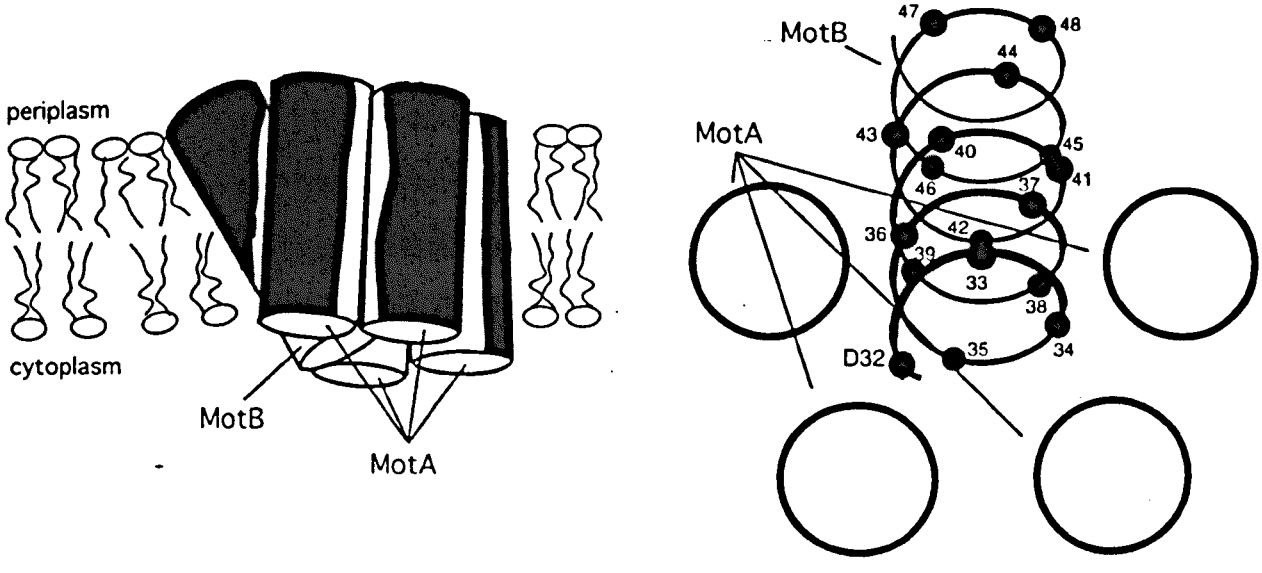


Fig 1.10 Possible structure of the MotA/MotB proton channel. Taken from Sharp *et. al.* 1995. Only the membrane spanning helices are shown. Left is shown the overall arrangement of the membrane-spanning segments. In the right the view from the cytoplasmic end of the channel is shown with the helical arrangement of residues displayed.



previously suggested (Blair & Berg, 1990), but in fact requires the presence of MotB (Stolz & Berg, 1991). This error was because in the original experiments, whereby Blair and Berg overexpressed MotA in *E.coli* and found vesicles containing MotA to be more permeable to protons than mutant variants of MotA known to prevent motility (Blair & Berg, 1990), the observed effect was due to the co-overexpression of a MotB-TetA fusion protein unfortunately produced as a result of the cloning procedure (Stolz & Berg, 1991). This fusion protein, containing the N-terminal 60 residues of MotB fused to the C-terminal portion of TetA resulting in the formation of the following peptide:

MKNQAHPIIV VKRRKAKSHG AAHGSWK**IAY ADFMTAMMAF FLVMWLISIS** SPKELIQIAE SASRVSVMTV KTSDTCSSRR RSQLVCKRMP GADKPVRARQ RVLAGVGAQP. The TetA part of the fusion (underlined above) is highly polar and was presumed to remain in the periplasm. Interestingly, similar fusions in which either alanine 39 is changed to valine or both alanines 29 and 31 are changed to threonine fail to cause the effect on growth and proton permeability (Stolz & Berg, 1991). These mutations, thought to reside within the membrane spanning helix of MotB (in bold in the fusion protein above) (Chun & Parkinson, 1987), have been shown to be dominant when expressed in full-length MotB (Blair *et al.*, 1991). It is interesting to note that this fusion protein contains a high level of potential proton binding residues (S, T, D), most of which would be in close proximity to the membrane and presumably the periplasmic side of the channel, possibly contributing to proton transport along with the membrane spanning helices of MotA (see below) or funnelling of protons into the channel. This may suggest the need for proton delivery into the channel by periplasmic proton conducting residues, possibly residing within MotB. The mechanism of proton conductance across the membrane via MotA/MotB had previously been hypothesised with Blair and co-workers suggesting that 10 of the 15 potential hydrogen bonding residues in MotA could contribute to a proton channel in the

centre of the four membrane spanning helices of MotA (Blair *et al.*, 1991). As the authors correctly note, it has been estimated that it requires 20 residues to transverse the membrane (Nagel *et al.*, 1980), and they subsequently put forward the idea that part of the proton conductance might involve water molecules (Blair *et al.*, 1991). It has been suggested that part of the pathway relies on residues on one of the intramembranal components of the basal body (Caplan & Kara-Ivanov, 1993). The MS-ring complex is a good candidate for this as the M-ring would be expected to be in close proximity to the torque generators in the cytoplasmic membrane (Stallmeyer *et al.*, 1989). This would presumably rely on the presence of proton binding residues in the membrane spanning helix that constitute the M-ring portion of the MS-ring complex (see Fig 1.5). Given the conserved nature of the process, i.e. proton conductance leading to flagellum rotation, it could be predicted that such proton binding features would be conserved in FliF's from bidirectional flagella and unidirectional flagella. As will be shown in chapter 5, this is not the case and the recent finding that MotA binds to FliG suggest that this protein may play a role in proton binding (Tang *et al.*, 1996). A greater discussion of this will be presented in chapter 5.

In the light of the evidence on the MotB-TetA fusion (Stolz & Berg, 1991), the idea of a channel consisting of MotA alone is not substantiated. Further evidence to support the idea that it is a complex of both MotA and MotB that controls proton conductance and flagellum rotation comes from the analysis of dominant MotB mutations (Blair *et al.*, 1991) and suppressors of them (Garza *et al.*, 1995; Garza *et al.*, 1996). Mutations affecting the large periplasmic domain of MotB, involved in anchoring MotB to the peptidoglycan (Chun & Parkinson, 1987; De Mot & Vanderleyden, 1994), which prevented the correct interaction of the MotA/B complex with the rotor components of the basal body, were not due to the sequestration of MotA away from the motor but possibly the poor interaction of MotB with MotA or the incorrect positioning of

the MotA/B complex with respect to the basal body (Blair *et al.*, 1991). The use of tryptophan scanning mutagenesis in the analysis of the interactions between MotA and MotB had also given some insight into the nature of the proton channel and has led to the idea that the proton channel is asymmetric (Sharp *et al.*, 1995) being made up of the 4 membrane spanning helices of MotA and the membrane spanning helix of MotB as shown in Fig 1.10. The authors also suggest that membrane spanning helices 1 and 2 of MotA that play the more important role in the interaction with MotB.

Several models of flagellum rotation have involved the delivery of protons to a cytoplasmically located component of the motor (for a review see Caplan and Kara-Ivanov, 1993) possibly FliG or FliM (Tang *et al.*, 1996) The authors therefore suggest that the tilted nature of the MotB protein in the channel may reflect the need for lateral delivery of protons to the cytoplasmic component, possibly via the aspartate residue at position 32 (Sharp *et al.*, 1995). The idea that the periplasmic domain of MotB, specifically the region predicted to be close to the membrane, is involved in the channelling of protons to the membrane is consistent with there being a high degree of proton binding residues present in this region (Stader *et al.*, 1986; Mirel *et al.*, 1992; Shah *et. al.*, 1995). This is similar to situation found in the MotB-TetA fusion which was shown to allow proton conductance only in the presence of MotA (Stolz & Berg, 1991). The idea does not correlate with the fact that no mutations affecting motility have been isolated in this region (Blair *et al.*, 1991), but this may simply reflect the low number of mutants isolated or the methods used to isolate such mutants. Direct evidence for the role of this domain requires site-directed mutagenesis.

The motor proteins from *R.sphaeroides* differ from those from bidirectional flagella and this may cause the unidirectionality of the motor;

The analysis of the motor proteins from the unidirectional flagellum of *R.sphaeroides* has shown some quite intriguing differences that may govern the unidirectionality of the motor (Shah *et. al.*, 1995; Shah & Sockett, 1995). The respective proteins retain the overall features of MotA possessing 4 membrane spanning helices and MotB possessing one, but contain several interesting differences. In the case of MotA, the degree of conservation is low, but appears to be clustered in the membrane spanning helices and the cytoplasmic loop, with a particularly high region of homology between all MotA proteins being centred around membrane spanning helix 4 (see Fig 3 in Shah and Sockett, 1995). This therefore suggests that helix 4, being the most conserved helix, plays a major role the function of MotA. The authors also suggest that the highly charged cytoplasmic domain may act as a 'gate' to control proton flux and has subsequently been shown to be essential for motility in *R.sphaeroides* (G.Günter, 1996 pers. comm.). The deletion of this region from MotA probably resulted in the disturbance of the structure of the MotA/MotB complex preventing torque generation and the exact role of the cytoplasmic loop in flagellum rotation is currently being investigated with site directed mutagenesis (D. Shah, 1996 pers. comm.).

The *R.sphaeroides* MotB protein appears to have limited homology with other MotB proteins which is almost exclusively restricted to the membrane spanning helix (Shah *et. al.*, 1995). The aspartate residue at position 32 in *E.coli* MotB, which has been implicated as being essential for either interaction with a cytoplasmic component or 'gating' of the channel (Blair *et al.*, 1991; Sharp *et al.*, 1995), is also highly conserved. This is consistent with the idea that the proton channel is a complex of both MotA and MotB. The most striking feature of the *R.sphaeroides* MotB protein is the histidine rich C-terminus where the motif HARCARGHGPDCRGHAQSRTHHHH is present (Shah *et.al.*, 1995), similar to the histidine repeat seen in the heavy chain of dynein, the force generating protein of eukaryotic flagella (Garber *et al.*, 1989). The

authors suggest that this may represent the point at which the protein binds to another protein, possibly interacting with the peptidoglycan, or that it may function in the stopping of the flagellum, again this theory is currently under investigation (D. Shah, 1996 pers. comm.).

As previously stated, it is the interaction of the force generating unit, possibly MotA/MotB complex, with a cytoplasmic component that is thought to result in flagellum rotation (for a review see Caplan and Kara-Ivanov, 1993). Analysis of MotB suppressor mutants have revealed that the interacting component may in fact be FliG (Garza *et al.*, 1995; Garza *et al.*, 1996). The authors also predicted that it may be the cytoplasmic component of MotA that interacts with FliG and these findings are consistent with the presence of mutations in FliG that affect motor function (Yamaguchi *et al.*, 1986; Lloyd *et al.*, 1996) and the recent finding that MotA binds to FliG (Tang *et al.*, 1996).

Given the information available on MotA/B and the switch proteins, along with the information gained throughout this project, a model for flagellum rotation was predicted. A detailed discussion of this model can be found in chapter 6 along with the rationale used to predict it.

14.3k The export apparatus

As stated above, the export and assembly of the flagellum is an intriguing process with respect to how it is controlled. In *R.sphaeroides* there is the added complexity of the presence of the photosynthetic apparatus and the flagellum must be positioned correctly with respect to this. During the course of this project, a component of the export apparatus was discovered and analysed. What follows is a brief description of previous studies on the export apparatus from *E.coli* and *S.typhimurium* and this should give the reader adequate background information to fully understand the work detailed in chapter 4.

Flagella are assembled in an ordered manner;

How cells export flagellar components is not well understood but numerous groups have shown that it proceeds in an ordered manner (Iino, 1969; Emerson *et al.*, 1970) and a number of structural intermediates can be isolated from mutants in various flagellar genes that are stable (Suzuki *et al.*, 1978; Suzuki & Komeda, 1981; Jones & Macnab, 1990; Kubori *et al.*, 1992) as shown in Fig 1.11.

Most flagellar components are exported by a dedicated apparatus;

Most bacterial proteins that are located outside the cytoplasmic membrane are exported to their position via the General Secretory Pathway (GSP). The reader is referred to the review by Pugsley (Pugsley, 1993) for a discussion of the GSP as an understanding of the processes involved is beyond the scope of and not essential for an understanding of this project. However, any relevant information on the GSP will be discussed below. Only three flagellar proteins are thought to be exported by the GSP, namely FlgH (L-ring), FlgI (P-ring) and FliP, by virtue of an N-terminal signal peptide (Homma *et al.*, 1987; Malakooti *et al.*, 1994). The rest are exported via a flagellum-specific export pathway.

Components pass throught the centre of the flagellum;

Flagellum specific export occurs via the central core of the growing flagellar structure (Namba *et al.*, 1989) and allows the addition of protein subunits onto the growing distal end of the flagellum (Iino, 1969; Emerson *et al.*, 1970). As stated above, the components are assembled in an ordered and controlled manner and the apparatus for the export of flagellar components is thought to reside on the cytoplasmic side of the MS-ring complex (Vogler *et al.*, 1991).

The export apparatus is a multi protein complex;

Many proteins have been identified that are thought to play a role in the flagellar specific export pathway, namely FliF; FliN; Fli H,I and J (Vogler *et al.*, 1991); FliK (Williams *et al.*, 1996); FliO, P, Q and R (Malakooti *et al.*,.

Fig 1.11 Morphological pathway of flagellum assembly. Taken from Kubori *et. al.* 1992.

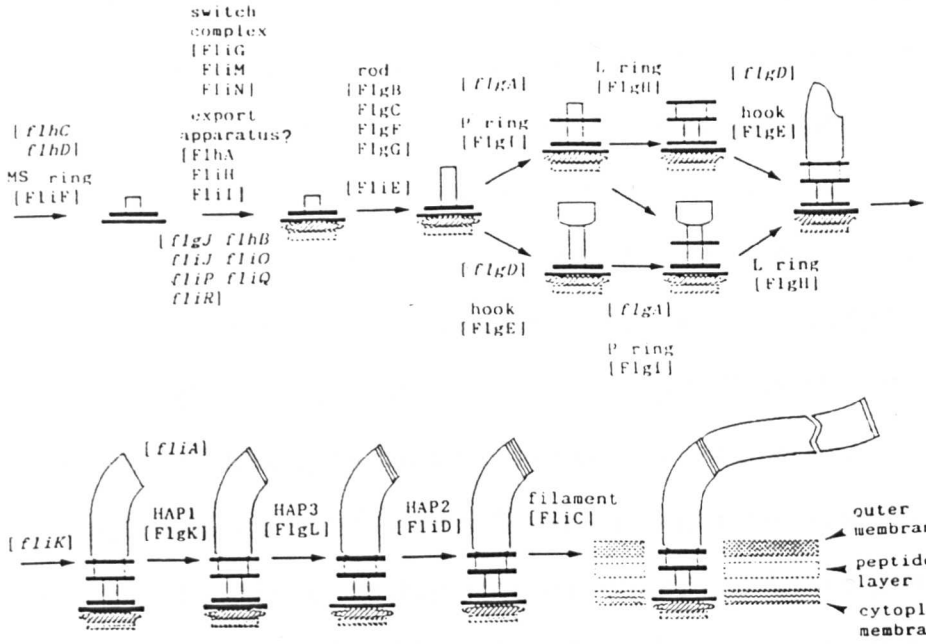
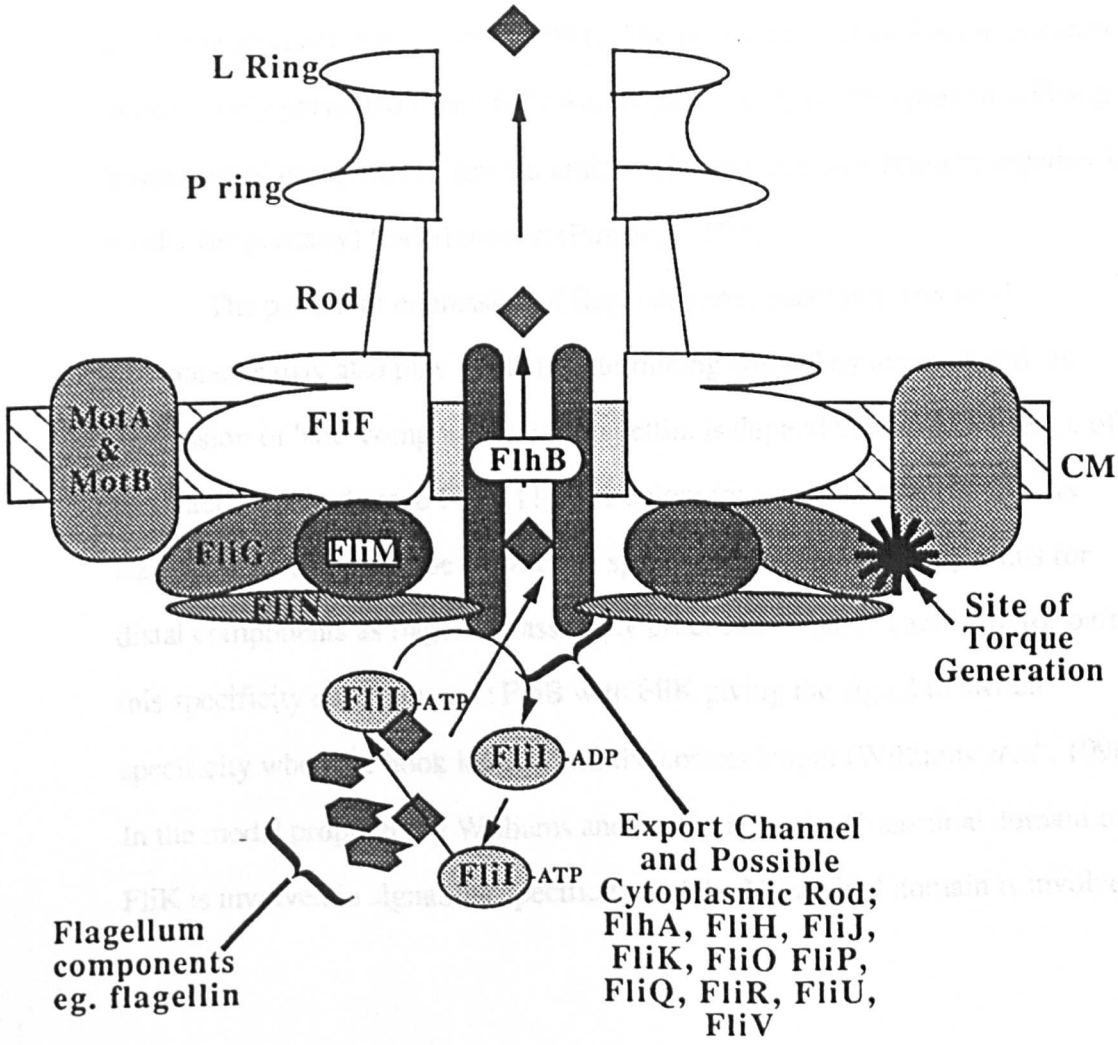


Fig 1.12 Possible structure and function of the flagellum-specific export channel. Modified from Tang *et. al* 1995.



1994); FliU and V (Doll & Frankel, 1993); FlhA and FlhB (Minamino *et al.*, 1994) and they are homologous to virulence export proteins (Harshey & Toguchi, 1996) (Table 1.1) which may suggest a common evolutionary ancestor

The export apparatus is specific for flagellar proteins;

The mechanism that decides when specific components of the flagellum are exported is not known but may be due to the recognition of 'cryptic' signal sequences by specific protein 'gate-keepers' that confer specificity to the export apparatus (Homma *et al.*, 1990; Homma *et al.*, 1990; Williams *et al.*, 1996). The nature of these signal sequences are unknown but there appears to be sequence homology between several exported components at their N- and C-termini, but whether or not they function in export is not known (Homma *et al.*, 1990; Homma *et al.*, 1990; Dingwall *et al.*, 1992). Work on flagellin and hook proteins have revealed regions that are required for export but there appears to be no common structural theme or primary sequence conservation (Kuwajima *et al.*, 1989; Kornacker & Newton, 1994). The possibility exists that the secretion signal is only generated in the fully folded (or partially folded) protein utilising a 'patch' signal composed of several amino acids that are only brought together in a fully (or partially) folded protein (Pugsley, 1993).

The pattern of expression of flagellar genes encoding structural components may also play a role in determining when they are exported as expression of 'late' components e.g. flagellin, is dependant on the presence of an intact basal body (see Fig 1.11) (see below for a greater discussion). As stated above, there may be a switch in specificity of the export apparatus for distal components as flagellum assembly proceeds. A good candidate for part of this specificity determinant is FlhB with FliK giving the signal to switch specificity when the hook has reached the correct length (Williams *et al.*, 1996). In the model proposed by Williams and co-workers, the C-terminal domain of FliK is involved in signalling specificity and the N-terminal domain is involved

in controlling hook length. They suggest that FliK alters the specificity of FlhB to switch from export of hook protein to the export of hook-associated proteins, flagellin, the filament cap and the negative regulator of late gene expression FlgM.

Export requires input of energy;

The energy for export procedure comes from ATP hydrolysis via the protein FliI (Vogler *et al.*, 1991). The ability of FliI to bind and hydrolyse ATP was originally deduced from its predicted protein sequence as it contained nucleotide binding motifs (Walker boxes) (Walker *et al.*, 1982). Direct evidence of ATP hydrolysis by FliI has recently been shown along with its ability to bind to flagellin (Silva-Herzog *et al.*, 1995). The authors also showed that upon binding of flagellin, the ATPase activity of FliI increases.

FliI is present at high concentrations in the cell;

The high stoichiometry of FliI (1,500 subunits per cell) may reflect the cyclic nature of its interaction with the export apparatus as shown in Fig 1.12. FliI may act as a chaperone to 'escort' flagellum components to the export apparatus which might explain its high stoichiometry. As will become apparent from chapter 4, the FliI component has been identified in *R.sphaeroides* during this project and also by another group independently (Ballado *et al.*, 1996). The features of the *R.sphaeroides* FliI homologue and other export components will be discussed in chapter 4.

1.5 Regulation of flagellar synthesis

The regulation of flagellar synthesis is another area of motility research that is fascinating as, as will be discussed below, the formation of the flagellum is tightly coupled to many signals. During this project the role of an alternative sigma factor was investigated and a putative flagellar gene regulator was identified. As a consequence it is necessary to give the reader an insight into the

Table 1.1 Similarities among bacterial flagellar proteins and virulence-factor export components. Taken from Harshey and Toguchi, 1996. Flagellar proteins listed at the top are from *E.coli* and *S.typhimurium* and have been implicated in flagellar export and assembly.

	Flagellar proteins							
Species	**FlhA**	**FlhB**	**FliF**	**FliI**	**FliN**	**FliP**	**FliQ**	**FliR**
Bacillus	-	FlhB	FliF	FliI	FliY	FliP	FliQ	FliR
Caulobacter	-	-	FliF	-	FliN	-	FliQ	FliR
Erwinia	HrpI	-	-	-	MopA	MopC	MopD	MopE
Pseudomonas	HrpO	HrpN	HrpI	HrpE	HrpQ	HrpT	HrpU	HrpC
Salmonella	-	SpaS	-	SpaL	SpaO	SpaP	SpaQ	SpaR
Shigella	Spa76	Spa40	MxiJ	Spa47	Spa33	Spa24	Spa9	Spa29
Xanthomonas	HrpC2	HrpB6	HrpB3	HrpB6	-	-	-	-
Yersinia	LcrD	YscU	YscJ	YscN	YscQ	TscR	YscS	YscT

Table 1.2 Flagellar gene transcriptional hierarchy in *S.typhimurium* and *E.coli* - examples of genes in the 4 classes.

Class	Gene	Function
1	*flhC flhD*	Activator of level 2
2	*fliF*	MS-ring complex
	fliM	Switch complex
	flhB	Export apparatus
3a	*fliD*	Filament 'cap'
3b	*fliC*	Filament protein
	motA	Motor protein

current knowledge of how the synthesis of flagella is regulated with the emphasis being on the role of alternative sigma factors and response regulators

Regulation of flagellar synthesis is tightly controlled;

The regulation of flagellar synthesis is a highly controlled and complex mechanism (reviewed in Macnab, 1992) as the synthesis of a flagellum is bioenergetically very costly to the cell. Consequently, organisms have developed systems whereby the synthesis of the flagellum is dependant not only on the nutritional status of the cell, but also the cell-cycle, temperature and whether or not the cell is under environmental stress (see reviews Macnab, 1992 and Brun, 1994). Most of the knowledge of how these regulation systems operate comes from work on *E.coli, S.typhimurium* and *C.crescentus*. The systems used in these organisms will be covered separately to attempt to simplify the process. The current state of knowledge on the regulation of flagellar formation in *R.sphaeroides* will also be covered to allow give some insight into how little is known in the photosynthetic organism *R.sphaeroides*.

1.5.1 *S.typhimurium* and *E.coli*

Flagellar genes are arranged in operons, 13 in *S.typhimurium* (Kutsukake *et al.*, 1988) and 14 in *E.coli* (Komeda *et al.*, 1980). These operons have been placed in classes (1, 2,3a and 3b) depending on when they are expressed. Level 1 is the master operon and expression of all the other classes is dependant upon its expression (see Macnab, 1992) See table 1.2 for examples of genes in each of the classes. The expression of level 1 genes *flhC* and *flhD* are themselves regulated by a number of mechanisms e.g. catabolite-repression e.g. cAMP levels (Adler & Templeton, 1967; Silverman & Simon, 1974), the presence of the heat shock proteins DnaK, DnaJ and GroEL (Shi *et al.*, 1992), and mutations in genes involved in phospholipid biosynthesis (which are required for cell division) inhibit motility (Komeda *et al.*, 1977; Fiedler & Rotering, 1988; Nishino *et al.*, 1993; Shi *et al.*, 1993).

Regulation involves an alternative sigma factor;

The expression of class 2 genes leads to the synthesis of the basal body-hook structure and the activator of class 3 genes sigma-F (σ^F), the product of the *fliA* gene (Ohnishi *et al.*, 1990). All class 3 genes in the flagellar cascade are proceeded by a consensus sequence for σ^F and epistasis experiments have revealed that these genes are dependant on *fliA* for their expression (Kutsukake *et al.*, 1990). Interestingly, most class 2 genes also have σ^F consensus sequences, but do not require *fliA* for expression (Helmann, 1991). This suggests that there are other, *fliA*-independent, promoters allowing transcription of these operons but does not exclude a role of σ^F (or a polymerase of related specificity) in transcription of these genes (Helmann, 1991).

The activity of the sigma factor is controlled;

The ability of FliA (σ^F) to activate class 3 genes is controlled by the anti-sigma factor FlgM which binds to FliA and prevents its interaction with RNA polymerase (Ohnishi *et al.*, 1992). FlgM functions to repress expression of class 3 genes until the assembly of a complete basal body-hook complex whereby it is exported and the expression of class 3 genes occurs to complete the formation of the flagellum (Hughes *et al.*, 1993; Kutsukake & Iino, 1994). It has been shown that the export protein FlhB is responsible for the inhibition of FlgM export until the completion of hook assembly (Kutsukake *et al.*, 1994). More recently, the proteins FliD, S and T have been shown to have a negative effect on the export of FlgM after the completion of hook assembly (Yokoseki *et al.*, 1996) which may suggest that they play some role in the specificity of the export apparatus.

FlgM may also regulate the number of flagella on a cell as FlgM depleted mutants have double the numbers of flagella as wild type cells (Kutsukake & Iino, 1994). A similar effect was also seen for the FliD, S and T mutants (Yokoseki *et al.*, 1996). In the uniflagellated organisms such as

R.sphaeroides and *C.crescentus*, the presence of a 'hyperactive' FlgM protein may be partly responsible for the presence of a single flagellum.

1.5.2 C.crescentus

The regulation of flagellum synthesis in *C.crescentus* is a very complex system involving cell-cycle determinants and a regulatory hierarchy similar to that in *E.coli* and *S.typhimurium*. As a consequence of this complexity, the reader is refereed to the recent reviews by Brun and Gober (Brun *et al.*, 1994; Gober & Marques, 1995) as what follows is a very simplified version of the complete model.

Genes are expressed in four levels and expression is linked to the cell-cycle;

As shown in Fig 1.14 a similar regulatory hierarchy exists in *C.crescentus* as in *E.coli* and *S.typhimurium*; the genes organised into 4 levels of expression, with expression of the lower levels being dependant on the genes at the higher level (Brun *et al.*, 1994). Until recently, the gene(s) required for the link between the cell cycle and flagellum biosynthesis i.e. at level 1, were not isolated. However, several of the class 2 promoters share a common motif (Stephens & Shapiro, 1993; Van Way *et al.*, 1993; Zhuang & Shapiro, 1995) and the features of this motif are also shared with genes known to be important in cell cycle control i.e. *ccrM* (an essential DNA methytransferase) (Stephens *et al.*, 1995) and *hemE* P_{strong} (lies within the origin of replication and acts as a key regulatory element for the cell type-specific initiation of DNA replication) (Marczynski *et al.*, 1995). The work by Quon and co-workers has identified an essential two-component signal transduction protein (CtrA), which is required for the regulation of many class 2 promoters, *ccrM* and *hemE* P_{strong} (Quon *et al.*, 1996). CtrA binds to the conserved motifs in the class 2 promoters (referred to above) and can act both as an activator or a repressor of these promoters. The state of flagellum assembly may determine the ability of CtrA

Fig 1.13 Regulation of flagellar gene expression in *S.typhimurium* and *E.coli*. Only the first gene in each operon is shown.

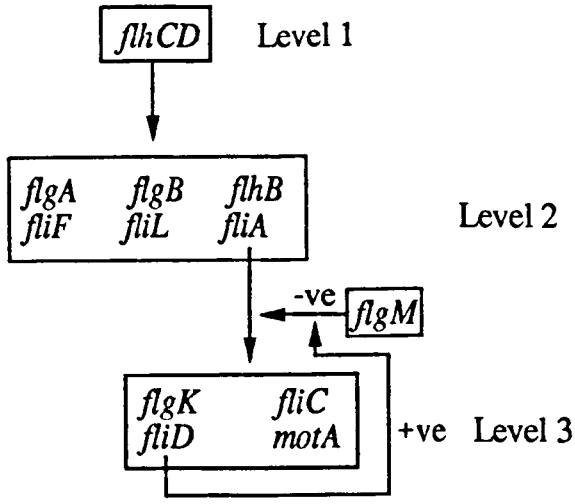
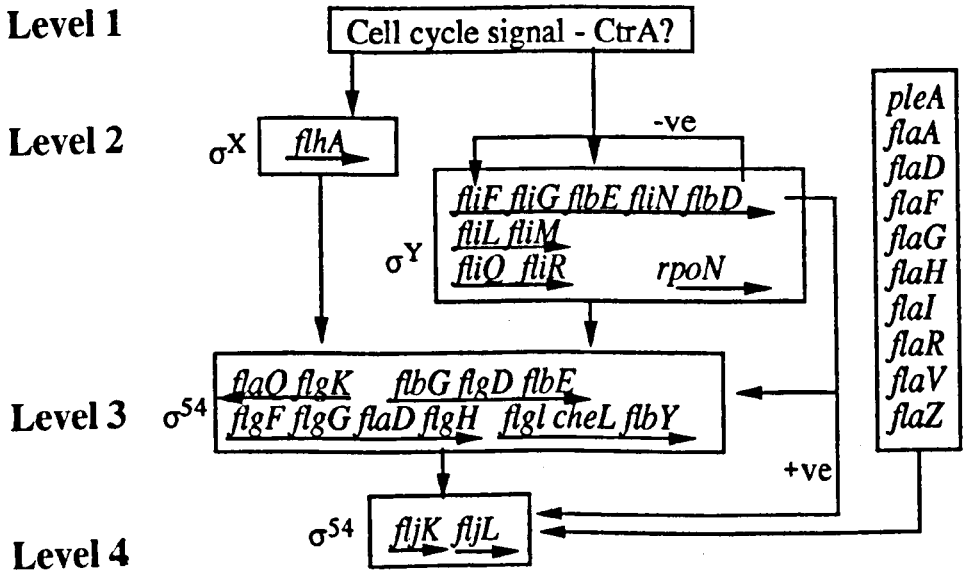


Fig 1.14 Regulation of flagellar gene expression in *C.crescentus*. Modified from Anderson *et al.*, 1995. Genes at the right of the diagram beginning with *pleA* have not been placed in the heirarchy, but are known to be required for production of the level IV flagellins.



to regulate flagellar gene expression, as seen for FlgM in *E.coli* and *S.typhimurium* (U. Jenal cited in Quon & Shapiro, 1996).

Alternative sigma factors play a major role;

Alternative sigma factors play a significant role in the regulation of flagellar gene expression in *C.crescentus*. As shown in Fig 1.14, three different sigma factors have been implicated in controlling different classes of promoter. The level 2 *fliF* (Van Way *et al.*, 1993), *fliL* (Stephens & Shapiro, 1993) and *fliQ* (Dingwall *et al.*, 1992) promoters share a promoter consensus that is thought to be recognised by a novel sigma factor, designated σ^Y (Benson *et al.*, 1994). The sequence of the *flhA* promoter is not similar to that from the other class 2 promoters and may be recognised by another specialised form of RNA polymerase (Benson *et al.*, 1994). The class 3 and 4 gene promoters have been most extensively studied and DNA sequence analysis originally showed that σ^{54} may be responsible for recognition of these promoters (Chen *et al.*, 1986; Minnich & Newton, 1987; Mullin *et al.*, 1987). The work of Mullin and co-workers showed that the conserved sequences and spacing of these promoters are absolutely required for their transcription *in vivo* (Mullin & Newton, 1989). The *C.crescentus rpoN* gene, encoding σ^{54}, is required for class 3 and 4 gene expression (Anderson *et al.*, 1995). σ^{54} has also been implicated in flagellar gene expression in *Pseudomonas putida* (Kohler *et al.*, 1989) as well as being involved in many other physiological functions (for a review see Merrick, 1993). A more in depth discussion of σ^{54} shall be given in chapter 3.

σ^{54} has many unique features;

σ^{54} containing RNA polymerase is unique among prokaryotic RNA polymerase holoenzymes as the formation of transcription-competent open complexes depends on the function of activation proteins that typically bind to specific DNA sequences elements (enhancer like elements) located upstream of the transcription start site (for reviews see Merrick, 1993; Morett & Segovia, 1993). Enhancer elements known as *ftr* (for flagellar transcriptional regulator)

have been identified approximately 100 bp from the transcription start sites many *C.crescentus* flagellar genes (see Fig 1.12) and are essential for the correct regulation of flagellar gene expression (Mullin & Newton, 1989; Mullin & Newton, 1993). FlbD, a homologue of the two-component regulator NtrC (Ramakrishnan & Newton, 1990), is essential for transcription of σ^{54} dependent promoters (Newton *et al.*, 1989) and it functions to activate transcription from the class 3 and class 4 promoters and represses transcription from the class 2 *fliF* promoter by its interactions with *ftr* sequence elements (Van Way *et al.*, 1993; Benson & Newton, 1994; Benson *et al.*, 1994; Mullin *et al.*, 1994; Wingrove & Gober, 1994; Wu *et al.*, 1995). A more detailed discussion of the mechanism of action and the function of FlbD shall be given in chapter 5.

In summary, the regulation of flagellar gene expression in *C.crescentus* involves similar mechanism of a regulatory hierarchy as seen in *E.coli* and *S.typhimurium* but with the additional complexity of a two-component regulatory system and multiple sigma-factors that result in the coupling of flagellar gene expression to the bi-phasic cell-cycle.

1.5.3 Regulation in *R.sphaeroides*

Flagellum formation in *R.sphaeroides* is regulated;

The regulation of flagellar gene expression in *R.sphaeroides* has not been studied a great deal, although there does appear to be some regulation with respect to environmental signals. The physical clustering of flagellar genes has been shown (Foster, 1991) which suggest that flagellar genes are arranged into operons as seen in other systems. This is further substantiated by the finding that MotA and MotB are arranged in a operon which is proceeded by a promoter region (Shah *et.al.* 1995; Shah & Sockett, 1995). As will become apparent from chapters 4 and 5 the existence of other motility operons in this organism has been demonstrated as a result of this project.

Fig. 1.15 Alignment of *ftr* sequences found in the promoter regions of *C.crescentus* flagellar genes. Re drawn from Mullin and Mullin, 1993. Numbering is relative to the transcription start site.

Class	Gene	Sequence			
2	*fliF*	CT	GGGT	AAATCCT	$GCCT^{+12}$
3	*flbG (ftr1)*	CT	CGGC	AAAAAGC	$GCCG^{-100}$
3	*flbG (ftr5)*	CG	CGGC	AACTCCC	$GTCC^{-134}$
3	*flaN (ftr2)*	CC	CGGC	GAAACTT	$GCCG^{+85}$
3	*flaN (ftr3)*	CT	CGGC	AAACCGC	$GCAA^{+119}$
3	*flgI*	CC	GGGC	AGAATCT	$GCCG^{-97}$
3	*flgF*	CC	CGGC	AAAACAC	$ATCG^{-121}$
3	*flgH*	GG	CGGC	GGCCATG	$GCCT^{-125}$
3	*flgK*	CT	CGGC	AGAAATT	$GCCG^{-100}$
4	*flgL*	CG	CGGG	CAAAACG	$GCCG^{+15}$
4	*flgJ*	AA	CGGC	GAAAATC	$GCCG^{+40}$
	Consensus	CB	CGGC	RRAHHBY	GCCD

Fig 1.16 The structure of the *R.sphaeroides motA/B* promoter region. Sequence courtesy of D.S. Shah. The inverted repeats are underlined with the start of MotA in bold.

Inverted repeats σ^{54} consensus sequence

CCGCCCG agagccgccgcga CGGGCGG ... 120 bp ... TGGCAC ggatc TTGC

..38 bp .. **ATG**
Met

Fig. 1.17 The structure of the *R.sphaeroides* flagellin gene (*fliC*) promoter region. Sequence courtesy of D.S. Shah. -35 and -10 represent the σ^F (σ^{28}) consensus sequence.

-35 -10

GCTAAA AGTTTCTCCGGCCGG CCGTTGA

Alternative sigma factors control expression in *R.sphaeroides*;

The role of alternative sigma factors in the control of flagellar gene expression has been implicated in two cases in *R.sphaeroides*; upstream of the *mot* operon, there is a consensus sequence for the alternative sigma factor σ^{54}(see Fig 1.16) with inverted repeats 120 bases further upstream (Shah & Sockett, 1995). The promoter region of the flagellin gene (fliC) has a σ^{28} consensus sequence (Shah *et al.*, 1996) (see Fig 1.17). Direct evidence for the roles of these sigma factors and the inverted repeats in flagellar gene regulation has not yet been shown. The presence of enhancer like elements (ELE's) upstream of MotA is suggestive that an enhancer binding protein (EBP) may interact with these ELE's. As stated in the abstract this project has revealed a putative member of the σ^{54} EBP family that may bind to the ELE's. A more in depth discussion of the role of ELE's and EBP's in gene activation will be given in chapter 5.

1.6 Aims of this project

As should now be apparent to the reader, little is currently known about the genes involved in flagellar formation and function in *R.sphaeroides*. So **why study motility in *R.sphaeroides* when so much progress has already been made in other systems?** In *E.coli* and *S.typhimurium*, when studying the effects of mutations for example on flagellum function on a single cell, what one observes is a mean effect on 5 flagella. In contrast, in *R.sphaeroides* such a study would give more accurate results as any effects observed would be the result of the mutation on a multi-enzyme complex present at one copy per cell and consequently more subtle effects on motility or flagellation would be easier to detect. As shall be shown in chapter 4 this is particularly important when studying the export apparatus. The complex process of photosynthesis also takes place in *R.sphaeroides* and much is known about the formation and function of the photosynthetic apparatus in this

organism (for a review see Kiley, 1988). The process of membrane differentiation that takes place when the cells are shifted from aerobic growth to anaerobic photosynthetic growth (see Kaplan *et al.*, 1983 for a review) adds another level of complexity to flagellar formation with respect to targeting. Also, the relative simplicity of the motor of *R.sphaeroides* i.e. its unidirectionality, may allow the elucidation of the mechanism of torque generation without the added complexity of switching. Consequently, *R.sphaeroides* is a good model for studying flagellum formation and function as the organism possesses many fascinating features of its life cycle and the motor is relatively simple compared to that from other systems.

The main aim of this project was to identify the genes encoding the rotor components of the motor (i.e. FliF and FliG). The motor components had already been identified (Shah *et. al.*, 1995; Shah & Sockett, 1995) and in order to fully understand the interactions that take place in the motor to allow torque generation, one must have identified the genes encoding all the components. The role of the alternative sigma factors and activators/repressors in the regulation of flagellum biosynthesis was also addressed.

As will be presented in chapters 3,4 and 5, molecular characterisation of motility mutants resulted in the identification of a motility gene operon that encoded structural components of the flagellum, the components of the rotor, a possible transcriptional regulator protein and a component of the flagellar specific export. The details of which are described in the following chapters.

Chapter 2

METHODS

To conserve space, the constituents of all reagents and solutions have been described in appendix 1.

2.1 Bacterial strains and plasmids

Bacterial strains and plasmids used as cloning vectors or host strains are listed in Table 2.1. Plasmids and strains constructed throughout this project will be discussed in the relevant chapters.

E.coli strains for plasmid maintenance and isolations were grown at 37°C in Luria Bertani medium (LB) (Maquat & Reznikoff, 1978) and grown in Terrific broth (Tartof & Hobbs, 1987) for the maintenance and isolations of cosmid clones. Strains for protein expression were grown either in LB (Maquat & Reznikoff, 1978) or 2 X YT (Pharmacia Biotech) (see appendix 1).

R.sphaeroides strains were grown either in succinate medium (Cohen-Bazire *et al.*, 1957 and Sistrom, 1960) (see appendix 1) at ambient temperature with constant illumination or shaken at 200rpm in LB (Maquat & Reznikoff, 1978) at 30°C.

Antibiotic selection was as follows; 50μg/ml of ampicillin, spectinomycin, streptomycin for both *E.coli* and *R.sphaeroides*, tetracycline was used at 25μg/ml for *E.coli* and 1μg/ml for *R.sphaeroides*, kanamycin was used at 25μg/ml for both *E.coli* and *R.sphaeroides*, naladixic acid was also used at 25μg/ml.

2.2 DNA Manipulation Techniques

2.2.1 Small Scale Isolation of plasmid DNA

Plasmid extractions were carried out using the alkaline lysis method of Birnboim & Doly (Birnboim & Doly, 1979). Essentially, this method involves the lysis of cells using a 1% SDS/0.2M NaOH solution which also eliminates protein and chromosomal DNA. The plasmid DNA is then purified

Table 2.1 Bacterial strains and plasmids

Strain/Plasmid	Genotype/Description	Reference
Strain		
Escherichia coli DH5α	*F'/endA1 hsdR supE44 thi-1 recA1 gyrA relA1 Δ(lacZYA-argF) U169 deoR (ϕ80 dlac Δ lacZ) M15*	Woodcock *et al.*, 1989
XL1-Blue	F'::Tn10 *ProA*$^+$*B*$^+$ *lacI*^{q} *Δ (lacZ) M15/recA1 endA1 gyrA46 thi hsdR17 supE44 relA1 lac*	Bullock *et al.*, 1987
S17-1	*Pro*$^-$, *Res*$^-$, *recA*; integrated plasmid RP4-Tc::Mu-Kn::Tn7	Simon *et al.*, 1983
Rhodobacter sphaeroides WS8	Wild-type	W.R. Sistrom
Rhodobacter sphaeroides WS8N	Wild-type: spontaneous Nal^R	Sockett, *et al.*, 1990
Plasmid		
pUC19	Cloning vector, Amp^R	Yanish-Perron *et al.*, 1985
pRK415-1	Broad-host range cloning vector, conjugatable, Tc^R	Keen *et al.*, 1988
pLA2917	Broad-host range cosmid vector, Kn^RTc^R	Allen & Hanson, 1985
pHP45Ω	Vector containing the Sp^R/Sm^R Ω cartridge	Prentki & Krisch, 1984

by a series of precipitation steps for washing. 1ml aliquots of o/n cultures of plasmid-containing cells were harvested by centrifugation. The supernatant was discarded and 100µl of lysis buffer (for extraction from *Rhodobacter sphaeroides*, lysozyme was added to 2µg/ml) (see appendix 1) and 200µl of fresh alkaline SDS (1% SDS/0.2M NaOH) were added to the pellet. The tubes were whirlimixed and kept on ice for 5min. 150µl of 3M sodium acetate pH5 (NaAc) was added, mixed in by inverting and kept on ice 15min. The tubes were then centrifuged at full speed for 5min. and the supernatant transferred to a fresh tube containing 400µl of isopropanol. This was placed at -20°C for 30min. The tubes were centrifuged for 3min. and the supernatant discarded. The pellet was resuspended in 100µl of 1 X TNE (see appendix 1) and then 120µl of isopropanol was mixed in by inversion. The tubes were centrifuged and the pellets were again washed with 100µl 1 X TNE and precipitated with isopropanol. The DNA was pelleted by centrifugation for 3min, the supernatant discarded and the pellet vacuum dried and resuspended in 50µl of SDW. This DNA was subsequentley used for restriction digestion (see below).

2.2.2 Large Scale Isolation of DNA

Primarily two methods of isolating large amounts of pure DNA were used. The first method imployed a large scale alkaline-lysis followed by caesium chloride (CsCl) density centrifugation to purify the DNA. The second method was the commercially available Qiagen Tip 500 kit. The methods are described in more detail below.

2.2.2a Large Scale Alkaline Lysis-CsCl density gradient centrifugation

500ml aliquots of o/n culture were harvested and the pellets were resuspended in 25ml of lysis buffer (see appendix 1). 50ml of alkaline SDS was added, mixed vigorously and the samples kept on ice for 10min. 37.5ml of NaAc was then mixed in and kept on ice for 30min. The cell debris was

pelleted for 15min. at 10,000 rpm in a Beckman JA-10 rotor, and the supernatant was filtered into fresh containers. 100ml of isopropanol was added to the filtered supernatant and the DNA was precipitated at -20°C for 1h and then spun down at 10,000 rpm. in a Beckman JA-20 rotor for 15min. The pellet was resuspended in 20ml 1 X TNE, 24ml isopropanol was added and placed at -20°C for 1h. The DNA was pelleted by centrifugation at 10,000 rpm for 10min. in a Beckman JA-20 rotor, and the pellet resuspended in 4ml of 1 X TNE ready for purification on CsCl gradients.

The DNA solution was added to 4.62g of CsCl and 0.5ml of a 10mg/ml solution of ethidium bromide (EtBr). This was transferred to 3.9ml Beckman tubes which were subsequently sealed. Centrifugation was carried out in a Beckman TLN 100 rotor at 100,000 rpm. for at least 3h but usually o/n. The DNA was visualised under ultraviolet light and the plasmid bands drawn out using a needle and syringe. This fraction was extracted several times with CsCl-saturated isopropanol to remove the ethidium bromide. The DNA containing fraction was then divided into 400µl aliquots and to each aliquot was added 500µl SDW, 100µl NaAc and 530µl isopropanol. These were shaken and the DNA pelleted by centrifugation The pellet was resuspended in 200µl 1 X TNE, 25µl NaAc and 500µl ethanol and placed at -20°C for 30min, spun down, dried and resuspended in 100µl SDW. This method was rotuinely used to purify cosmid DNA in large quantities as other methods were found to give low yields of DNA or pure quality DNA.

2.2.2b Qiagen Tip 500

The Qiagen plasmid purification procedure is based on a modified alkaline lysis procedure, followed by binding of the plasmid DNA to an anion-exchange resin under appropriate conditions. RNA, protein and low molecular weight impurities are removed by a medium salt wash. Plasmid DNA is eluted in a high salt buffer, and concentrated and desalted by isopropanol precipitation.

150 - 500ml of cells harbouring the plasmid were used to isolate the plasmid according to manufacturers instructions.

2.2.3 Isolation of *R.sphaeroides* genomic DNA

The isolation of genomic DNA from *R.sphaeroides* strain was carried out by a modification of the method of Giuliano and co-workers (Giuliano *et al.*, 1988). Essentially, 1ml of a stationary phase *R.sphaeroides* strain was harvested and frozen at -80°C. The frozen pellet was then thawed by the addition of pre-warmed (65°C) chromoprep buffer (see appendix 1) and the solution incubated at 65°C for a further 15 min. 100µg of proteinase K (Sigma) was added and the solution incubated at ambient temperature for 15 min. Following two extractions with phenol:chloroform, the DNA was precipitated with 1 volume of cold (-20°C) ethanol and the DNA pelleted by centrifugation. After rinsing of the DNA with 70% ethanol (-20°C) the pellet was dried and resuspended in SDW.

2.2.4 Agarose Gel Electrophoresis

Electrophoresis was performed in 0.5 X TBE gels containing 0.10µg/ml of EtBr unless stated otherwise, with varying percentages of agarose depending upon the size of the DNA of interest. Gels were run at a constant voltage of 100V. The DNA was visualised under ultraviolet light and photographed through an orange filter. Size markers were either Lambda DNA digested with *Bst*E II or pUC19 digested with *Hae* III depending on the size of the fragments of interest.

2.2.5 Isolation of DNA Restriction Fragments

The digested DNA was electrophoresed on an agarose gel. The band of interest was excised and the DNA was separated from the agarose using the Geneclean kit (Bio101 Inc.) as described by the manufacturer.

2.2.6 Restriction Digestion of DNA

Restriction enzymes and buffers were obtained from several commercial sources, mainly Northumbria Biologicals (NBL), Gibco (BRL), Promega and New England Biolabs (NEB). Analytical digests were carried out in 20µl volumes with 2.5µl of small scale isolated DNA and 5 units of enzyme in appropriate buffer conditions. The reaction tubes were incubated at the appropriate temperatures (usually 37°C) for 2h.

2.2.7 Ligations

0.5-5µg of digested DNA of the plasmids to be ligated were resuspended in 1 X ligation buffer (NEB) and mixed in the appropriate combinations such that the final volume was 25µl. 1 unit of T4 DNA ligase (NEB) was added and the ligation left to occur at 16°C overnight. 10µl of this was used in the transformation of bacteria.

2.2.8 Transformations

E.coli strains were made competent by an adaptation of the procedure of Meyer *et al.*, (Meyer *et al.*, 1977). 0.5ml of o/n culture was added to 50ml of LB and grown up to OD_{600} of 0.5. The cells were harvested, resuspended in 2ml ice-cold 100mM $CaCl_2$, left on ice for 30min., centrifuged again, resuspended in 0.5ml of 100mM $CaCl_2$ and kept on ice for at least 1h prior to use.

For transformation 100µl of competent cells were added to 0.05-0.3µg of DNA, kept on ice 1h and then heat-shocked at 42°C for 2min. and made up to 1ml with LB and incubated at 37°C 1h. 0.1 ml aliquots were spread on selective medium.

2.2.9 Modification of Restriction Ends

In many constructions, when sticky end ligations could not be carried out due to incompatible ends, plasmids or fragments with sticky ends had to be blunt ended. Two methods were used as described (Maniatas *et al.*, 1989).

2.2.9a Filling in 3' recessed ends using Klenow enzyme

2.5µl of 10 X NEB buffer 3 (NEB), 1µl of 2mM dATP ,dGTP, dCTP, dTTp and 0.25µl Klenow enzyme (BRL) were added to 20µl of DNA solution. The mixture was made up to 25µl with SDW and incubated at room temperature for 30min. After heat inactivation of enzyme at 65^{o}C for 15min, the DNA was precipitated with NaAc and isopropanol, spun down and dried. The DNA was then ready for blunt-ended ligation

2.2.9b Removal of "sticky" ends using Mung Bean Nuclease (MBN)

5µl of 10 X Mung bean nuclease buffer reaction buffer (New England Biolabs) and 0.25µl of MBN enzyme (2.5 units) were added to 35µl of DNA solution. The reaction was incubated at 37°C for 30min. The enzyme was then heat inactivated for 15min. at 65°C. The DNA was precipitated with 1/10 vol 3M NaAc pH 5.0 and 1.2 vol isopropanol, spun down and dried.

2.2.10 DNA Sequencing

DNA sequencing was carried out using the chain termination method of Sanger and co-workers (Sanger *et al.*, 1977). The commercially available Sequenase kit (United States Biochemical Corporation) was used. Essentially the technique involves the *in vitro* synthesis of a DNA strand by a modified version of T7 DNA polymerase. Synthesis is primed only at the site of annealing of a synthetic primer on the template. The synthesis reaction is terminated by the incorporation of 2',3'-dideoxynucleoside 5'-triphosphates (ddNTPs) which lack the 3'-hydroxyl group essential for strand elongation. When a mixture of the four dNTPs and a ddNTP is used, a population of chains

terminated at every site where the ddNTP can be incorporated will be generated. Four separate reactions each with a different ddNTP are carried out and so complete sequence data is obtained.

Due to the high guanine/cytosine (G/C) ratio in the *R.sphaeroides* genome, sequencing artefacts were a serious problem. It was found that the use of 7-deaza-dGTP and dimethyl-sulfoxide (DMSO) eliminated many of the problems associated with sequencing DNA with a high G/C content. With these modifications, typically 250 bp could be obtained per reaction.

Primers used throughout were either M13 universal (Gibco), M13 reverse (Gibco) or custom made primers which were synthesised by Cruachem Ltd. Typically 18 mers were synthesised and used in sequencing.

2.2.10a Double-stranded DNA template denaturation/primer annealing

Approx. 3µg of DNA was denatured at 42°C by incubation with 1µl of 1M NaOH and 1µl (0.8 - 25 pmol) of primer for 5 min. The NaOH was neutralised by the addition of 1µl of 1M HCl. The addition of 2µl of Sequenase reaction buffer (see appendix 1) and 0.5µl of DMSO (SIGMA), followed by incubation at 42°C for 5 min. allowed primer-template annealing to proceed.

2.2.10b Labelling and termination reactions

1 µl of 0.1M DTT was added to the annealing reaction, followed by 2µl of 1:5 diluted labelling mix (diluted in SDW) (see appendix 1), 0.5 µl of S^{35}dATP, 1µl DMSO and 2µl of 1:8 diluted Sequenase enzyme (diluted in enzyme dilution buffer). The labelling was allowed to proceed at 20°C for 5 min. Subsequently, 3.5µl of the labelling reaction was aliquoted into four tubes containing 2.5µl of the termination mixes (see appendix 1) and 0.28µl of DMSO. Termination was allowed to proceed for 5 min. at 42°C followed by the addition of 4µl of Stop buffer to each tube. Reactions were stored at -80°C until required.

2.2.10c Electrophoresis of sequencing reactions

6 % polyacrylamide gels were used for the resolution of the sequencing products. Gels were prepared using the Sequagel sequencing system (Flowgen). Prior to loading the sequencing products, the gels were run for 30 min. at 50 Watts in 1 X TBE (see appendix 1). Samples were boiled for 3 min. before loading and electrophoresed for a total of 4 h, with a second loading being performed 2 h. after the first to allow overlapping sequence to be obtained. Once completed, the gels were fixed in 10%(v/v) methanol/ 10% (v/v) acetic acid for 30 min. dried under vacuum onto filter paper and exposed to autoradiography film.

2.2.11 Computer analysis of sequencing data

Analysis of DNA sequence was carried out using the GENEJOCKEY sequence analysis program (Biosoft, Cambridge) and the University of Wisconsin genetics computer group series of programs (GCG) (Deveraux *et al.*, 1984) held at Daresbury, UK. The specific programs within the GCG packages will be described in more detail in the relevant chapters.

2.2.12 Polymerase Chain Reaction

The polymerase chain reaction enables the amplification of specific regions of DNA *in vitro*. Two convergent primers are used to prime DNA strand synthesis on a template (Mullis *et al.*, 1986). The thermostable *Taq* polymerase from *Thermus aquaticus* is used for chain elongation. Amplification requires the cyclic incubation of template, primers, dNTPs and enzyme at high temperatures (90-96°C) to denature the template, then at 30-60°C (depending on primer) to enable the primers to anneal to the denatured DNA and finally at 68-72°C for optimal strand elongation by *Taq* polymerase. After a number of cycles the predominant DNA species in the reaction tube will be the region between the primers.

Typically reactions were set up as follows: 0.01-0.1µg of template DNA was mixed with 50pmol of each primer, 5µl of 10 X *Taq* buffer (Boehringer Mannheim), 8µl of 1.25 mM dNTP's, 6µl of 20mM $MgCl_2$, 5µl DMSO, 2.5 units of *Taq* polymerase (Boehringer Mannheim) and SDW up to 50µl. The samples were overlaid with mineral oil and the DNA was denatured for 5 min. at 96°C followed by 25 cycles of 96°C 1min., 55°C 1min. and 72°C 3min. This was followed by a final 72°C step for 5 min. Cycling was carried out using a Techne Progene thermal cycler with the ramping value set to maximum.

2.2.13 Conjugative transfer of Plamids

The transfer of pARO191(Park, 1990), pRK415-1 (Keen *et al.*, 1988), pLA2917 (Allen & Hanson, 1985) and pSUP202 (Simon *et al.*, 1983) based clones into *R.sphaeroides* strains was carried out via diparental mating using *E.coli* S17-1 as the donor strain (Simon *et al.*, 1983) essentially as described elsewhere (Moore & Kaplan, 1989). 1 ml of aerobically grown *E.coli* donor strains, and *R.sphaeroides* recipient strains were harvested by centrifugation at 6000 rpm. The pellets were washed free of residual antibiotics with LB and resuspended in 100 µl of LB and mixed together. 20µl aliquots were then spotted onto sterile nylon filters (Life Technologies) which had been placed onto the surface of an LB agar plate. The filters were then incubated at 30°C for 6 hours, and the bacteria subsequently removed into 1 ml of LB. The filters were then mixed vigorously and aliquots of the cell suspension plated onto selective LB agar plates.

2.2.14 Southern Blotting

Southern blot analysis (Southern, 1975) was carried out using DNA immobilised on Nylon membrane probed with DNA probes labelled with biotin-14-dATP. The hybridisation was visualised using a streptavidin - alkaline

phosphatase conjugate followed by incubation with the photogenic substrate 4-methoxy-4-(3-phosphate phenyl)-spiro[1,2-dioxetane-3,2'-adamantane] (PPD).

Restriction digested DNA was electrophoresed on agarose gels followed by 30 min. - 1 hour incubation in Southern denaturing solution (see appendix 1). Gels were subsequently incubated for 30min. - 1 hr. in Southern neutralisation solution (see appendix 1) and the DNA transferred o/n by capillary action to Photogene Nylon membrane (Life Technologies) (Maniatis *et al.*, 1989). Following transfer, the nylon membranes were washed for 5min. in 2 X SSC (see appendix 1) and subsequently baked for 2 hours at 80°C under vacuum.

Probes were synthesised using restriction fragments isolated by the Geneclean method described above and labelled using the Bionick (Life Technologies) labelling system. Briefly, the DNA (in 40µl of SDW) was mixed with 5µl of dNTP solution and 5µl of Enzyme solution. This was incubated for 1 hour at 16°C followed by mixing with 5µl of Bionick stop solution.

Hybridisation was carried out at 65°C following 3 h. prehybridisation in hybridisation solution (see appendix 1). The probe was mixed with 1ml of hybridisation solution, boiled for 5 min., and added to the Nylon membranes. Post-hybridisation washes were as: 2 X 5min. 65°C 5 X SSC, 0.5% (w/v) SDS; 30min. 0.1 X SSC, 1%(w/v) SDS usually at 55°C but temperature depended on stringency required; 1min. TBS-Tween 20 (see appendix 1). Membranes were blocked for 1 h. in Southern blocking solution at 65°C, followed by a 10min. incubation with a 1:1000 dilution of the streptavidin-alkalinephosphatase conjugate (Life Technologies). This was followed by 2 X 15 min. washes in TBS-Tween 20, and 1 hour in 1 X Final wash solution (see appendix 1). Membranes were then removed and incubated with PPD (Life Technologie) for 1-3 h. at room temperature followed by exposure to autoradiography film to visualise hybridising bands.

2.3 Protein Analysis Techniques

2.3.1 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The method of Laemmli (Laemmli, 1970) was used to analyse proteins via separation of polyacrylamide gels. 12 % gels were used unless otherwise stated and were prepared using the Hoeffer minigel system as follows:

The minigel gel casting apparatus was cleaned and assembled and a resolving gel poured. The resolving gel consisted of 10ml 30% Easi Gel acrylamide mix (Scotlab), 6.3 ml 1.5M Tris (pH8.8), 0.25 ml 10% SDS, 0.25 ml of 10% (w/v) ammonium persulphate (Sigma) , SDW to bring the volume up to 25 ml and 10µl of TEMED (Sigma). The resolving gel was overlaid with water saturated butanol and allowed to set for 30 min. Following this, the casting apparatus was dissembled, the gels washed with SDW, the combs inserted and the stacking gel poured. The stacking gel consisted of 0.83ml of 30 % Easi Gel acrylamide mix (Scotlab), 0.63ml of 1M Tris (pH6.8), 50µl of 10% SDS, 50µl of 10% (w/v) ammonium persulphate, SDW to make the volume up to 5ml and 5µl TEMED. After 30 min. the combs were removed, the upper and lower reservoirs filled with 1 X Tris-glycine electrophoresis buffer and the wells washed with 1 X Tris-glycine electrophoresis buffer (see appendix 1).

Samples were typically prepared using 2 X SDS gel loading buffer to give a final of 1 X, boiled for 5 min. and loaded onto the gel, along with a suitable marker (usually 10KDa ladder (Gibco BRL) or Rainbow markers (Amersham)). The gels were run at a constant 30 mA until the dye front reached the bottom and the gels removed from their glass plates and either stained in Coomassie blue stain (see appendix 1) or used in Western blotting.

2.3.2 Glutathione-S-transferase Fusion Protein Over-expression and Purification

The FliF protein from *R.sphaeroides* was overexpressed and purified using the glutathione S-transferase (GST) gene fusion system (Pharmacia Biotech) first introduced by Smith and Johnson (Smith & Johnson, 1988). The gene of interest is first cloned into an expression vector, downstream of the gene encoding *Shistosoma japonicum* GST. The GST-gene fusion is under the control of the strong, inducible expression promoter Ptac hence, after induction of cells harbouring the GST-gene fusion by IPTG the fusion protein is expressed at high levels. The fusion protein is then purified from the lysate via the high affinity of GST for glutathione sepharose 4B. The GST protein can subsequently removed from the fusion by cleavage by a site specific protease.

2.3.2a Cloning of the *fliF* gene into the GST fusion vector

The 2.2 Kb *Bam*HI fragment containing the *fliF* gene (see chapter 5) was cloned into the *Bam*HI site of pGEX3X (Pharmacia) in the correct orientation to allow the fusion of GST and FliF. The junction between the fusion was subsequently sequenced to confirm the correct reading frame.

2.3.2b Small scale expression of fusion protein

E.coli DH5α cells harbouring the fusion construct (pGEX-FliF) were grown o/n at 37°C in 2 X YT containing 100μg/ml of ampicillin. The following day, 2 X 3ml cultures of 2 X YT, 100μg/ml of ampicillin were inoculated with 100μl of the o/n starter culture and grown for 1 h. at 37°C. 100 mM IPTG (Pharmacia) was subsequently added to one of the two cultures and the incubation continued for a further 3 h. Cells were harvested and resuspended in 300μl 1 X PBS (see appendix 1). 10μl was removed as a whole cell sample and the remaining sample was sonicated using an probe sonicator at full power for 30 sec in 10 sec bursts. The bacterial lysates were cleared by centrifugation at 15,000 rpm for 10 min. in a Beckman JA-20 rotor and the supernatants removed. The subsequent pellet was resuspended in 300μl of 1 X PBS and the samples (whole cell, soluble and insoluble) analysed by 8% SDS PAGE.

Subsequent analysis of the samples revealed that all of the FliF fusion protein was insoluble. Growth at 30°C, shorter induction time, lower levels of IPTG, induction at a higher cell density have all been known to increase the solubility of certain GST-fusion proteins (Pharmacia) but did not increase the solubility of the GST-FliF fusion protein. The use of detergents to solubilise the GST-FliF fusion protein was therefore examined.

2.3.2c Solubilisation of the GST-FliF fusion protein by sarkosyl treatment

Frangioni and Neel (Frangioni & Neel, 1993) first reported the use of the anionic detergent sarkosyl for solubilisation of GST-fusion proteins. Briefly, the method involves lysozyme treatment of bacteria expressing the fusion protein, followed by sonication in the presence of sarkosyl which renders the fusion protein soluble. The sarkosyl is then removed by the addition of Triton X-100 which causes the sequestering of the sarkosyl into mixed detergent micelles with the Triton X-100, allowing the binding of the GST-fusion protein to glutathione sepharose.

2.3.2d Small scale purification of GST-FliF fusion protein

Two 10ml o/n starter cultures were used to inoculate 2 X 90ml of 2 X YT, 100µg/ml of ampicillin. Cultures were subsequently incubated for 1 h. at 37°C followed by induction for 3 h. using 100mM IPTG. The cells were then harvested and resuspended in 2ml of cold STE (see appendix 1). 20µl of a 10mg/ml lysozyme solution and 15µl of 1M DTT was added, the samples mixed and incubated on ice for 15 min. Sarkosyl was then added to give a final concentration of 0.7%. The samples were subsequently sonicated for 30 seconds in 10 seconds bursts and the samples incubated on ice for 15 min. Following this the sonicates were cleared by centrifugation at 10,000 rpm in a Beckman JA-20 rotor for 30 min. Triton X-100 was then added to a final concentration of 2% and 50µl of a 50% glutathione sepharose 4B slurry (Pharmacia) added to 400µl aliquots of sonicate. The sonicate was then

incubated at room temperature for 30 min. with constant mixing followed by centrifugation at 2,000 X g for 5 min. to pellet the sepharose. Non-specifically bound proteins were washed off the sepharose using 1 X PBS, usually 8 X 1ml washes were sufficient to remove most non-specific proteins. The purified FliF-GST fusion protein was eluted using 100 µl of 20mM reduced Glutathione (Boehringer Manhein), 100 mM Tris pH8.8 o/n at room temperature.

2.3.2e Large scale purification of GST-FliF fusion protein

A 150 ml o/n starter culture was used to inoculate 1350 ml of 2 X YT containing 100mg/ml of ampicillin. After incubation for 1h at 37°C, the cells were induced using 1mM IPTG for 3h. The cultures were harvested and resuspended in 10ml of STE followed by the addition of 100µl of 10mg/ml of lysozyme and 50µl of 1M DTT. Following incubation on ice for 15min, sarkosyl was added to a final concentration of 0.7% and the cells sonicated for 1min. in 15 sec bursts. The supernatants were cleared by centrifugation and 8ml of 10% sarkosyl in STE and 1ml of 1 X PBS added.Triton was added to a final concentration of 2%, the supernatants incubated at room temperature for 30 min. and 2 ml of glutathione sepharose added. Incubation was continued at room temperature for 30 min. Following this, the sepharose was pelleted by centrifugation at 2,000 X g 5 min. Non-specifically bound proteins were washed away with 1 X PBS, typically 10 X 15ml washes was sufficient. The purified GST-FliF fusion protein was eluted in 3ml of elution buffer (see above) o/n at room temperature and repeated again using a further 3ml for 5 hours. The protein was stored at -80°C until required.

2.3.3 Histidine Tagged Fusion Protein Overexpression and Purification

The purification of histidine tagged proteins was carried out using the Qiaexpressionist system from Qiagen. This method relies on the high affinity of 6 histidine residues for nickel agarose (Ni-NTA). The protein of interest is

placed under the control of a high level expression promoter (in this case the T5 phage promoter was used) with a region encoding 6 histidines at the C or N terminus. The high affinity of the 6 his domain for the Ni-NTA resin allows the purification of fusion proteins under very high stringency and also under denaturing conditions.

2.3.3a Cloning of the *fliG* gene into the his tag protein overexpression vector

The *R.sphaeroides fliG* gene (see chapter 5, Fig 5.21) was amplified by PCR using the primers FliG - F (GAA GAT CTA CCA CAG CAG CCG CCA CC), which engineered a *Bgl*II site at the beginning of *fliG* as well as removing the initial methionine residue, and the pUC forward primer (CCC AGT CAC GAC GTT GTA AAC G) using the clone p5.5 as the template DNA. The conditions used have been described above. The subsequent PCR product was digested with *Bgl*II and *Xma*I and cloned into pQE 30 (Qiagen) digested with *Bam*HI and *Xma*I. The construct was confirmed by DNA sequencing.

2.3.3b Small scale purification of the FliG-6His protein

A 10ml o/n starter culture of cells containing the overexpressing construct were used to inoculate 30ml of 2 X YT containing 100µg/ml of ampicillin. The culture was incubated for 30 min. at 37°C and IPTG added to a final concentration of 1mM. After 2 h. of incubation at 37°C the cells were harvested and resuspended in 2ml of 1 X binding buffer (ice cold) (see appendix 1). Samples were then sonicated using a probe sonicator on full power for 3 X 15 sec bursts. The supernatants were cleared by centrifugation and the pellet resuspended in 1 X binding buffer.

Subsequent analysis of the samples prepared above by SDS-PAGE revealed that the sonication process released very little intracellular protein and the pellet contained a large proportion of the fusion protein. As a first attempt, the fusion was treated as insoluble to discover if it would purify on Ni-NTA resin. The pellet was therefore solubilised in 1.5ml of 1 X binding buffer

containing 6 M urea by incubation for one hour on ice with occasional vigorous mixing. The unsolubilised protein was removed by centrifugation and the solubilised protein loaded onto a 1ml bed-volume Ni-NTA resin column (Novogen), which had been prepared by washing with 3 column volumes of SDW, 5 volumes of 1 X charge buffer (see appendix 1), 3 volumes of 1 X binding buffer and 3 volumes of 1 X binding buffer containing 6M urea. After the supernatant had passed through the column, the resin was washed with 10 volumes of 1 X binding buffer containing 6M urea, followed by 6 volumes of 40mM imidazole buffer (see appendix 1). The purified protein was eluted from the column with 3 volumes of 1 X elution buffer containing 6M urea.

2.3.3c Large scale isolation of the FliG-6His protein

A 200ml o/n starter culture was used to inoculate 200ml of LB containing 100μg/ml of ampicillin. This was incubated at 37°C for 30min. prior to induction using 1mM IPTG. Induction was allowed to proceed for 2h after which the cells were harvested and burst using a French Press at 2000 psi. The solution was cleared by centrifugation and the supernatant used in column chromatography to purify any soluble FliG using a 2ml bed volume column essentially as described above in section 2.3.3b The FliG-6His protein was purified in batches using 5ml of supernatant per batch.

The majority of FliG was found to reside in the insoluble fraction and was subsequently extracted by the resuspension the insoluble protein in 8ml of 1 X binding buffer containing 6M urea, followed by vigorous mixing and incubation on ice for 1 hr. The solution was cleared by centrifugation and the solubilised protein purified in 4 ml batches using a 2ml bed volume Ni-NTA column essentially as described above in section 2.3.3b Subsequent analysis of the unsolubilised protein revealed that much of the FliG remained insoluble and the insoluble protein pellet was extracted a further 3 times using the method described, to purify additional FliG.

2.3.4 Preparation of Antigens for Antibody Production

For the preparation of anti-FliF antibody, approximately 200µg of FliF-GST fusion protein was electrophoresed on a preparative SDS-PAGE gel (as described above), the gel stained for 15 min. in Coomassie stain (see appendix 1) and the protein band excised. The gel slice containing the protein was subsequently frozen in liquid nitrogen and the gel slice ground using a sterile mortar and pestle until a fine powder was produced. The powder was resuspended in 250µl of SDW, mixed with Freunds adjuvant and injected subcutaneously into a New Zealand White rabbit. Booster injections of 100µg of FliF-GST fusion protein were administered every 28 days after the primary injection and blood samples taken before each booster.

2.3.5 Antibody Purification

The purification of antibodies from crude serum, and the removal of non-specific antibodies was used to purify antibodies directed against the GST-FliF fusion product. Several methods of antibody purification were used throughout the course of this project which are all as described previously (Maniatas *et al.*, 1989):

2.3.5a Purification of antibody using immobilised antigen

This method relies on the immobilisation of purified antigen onto nitrocellulose followed by the binding of antibodies to the antigen. The bound antibodies are then eluted from the antigen and are used for further analysis.

100µg of purified GST-FliF fusion protein was electrophoresed by SDS-PAGE on a 6% gel. Following electrophoresis the protein was transferred onto Hybond Super-C nitrocellulose (Amersham) as described for Western blot analysis. A 1cm strip was then cut from the membrane and stained with amido black to visualise the immobilised antigen. This was subsequently aligned with the original membrane and the strip of membrane containing the unstained

antigen excised and blocked in 1% (w/v) BSA in 1 X PBS o/n. The membrane was subsequently washed for 20 min. in 150 mM NaCl and again in 1 X PBS. Following several more washes in 1 X PBS, the bound antibody was eluted in 0.2 M Glycine pH 2.8, 1mM EGTA (500 µl) for 30 min. at room temp. The eluted antibody was removed and added to 50µl of 1M Tris base to neutralise the effects of the glycine, and 10 X PBS added to a final of 1 X. The purified antibody was subsequently used for Western blot analysis.

2.3.5b Serum pre-adsorption

The removal of non-specific antibodies by pre-adsorption on bacterial cell extracts is a commonly used method of purifying antibodies to satisfactory specificity for most applications. Basically, an acetone extract of whole cell proteins from a mutant lacking the antigen are prepared and added to the crude serum. This causes the non-specific antibodies to complex with the acetone extracted proteins and these are subsequently removed by centrifugation.

An acetone extract of the mutant Nm7 (see chapter 5) was prepared by sonication of whole cells followed by the addition of 4 volumes of cold acetone. The proteins were allowed to precipitate o/n on ice, followed by centrifugation to pellet the precipitated protein. The protein pellet was resuspended in 1ml of cold acetone and centrifuged again. The acetone was evaporated by vacuum drying for 20 minutes and the acetone extracted protein stored at -20°C until needed.

The acetone extract was added to the crude serum to a final concentration of 1% (w/v) followed by incubation on ice for 6h to o/n. The antigen-antibody complex was removed by centrifugation and the supernatant used as the pre-adsorbed serum.

2.3.5c Ammonium sulphate precipitation - DEAE chromatography

The third method used to purify antibodies was ammonium sulphate precipitation followed by DEAE cellulose chromatography. Briefly, the antibodies are precipitated from solution using ammonium sulphate (50% w/v

saturation) and the remaining proteins in the precipitate removed, after extensive dialysis, by binding to DEAE cellulose at pH 6.5 which allows the antibodies to remain in solution.

2 ml of crude serum was mixed with saturated ammonium sulphate to give a final saturation of 25% (w/v), incubated on ice for 3 h. and the protein pelleted by centrifugation. This step was to remove any proteins that precipitate below 50% saturation. Ammonium sulphate was added to the remaining supernatant to a final saturation of 50% and the solution incubated on ice for 3h. The precipitated protein was pelleted by centrifugation, the pellets resuspended in 20mM sodium phosphate buffer (pH 6.5) (see appendix 1) and dialysed against 20mM sodium phosphate buffer (pH 6.5) o/n at 4°C.

DEAE cellulose (Sigma) was prepared by washing in 0.5N sodium hydroxide followed by 0.5N hydrochloric acid. The cellulose was subsequently washed extensively in 20mM sodium phosphate buffer (pH 6.5) until the pH reached 6.5. 1ml of prepared DEAE cellulose was mixed with 1ml of 50% ammonium persulphate precipitated serum which had been dialysed against 20mM sodium phosphate buffer and incubated at room temperature for 1 h. This was followed by centrifugation to pellet the cellulose and any associated proteins, the supernatant was used as purified antibody (1in 200 dilution) or stored at -80°C until required.

2.3.6 Western Immunoblot Analysis

Essentially, proteins are electrophoresed using SDS-PAGE and transferred to nitrocellulose membrane using horizontal transfer apparatus. After extensive blocking of non-specific binding sites, the membrane is incubated with the primary antibody, washed, and incubated with an enzyme coupled secondary antibody directed against the idiotype of the primary antibody. Alkaline-phosphatase conjugated antibody was used in this study and the presence visualised using the chromogenic substrates NBT and BCIP.

Following SDS-page the gel was removed and soaked in transbot buffer for 10 min. Tranfer of the proteins to Hybond-C super nitrocellulose (Amersham) was then carried out using a 2051 midget multiblot electrophoretic transfer unit (LKB) in transblot buffer for 1 h. at 100V. Following transfer, the filters were blocked o/n by incubation in 1% (w/v) non-fat dried milk in 1 X PBS-Tween 20. Filters were then washed for 5 min. in 0.3 % (w/v) non-fat dried milk in PBS-Tween 20 followed by incubation with primary antibody. Typically, anti-flagellin antibody (Sockett & Armitage, 1991) was diluted 1:1000 and anti-FliF antibody (see chapter 5) diluted 1:200 in 1% (w/v) non-fat dried milk, 1% (w/v) BSA in 1 X PBS-Tween 20 and incubated for 2 h. Following primary antibody binding, the filters were washed three times in 0.3% (w/v) non-fat dried milk in PBS-Tween 20 for 5 min. Filters were then incubated for 1.5 h. with the secondary antibody, anti-rabbit IgG alkaline phosphatase conjugate (Sigma), which had been diluted 1:1000 in 1% (w/v) non-fat dried milk in 1 X PBS-Tween 20. Two washes in 0.3% (w/v) non-fat dried milk in 1 X PBS-Tween 20 followed the secondary antibody binding, and was followed by 2 further washed with 1 X PBS and two washed with 1 X AP buffer. The antibody binding was then visualised by incubation with 2.5 mg of BCIP (Sigma) (dissolved in 100µl of dimethyl-formamide) and 5 mg of NBT (Sigma) (dissolved in 100µl of 70% (v/v) dimethyl-formamide) in 15ml of AP buffer (see appendix 1). Typically, 5 min. incubation was required for visualisation.

2.3.7 Cell Fractionation

Cellular fractionations were carried out essentially as described (Tai & Kaplan, 1985). Briefly, cells were sphaeroplasted using lysozyme to release the outermembrane and periplasm. The sphaeroplasts, osmotically stabilised in sucrose, are then lysed by a combination of osmotic and cold shock to release the cytoplasm and cytoplasmic membrane.

100 ml of log phase cells were harvested and resuspended in 5ml of 0.1M Tris (pH 8), 20% (w/v) sucrose at 37°C. This cell suspension was shaken at 200 rpm for 10 min. Following this the cells were sphaeroplated by the addition of 225 µl of a 2mg/ml lysozyme (Sigma) and 100µl of 0.5 M EDTA (pH 8). Sphaeroplast formation was monitored by removing two 1µl aliquots of the suspension, adding one to 20µl of Tris-sucrose solution and the other to 20 µl of SDW. When no cells were observed in the SDW treated sample by light microscopy (typically 30 min.) the sphaeroplasts were harvested at 10,000 rpm. for 10 min. The supernatant was removed and centrifuged at 50,000 rpm in a Beckman mini-ultra centrifuge in a TLA 100 rotor, the supernatant was removed as the periplasmic fraction and the pellet resuspended in 3ml of 0.1M Tris pH8.0 and kept as the outer membrane fraction. 5ml. of cold 0.1M Tris pH8.0 was added to the pelleted sphaeroplasts using a paintbrush to gently resuspend the pellet. $MgCl_2$ was then added to a final concentration of 1mM and 3µg of DNase added (Pharmacia). The suspension was incubated at 37°C for 15 min. to digest the chromosomal DNA. Unbroken cells were removed by centrifugation at 10,000 rpm for 10 min. and the supernatant centrifuged in an Beckman ultra centrifuge at 50,000 rpm for 1 h in a TLA 100 rotor. The supernatant was removed and kept as the cytoplasmic fraction and the pellet washed with 1ml of 0.1M Tris pH 8.0; 0.1% SDS. The washed pellet was subsequently centrifuged again at 50,000 rpm for 1 h., resuspended in 0.5 ml of 0.1M Tris pH 8.0; 0.1% SDS and kept as the cytoplasmic membrane fraction.

2.3.8 Isolation of flagellar filaments

The isolation of external flagellar filaments was achieved via the shearing of an aliquot of *R.sphaeroides* cell that had previously been harvested and washed in TE buffer. Essentially, the OD_{600} was measured for the strains to be analysed and the cells harvested and washed in TE buffer. The cells were

subsequently resuspended in a suitable volume of TE which gave equal cell numbers for the strains to be analysed and 100µl removed and kept for use as a whole cell preparation. The resuspended cells were subsequently sheared by 20 passages through a 25 gauge cannulum, followed by centrifugation to pellet the cells. The supernatant contained the purified flagellar filaments and was subsequently used in Western immunoblot analysis.

2.3.9 Estimation of protein concentrations

The concentration of proteins within samples was analysed using the Bio-Rad protein assay kit according to manufacturers instructions.

2.4 Cellular Studies

2.4.1 Analysis of Motility

Analysis of motility was by two methods; 1) phase contrast light microscopy of cells grown in constant illumination in succinate medium and 2) swarm plates (Sockett, 1986).

2.4.2 Chemotaxis Assay - Plug Plate Method

Chemotaxis was studied using a modified version of the plug plate method as described (Tso & Alder, 1974). Typically 240ml of photosynthetically grown until OD_{600} 1 - 1.5 , cells were harvested by centrifugation at 5,000 X g for 5 min. and gently resuspended in 60ml of 10 mM HEPES pH 7.0. 60ml of 0.7% agar in 10mM HEPES pH 7.0 was subsequently added and the suspension poured into petri dishes. After setting, holes were cut in the solidified suspension and plugs of varying concentration of chemoattractants, made in 10mM HEPES pH7.0 1.5% agar, were inserted into them. Typically sodium acetate, fructose, sodium pyruvate, sodium propionate and HEPES (negative control) were used at concentrations ranging from 50mM to 1mM. Plates were then incubated at room temperature with

constant illumination and the accumulation of cells around the plugs was examined for, typically after 2 h.

Chapter 3

Studies on the role of the alternative sigma factor σ^{54} on flagellar gene regulation in *R.sphaeroides* WS8.

3.1 Introduction

As mentioned in the introduction, the "alternative" sigma factor σ^{54} controls flagellar gene expression in various organisms (see Merrick, 1993 for a review). σ^{54} also controls the expression of other metabolic functions, almost all of which are non-essential to life of the cell (Merrick, 1993). Amongst others, the major metabolic functions that it controls in most organisms is nitrogen fixation and nitrate utilisation (Merrick, 1993), as a consequence, σ^{54} is often referred to as σ^{N}, with the gene encoding it being known as *rpoN* (the N standing for nitrogen/nitrate). Throughout this chapter I will use the σ^{54} nomenclature, although the 54 suffix is to denote the molecular weight in kilo Daltons of the sigma factor and many organisms have larger or smaller σ^{54} sigma factors but are still refereed to as σ^{54}. As a detailed description of sigma factors did not fit into chapter 1, I include such a discussion now in order to show the diverse functions of the protein and illustrate the evidence that suggest that they are involved in flagellar gene regulation.

The presence of a σ^{54} homologue has been demonstrated in over 18 genera and is required for the expression of a wide variety of genes involved in many metabolic functions including pilin synthesis, xylene catabolism, dicarboxylic acid transport and nitrogen fixation (see Merrick, 1993 for a review). One of the major differences between σ^{54} and the major vegetative sigma factor σ^{70} is the method of promoter recognition and gene activation. Unlike σ^{70}, σ^{54} does not bind to the typical -35, -10 promoter sequence, instead they recognise a relatively precise consensus sequence of TGGCAC-

N5-TTGC located between -26 and -11 (Morett & Buck, 1989). The GG and GC motifs underlined are almost always invariant and the spacing between the motifs is essential i.e. changes of only 1bp will cause inactivation of the promoter. σ^{54} also has the unique ability to recognise and bind to promoters in the absence of core RNA polymerase (Buck & Cannon, 1992; Cannon *et al.*, 1993) via a DNA binding domain at its C-terminus. σ^{54} requires the presence of an activator protein in order to catalyse the formation of open complexes and hence transcriptional activation when bound to RNA polymerase (Sasse-Dwight & Gralla, 1988; Popham *et al.*, 1989). The activator protein binds to enhancer like elements (ELE's) usually located 100 to 200 bp upstream from the promoters they regulate but these ELEs can be moved up to 1Kb away and still function (Kutsu *et al.*, 1989; Morett & Segovia, 1993). It is thought therefore that the interaction of the transcriptional activator with the promoter bound complex requires looping out of the intervening DNA so as to bring the activator close to the sigma factor to allow interaction (Kutsu *et al.*, 1989) and in some cases this DNA-bending is facilitated by integration host factor (IHF) or A/T rich DNA (see Perez-Martin *et al.*, 1994). A more detailed description of the activator proteins will be given in chapter 5 as during this project an activator was analysed.

σ^{54} has been implicated in the control of flagellar gene expression in *C.crescentus* (Anderson *et al.*, 1995) and *P.putida* (Kohler *et al.*, 1989). The activator protein involved in regulating σ^{54} dependant promoters in *C.crescentus* is FlbD and it has been shown to be regulated with respect to the cell cycle (Newton *et al.*, 1989; Brun *et al.*, 1994). As I shall show in chapter 5, during my PhD a homologue of FlbD was identified in *R.sphaeroides* WS8. A σ^{54} has been identified in another strain of *R.sphaeroides* known as HR but a specific role for it has not been determined (Meijer & Tabita, 1992). The presence of a σ^{54} consensus sequence upstream of the *motA/B* operon (see chapter 1, Fig 1.16) and also the presence of ELE's is was predicted that a

portion of the flagellar genes in *R.sphaeroides* may be under the control of σ^{54} (Shah & Sockett, 1995). This is further strengthened by the findings of chapter 5, whereby upstream of the operon encoding FliF is a consensus sequence that resembles the σ^{54} consensus sequence. The aim of this part of the project was to determine whether or not σ^{54} was in fact the sigma factor responsible for regulating flagellar gene expression in *R.sphaeroides* WS8 as the previous study in *R.sphaeroides* HR failed to obtain a phenotype for an *rpoN*$^-$ strain (Meijer & Tabita, 1992). The authors tested for a nitrogen fixation phenotype and not motility. The authors attempted to explain the lack of a phenotype by showing the presence of a second copy of *rpoN* in *R.sphaeroides* HR. As shall be demonstrated in this chapter, the presence of a second copy of *rpoN* in *R.sphaeroides* is questionable and the phenotype of an *rpoN*$^-$ *R.sphaeroides* WS8 strain is also presented.

3.2 Results

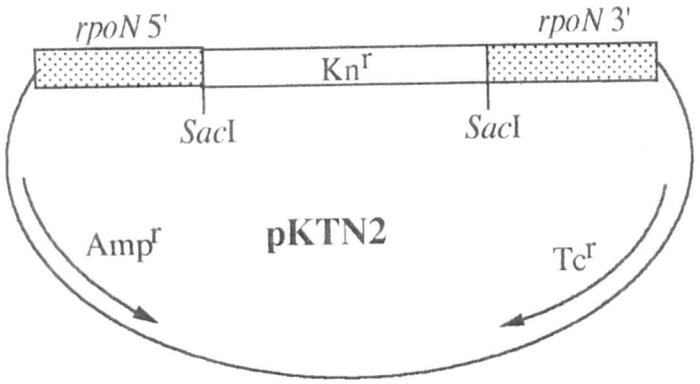
3.2.1 Gene replacement mutagenesis using *R.sphaeroides* HR *rpoN*.

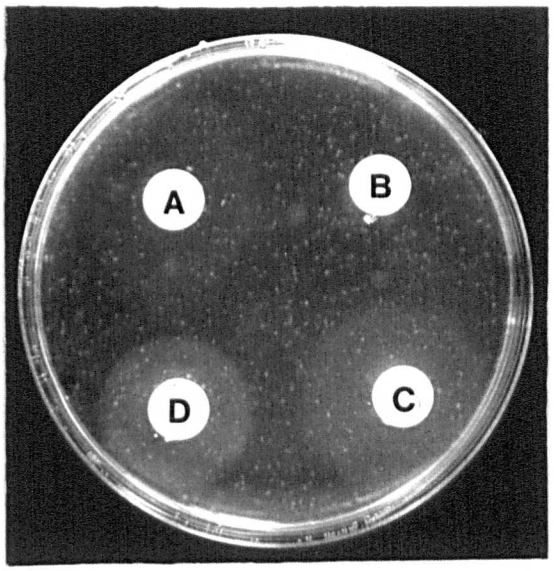
Initial attempts to create an *rpoN*$^-$ strain of *R.sphaeroides* WS8 centred around the use of the homologue from HR. The *rpoN* gene from *R.sphaeroides* HR was obtained from Meijer and Tabita already inactivated by the presence of a kanamycin resistance conferring cartridge at an internal site as demonstrated in Fig 3.1. The inactivated gene was carried on the suicide plasmid pSup202 (Simon *et al.*, 1983) and was called pKTN2 (Meijer & Tabita, 1992). The pKTN2 construct was introduced into *R.sphaeroides* WS8N by conjugation and exconjugants isolated. Since pSup202 cannot replicate in *R.sphaeroides*, Kn^r can only be obtained via the integration of the plasmid into the chromosome. In cases whereby an even number of cross-over events has occurred i.e. one on either side of the wild-type copy of the gene in the

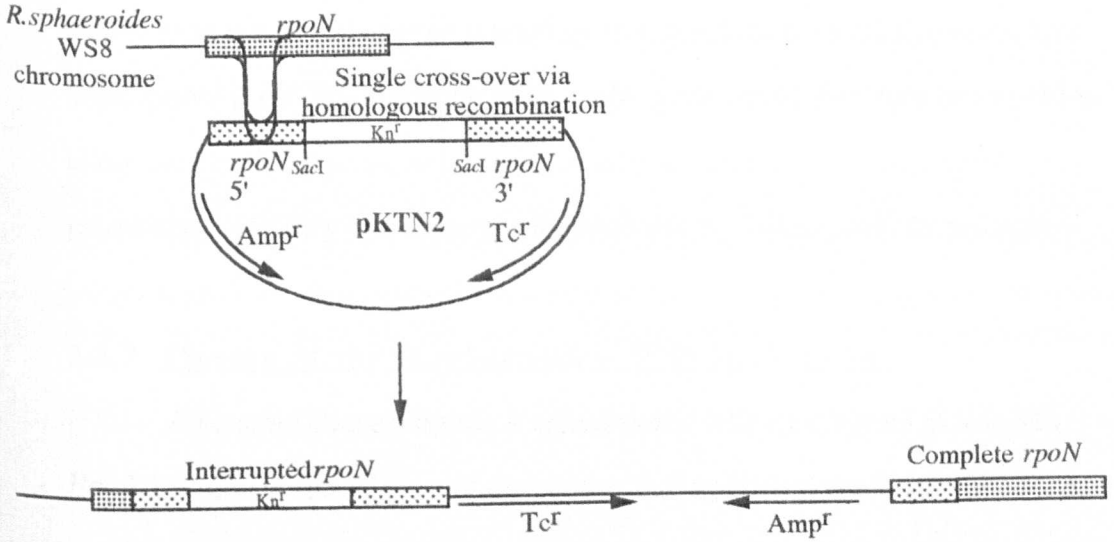
Fig. 3.1 Restriction map of pKTN2. The genes encoding resistance to tetracycline, ampicillin and kanamycin are shown as Tc^r, Amp^r and Kn^r respectively.

Fig 3.2 Chemotaxis 'plug plate' assay of WS8::*rpoN*X. 50mM of attractants have been used. (A) Sodium acetate, (B) HEPES (C) Sodium propionate and (D) Sodium pyruvate. Positive taxis is shown by the accumulation of cells around the 'plug' of attractant.

Fig 3.3 Possible cross over event leading to the integration of pKTN2 into the chromosome of *R.sphaeroides* WS8. The production of two copies of *rpoN* is demonstrated. For the purpose of simplicity, the 5' region is shown to be the region that combines although this may not be the case.


rpoN 5'
rpoN 3'
Kn^r
SacI
SacI
Amp^r
pKTN2
Tc^r


A
B
D
C


R. sphaeroides
WS8
chromosome
rpoN
Single cross-over via
homologous recombination
Kn^r
rpoN SacI
5'
SacI rpoN
3'
Amp^r
pKTN2
Tc^r
Interrupted rpoN
Kn^r
Complete rpoN
Tc^r
Amp^r

chromosome, integration results in the replacement of the wild-type copy with the inactivated copy, loss of the plasmid resulting in tetracycline sensitivity.. Screening of over 5,000 independently isolated exconjugants for tetracycline sensitivity failed to reveal the presence of any *R.sphaeroides* WS8 exconjugants that had lost the plasmid. Consequently, all of the exconjugants isolated were single cross-over mutants that possessed an integrated copy of pSup202 and were referred to as WS8::*rpoN*X.

3.2.2. Characterisation of the *rpoN* single cross-over mutants, WS8::*rpoN*X.

WS8::*rpoN*X mutants found to be motile over there entire life cycle, showing identical characteristics to wild-type cells. To determine if the chemotactic response to various compounds (detailed in Fig 3.2) was affected a chemotaxis assay was carried out using the plug plate method. As shown in Fig 3.2, the chemotactic response was normal, identical to wild-type WS8. This lack of a motility phenotype may have been due to the presence of two copies of *rpoN* within the genome of WS8 as can occur during the integration process (Fig 3.3). This may have resulted in an inactivated copy and a normal copy of *rpoN* being produced as demonstrated in Fig 3.3. This lack of ability to isolate a double cross-over mutant was probably due to differences in the DNA sequence within *rpoN* so that only one portion of the *R.sphaeroides* HR *rpoN* gene possessed enough homology to recombine with the *R.sphaeroides* WS8 *rpoN* gene. The authentic WS8 *rpoN* gene would therefore be needed to allow the knock out strain to be constructed and as a consequence it was necessary to identify and clone the *R.sphaeroides* WS8 *rpoN* homologue.

3.2.3 Cloning of the *R.sphaeroides* WS8 *rpoN* gene.

A plasmid containing the *R.sphaeroides* HR *rpoN* gene as a 1.2Kb *Bam*HI fragment (pSNT4), was digested with *Bam*HI and the fragment used to

probe a genomic DNA from *R.sphaeroides* WS8 under low stringency as shown in Fig 3.4. The probe strongly hybridised to a *Bam*HI fragment in *R.sphaeroides* . WS8 genomic DNA of approximately 1Kb in size (Fig 3.4, lane 3) and also to a very large (>14.14Kb) *Eco*RI fragment (Fig 3.4, lane 2). It was also found to hybridise weakly to a 3.5Kb *Bam*HI fragment in genomic DNA as well as the 0.702Kb fragment from a λ *Bst*EII digest. These poorly hybridising bands may represent other DNA-binding proteins. In order to clone the *rpoN* gene from *R.sphaeroides* a cosmid library of *R.sphaeroides* genomic DNA in the conjugative vector pLA2917, (Sockett & Armitage, 1991) was probed with the pSNT4 probe. The probe was found to hybridise to two cosmid clones known as cosmid 146 and cosmid 360. As shown in Fig 3.5, the hybridising *Bam*HI fragments (Fig. 3.5, lanes 2 & 3) were of the same size of that from WS8 genomic DNA (Fig 3.4, lane 3).

In order to determine if two *rpoN* genes were present or that the cosmids overlapped, these cosmids were isolated and digested with restriction enzymes and probed once more with the pSNT4 probe (Fig. 3.6). The *R.sphaeroides* HR *rpoN* gene probe was again found to hybridise to *Bam*HI fragments of approximately 1Kb (Fig 3.6, lanes 3 & 7). It was also found to hybridise to *Eco*RI fragments of 11Kb (Fig 3.6, lanes 1 & 5) and *Pst*I fragments of 3.5Kb (Fig. 3.6, lanes 2 & 6). The *Hin*dIII digest of the cosmids showed the only difference in the hybridising pattern with cosmid 146 giving a fragment of approximately 9.5 Kb (Fig 3.5, lane 4) whereas in cosmid the hybridising fragment was >14.14 Kb in size. This suggested that cosmids 146 and 360 did overlap, but that cosmid 146 contained an additional *Hin*dIII site, possibly from the cosmid vector arm and possibly covered a different region of the genome.

3.2.4 Attempts to construct an *rpoN*⁻ *R.sphaeroides* WS8 strain.

In order to construct an *rpoN*$^-$ strain of *R.sphaeroides* WS8 it was necessary to clone the *rpoN* gene into a suitable cloning vector to allow manipulation. Attempts were made to isolate the *Eco*RI, *Pst*I, *Bam*HI and *HindIII* fragments into the high copy number cloning vector pUC19 (Vieira & Messing, 1982), but failed. Subsequent use of medium copy number vectors and low copy number vectors such as pACYC 177 (Chang & Cohen, 1978) and pLA2917 (Allen & Hanson, 1985) also failed to result in the isolation of any positive clones. It was possible that the *rpoN* gene product was toxic to the host *E.coli* strain and several strains were used also in combination with high, medium and low copy number plasmids, but again failed to give any positive clones.

It was therefore necessary to clone the *rpoN* gene in two fragments to prevent the toxicity of the product on the host cell and then clone them onto either end of an omega cartridge to allow homologous recombination. By restriction and Southern blot analysis, it was found that the *R.sphaeroides* WS8 *rpoN* gene possessed a *Sma*I restriction site internal to it (data not shown) and that cleavage with *Bam*HI and *Sma*I resulted in the production of two fragments of approximately 550bp and 450bp which hybridised to the HR *rpoN* gene. The cloning of these fragments in the high copy number vector pUC19 was attempted and again problems with toxicity were observed in numerous strains. The use of the medium copy number plasmid pACYC177 allowed the cloning of these fragments, but the construction of a plasmid that would allow the mutagenesis of the *rpoN* gene in the chromosome of *R.sphaeroides* WS8 was not possible due to the combination of time constrictions, the instability of the clones, the lack of convenient restriction sites and the apparent toxicity of the gene fragments or partial products.

Fig 3.4 (Left) Southern blot analysis of *R.sphaeroides* WS8 genomic DNA probed with the *rpoN* gene from *R.sphaeroides* HR. 1, pSNT4 digested with *Bam*HI; 2 and 3 *R.sphaeroides* WS8 genomic DNA digested with *Eco*RI and *Bam*HI respectively; M, Lambda DNA digested with *Bst*EII. Sizes of markers are given in Kb.

Fig 3.5 (Right) Southern blot analysis of cosmids 146 and 360 probes with pSNT4. 1, pSNT4 digested with *Bam*HI; 2 and 3, cosmids 146 and 360 digested with *Bam*HI. Sizes are in Kb.

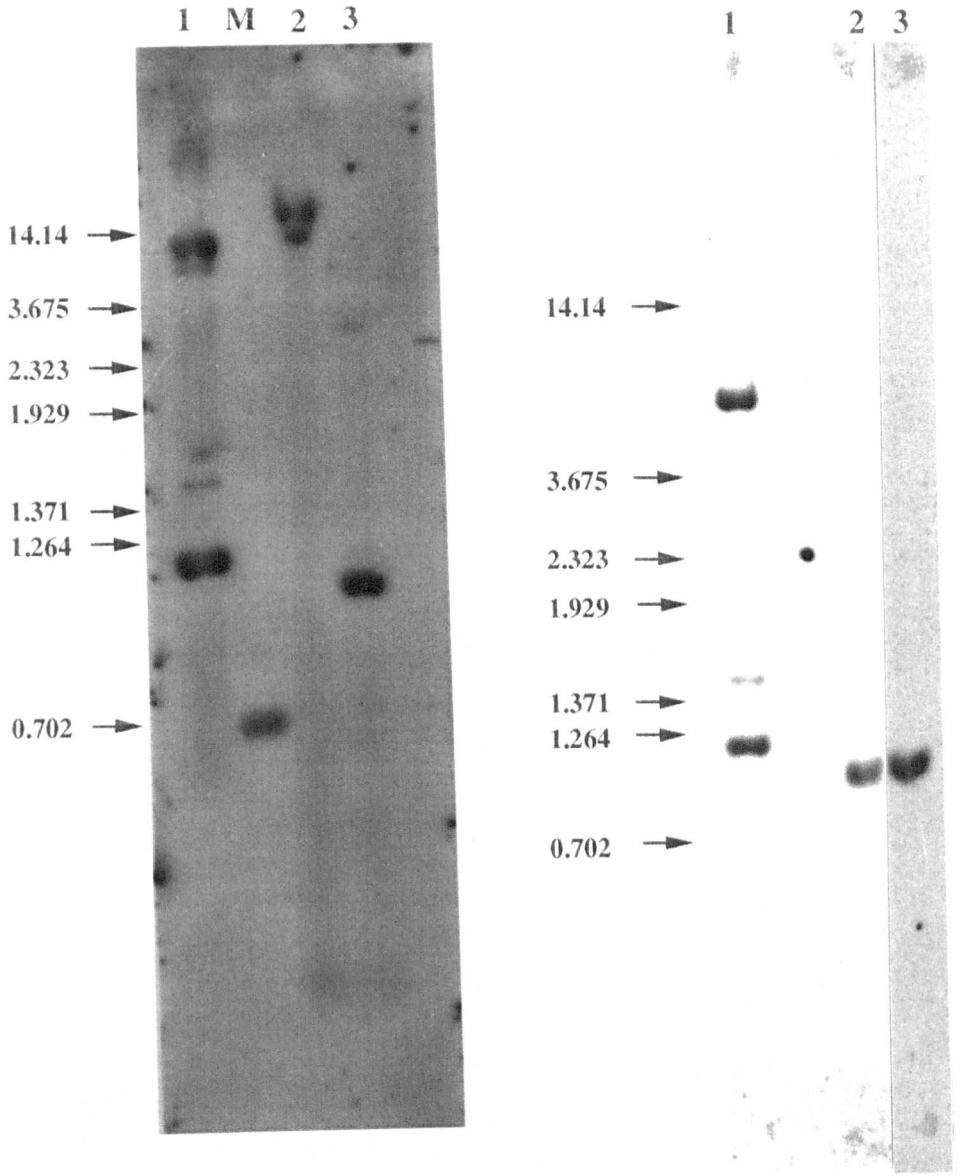
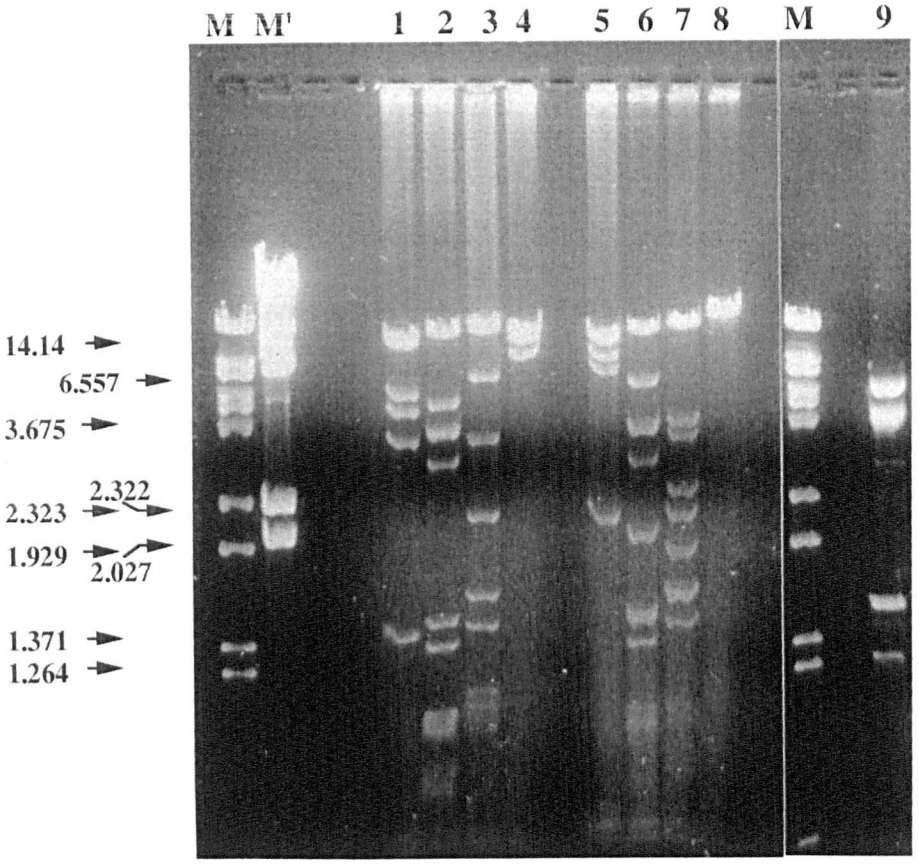
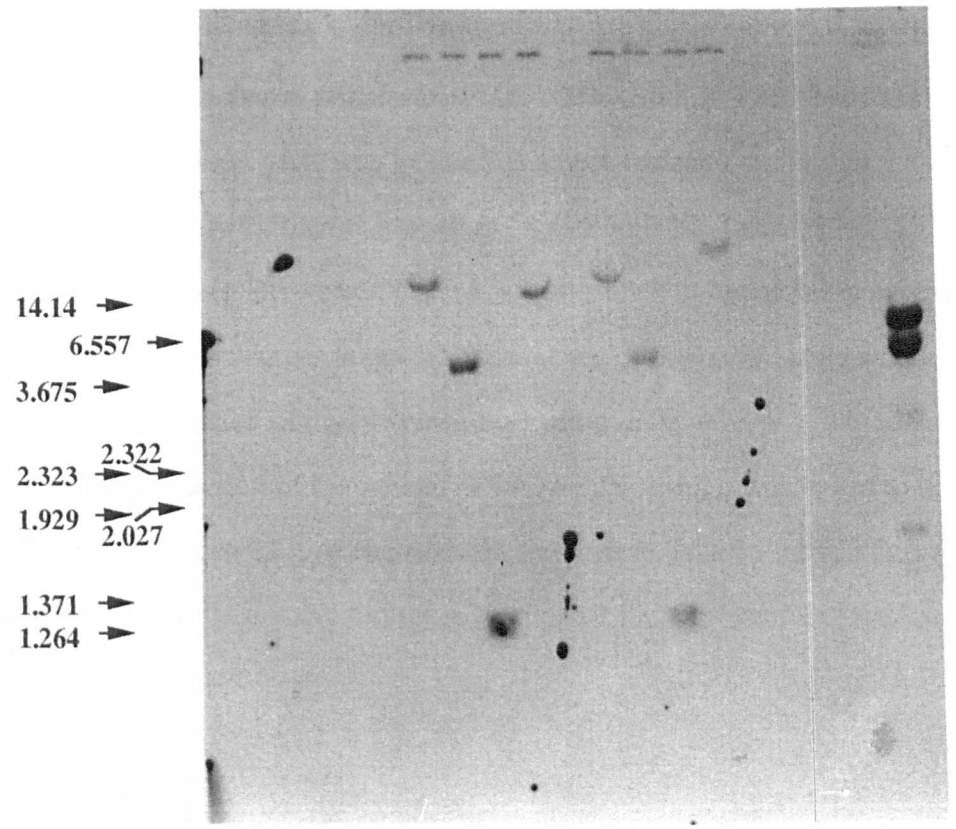

1 M 2 3
14.14
3.675
2.323
1.929
1.371
1.264
0.702
1 2 3
14.14
3.675
2.323
1.929
1.371
1.264
0.702

Fig 3.6. Restriction and Southern blot analysis of cosmids 146 and 360 using the *R.sphaeroides* HR *rpoN* gene as the probe. Top is the fragment pattern given when the cosmids are digest with *Eco*RI (lanes 1 & 5), *Pst*I (lanes 2 & 6), *Bam*HI (lanes 3 & 7) and *Hin*dIII (lanes 4 & 8). Lanes 1 to 4 are cosmid 146 and lanes 5 to 8 show cosmid 360. M and M' are λ *Bst*EII and λ *Hin*dIII respectively. Lane 9 contained pSNT4 digested with *Bam*HI and *Sac*I. Bottom shows the result of Southern blot analysis using the pSNT4 probe. Molecular weight values are in Kb.


M M' 1 2 3 4 5 6 7 8 M 9
14.14
6.557
3.675
2.322
2.323
1.929
2.027
1.371
1.264


14.14
6.557
3.675
2.322
2.323
1.929
2.027
1.371
1.264

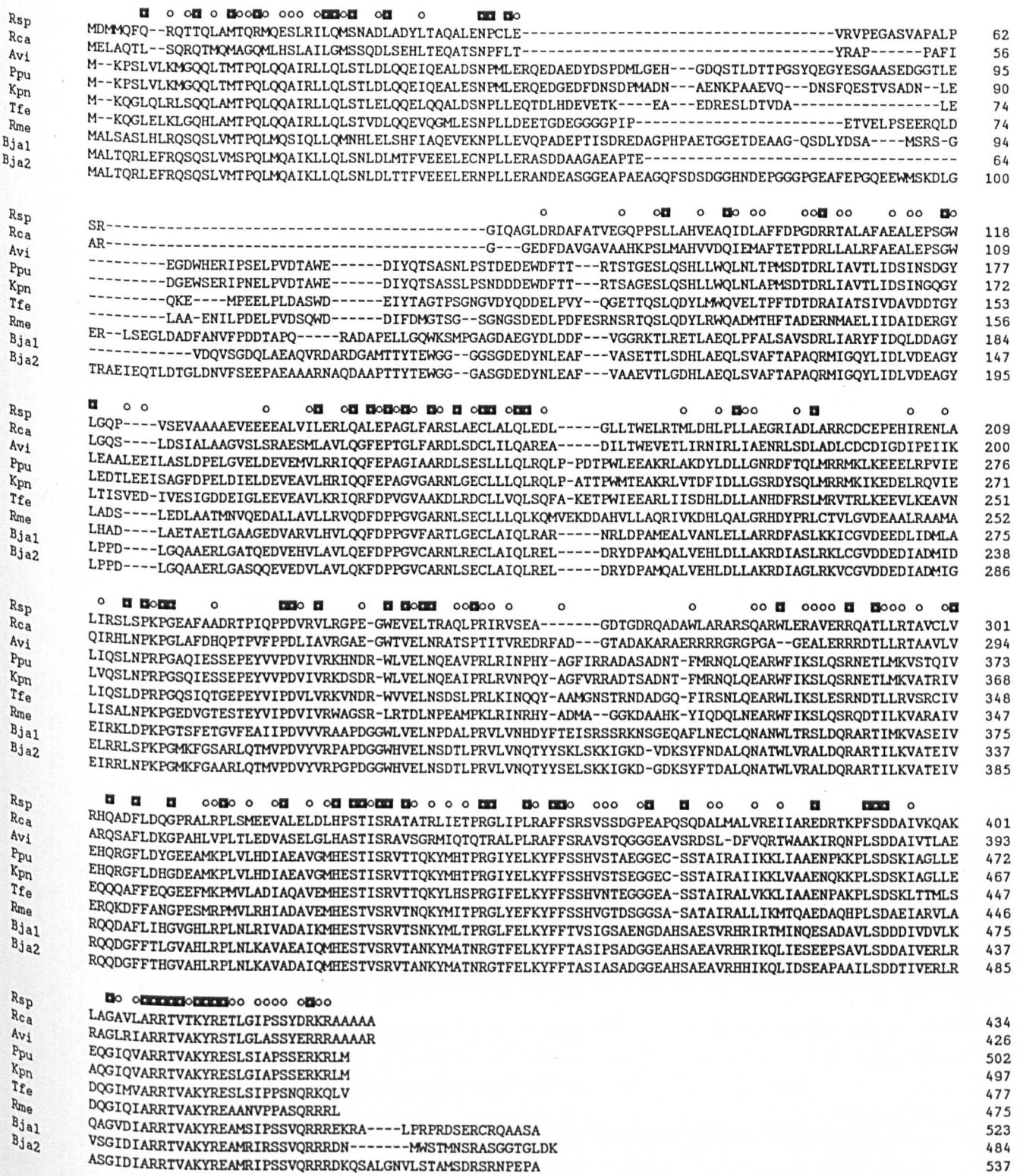
3.3 Discussion

The *R.sphaeroides* WS8 *rpoN* gene differs from the HR *rpoN* gene as HR is a heat resistant strain;

The use of the *R.sphaeroides* HR *rpoN* gene for the construction of a gene replacement mutant of *R.sphaeroides* WS8 led to the isolation of only single cross-over insertions. This suggested that a region of the HR *rpoN* gene is substantially non-homologous to the WS8 *rpoN* gene so as to prevent recombination or that *rpoN* is essential under the growth conditions used. Based on sequence alignments of many of the RpoN homologues identified in many bacterial species (see Fig 3.7), it seems plausible that the N-terminal region would be the more variable domain as it appears to have the largest number of insertions and deletions between bacterial species. It would be predicted that the level of homology between different strains of the same species would be more than sufficient to allow recombination to occur. For the photosynthesis regulatory protein PpsR, strain to strain variation is only in the order of 3% i.e. the genes from two different strains of *R.sphaeroides* are 97% identical at the DNA sequence level (Gomelsky & Kaplan, 1995). This therefore poses the question: **Why would there by a high degree of variation in the *rpoN* genes from *R.sphaeroides* HR and *R.sphaeroides* WS8?** *R.sphaeroides* HR was isolated as a heat-resistant strain that possessed the ability to tolerate growth at 42°C (Meijer & Tabita, 1992) . The adaptations that must have taken place to allow growth at higher temperature may have involved changes in the DNA sequence of the *rpoN* gene which prevented the isolation of double cross-over mutations.

The production of the mutant WS8::*rpoN*X, which was found to be motile and possess wild-type chemotactic responses, gave no insight into whether or not σ^{54} is involved in the regulation of flagellar gene expression in

Fig 3.7 Protein sequence alignment of the *R.sphaeroides* HR RpoN protein (Rsp) with those from other bacteria. Taken from Meijer and Tabita, 1992. See Meijer and Tabita, 1992 for references for sequences. Abbreviations: Rca, *R.capsulatus*; Avi, *Azotobacter vinelandii*; Ppu, *P.putida*; Kpn, *Klebsiella pneumoniae*; Tfe, *Thiobacillus ferrooxidans*; Rme, *Rhizobium meliloti*; Bja1, *Bradyrhizobium japonicum* RpoN1; Bja2, *B.japonicum* RpoN2 . ■ represents residues that are identical with O denoting conservative substitutions.


Rsp
Rca
Avi
Ppu
Kpn
Tfe
Rme
Bja1
Bja2

R.sphaeroides WS8. This was due to the fact that the mutants were, presumably, diploid for *rpoN*. As stated above, the degree of similarity between the *R.sphaeroides* HR RpoN protein and others appears to be greatest at the C-terminus, and consequently the single cross-over probably occur via the 3' region of *rpoN*. This is contrary to the event which is shown in Fig. 3.3 and would presumably result in an intact copy of *rpoN* being present 5' to the interrupted version. This would explain the lack of apparent lack of phenotype unless the interrupted version produced a dominant effect, which obviously was not the case.

***R.sphaeroides* WS8 may not contain two copies of *rpoN*;**

The subsequent isolation and cloning of the *R.sphaeroides rpoN* gene using the HR *rpoN* gene as a probe, confirmed that the genes shared enough homology to allow hybridisation. The finding that two hybridising fragments were seen in chromosomal DNA of *R.sphaeroides* WS8 (Fig. 3.4, lane 3) was suggestive of their being two copies of *rpoN* in *R.sphaeroides* WS8 as has been reported for *R.sphaeroides* HR (Meijer & Tabita, 1992). However, the fact that the hybridisation was very poor under low stringency conditions, and that the hybridisation was not seen when under high stringency (data not shown) suggests that the hybridising fragment may simply represent another DNA binding protein with conserved regions. Meijer and Tabita also showed that a second hybridising fragment in *R.sphaeroides* HR (Fig. 6 of Meijer and Tabita, 1992), but they found that the hybridisation still occurred under high stringency conditions. The quality of the hybridisation was again poor (see Fig. 6 of Meijer and Tabita, 1992). The authors correctly note that duplication of genes in *R.sphaeroides* is not without a precedent as the Calvin cycle CO_2 fixation genes have been seen to be duplicated (Gibson & Tabita, 1988) and that *B.japonicum* has duplicate copies of *rpoN* (Kullik *et al.*, 1991). However, taking into consideration the points illustrated below, it is questionable whether or not duplicate copies exist: 1) The hybridisation seen in

R.sphaeroides HR and WS8 to the second fragment was poor; 2) The method Meijer and Tabita used to clone *rpoN* was functional complementation of an *rpoN*$^-$ mutant of *R.capsulatus* with a cosmid library and, assuming complete coverage of the chromosome within the library, if two copies were present this procedure would have isolated both copies; 3) The failure of this study to reveal two different copies of *rpoN* as the cosmids overlapped; and 4) The lack of further reports from Meijer and Tabita confirming the presence of two copies.

rpoN is difficult to manipulate

The problems encountered with the manipulation of the cloned *rpoN* gene, prevented me constructing a *R.sphaeroides* WS8 *rpoN*$^-$ strain. The observed lethality of the *rpoN* gene is consistent with the finding that another *R.sphaeroides* probable DNA-binding protein (Torf) is also unstable in *E.coli* (see chapter 5). The instability of the *rpoN* clone from *R.sphaeroides* WS8 was in contrast to the apparent stability of the *R.sphaeroides* HR clone. However, in my hands, the pSNT4 clone, containing the complete *rpoN* gene from *R.sphaeroides* HR (Meijer & Tabita, 1992), was also unstable. Whether or not Meijer and Tabita also experienced such instabilities is unknown, but it was not reported (Meijer & Tabita, 1992). Of interest is the fact that the instability of the *rpoN* clone in pUC19, or rather the fact that no clones could be isolated even when the direction of transcription of *lacZ* was opposite to the orientation of *rpoN*, may have been due to the fact that pUC19 has been reported to contain a 'cryptic' promoter which drives expression of inserted genes in the opposite direction to *lacZ* (Errington, pers. comm.).

In summary, in this part of the study, the role of the alternative sigma factor σ^{54} in flagellar gene regulation in the motile strain of *R.sphaeroides*, WS8 was investigated. Attempts were made to use the *rpoN* gene from another strain of *R.sphaeroides* for gene replacement mutagenesis but failed, probably due to sequence variation at the 3' region of the gene. The *R.sphaeroides* WS8 *rpoN* gene was subsequently cloned on two overlapping cosmids but the

subcloning of the *rpoN* gene onto a suitable cloning vector was hampered by its apparent toxicity to the host cell. This was partly overcome by cloning the gene in two fragments, but these again were seen to be unstable. Consequently, the role of σ^{54} in flagellar gene regulation in *R.sphaeroides* still remains unknown and requires further investigations as it is obvious, in the light of the presence of the consensus sequence being present upstream of the *mot* operon (Shah & Sockett, 1995) and the *torf* operon (see chapter 5), that it may play a major role in flagellar gene regulation.

Chapter 4

Analysis of a transposon mutant with low levels of motility

4.1 Introduction

Due to the extracellular and extracytoplasmic nature of many of the proteins in the flagellum, they must be exported in a highly controlled manner so as to allow the correct assembly of the flagellum. The flagellar specific export apparatus is responsible for this process and a detailed description of it can be found in the introduction. The aim of this part of the project was to characterise a motility mutant of *R.sphaeroides* that had a defect that led to low levels of motility. This mutant, as will be shown, had an export defect, but it was possible that the phenotype was due to the production of a truncated motor protein which interfered with flagellum formation and this is why it was examined. The phenotype of this mutant was intriguing as it suggested that the interrupted gene was not essential for flagellar assembly but did play some role.

The mutant (M18) was originally isolated by Foster (Foster, 1991) by Tn*pho*A mutagenesis as a non-motile strain. Foster claimed that the mutant was in fact completely non-motile and non-flagellate, with the lesion of the mutation being located to a 5.5Kb *Eco*RI fragment of a cosmid known as cosmid 140. As will be shown in this chapter, these findings were incorrect.

4.2 Results

4.2.1 Characterisation of the mutant M18.

The *R.sphaeroides* mutant M18 was isolated after Tn*pho*A mutagenesis (Sockett, 1988) as a kanamycin resistant colony which failed to form a diffuse swarm on a semisolid tryptone-yeast extract plate (Sockett & Armitage, 1991). Foster first characterised the mutant and classified it as non-motile and non-

flagellated (Foster, 1991). Western blot analysis of sheared flagellar filaments with anti-flagellin antibody, carried out in this project, revealed the presence of lowered levels of exogenous flagellin in M18 (Fig 4.1 lane 2), compared to wild type WS8 (Fig 4.1 lane 1). This suggested that the defect in M18 might lie in the export or assembly of the flagellin protein monomers. Examination of M18 by phase contrast microscopy revealed that it was motility impaired; approximately 0.1% of wild type levels of motility i.e. 1 in 1000 cells were motile and behaved with wild type characteristics. These data confirmed that Foster's original classification was incorrect.

Southern blotting with a Tn*phoA* probe (*Hin*dIII fragment from pUI800 (Moore & Kaplan, 1989)) was used to determine the presence of Tn*phoA* in the chromosome of M18 (Fig. 4.2). A single hybridising band of approximately 3.5 Kb was seen in *Hin*dIII digested M18 chromosomal DNA (Fig 4.2 lane 2), no hybridisation was seen with wild-type WS8 DNA (Fig 4.2 lane 1). This band represents the internal Tn*phoA* *Hin*dIII fragment. No Pho A phosphatase activity was observed either on plates or by Western blot (Sockett, 1988) analysis using anti-alkaline phosphatase antibody (Moore & Kaplan, 1989). This suggested that the transposon had not inserted in frame within an interrupted gene.

4.2.2 Cloning and study of the wild type motility gene.

To determine the nature of the lesion, genomic M18 DNA upstream of the transposon insertion site was cloned into pUC19 as a 7Kb *Sal*I restriction fragment (pM18S) by selecting for Tn*pho*A-encoded kanamycin resistance. This clone was used to probe digests of a cosmid clone of *R.sphaeroides* genomic DNA (Sockett & Armitage, 1991) known as cosmid 140 (Fig 4.3). The probe was found to hybridise to two *Bam*HI fragments of 2.15 and 1 Kb (lane 1), an approximately 5.9Kb *Eco*RI fragment (lane 2), an approximately

9Kb *Hin*dIII fragment (lane 3), two *Pst*I fragments of 1.9 and 1.7Kb (lane 4), a 7.6 Kb *Nru*I fragment (lane 5) and a 3.4 Kb *Sal*I fragment (lane 6). Restriction maps of pM18S and the 5.9 Kb *Eco*RI fragment to which it hybridised were compared and this showed that the site of transposon insertion in M18 lay within a 2.8 Kb *Eco*RI to *Sal*I fragment (Fig 4.4), approximately 100 to 150 bp upstream of a *Bam*HI site.

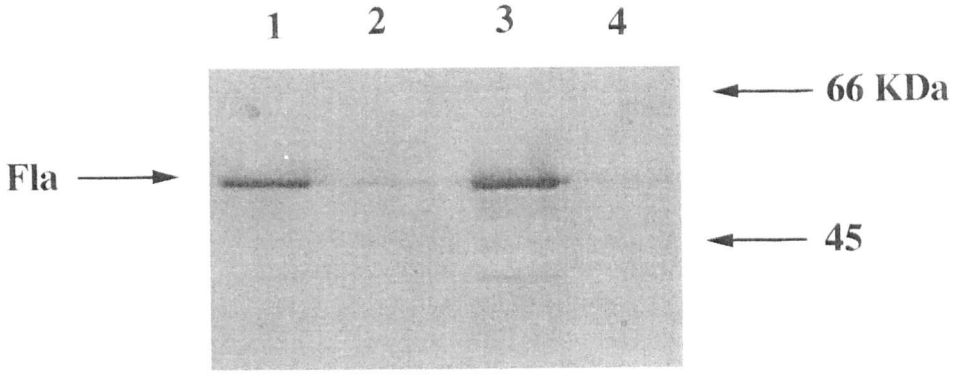
When the cosmid clone (cosmid 140), containing 35 Kb of *R.sphaeroides* DNA, was conjugated into M18, it was found by microscopy to complement M18 back to wild type levels of motility. In addition, Western blot analysis with anti-flagellin antibody demonstrated that complemented M18 had wild type levels of exogenous flagellin (Fig 4.1 lane 3). The 1 Kb *Bam*HI restriction fragment from the 2.8Kb *Eco*RI to *Sal*I fragment (Fig 4.4) was subsequently used as a probe in Southern blot analysis of digests of M18 and WS8 genomic DNA (Fig 4.5). It was found that the probe hybridised to fragments of 1.5 Kb and 0.8 Kb in *Sma*I digested WS8 genomic DNA (Fig 4.5 lane 1) and fragments of 3.8 Kb and 0.8 Kb in *Sma*I digested M 18 genomic DNA (Fig 4.5 lane 2). The 0.8 Kb fragment in both WS8 and M 18 genomic digests represents the 3' *Sma*I fragment, whereas the other fragment (1.5 Kb in WS8 and 3.8 Kb in M 18) extends 5' from the most left-hand *Sma*I site within the 1Kb *Bam*HI fragment, to a 5' *Sma*I site upstream of the *Eco*RI site (Fig 4.4) in the case of WS8, and to the *Sma*I site within Tn*phoA* in the case of M 18.

4.2.3 DNA sequence analysis.

The nucleotide sequence of clones around the site of Tn*phoA* insertion were determined on both strands by plasmid sequencing using subclones and custom synthesised primers as shown in Fig 4.4. A single open reading frame was found and the DNA and deduced amino acid sequence are shown in Figs 4.6 and 4.8 respectively. The open reading frame encodes a protein of 46.8

Fig 4.1 Western immunoblot analysis of sheared flagellar filaments from: 1) WS8, 2)M18, 3) M18:cos 140, and 4) M18:pLA29-17, probed with anti-flagellin antibody. The band of approximately 55KDa (marked Fla) represents flagellin. In each case 10µl of sheared flagellar filaments, as prepared by the method described in chapter 2, were loaded. KDa, kilodaltons.

Fig 4.2 Southern blot analysis of *Hin*dIII digested genomic DNA from: 1) WS8, and 2) M18 probed with the internal *Hin*dIII probe from Tn*pho*A (Table 2.1). Lane 3 contains pUI800 - the cosmid vector as a control. Kb, kilobase pairs. Genomic DNA was prepared by the method described in chapter 2.


1
2
3
4
66 KDa
Fla
45

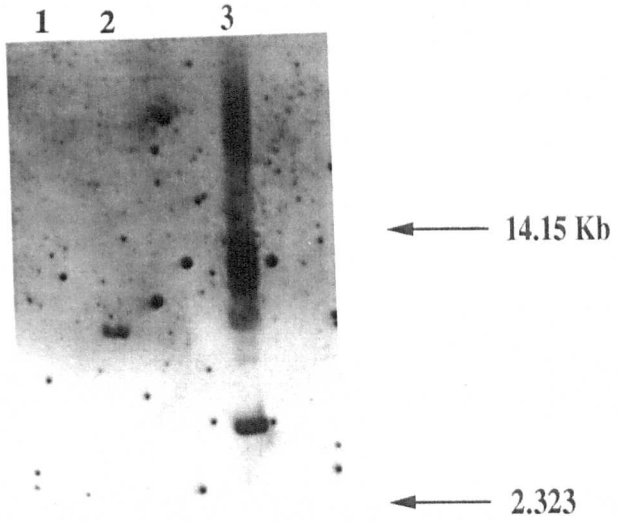
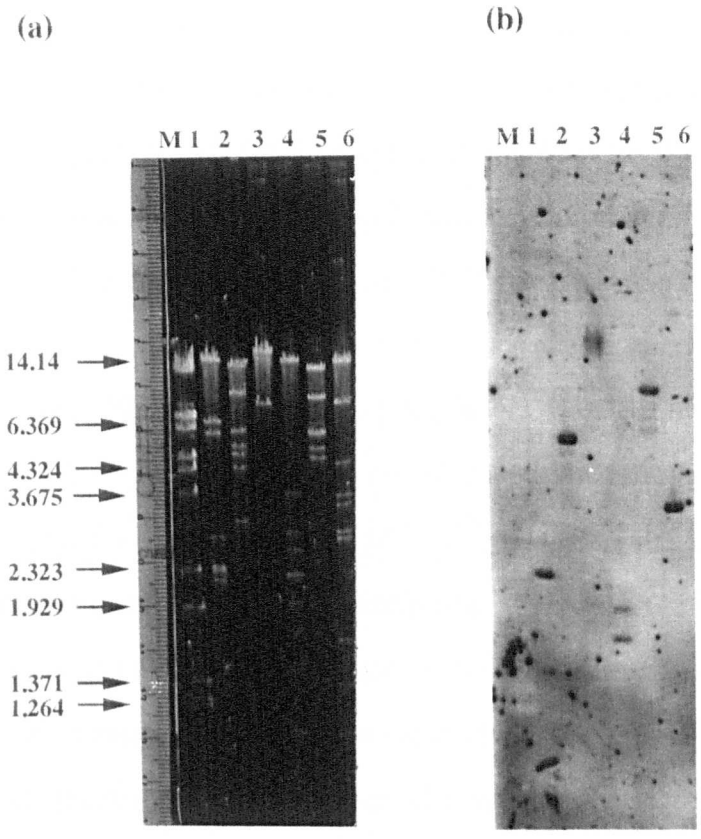

1
2
3
14.15 Kb
2.323

Fig 4.3 Southern blot analysis of Cosmid 140 digests probes with pM18S. (a) Restriction digested cos 140 DNA, lanes are as follows; M) Lambda *Bst*EII marker, 1) *Bam*HI, 2) *Eco*RI, 3) *Hin*dIII, 4) *Pst*I, 5) *Nru*I and 6) *Sal*I. (b) Southern blot of (a) using pM18S. Sizes of the markers are shown arrowed in kilobase pairs

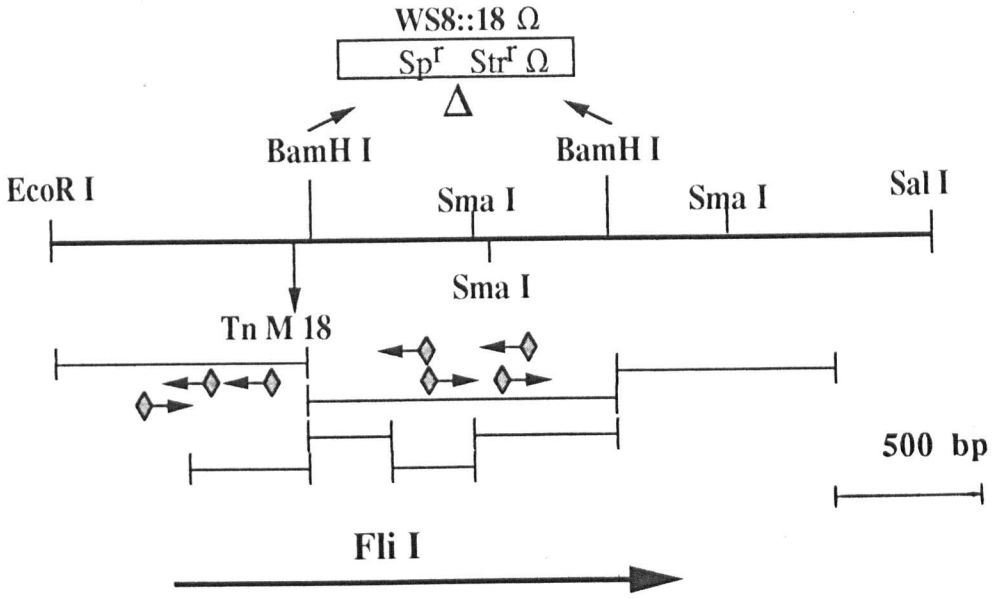

(a)
(b)
M 1 2 3 4 5 6
M 1 2 3 4 5 6
14.14
6.369
4.324
3.675
2.323
1.929
1.371
1.264

KDa which was found to have extensive homology to the FliI protein from *S.typhimurium* (Vogler *et al.*, 1991) (Fig 4.8) and various ATPases by FASTA searches (Deveraux *et al.*, 1984) of the Swissprot protein database held at Daresbury, UK. It was found to have 49.5% identity and 68% similarity to *S.typhimurium* FliI and 41.8 % identity and 62% similarity to HrpB6 from *Xanthamonas campestris* (Fenselau *et al.*, 1992) using the GAP program from the GCG package (Deveraux *et al.*, 1984). There is a potential Shine-Dalgarno ribosome binding sequence (Shine & Dalgarno, 1974) 8 bases upstream from the ATG 'start' of the open reading frame (marked SD Fig 4.6).

The site of Tn*phoA* insertion was mapped to 123 bp upstream of the *Bam*HI site by comparison with DNA sequence of pM18S. This site (arrowed in Figs 4.6 & 4.8) corresponds to amino acid 58 in the FliI protein. The ATP binding motifs (Walker boxes) (Walker *et al.*, 1982) present in the *S.typhimurium* FliI and other ATP requiring enzymes appear to be conserved in the *R.sphaeroides* FliI protein (marked 1 & 2 in Fig 4.8). Sequence analysis across the transposon-*fliI* region revealed a possible start point for expression of a shortened *fliI* product in M18 (Fig 4.7). This was a GTG initiation codon is situated 15bp from the end of the Tn*pho*A DNA, 9 base pairs downstream from a potential Shine-Dalgarno sequence (Shine & Dalgarno, 1974) , and in the correct reading frame for the production of a hybrid FliI protein. This hypothetical hybrid protein would begin with the sequence VYKSQ fused to *R.sphaeroides* FliI from amino-acid 59 through to the end. So, it was possible that a shortened FliI, which was partially functional, was producing the low levels of motility in M18 as seen in an *E.coli fliN* mutant (Tang *et al.*, 1995). To determine if this was the case, a more substantial deletion of the *fliI* gene was required.

Fig 4.4 Partial restriction map of the 2.8Kb *Eco*RI to *Sal*I fragment showing the site of Tn*pho*A insertion (arrowhead Tn M18) and the FliI open reading frame (indicated as an arrow). The construction of the *fliI* partial deletion strain (WS8::18Ω) is also shown with Δ representing the deletion of the internal *Bam*HI fragment and its replacement with the spectinomycin (Sp^r) and streptomycin (Str^r) resistance conferring omega cartridge (Ω) from pHp45Ω. Shaded diamonds represent custom made oligonucletides used for sequencing. Subclones made for sequencing are depicted underneath as ⊢——⊣.

Fig 4.5 Southern blot analysis of *Sma*I digested genomic DNA from; 2)WS8, and 3) M18, using the internal 1Kb *fliI Bam*HI fragment (Fig 4.4) as the probe. 1) pNM18B digested with *Bam*HI. Kb, kilobase pairs.


WS8::18 Ω
Sp^r Str^r Ω
Δ
BamH I
BamH I
EcoR I
Sma I
Sma I
Sal I
Sma I
Tn M 18
500 bp
Fli I

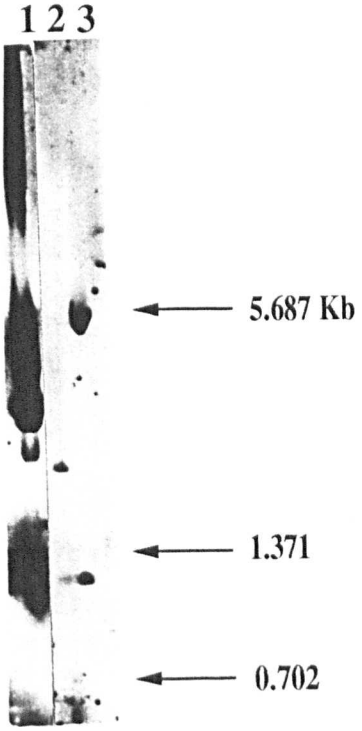
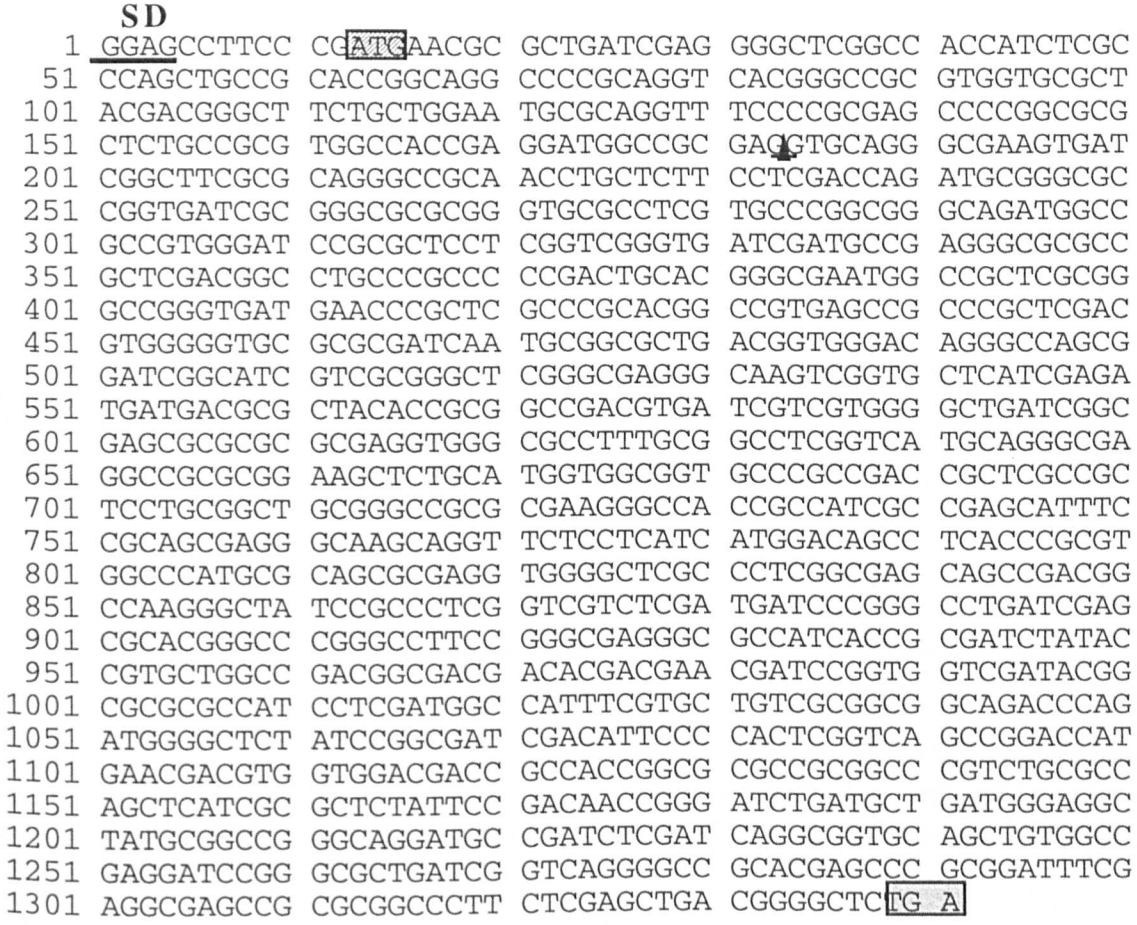
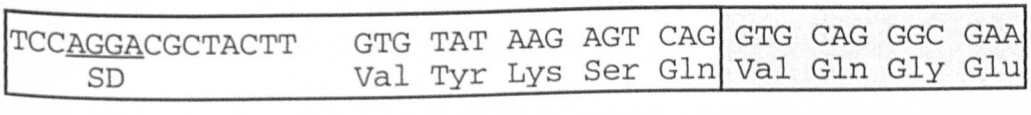

1 2 3
5.687 Kb
1.371
0.702

Fig 4.6 DNA sequence of the region containing the *fliI* gene from *R.sphaeroides*. SD, Shine-Dalgarno ribosome binding site. Shaded boxes represent the start and stop codons of *fliI*. The transposon insertion site is marked with an arrow.

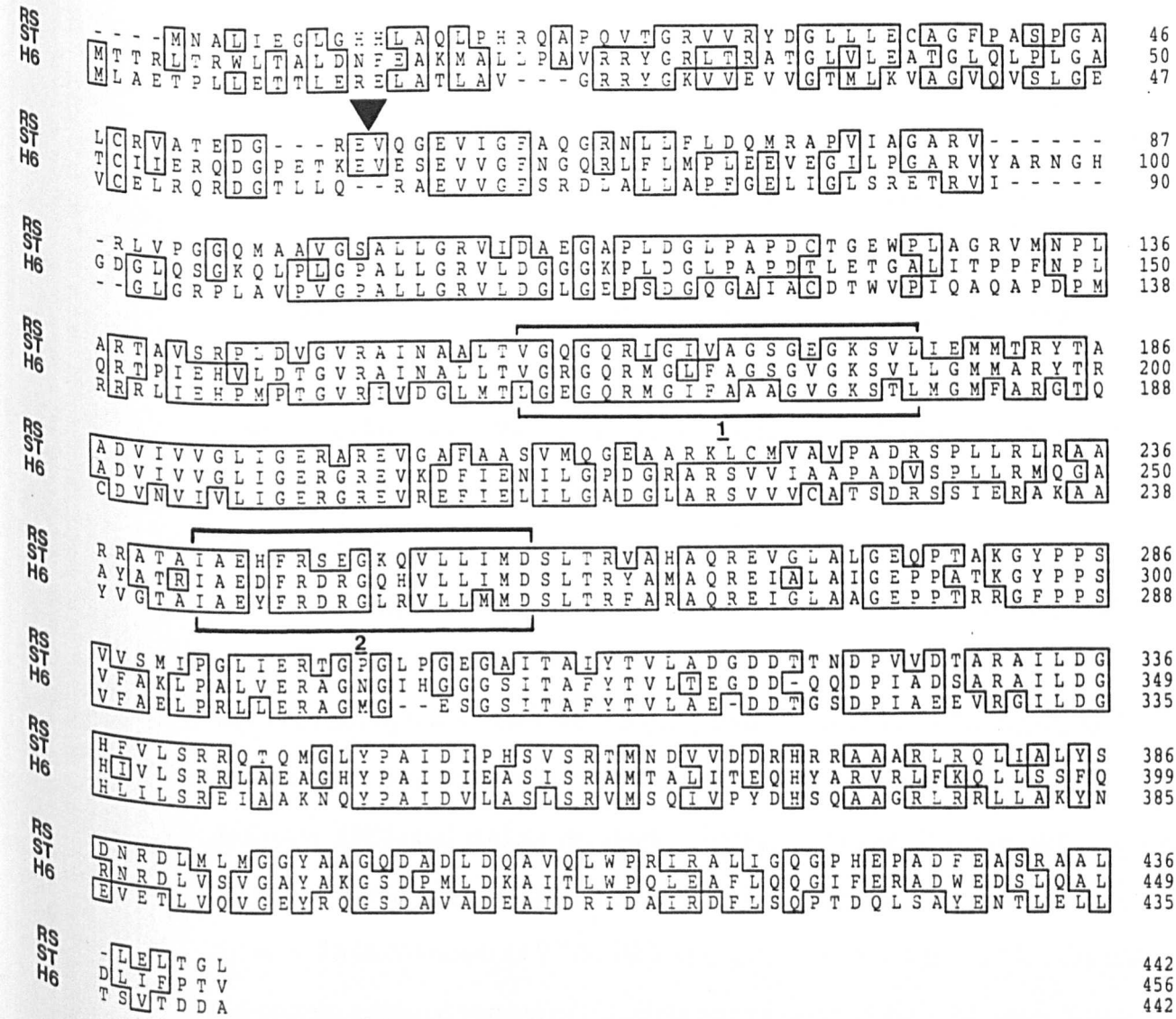
Fig 4.7 Nucleotide sequence of the potential GTG 'restart' point in M18 showing the junction between Tn*pho*A and the *R.sphaeroides* DNA in the chromosome of M18. The potential Shine-Dalgarno (SD) (Shine & Dalgarno, 1974) is underlined and the *fliI* sequence is shown in the shaded right hand half of the box.





Tn*pho* A Sequence *fli*I Sequence

Fig 4.8 Protein sequence alignment of the predicted *R.sphaeroides* FliI (RS) protein sequence with *S.typhimurium* FliI (ST) (Vogler *et al.*, 1991) and HrpB6 from *X.campestris* (H6) (Fenselau *et al.*, 1992). The sequence alignment was created using the PILEUP program and conserved or identical residues boxed using the PRETTYBOX program, both from the GCG package (Deveraux *et al.*, 1984). The conserved nucleotide-binding domains (Walker *et al.*, 1982) are shown as horizontal brackets 1 and 2. The site of Tn*phoA* insertion in the mutant M18 is arrowed on the *R.sphaeroides* sequence.


RS
ST
H6
1
2

4.2.4 Construction of a chromosomal deletion of *fliI*.

The 2.8 Kb *Eco*R I to *Sal* I fragment from cosmid 140 (Fig 4.4) was cloned into the suicide plasmid pARO191 (Park, 1990) and the 1Kb *Bam*HI fragment replaced with an omega cartridge (Fig 4.4), conferring spectinomycin and streptomycin resistance and containing transcriptional and translational terminators in all possible reading frames from pHp45 Ω (Prentki & Krisch, 1984). This construct was then introduced into *R.sphaeroides* WS8-N by diparental mating. The *fliI* gene replacement could therefore occur via the homologous recombination of the regions flanking the Ω cartridge with the wild type copy of the gene in the chromosome. An even number of recombination events (i.e. one either side of the Ω) would result in the integration of the deleted *fliI* into the chromosome and loss of the suicide vector. Exconjugants with the *fliI* gene replaced by the deleted version were found by testing for the spectinomycin and streptomycin resistance encoded by the Ω cartridge (Prentki & Krisch, 1984), and the absence of pARO191-derived kanamycin resistance (Park, 1990). Further confirmation was required to show that the *fliI* gene *Bam*HI fragment had been deleted by homologous recombination, and that pARO 191 was not present in the chromosome. This was obtained by using the 1 Kb *Bam*HI fragment from the 2.8 Kb *Eco*RI to *Sal*I fragment (Fig 4.4), the *Bam*HI fragment carrying the Ω cartridge (Prentki & Krisch, 1984) and pARO 191 (Park, 1990) as probes in Southern blot analysis of genomic DNA isolated from several *fliI* deletion isolates (data not shown). Isolates known as WS8::18Ω were found not to contain pARO191 but did contain a deleted version of *fliI*. Subsequent microscopic analysis of liquid WS8::18Ω cultures showed that they too were partially motile, 1 in 1000 cells swam, as in the M18 strain.

4.2.5 Electronmicroscopic analysis of M18 and WS8::18Ω

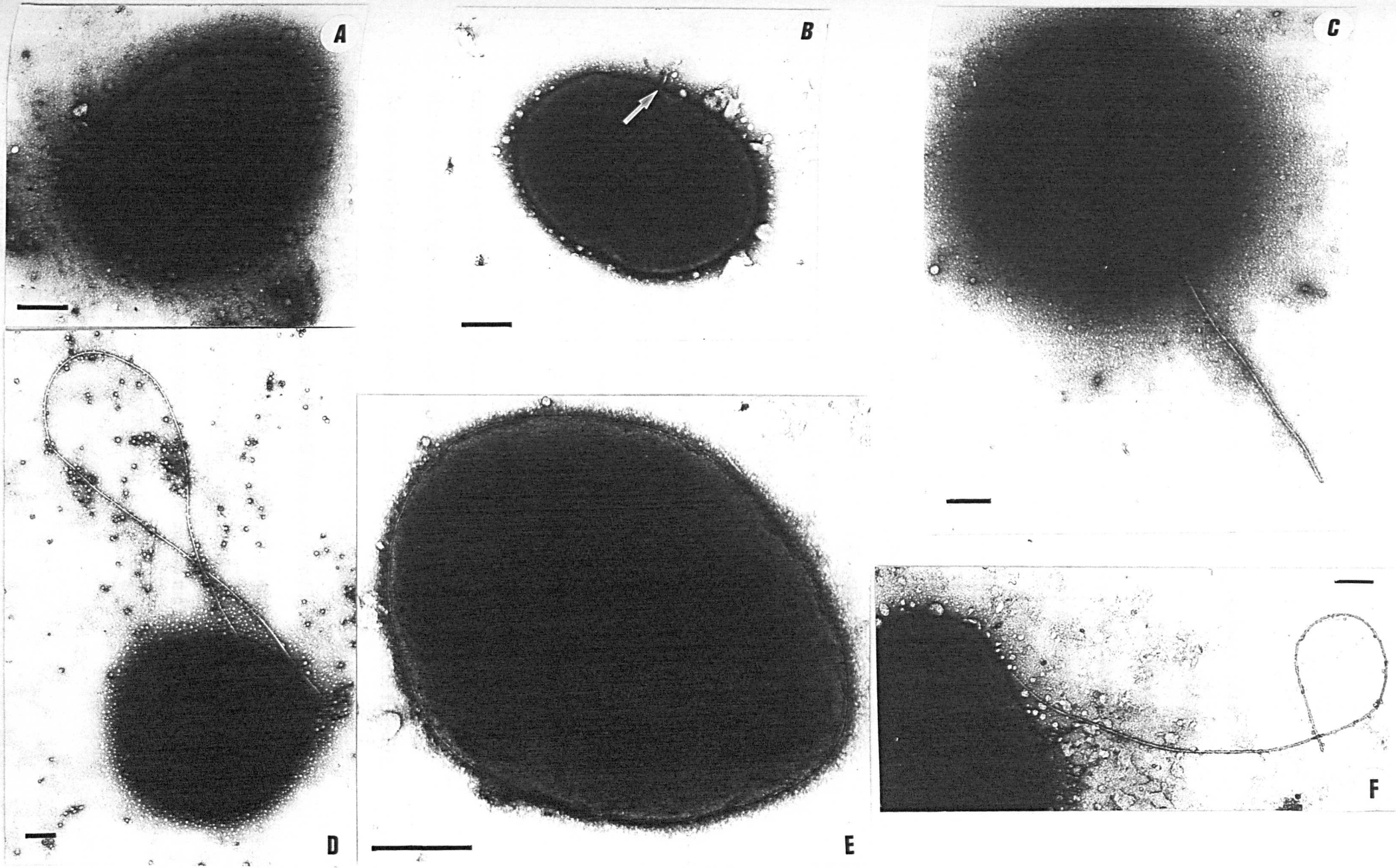
The presence of very low levels of motile cells and extracellular flagellin in M18 cultures suggested a defect in flagellin export causing slow export of flagellin subunits, with only a few cells reaching the threshold filament length required for motility. In order to determine if such a situation was in fact the case, electronmicroscopic analysis of liquid cultures of M18 and WS8::18Ω was performed. It was found that of the 200 cells of M18 and WS8::18Ω that were observed under the electron-microscope after being negatively stained, only one cell from each had a full-length (3μM) flagellum. Less than 5% of the other cells showed the presence of a very short flagellar filament or a 'stub' (50 - 1000nm long see Fig 4.9) but there was no obvious gradation in flagellar lengths between cells. This precludes the idea that slow export of flagellin subunits was occurring, giving only a few cells with a flagellum above the threshold length required for motility.

4.3 Discussion

R.sphaeroides has a FliI homologue;

Analysis of a motility mutant (M18) has led to the identification and sequencing of the *R.sphaeroides fliI* gene. There are two lines of evidence which confirm the identity of *fliI*: First, the homology of the *R.sphaeroides* FliI protein, to the FliI protein from *S.typhimurium* (Vogler *et al.*, 1991), and other transport proteins such as HrpB6 from *X.campestris* (Fenselau *et al.*, 1992); 49.5% and 41.8% identity respectively (Fig 4.8). Second, the presence of low levels of motile cells and low levels of extracellular flagellin in M18 and WS8::18Ω cultures was suggestive of a defect in flagellin export. One possibility was that slow export of the subunits, resulted in only a few cells reaching the threshold filament length required for motility. This was subsequently shown not to be the case as M18 and WS8::18Ω show no

Fig 4.9 Electronmicroscopic analysis of M18 and WS8::18Ω. Cells were negatively stained with 2% uranyl acetate pH 7. A) M18 lacking a flagellum; B) M18 showing very small stub (arrow); C) M18 showing medium length filament; D) M18 X cosmid 140 showing full length flagellum; E) WS8::18Ω lacking a flagellum; F) WS8::18Ω showing medium length flagellum. Bar represents approximately 200nm.


A
B
C
D
E
F

obvious gradation of filament lengths precluding the idea of slow flagellin export.

FliI is not essential for flagellum formation in *R.sphaeroides*;

The fact that the mutant population is not completely non-motile, and contains rare cells with a full length flagellum, poses many questions. One possible reason for this low level of motility may be reversion or rather excision of the transposon leaving an intact *fliI* gene. This is unlikely as repeated experiments have shown that it was impossible to isolate a culture of M18 which possesses higher (or lower) levels of motility than the original parent culture. The possibility also existed that a translational restart had occurred within the transposon of M18, producing a hybrid protein, which retained low levels of activity allowing the export of low levels of flagellin. This possibility was eliminated by the construction of a *fliI* null, deletion strain, WS8::18Ω, which was found to have the same phenotype as M18. These observations confirm that the low-motility phenotype of M18 is characteristic of a *R.sphaeroides fliI* mutant and that the FliI protein is involved in flagellar export. The rare motile cells in *fliI* mutant cultures may have been the result of a phase variation phenomenon, with resulted in the expression of genes that are not normally expressed (e.g. Lederberg & Iino, 1956). Such variation could cause either the expression of an alternative flagellin gene whose product is exported independently of FliI and assembled poorly, or the expression of an alternative export gene that replaces *fliI*. Both of these variations could explain the 'rare motile' phenotype. However, there is no evidence of duplicate flagellin genes (Shah *et al.*, 1996) or *fliI* genes as Southern blot analysis of WS8 genomic DNA using an internal *fliI Bam*HI fragment (see Fig 4.4) revealed only two hybridising *Sma*I fragments (Fig 4.5) as predicted if only a single *fliI* gene is present. The phase variation event that the 'rare motile' cells may be analogous to the event that takes place in WS8 that allows the formation of 'petite' colonies that possess the ability to rotate their flagella ccw instead of the

normal cw direction (Packer & Armitage, 1993). In such 'petite' variants, the flagellum functions in a similar manner to wild type cells i.e. stopping and starting, but the direction of rotation of the motor is reversed (Packer & Armitage, 1993). Such variants do not breed true and therefore the effect may not be a stable genetic rearrangement. It would have been very interesting to determine the direction of rotation of the flagella from the 'rare motile' cells of M18 and WS8::18Ω but motor direction is unlikely to be due to filament structure.

Another possible explanation for the observed effect is that FliI could be involved in altering the specificity of the export apparatus in conjunction with FliK and FlhB. It may be that the phase variation described above is not at the genetic level but is at the protein level i.e. FlhB could adopt a conformation that would allow the export of hook and filament proteins. It may be that in the absence of FliI, the FlhB protein cannot readily adopt the correct conformation to allow the export of these proteins but does so in the 'rare motile' cells.

Other components of the export apparatus are also not essential for flagellum formation in *R.sphaeroides*;

The identification of an *R.sphaeroides fliJ* homologue, another component of the export apparatus (see section 1.4.3k), downstream of *fliI* (Pollitt, 1996) and the finding that the phenotype of a *fliJ* null mutant was similar to the phenotype of a *fliI* null mutant suggests that several components of the export apparatus of *R.sphaeroides* are not essential to flagellar formation. A FliM and N null mutant has also been seen to form small filaments as M18 (Sockett, 1996). This may be a reflection of the presence of only a single flagellum in *R.sphaeroides* i.e. as stated above the enteric bacteria require approximately 5 flagella to be motile and therefore require a specialised apparatus to export such a large amount of protein whereas in *R.sphaeroides*, less protein has to be exported to form the flagellum as there is only a single

flagellum and consequently the export apparatus is not as essential as in the enteric bacteria.

A similar operon structure may exist in *R.sphaeroides* as is seen in other organisms;

The finding that *fliJ* is downstream of *fliI* is suggestive that a similar operon structure exists in *R.sphaeroides* flagellar genes as is seen in the enteric bacteria (Homma *et al.*, 1988; Vogler *et al.*, 1991). This is further strengthened by the finding that in *R.sphaeroides* strain 241, a *fliH* homologue was identified upstream of *fliI* (Ballado *et al.*, 1996). If such a case is true then it could be predicted that *fliF* and *fliG* would be upstream of this region with *fliM* and *fliN* being located downstream. As will be shown in subsequent chapters, this was found to be the case.

FliI is *S.typhimurium* is essential to flagellum assembly;

The work of Vogler and co-workers who studied post-shearing filament re-growth and motility of a temperature sensitive *fliI* mutant of *S.typhimurium* has shown that in *S.typhimurium* FliI is responsible for flagellar filament growth (Vogler *et al.*, 1991). They found that 10% of cells were motile at the non-permissive temperature. These may have been due to incomplete shearing or because *fliI* was not completely inactivated at the non-permissive temperature. However, a *fliI* null strain of *S.typhimurium* was found to be completely non-motile and non-flagellate by high intensity dark-field microscopy (Dreyfus *et al.*, 1993). This contrasts with my observations in *R.sphaeroides* where external flagellin was determined by Western blotting (Fig 4.1) and electronmicroscopy (Fig 4.8). The difference in 'flagellar detection methodology' may explain apparent differences in *fliI* mutant phenotypes as dark-field microscopy would not visualise short flagellar stubs if present in *S.typhimurium*. It may be that in the *S.typhimurium fliI* mutants, enough flagellin 'leakage' occurred to form a few flagellar 'stubs' but these were not seen under the dark-field microscope. Whether or not *fliI* null strains

of *S.typhimurium* actually form flagellar 'stubs' awaits direct proof via Western blot analysis of sheared flagellar filaments.

What is the minimum length of flagellar filament required for motility?

The minimum length of flagellar filament required to propel a *R.sphaeroides* cell would, in theory be ascertainable from electronmicrographs of the *R.sphaeroides fliI* mutants M18 and WS8::18Ω. The length of the filament observed on wild-type cells is estimated to be approx. 10 μm (±5) (Sockett, 1986) compared to the length of the *S.typhimurium* filament of 10-15 μm (Iino, 1974). It seems reasonable to assume that the minimum length of filament required for motility would not differ markedly from that observed in a wild-type population; the export of flagellin subunits is an energy consuming process (Dreyfus *et al.*, 1993) and export of subunits additional to what is required for motility would put an unnecessary energy requirement on the cell. However, the length of the flagella from M18 and WS8::18Ω (50nm to 3μm) appears to be less than that of the wild type cells. It is important to note that the length of the filaments from the *fliI* mutants may not be representative as only a small number of cells were observed. Consequently, in order to fully determine the average length of the filaments from *fliI* mutant strains a more detailed examination or a large number of flagellated cells would be required. I was unable to perform such a study due to a combination of time limitations and the poor quality of the micrographs I could obtain.

What is the function of FliI in the export procedure?

FliI is an ATPase

It is interesting to note that the residues implicated as being catalytically important in the function of *S.typhimurium* FliI via mutagenesis (Dreyfus *et al.*, 1993), namely Lys-188, Asp-272 and Tyr-363, are absolutely conserved in the *R.sphaeroides* FliI protein (Fig 4.8). The nucleotide binding motifs (Walker boxes) (Walker *et al.*, 1982) also appear to be very highly conserved and these

two facts may indicate that the *R.sphaeroides* FliI protein may possess the capability to bind (and possibly hydrolyse) ATP as does the FliI protein from *S.typhimurium* (Dreyfus *et al.*, 1993; Silva-Herzog *et al.*, 1995). However, this is yet to be proved conclusively and awaits the purification of FliI from *R.sphaeroides* in a native state, but the finding that a GST-FliI fusion protein is deleterious to *E.coli* cells when overexpressed (Pollitt, 1996) may suggest that it does act as an ATPase or that it has other activities, but again this requires further investigations.

S.typhimurium FliI can bind to the flagellar filament subunits (flagellin) and the rate of ATP hydrolysis increases upon such binding (Silva-Herzog *et al.*, 1995). It will be interesting to discover if FliI has the capabilities to bind any of the other flagellum components that require the flagellar specific export e.g. the hook protein, and determine what portions of the FliI protein participate in this binding. It is known that although there are conserved sequences at the N-terminal of exported flagellar proteins from *E.coli*, *S.typhimurium* and in the hook protein from *C.crescentus*, these do not function in secretion (Ohta *et al.*, 1985; Homma *et al.*, 1990; Homma *et al.*, 1990). However, Kuwajima and co-workers, and Kornacker have shown that regions in the N-terminus of *E.coli* flagellin and *C.crescentus* hook protein are required for export (Kuwajima *et al.*, 1989; Kornacker & Newton, 1994). The authors suggested that the secretion signal may be a conformational one for all exported axial proteins i.e. related to tertiary structure rather than primary sequence and that this would be a similar situation to that which is found in the Yop virulence family from *Yersinia* (Michiels *et al.*, 1991).

FliI may also act as a chaperone;

The finding that FliI can bind to flagellin (Silva-Herzog *et al.*, 1995) may implicate it in a chaperone like role. Chaperone proteins facilitate the correct export and folding of proteins (Gething & Sambrook, 1992; Cyr *et al.*, 1994) and it would seem feasible that the flagellar specific export pathway would

possess such chaperones. However, the previous finding that deletion of the genes encoding DnaK, DnaJ and GrpE fro *E.coli*, which are members of the Hsp70 family of chaperones, results in non-motile cells may suggest that these chaperones are the ones involved in the export of flagellar components (Shi *et al.*, 1992).

FliI may act as a 'multi-chaperone';

There appears to be good homology between the flagellar specific export components and virulence determinant export components, and FliI has been noted to have good homology to a type III secretion protein from *Yersinia* known as YscN (Harshey & Toguchi, 1996). YscN has been proposed to have ATPase activity and supply the energy required for secretion of Yop virulence determinants (Woestyn *et al.*, 1994). YscN is not, however, implicated in functioning as a chaperone, instead other proteins have been implicated (Wattiau *et al.*, 1996). These proteins thought to act as chaperones for secretion of virulence determinants are from the Syc family and include SycE, SycH and SycD all of which are thought to chaperone different proteins (Wattiau *et al.*, 1996). As with flagellar proteins, the signal for secretion of the virulence determinants is not 'obvious' i.e. sequence or structural similarities between exported proteins (Wattiau *et al.*, 1996), and this may suggest a similar mechanism of interaction between FliI and flagellin as between the Syc proteins and virulence determinants. The binding site of one of these proteins has been extensively analysed, namely SycE binding to YopE and it has been found to be localised to the N-terminal 98 residues of YopE and has been shown to be distinct to the secretion signal (Wattiau *et al.*, 1996). The existence of dedicated chaperones for each exported flagellar protein may be eliminated by the functions of FliI, which include energising the export apparatus (Silva-Herzog *et al.*, 1995) as YscN does in *Yersinia* (Woestyn *et al.*, 1994) and possibly as a chaperone (Silva-Herzog *et al.*, 1995) as does the Syc family in *Yersinia*, and also by the function of FlhB, which appears to confer specificity to the export

apparatus (Williams *et al.*, 1996). Given this and the level of homology between other flagellar genes and virulence exporters it may therefore be the case that the flagellar specific apparatus is a more sophisticated type III secretion system and that its ancestor was a virulence export system which subsequently became specialised for the flagellum and evolved into what we now see.

What is the subcellular location of FliI in *R.sphaeroides*?

The identification of the *R.sphaeroides* FliI protein may explain previous findings from studies carried out on *Rhodopseudomonas palustris*, a relative of *R.sphaeroides*, where a cytoplasmic polar organelle associated with the flagellar apparatus was observed. This organelle, was shown to have ATPase activity which was strong enough to allow cytochemical visualisation (Tauschel, 1987). This polar organelle may therefore represent the flagellar-specific export apparatus, with the FliI component producing the ATPase activity observed. The stoichiometry of FliI in *S.typhimurium* has been estimated as approximately 1500 copies per cell (Dreyfus *et al.*, 1993) which would equate to several hundred copies per flagellum. It would therefore seem reasonable to assume that, if a similar stoichiometry existed in *Rhodopseudomonas palustris* as in *S.typhimurium*, then visualisation using cytochemical technique would be quite feasible. Whether or not FliI remains attached to the base of the flagellum, associated with the export apparatus is not known. Dreyfus and co-workers (Dreyfus *et al.*, 1993) however, have suggested that, due to the high stoichiometry, that the interaction may be a transient one with a reaction cycle involving (i) binding to the flagellar base of both the substrate protein that is to be exported (i.e. flagellin) and ATP-FliI, (ii) ATP hydrolysis accompanied by protein translocation into the axial channel, and (iii) release of the ADP-FliI. This cycle is illustrated in Fig 4.10. This idea is consistent with the findings of Katayama and co-workers (Katayama *et al.*, 1996) using stereo-photogrammetry to study the cytoplasmic components associated with the flagellum in *S.typhimurium*. They found a previously

unseen rod like structure with a diameter of 9nm and height of 11nm (Fig 4.10). They proposed that this may be part of the export apparatus and that the central hole seen in the rod was the export channel. If FliI was a permanent part of this export apparatus, assuming several hundred copies per flagellum, you would expect a larger export apparatus to have been seen. The fact that this structure has never been visualised before may indicate some intrinsic instability of the association of the export apparatus with the flagellar basal body. Direct evidence for the polar localisation of FliI in *R.sphaeroides* awaits production of a FliI antibody and cytochemical studies as carried out by Tauschel.

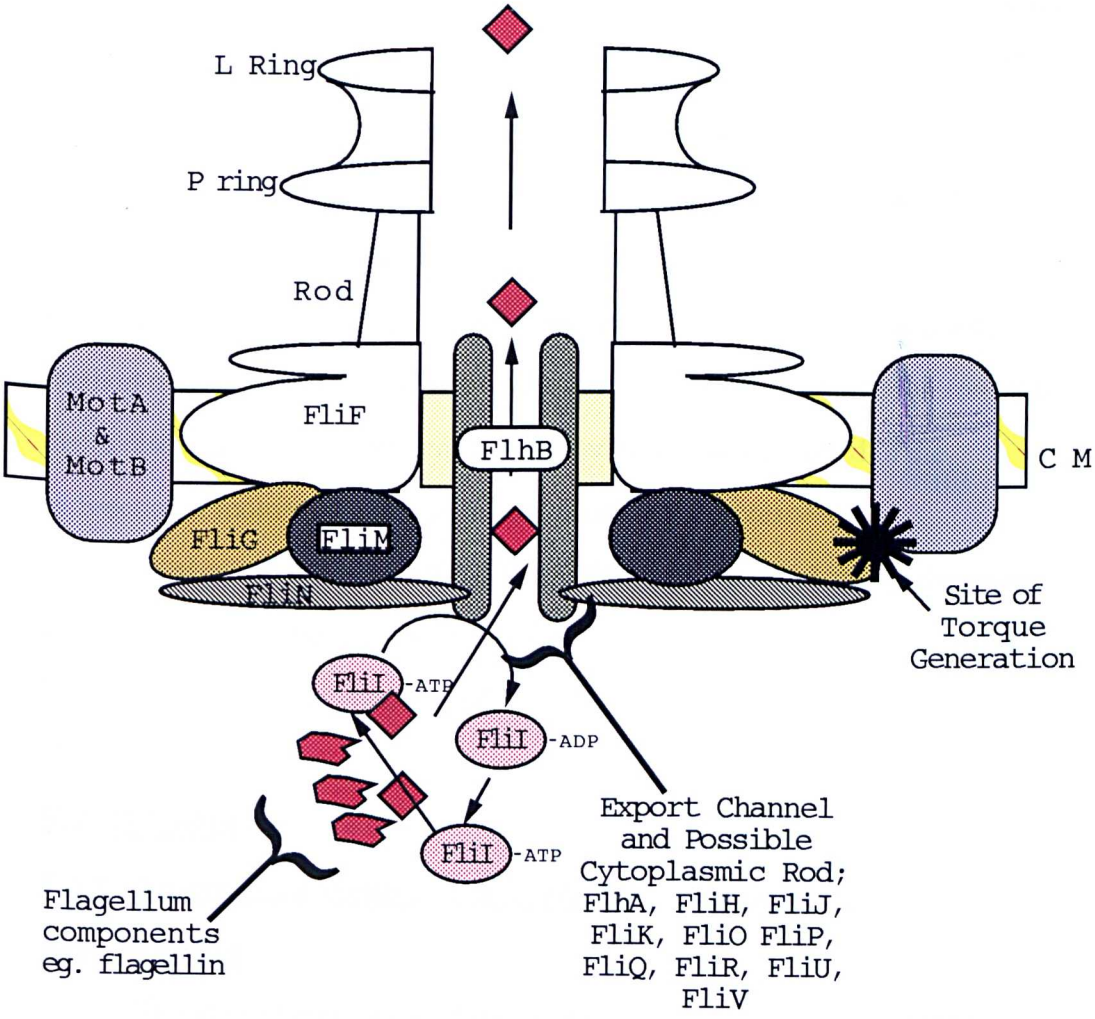
What is the nature of the export channel?

The observed rod like structure present in the central core of the cytoplasmic component of the flagellum is thought to consist of several of the export components (Fig 4.10) (Tang *et al.*, 1995; Katayama *et al.*, 1996; Macnab, 1996). The exact role of most of these proteins has not been assigned yet. Two proteins however have been extensively analysed; namely FliK and FlhB (Kawagishi *et al.*, 1996; Williams *et al.*, 1996) with the result being that the authors suggest that FlhB is the 'gate keeper' of the export apparatus conferring a specificity on the export procedure (Williams *et al.*, 1996). They suggest that FliK signals to FlhB by an unknown mechanism, so as to alter its specificity and allow export of later flagellar components. Whether or not FlhB confers specificity by binding to flagellar components is not known, but it would seem unlikely that a single protein could recognise so many different proteins that lack an 'obvious' signal sequence and it may be that FliI also participates in conferring specificity.

In conclusion, the aim of this part of the study was to characterise a mutant of *R.sphaeroides* which possessed low levels of motility with the aim of determining if it was mutated in a motor gene. The phenotype of this mutant could have been due to the truncation of a flagellar motor protein but was subsequently found to be due to the interruption of a component of the flagellar

specific apparatus. This component, identified as FliI, is proposed to be the energy supplier for the export procedure and may well function as a molecular chaperone to escort flagellum components to the base of the flagellum where they are exported to form a functioning flagellum. The identification of this protein in the uniflagellated bacterium *R.sphaeroides* may allow targeting studies of cytoplasmic components of the flagellum to be carried out as this is the first cytoplasmic component of the *R.sphaeroides* flagellum to be identified. Due to time constrictions, a more detailed analysis of FliI was not possible, but this work will presumably form the basis of a more detailed analysis of the export apparatus.

Fig 4.10 Schematic representation of the flagellum specific export procedure. Modified from ref. Tang *et. al.*, 1995.


L Ring
P ring
Rod
MotA
&
MotB
FliF
FlhB
C M
FliG
FliM
FliN
Site of
Torque
Generation
FliI -ATP
FliI -ADP
FliI -ATP
Export Channel
and Possible
Cytoplasmic Rod;
FlhA, FliH, FliJ,
FliK, FliO FliP,
FliQ, FliR, FliU,
FliV
Flagellum
components
eg. flagellin

Chapter 5

Cloning and analysis of the *fliF* and *fliG* genes from *R.sphaeroides* WS8.

5.1 Introduction

Previous work on *R.sphaeroides* has revealed the presence of the motor proteins MotA and MotB (Shah & Sockett, 1995; Shah & Sockett, 1995). The aim of this part of the project was to clone the genes encoding the remaining structures of the motor, namely the genes encoding the FliF and FliG proteins. This would enable us to characterise the interactions that take place in the unidirectional motor of *R.sphaeroides* and possibly give some insight into the method of torque generation. The reader is referred to the introduction for an in depth discussion of the functions of these proteins as determined in the bidirectional motors of *S.typhimurium, E.coli* and *C.crescentus.*

During the cloning of the *fliF* and *fliG* genes several approaches were used and these will be covered in the order that they were undertaken so as to highlight the problems encountered.

5.2 Results

5.2.1 Attempts to clone *fliF* and *fliG* by heterologous hybridisation

The use of heterologous hybridisation has been well documented as a method for cloning genes from the photosynthetic organism *R.sphaeroides* e.g.(Ward *et al.*, 1995) and it was assumed that the use of the *fliF* and *fliG* genes from *S.typhimurium* and *C.crescentus* might allow the cloning of the *R.sphaeroides* homologues.

The plasmid containing the *S.typhimurium fliF* and *fliG* genes (pAMH3) was obtained from R. Macnab (Jones *et al.*, 1989; Kihara *et al.*, 1989) and the *C.crescentus fliF* gene obtained on plasmid pGir174 from A. Newton. The *fliF* and *fliG* genes were used to probe restriction digests of *R.sphaeroides* genomic DNA, and cosmids 19, 140 and 523 which have been shown previously to contain motility genes (Sockett & Armitage, 1991; Shah & Sockett, 1995; Shah & Sockett, 1995; Shah *et al.*, 1996). This approach failed to reveal any hybridising fragments under both high and low stringency conditions (data not shown). The probes were used (independently) to probe a cosmid library of *R.sphaeroides* genomic DNA under low stringency conditions (Sockett & Armitage, 1991). Initial screenings revealed several positive clones but these were subsequently shown to be hybridising to the vector sequences or contaminating genomic DNA from *E.coli* (data not shown) so this approach was deemed of no use.

5.2.2 Attempts to clone *fliF* using functional complementation

Transposon mutagenesis and complementation is one of the more commonly used methods for the isolation of genes, however, the phenotype of a *fliF* or *fliG* mutant is non-flagellate, which is identical to the phenotype of any basal-body component mutation (see Kubori, 1992). It was beyond the scope of this project to sequence the site of mutation of every transposon mutant of *R.sphaeroides* that is non-flagellated and a more directed approach was required. An attempt was made to complement a number (20) of non-flagellate *R.sphaeroides* Tn*pho*A mutants that had been previously isolated (Sockett, 1988; Foster, 1991) with the *C.crescentus fliF* gene. It was decided not to use the *S.typhimurium* genes as the codon bias of *R.sphaeroides* is significantly different than that from *S.typhimurium*, as the G/C content of the DNA from *R.sphaeroides* is in the order of 70%.

The *C.crescentus fliF* gene was cloned into the broad host-range vector pRK415 (Keen *et al.*, 1988) in two orientations, to allow expression of the gene from the external tetracycline resistance gene promoter and from the internal *C.crescentus fliF* promoter. These constructs were subsequently conjugated into each of the 20 non-flagellate Tn*phoA* mutants and the mutants examined for motility by phase contrast microscopy and swarm plates. The *C.crescentus fliF* gene did not complement any of the non-flagellate mutants back to motility and this was not taken further.

5.2.3 Attempts to clone *fliF* and *fliG* using insertional inactivation

As the previous two approaches failed, another approach was undertaken that assumed that the genes encoding FliF and FliG would be clustered on the chromosome with other motor genes, namely *motA* and *motB*, as the proteins they encode are physiologically clustered in the flagellar basal body. This is not found in any other organisms studied to date as the *motA* and *motB* genes from *E.coli* are clustered with chemotaxis genes (Slocum & Parkinson, 1983), but the structure of the *motA/B* operon had not been investigated enough to determine if any flagellar genes lay upstream of downstream.

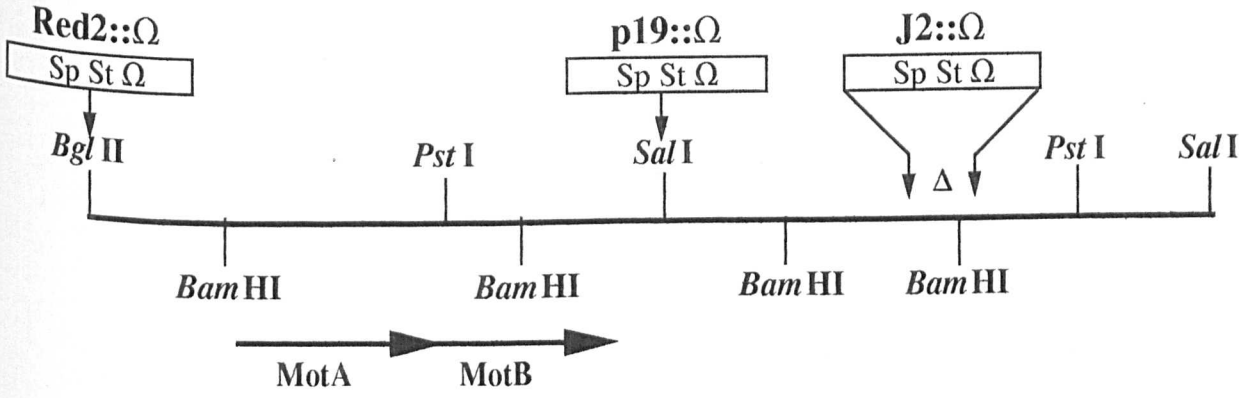
The structure of the *motA/B* operon is shown in Fig 5.1, and the insertions created are also depicted. The methods used to create these insertions is similar to that used in chapter 3 &4 to create the *rpoN* and *fliI* mutations. It is therefore not necessary to demonstrate how they were constructed, however, the production of 19::Ω utilised a novel method of creating a suicide vector and will therefore be discussed below.

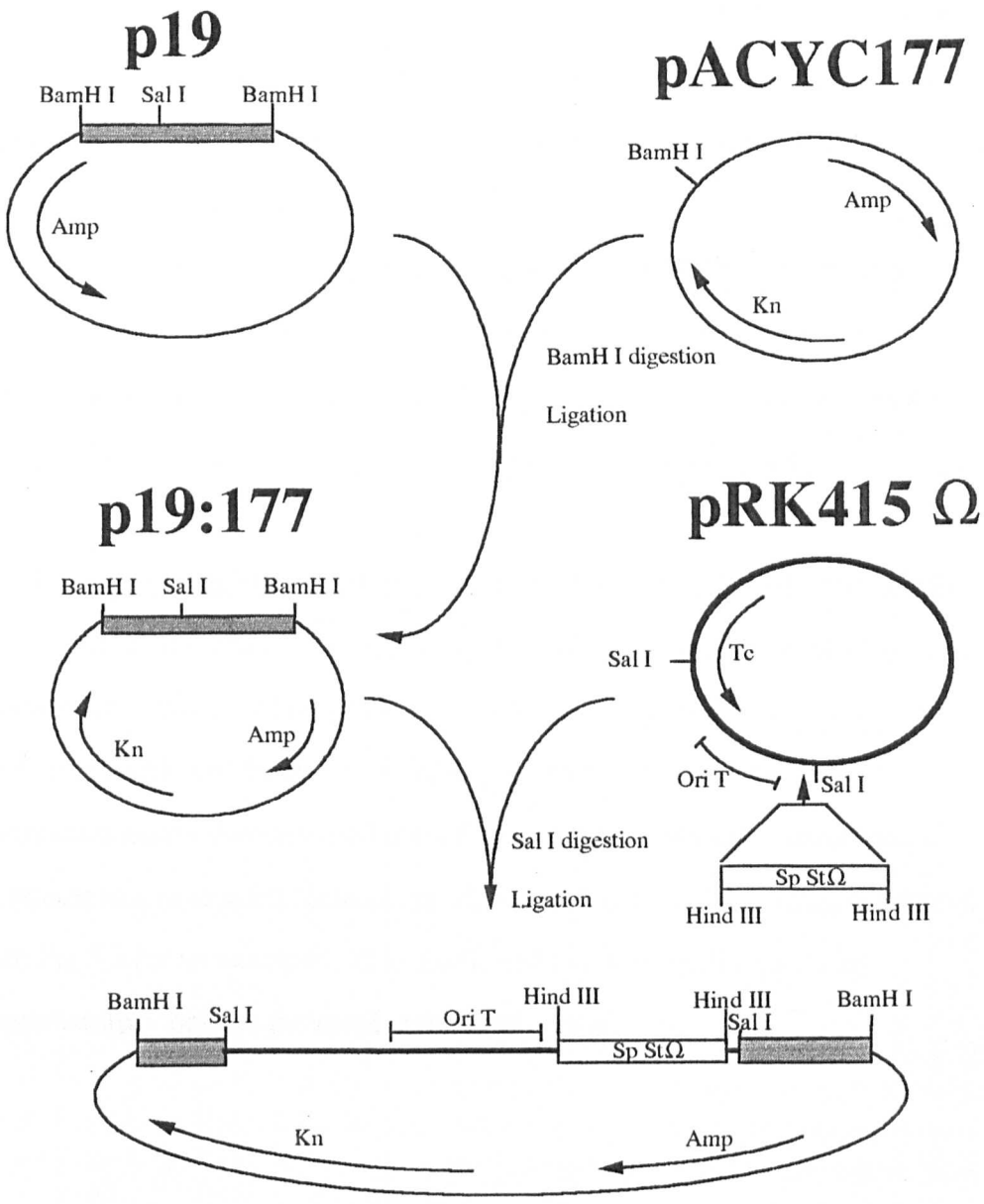
5.2.3a Production of 19::Ω

The *Bam*HI fragment downstream of the *motA/B* operon (p19 Fig 5.1) was identified by Southern blot analysis of cosmid 19 digests using the 600bp

Fig 5.1 The structure of the *motA/B* operon. The MotA/B open reading frames are shown underneath the restriction map as arrows. The sites of the Ω cartridge insertions created are shown. Δ represents a deletion created during the construction of J2::Ω.

Fig 5.2 Construction of the construct p19:177Ω used for creating the Ω insertion strain 19::Ω. The ampicillin resistance gene is shown as Amp, kanamycin resistance as Kn, tetracycline resistance as Tc, spectinomycin resistance as Sp and streptomycin resistance as St. The shaded box represents the region cloned from the *motA/B* operon. OriT is the origin of transfer from pRK415.


Red2::Ω
Sp St Ω
p19::Ω
Sp St Ω
J2::Ω
Sp St Ω
Bgl II
Pst I
Sal I
Δ
Pst I
Sal I
Bam HI
Bam HI
Bam HI
Bam HI
MotA
MotB


p19
BamH I
Sal I
BamH I
Amp
pACYC177
BamH I
Amp
Kn
BamH I digestion
Ligation
p19:177
BamH I
Sal I
BamH I
Kn
Amp
pRK415 Ω
Sal I
Tc
Ori T
Sal I
Sal I digestion
Ligation
Sp StΩ
Hind III
Hind III
BamH I
Sal I
Ori T
Hind III
Hind III
Sal I
BamH I
Sp StΩ
Kn
Amp

p19:177 Ω

*Bam*HI to *Sal*I fragment from the 3' region of the *motA/B* operon (Fig 5.1). This fragment was cloned into pUC19 and then pACYC177 (Fig 5.2) to use the kanamycin resistance gene (from pACYC177) as a marker for a single cross-over event. To facilitate the conjugative transfer of the construct (p19:177) from *E.coli* into *R.sphaeroides*, the origin of transfer of pRK415 (Keen *et al.*, 1988) was cloned into the unique *Sal*I site internal to p19:177 along with the Ω cartridge from pHp45Ω (Prentki & Krisch, 1984), which had been previously cloned into the *Hin*dIII site of pRK415 (Fig 5.2). This plasmid, p19:177Ω (Fig 5.2), was a suitable construct to allow the construction of 19::Ω as it contained an Ω cartridge to terminate transcription and translation, an origin of transfer to allow conjugation into *R.sphaeroides*, and kanamycin resistance to allow screening for a single cross-over event.

p19:177Ω was conjugated into *R.sphaeroides WS8N* and double cross-over mutants isolated as kanamycin sensitive, spectinomycin and streptomycin resistant colonies. Mutants (19::Ω) were isolated and analysed using phase contrast microscopy and plug plate chemotaxis assays.

5.2.3b Characterisation of the mutants Red2::Ω, 19::Ω and J2::Ω

All three mutants were grown photosynthetically and their motility characteristics examined by phase contrast microscopy. All three strains were seen to be motile and behave with wild-type characteristics. Plug plate chemotaxis assays demonstrated that all three mutants possessed chemotactic responses to amino acids identical as wild-type *R.sphaeroides* (data not shown - see Fig 3.2 for an example). This confirmed that no flagellar genes lay immediately 5' or 3' to the *motA* and *motB* genes.

5.2.4 The use of 'operon analogy' to clone *fliF* and *fliG*

Due to the failure of all cloning strategies previously attempted, a new strategy based on 'operon analogy' was used as a final method to clone *fliF* and *fliG*. This strategy relies on the fact that in *S.typhimurium*, *B.subtilis* and *E.coli* the *fliF* and *fliG* genes are upstream of the *fliI* and *fliK* genes (Zuberi *et al.*, 1991; Kawagishi *et al.*, 1992) and that a similar case may exist in *R.sphaeroides*. Prior to this project, it had been shown that a mutant with the characteristics of a *fliK* mutant (i.e. long hook structure) mapped to cosmid 140 (Saunders, 1993). Upstream of this region was shown to be a region of DNA that hybridised to the *fliI* gene from *S.typhimurium* (Pollitt, 1996), which was subsequently shown to be the *fliI* homologue from *R.sphaeroides* (see chapter 4). This order was identical to that in *S.typhimurium* and so this therefore confirmed that at least part of the operon was present in the same structure as in other systems, and the finding that a non-flagellate mutant (Nm7) mapped to a region upstream of *fliI* and *fliK* (Foster, 1991) made it a good candidate for a *fliF* or *fliG* mutant. As will be shown in this section, this characterisation of the mutant Nm7 showed that this was in fact the case.

5.2.4a Characterisation of Nm7

Foster had previously shown that NM7 was complemented back to wild-type levels of motility by a 5.5Kb *Eco*RI fragment of cosmid 140 (Foster, 1991), and that this fragment lay upstream of the 5.9Kb *Eco*RI fragment to which *fliI* and *fliK* were mapped (this study, Foster, 1991 and Saunders, 1993). In this work I have subsequently shown that the 5.5Kb *Eco*RI fragment and cosmid 140 did not complement Nm7 and that the findings of Foster were incorrect.

5.2.4b Isolation of the wild-type motility gene interrupted in Nm7

To further characterise Nm7, the site of transposon insertion in Nm7 had to be examined and the region of DNA flanking the transposon sequenced. The region of DNA flanking the site of the transposon insertion was cloned by digesting chromosomal DNA from Nm7 with *Sal* I and ligating to *Sal* I digested pUC19. Selection for the Tn*pho*A-encoded kanamycin resistance allowed the isolation of pNM7S which contained 600bp of Nm7 genomic DNA in addition to the kanamycin resistance gene from Tn*pho*A. The *Sal* I to *Dra*I fragment of this clone, which lacked the kanamycin resistance gene, was used to probe digests of Nm7 genomic DNA, cosmid 140, cosmid 19 and cosmid 523 as shown in Fig 5.3. The difference in size between the fragments seen to hybridise in Nm7 and wild-type WS8 genomic DNA was due to the restriction sites within Tn*pho*A. The fact that the hybridising fragments seen in WS8 genomic DNA were different in size to those in cosmid 140 digests suggested that the site of mutation lay off the end of the region of DNA covered by cosmid 140. This was consistent with the finding that the 5.5Kb *Eco*RI fragment of cosmid 140 contained the left-hand arm of the cosmid and consequently the end of the cosmid (Foster, 1991).

To clone the wild-type motility gene it was necessary to identify an additional cosmid(s) that overlapped with cosmid 140 and contained the upstream region of DNA from the *R.sphaeroides* WS8 genome. The pNM7S probe was found to hybridise to two groups of eight cosmids, number 89 and number 95 to give a 2.2Kb hybridising fragment (data not shown). The cosmids within these groups were isolated and one cosmid, number 709 was found to contain no inserted DNA and was therefore discarded. The remaining cosmids (number 705 - 712 and 753 - 760) were digested with *Bam*HI and probed using the pNM7S clone. The probe hybridised to two cosmids, numbers 711 and 753 (Fig 5.4, lanes 7 and 9 respectively). Both cosmids shared a 2.2Kb *Bam*HI fragment that hybridised to the pNM7S probe but the degree of overlap of these cosmids could not be ascertained from this.

To determine the degree of overlap a number of restriction digests were carried out and they were analysed by Southern blotting with various probes, namely pNM7S *Sal* I to *Dra* I fragment, a 1Kb *Bam*HI fragment from *fliI* (see chapter 4, Fig 4.4), pLA2917 left-hand probe and pLA2917 right-hand probe. An example of the results can be found in Fig 5.5 as well as an example of how the overlap was detected. These results showed that cosmids 711 and 753 overlapped with cosmid 140 by approximately 4.5Kb and that both cosmids 711 and 753 did not contain *fliI* (See Fig 5.6). Comparison of digests of pNM7S and cosmid 140, showed that the site of Tn*pho*A insertion in Nm7 lay approximately 120bp upstream of the 5' *Bam*HI site on cosmid 140 (TnNM7 on Fig 5.6).

5.2.4c Complementation analysis of Nm7

As stated in section 5.2.4a, the 5.5Kb *Eco*RI fragment from cosmid 140 and cosmid 140 itself were unable to complement Nm7 back to wild-type motility. This suggested that the gene/genes interrupted in Nm7 overlapped with the 5' end of cosmid 140. To test this theory, the 5' *Sal* I fragment overlapping cosmids 140, 711 and 753 (see Fig 5.6) was cloned into pRK415 in two orientations with respect to the vector tetracycline promoter. These clones (pIG10a and pIG10b) were conjugated into Nm7 and exconjugants were found to be non-motile. Cosmids 711 and 753 were conjugated into Nm7 and found to complement Nm7 very poorly i.e. approx 1 in 1000 cells were motile and behaved with wild-type characteristics. As this phenotype was similar to that observed in a *fliI* mutant strain (see chapter 4), and the *fliI* gene lay downstream of the region interrupted in Nm7, it was possible that the gene/genes interrupted in Nm7 was/were in an operon with *fliI*, and that the lack of complementation was due to the polar effects of the transposon

Fig 5.3 Southern blot analysis using the transposon flanking clone of Nm7. Left is a photograph of the digested DNA and right showing the corresponding blot. Lanes 1 and 4 contains WS8 genomic DNA; 2 and 5 Nm7 genomic DNA; 3 and 6 genomic DNA from M18 (see chapter 4) as an additional control; lanes 1,2 and 3 were digested with *Eco*RI; lanes 4 , 5 and 6 were digested with *Sal* I; lane 7 contained cosmid 19 DNA digested with *Sal* I; lanes 8,9 and 10 contained cosmid 140 DNA digested with *Bam*HI, *Eco*RI and *Sal* I respectively; lane 11 contained cosmid 523 digested with *Sal* I and *Eco*RI and lane 12 contained cosmid 688 DNA digested with *Bam*HI. Cosmids 523 and 688 have been shown to contain unidentified motility related genes (Sockett, Pers. Com.) M is λ *Bst*EII digested DNA. Sizes are in Kb. The probe hybridised to an 8.3Kb *Eco*RI fragment and a 3.9Kb *Sal*I fragment from WS8 genomic DNA (lanes 1 & 4 respectively) and also cosmid 140 ; 2.3Kb *Bam*HI fragment (lane 8), 5.5Kb *Eco*RI fragment (lane 9) and a 4.2Kb *Sal* I fragment (lane 10). It was found to hybridise to a 4.6Kb *Eco*RI fragment and a 6Kb *Sal* I fragment in Nm7 genomic DNA (lanes 2 and 5 respectively).

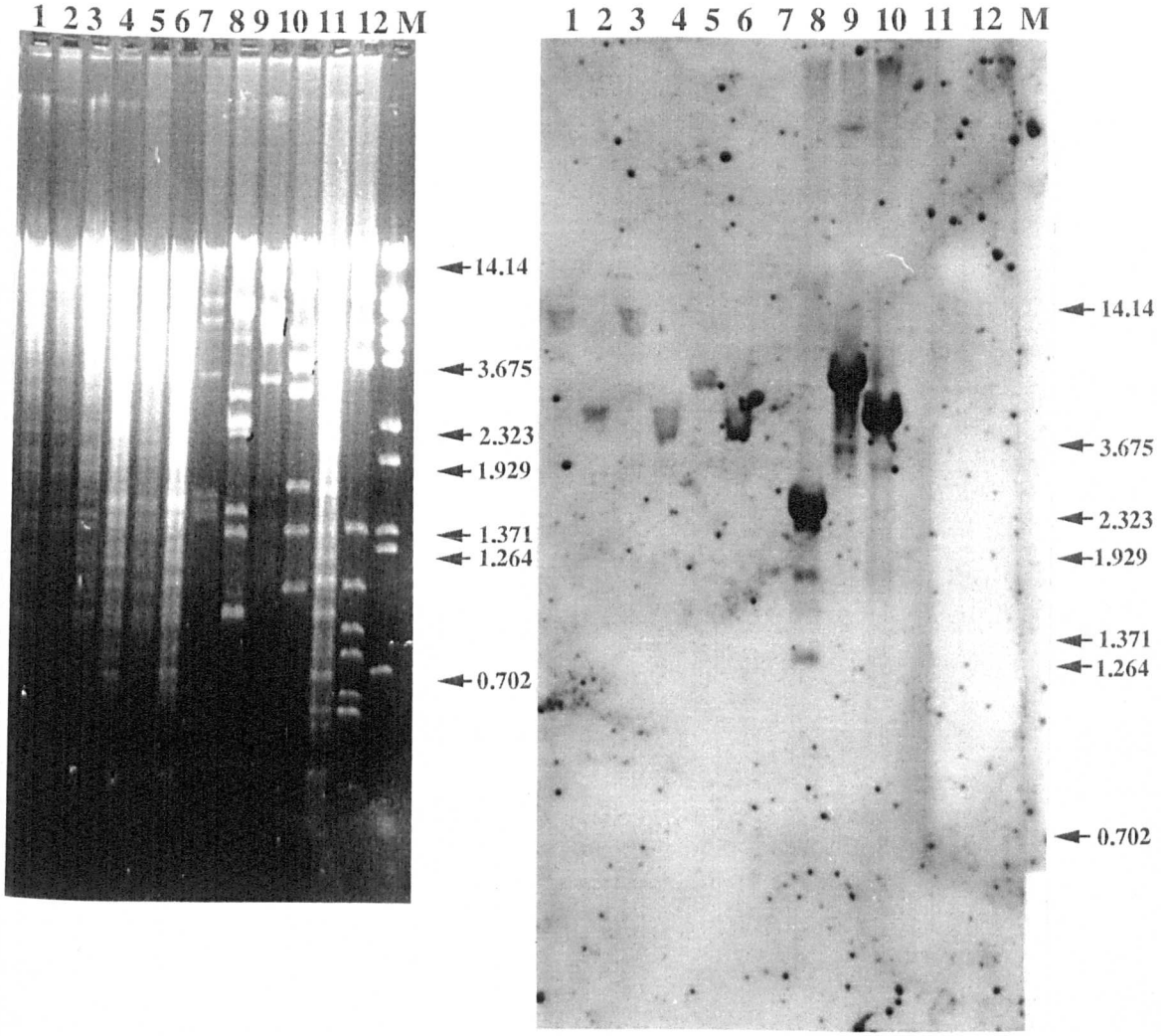
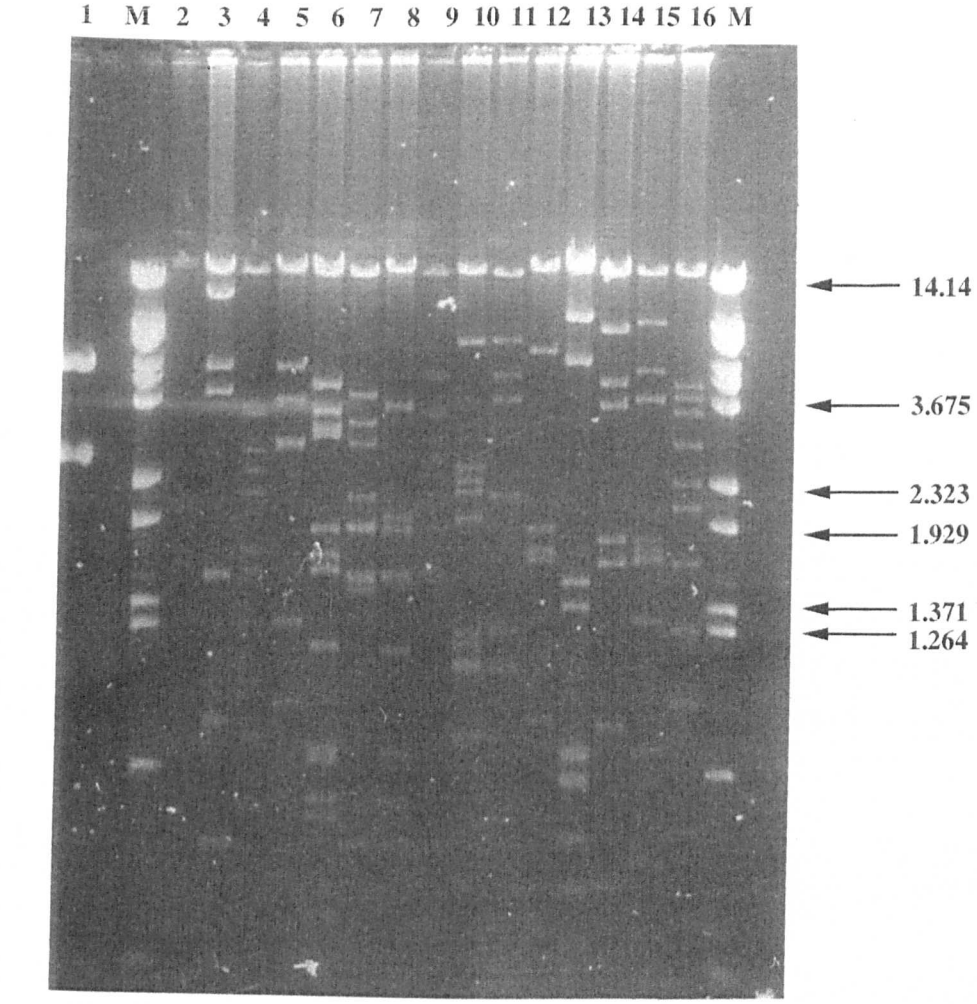

1 2 3 4 5 6 7 8 9 10 11 12 M
14.14
3.675
2.323
1.929
1.371
1.264
0.702
1 2 3 4 5 6 7 8 9 10 11 12 M
14.14
3.675
2.323
1.929
1.371
1.264
0.702

Fig 5.4 Southern blot analysis of cosmid groups 89 and 95 with pNM7S *Sal*I to *Dra*I fragment. Lane 1 contained the cosmid 140 5.5Kb *Eco*RI fragment cloned in pUC19; 2 - 8 contained cosmids 705, 706, 707, 708, 710, 711 and 712 respectively; lanes 9 - 16 contained cosmids 753, 754, 756, 757, 758, 759, 760 and 761 respectively. Cosmids were digested with *Bam*HI. M was λ*Bst*EII digested DNA. Sizes are in Kb.


1 M 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 M
14.14
3.675
2.323
1.929
1.371
1.264

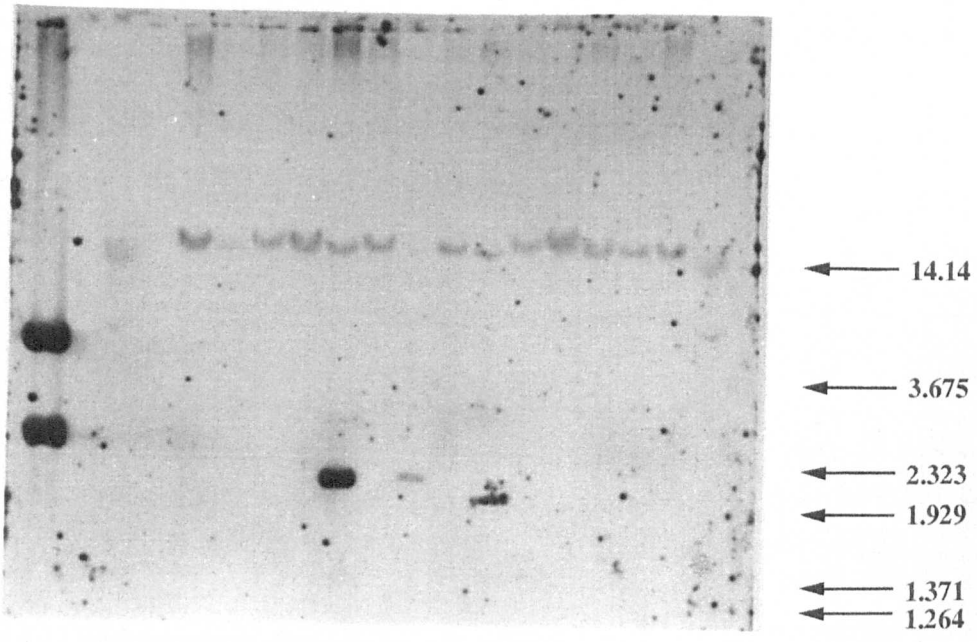

14.14
3.675
2.323
1.929
1.371
1.264

Fig 5.5 Southern blot analysis of cosmids 140, 711 and 753. A) Restriction digests of cosmids 711 (lanes 1,4 & 7), 140 (lanes 2,5 & 8) and 753 (lanes 3, 6 & 9) with *Sma*I (lanes 1-3), *Sma*I and *Bst*EII (lanes 4-6), and *Bst*EII and *Eco*RI (lanes 7-9). B - E are the results of Southern blot analysis of A) using; B) pNM7S probe (lacking the kanamycin resistance gene); C) *fliI* internal *Bam*HI fragment (see chapter 4); D) pLA2917 cosmid right-arm probe and E) pLA2917 cosmid left-arm probe. For simplicity only lanes required to illustrate the points described below are shown. Also, due to variability in printing size of blot photographs, some of the hybridising fragments in the Southern blot analysis do not perfectly correspond with their counterpart in other blots or A, but they do in fact correlate on the original films. As can be seen in B lane 2, the pNM7S probe hybridised to a fragment of approx. 3.55Kb, the pLA2917 left-arm probe also hybridised to this fragment (E, lane 2) demonstrating that this is the end of the cosmid. In B, lanes 1 & 3 hybridised to fragments of approx. 2Kb, the same as seen *R.sphaeroides* genomic DNA (data not shown) demonstrating that cosmids 711 and 753 contain the wild-type copies of the interrupted region in Nm7. From C, lanes 1, 4 & 7, we can see that the *fliI* probe does not hybridise to the same size fragments in cosmid 711 as in cosmid 140 (lanes 2, 5 & 8) and also that cosmid 753 did not hybridise at all. As cosmid 140 contained the wild-type copy of *fliI* (see chapter 4), this confirms that cosmid 711 contains only part of *fliI* and 753 did not contain it at all. The fact that the fragments in cosmid 711 seen to hybridise to the *fliI* probe, also hybridise to the pLA2917 right-arm probe (D, lanes 1, 4 & 7) confirm that in cosmid 711, *fliI* is at the right-arm of the cosmid with only part of it present. D & E both demonstrate that, in cosmid 753 (lane 9 in D and 6 in E), duplicate copies of the cosmid vector may be present, as cosmid 711 and 140 contain less hybridising fragments i.e. in D cosmid 711 and 140 hybridise once and 753 twice and in E cosmids 711 and 140 hybridise twice and 753 three times. Sizes are in Kb.

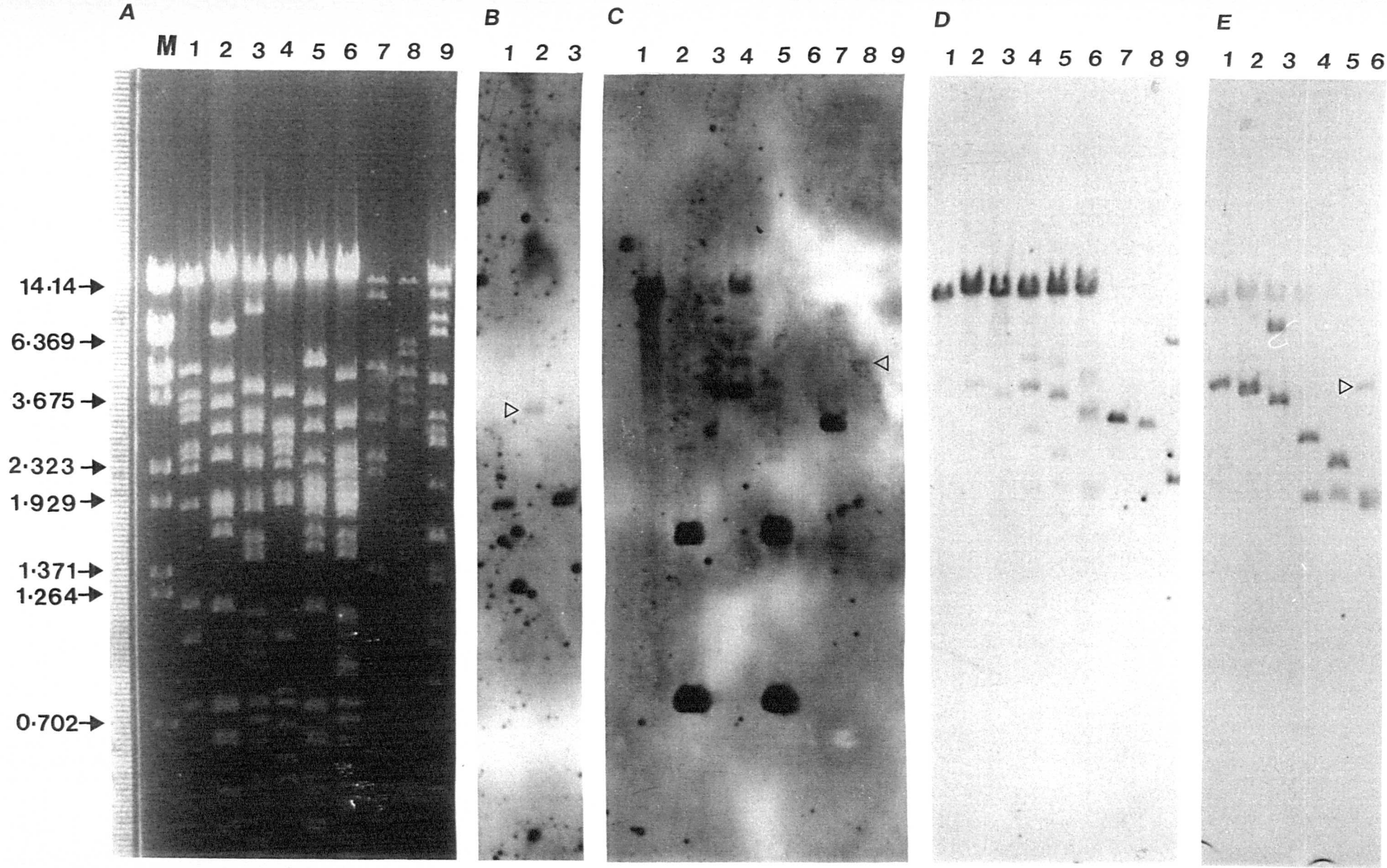
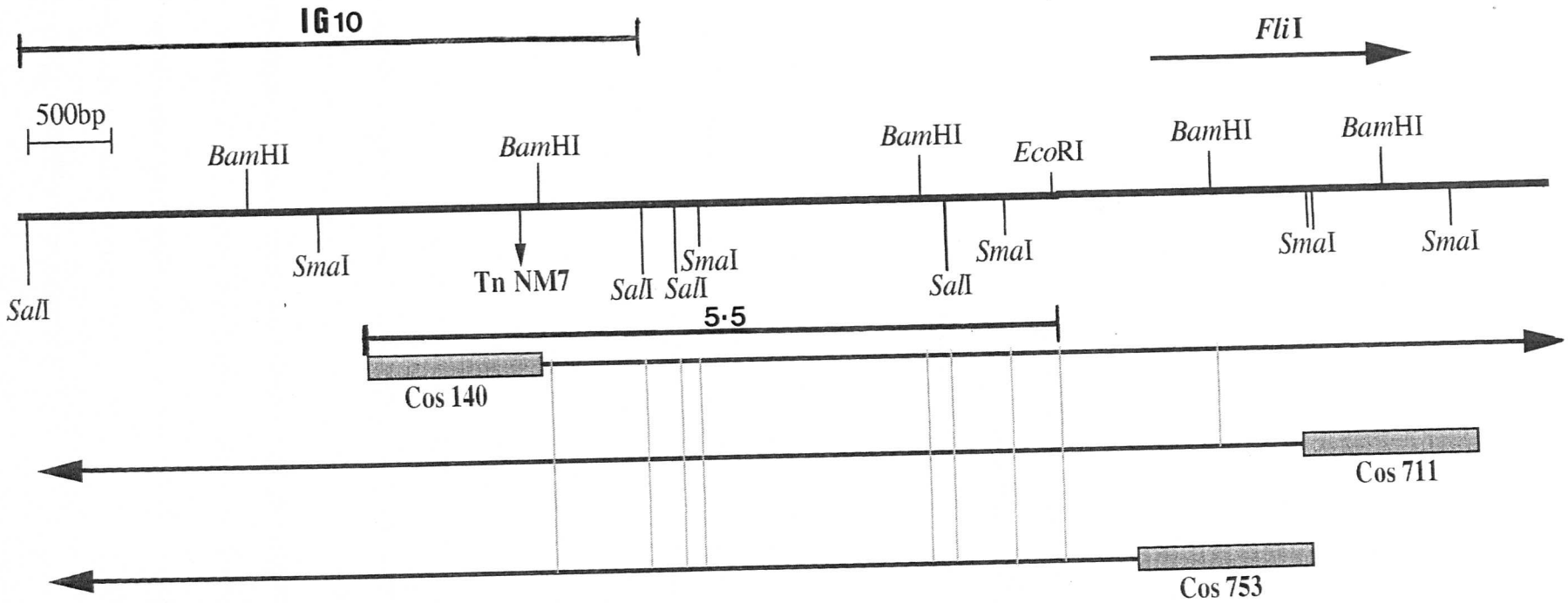

A
M 1 2 3 4 5 6 7 8 9
14·14
6·369
3·675
2·323
1·929
1·371
1·264
0·702
B
1 2 3
C
1 2 3 4 5 6 7 8 9
D
1 2 3 4 5 6 7 8 9
E
1 2 3 4 5 6

Fig 5.6 Restriction map of cosmids 140, 711 and 753 showing degree of overlap between them. The cosmid arms are shown as shaded boxes. The site of Tn*pho*A insertion in Nm7 is shown as Tn NM7. The regions contained in the plasmids pIG10 (a & b) and p5.5, described in the text, are also shown.


IG10
FliI
500bp
BamHI
BamHI
BamHI
EcoRI
BamHI
BamHI
SmaI
SmaI
SmaI
SmaI
SalI
Tn NM7
SalI
SalI
SalI
5·5
Cos 140
Cos 711
Cos 753

on *fliI*. The lack of an intact *fliI* gene on cosmids 711 and 753 may explain the phenotype observed as the cell would not contain FliI and subsequently show characteristics of a *fliI* mutant. Electronmicroscopic analysis of Nm7 and Nm7X cosmids 711 and 753 also showed similar structures to those present in the *fliI* mutants studied in chapter 4 (Fig 5.7 and Fig. 4.9).

5.2.4d DNA sequence analysis of the region interrupted in Nm7

The DNA sequence of regions flanking the site of transposon insertion in Nm7 was determined on both strands using subclones and custom synthesised primers as shown in Fig 5.8. The 4618bp contig was translated in all possible reading frames and the open reading frames used to search the Swissprot protein data base at Seqnet, Daresbury, Uk, using the FASTA program from the GCG package (Deveraux *et al.*, 1984). Four open reading frames were identified that possessed homology to previously isolated flagellar related proteins (Fig 5.8).

The site of Tn*pho*A insertion in NM7 was found to lie within a gene encoding a polypeptide of 60.7KDa, by DNA sequencing of pNM7S, which had extensive homology to previously studied FliF proteins (Zuberi *et al.*, 1991; Ueno *et al.*, 1992; Ramakrishnan *et al.*, 1994; Matsumurra, 1995; Arora *et al.*, 1996) and will be referred to as RSFliF herewith. Upstream was another open reading frame, with a predicted MW of 11.26 KDa with homology to previously studied FliE proteins (Zuberi *et al.*, 1991; Muller *et al.*, 1992; Arora *et al.*, 1996) and will be referred to as RSFliE herein. Further upstream of *fliF* was an open reading frame which encoded a protein of 40KDa with good homology to response regulators such as FlbD and FleR (Ramakrishnan & Newton, 1990; Richings *et al.*, 1995) which are known to be transcriptional regulators of flagellar gene expression in *C.crescentus* and *P.aeruginosa* respectively. This open reading frame will be referred to as Torf. Downstream of the gene encoding RSFliF was an open reading

Fig 5.7 Electron micrographs of Nm7 containing cosmid 753 (A) and cosmid 711 (B & C). Cell were negatively stained with 2% uranyl-acetate pH 7. Bar = 200 nm

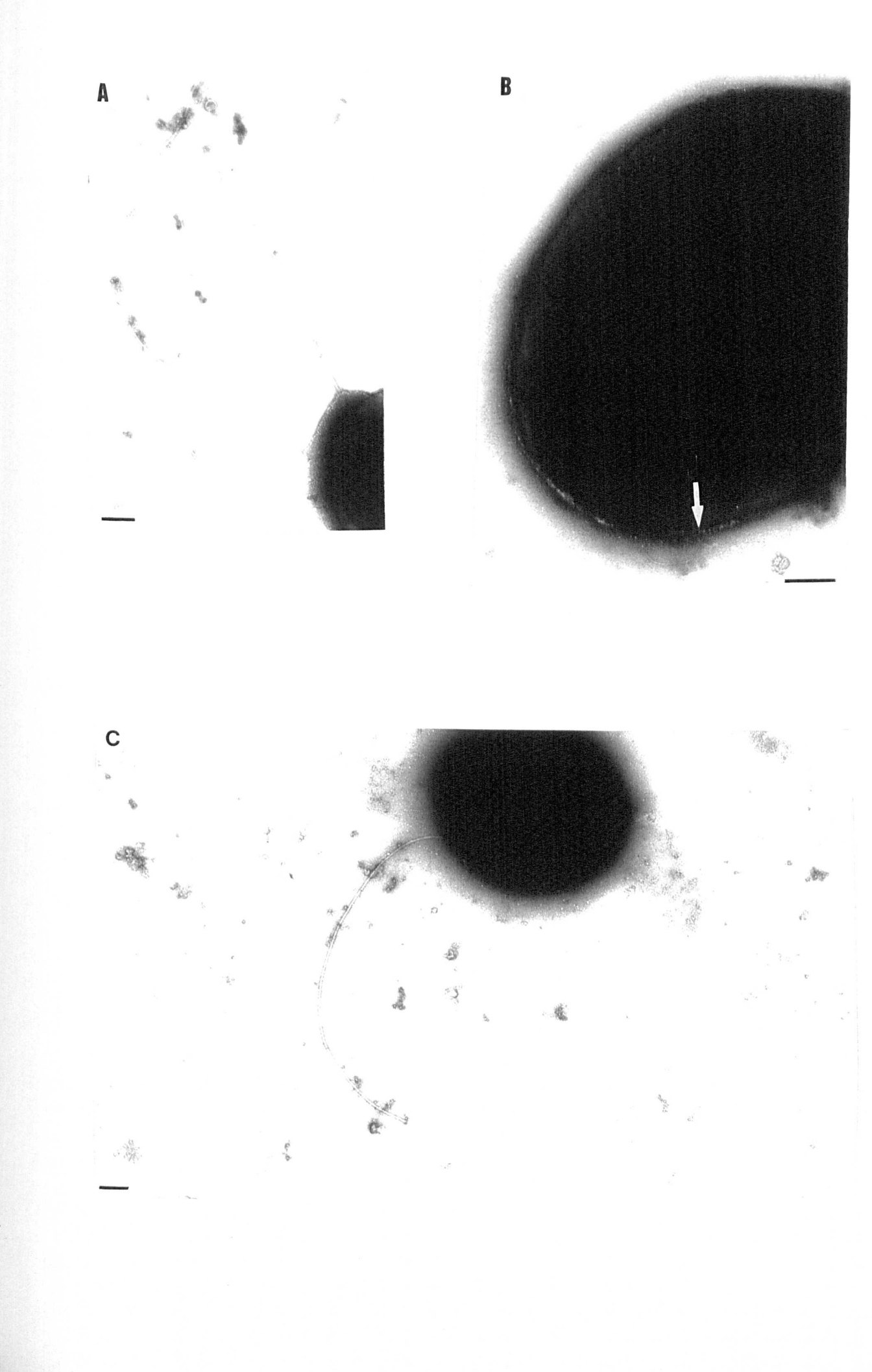
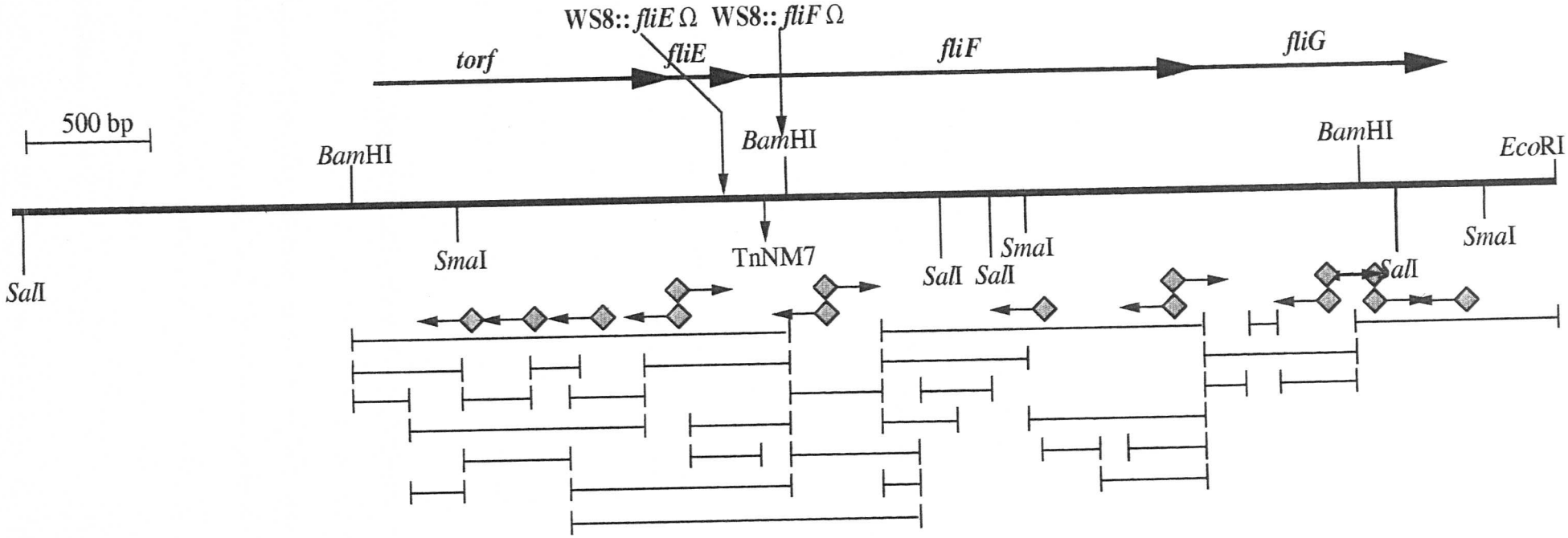

A
B
C

Fig 5.8 Sequencing strategy used to determine the interrupted gene in Nm7. The site of Tn*phoA* insertion in Nm7 is shown as Tn NM7 and the site of interposon insertion in the mutants WS8::*fliE*Ω and WS8::*fliF*Ω is also shown. Subclones used for sequencing are shown under the restriction map as ⊢——⊣ and custom made primers are shown as shaded diamonds with the arrow denoting the direction of priming of the primer.


WS8::fliE Ω
WS8::fliF Ω
torf
fliE
fliF
fliG
500 bp
BamHI
BamHI
BamHI
EcoRI
SalI
SmaI
TnNM7
SalI
SalI
SmaI
SmaI
SalI

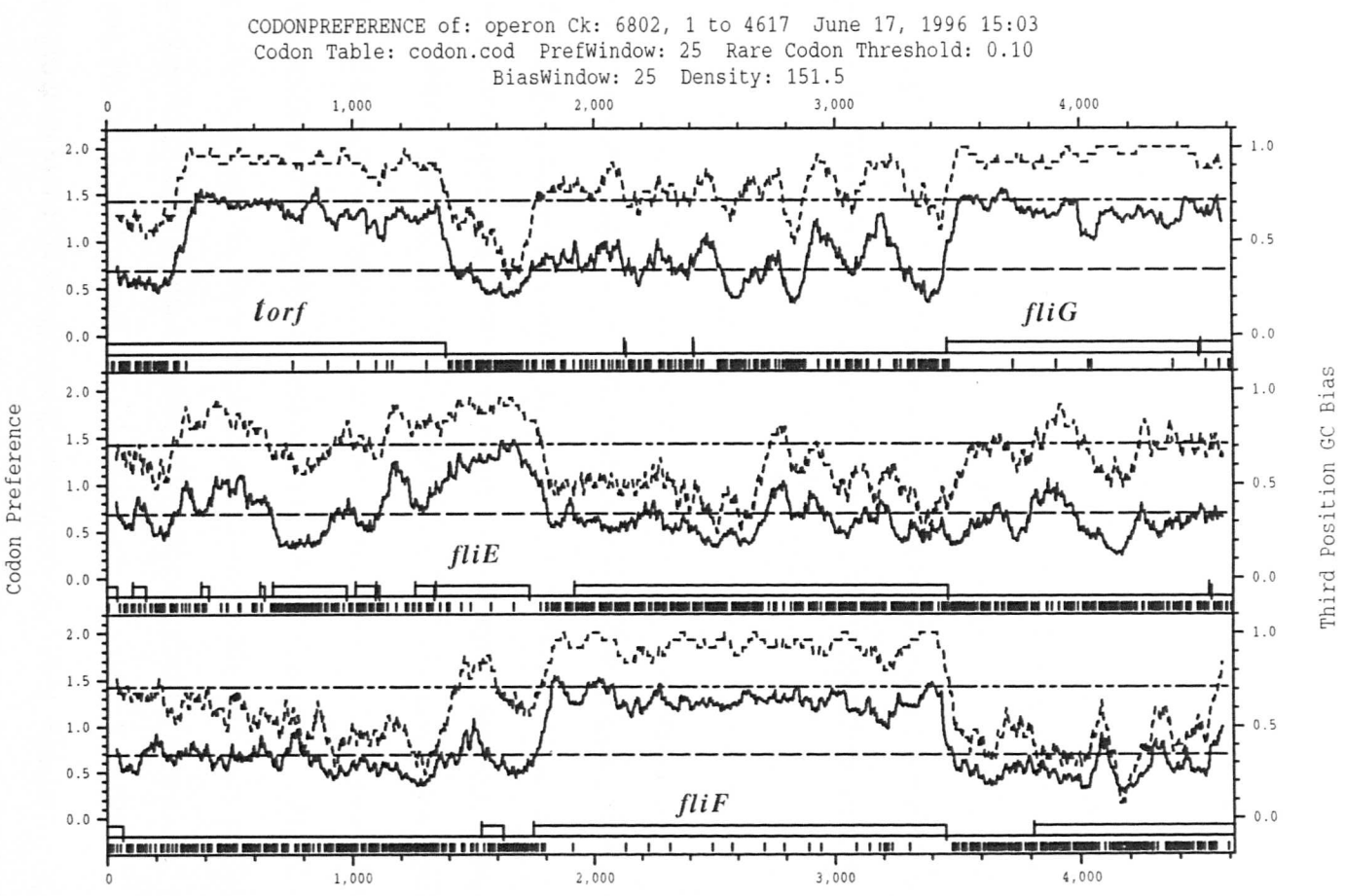
frame encoding a protein of 37.2 KDa with homology to previously studied FliG proteins (Kihara *et al.*, 1989; Zuberi *et al.*, 1991; Ramakrishnan *et al.*, 1994; Arora *et al.*, 1996) to be referred to as RSFliG. The sequences were submitted to the EMBL nucleotide sequence database under the following accession numbers: *torf* X98694; *fliE* X98693; *fliF* X98692 and *fliG* X98691. The characteristics of these open reading frames will be discussed separately below.

The codon usage of the genes was examined and compared to previously studied *R.sphaeroides* genes. Using the CODONFREQUENCY program of the GCG package (Deveraux *et al.*, 1984) a table was produced using the *R.sphaeroides* genes in the EMBL database held at Seqnet, Daresbury, UK. This table can be found in appendix 2. The CODONPREFERENCE program (Deveraux *et al.*, 1984) was used, with the table to determine if the codons within the open reading frames confer to the bias of previously isolated *R.sphaeroides* genes. As can be seen from Fig 5.9, the codon bias and the third position G-C bias of all the open reading frames is similar to other *R.sphaeroides* genes. The *torf* gene is in frame with the upstream DNA i.e. it is possible that the gene extends further upstream. However, based on the high number of rare codons within this region (Fig 5.9), the presence of a promoter sequence in this region and the homology to other regulator (see later), I have predicted that the gene does not extend further upstream.

5.2.4e Characteristics of the *torf* gene from *R.sphaeroides*

The DNA from the region containing the *torf* gene and the upstream region is shown in Fig 5.10. The ATG initiation codon for the Torf open reading frame is preceded by a potential Shine-Dalgarno ribosome binding site (Shine & Dalgarno, 1974) of GAGG. 30 base pairs upstream of this region is a σ^{54} consensus sequence (Fig 5.10). The sequence is identical to the *nifH* promoter from *R.capsulatus* which

Fig 5.9. Codonpreference plot indicating translated regions of the *R.sphaeroides* sequence. A statistical plot of the codon utilisation (lower line) and third place G-C bias (upper dotted line) for each reading frame is shown in the three graphs. The corresponding open reading frames are shown above the corresponding open boxes which denote potential open reading frames. The presence of rare codons is illustrated by vertical dashes below the open reading frames.

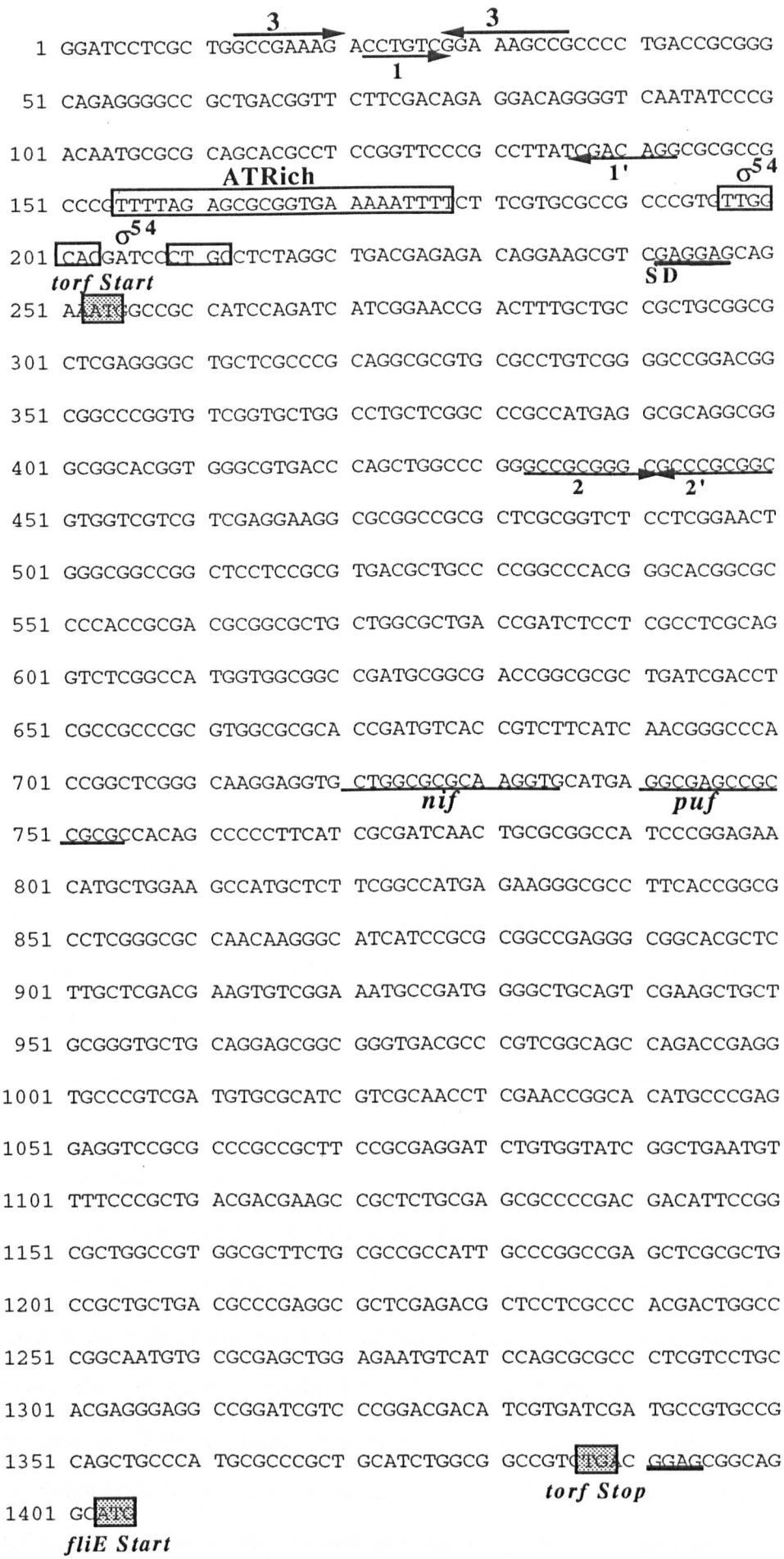

CODONPREFERENCE of: operon Ck: 6802, 1 to 4617 June 17, 1996 15:03
Codon Table: codon.cod PrefWindow: 25 Rare Codon Threshold: 0.10
BiasWindow: 25 Density: 151.5
Codon Preference
Third Position GC Bias
0
1,000
2,000
3,000
4,000
2.0
1.5
1.0
0.5
0.0
1.0
0.5
0.0
torf
fliG
fliE
fliF

also has been shown to be active *in vivo* (Foster-Hartnett & Kranz, 1992). There is a series of inverted repeats and other repeats upstream of and within the *torf* gene. These repeats (number 1, 2 and 3 in Fig 5.10) do not show any similarity with any previously discovered transcriptional control elements. The promoter region has a very high A/T content i.e. 36.5% whereas the remainder of the operon contains only 28.3% A/T.

Regions within the *torf* gene sequence share homology to regulatory sequences previously studied. These sequences, labelled *nif* and *puf* in Fig 5.10 share homology to the *K.pneumoniae* σ^{54} consensus sequence (Arnold *et al.*, 1988) and region of DNA known to be the binding site of a transcriptional regulator in *R.sphaeroides* present under aerobic conditions (Shimada *et al.*, 1993).

The predicted protein sequence of the *torf* gene shows a high level of homology to members of the response regulators in the 'two-component' systems reviewed in Stock *et. al.* 1989 (Stock *et al.*, 1989) as shown in Fig 5.11. More specifically, Torf shows homology to the members of the σ^{54} enhancer binding proteins (EBP's) (for reviews see Morett & Segovia, 1993; Shingler, 1996). The fact that it is controlled by a σ^{54} promoter may suggest that it is autoregulated. There is a high degree of conservation of the conserved clustered residues noted by Morett and Segovia (Morett & Segovia, 1993) (1-7 in Fig 5.11). It is apparent from the alignment shown in Fig 5.11, that the Torf protein does not contain a helix-turn-helix DNA binding domain at its C-terminus. No DNA-binding motif was detected using the HELIX-TURN-HELIX program in the GCG package (Deveraux *et al.*, 1984). Residues important for the function of the chemotaxis protein CheY (shown in Fig 5.11), which is related to the EBP's that are regulated by phosphorylation, namely D12, D13, D57, T87 and K109 appear to be poorly conserved in the Torf protein. The residue involved in receiving the phosphate group (D57) although

Fig 5.10 DNA sequence of the *R.sphaeroides torf* gene, showing the upstream promoter region. Inverted repeats are shown with an arrow underneath (1, 1', 2 & 2') as well as other repeated sequences (3). The potential ribosome binding site is shown as SD, with the potential σ^{54} promoter sequence boxed. AT Rich denotes the A/T rich region described in the text. *nif* and *puf* denote potential regulatory sequences described in the text.


3
3
1
1'
ATRich
σ54
σ54
torf Start
SD
nif
puf
2
2'
torf Stop
fliE Start

conserved in various other proteins in the alignment is not conserved in Torf.

The Torf protein has the highest degree of homology with FleR and FlbD, as demonstrated in Fig 5.12, both of which are known to be flagellar gene regulators (Ramakrishnan & Newton, 1990; Richings *et al.*, 1995) which is suggestive of Torf also controlling flagellar gene expression in *R.sphaeroides*. As shown in Fig 5.13, the degree of homology of all EBP's is mainly restricted to the central domain. It thought that the N-terminal regions are involved in receiving the signal for activation (e.g. see Shingler, 1996) and given the low level of homology in this region, it is reasonable to assume that Torf is involved in sensing a different signal to FleR and FlbD. This N-terminal domain has homology with various other proteins in the Swissprot protein database held at Seqnet, Daresbury, Uk, but there are no features conserved in all these proteins (see Table 5.1). These will be discussed in more detail in the discussion.

The predicted secondary structure of the Torf protein was determined using the PEPTIDESTRUCTURE program and displayed using the PLOTSTRUCTURE program, both from the GCG package (Deveraux *et al.*, 1984). The results can be seen in Fig 5.14. It is predicted to contain a large degree of alpha helical structures as predicted using Chou and Fasman and Garnier and Robson (Chou & Fasman, 1978; Garnier *et al.*, 1978). This structural information will be discussed with reference to homology to other proteins in the discussion.

To further characterise the *torf* gene from *R.sphaeroides*, several attempts were made to produce gene replacement mutants. However, a suitable construct to allow this could never be isolated as it appeared that, when present on its own, the *torf* gene is unstable i.e. undergoes deletions and rearrangements. This may suggest that the gene was functional in *E.coli* and the re-arrangements were due to its toxicity to the cell. The lack of a mutant in the *torf* gene and consequently the lack of a phenotype, prevents me from

Fig 5.11 Protein sequence alignment of Torf with members of the two-component systems. The alignment was created using the GCG program PILEUP and the consensus sequence determined using the PRETTY program, also from the GCG package (Deveraux *et al.*, 1984).The protein aligned with Torf are as follows: NtrCRC, *R.capsulatus* NtrC (Jones & Haselkorn, 1989); NtrC, *Rhizobium meliloti* (Szeto *et al.*, 1987); HydG, *E.coli* (Blatiner *et al.*, 1993); YfhA, *E.coli* (Lui & Magasanik, 1993); FleR, *P.aeruginosa* (Richings *et al.*, 1995); FlbD, *C.crescentus* (Ramakrishnan & Newton, 1990); NifA, *K.pneumoniae* (Arnold *et al.*, 1988). NH, CENTRAL and COOH refer to the conserved domains of EBP's with 1-7 representing the conserved clusters of residues noted by Morett and Segovia (Morett & Segovia, 1993). ∇ denotes residues important in CheY function as reviewed in Volz, 1993 (Volz, 1993). See the discussion for a more detailed description of the role of the clusters and the CheY functional residues.

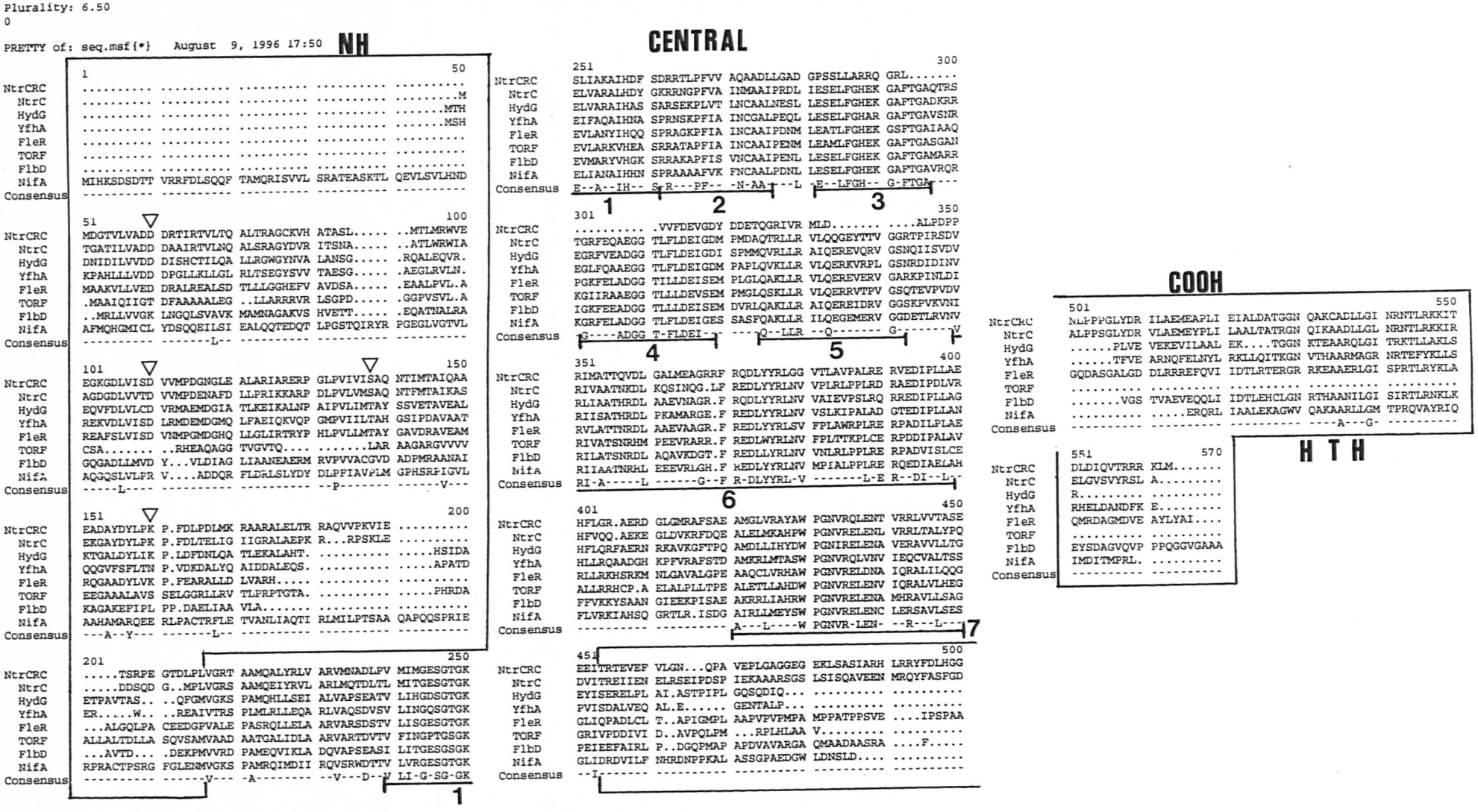
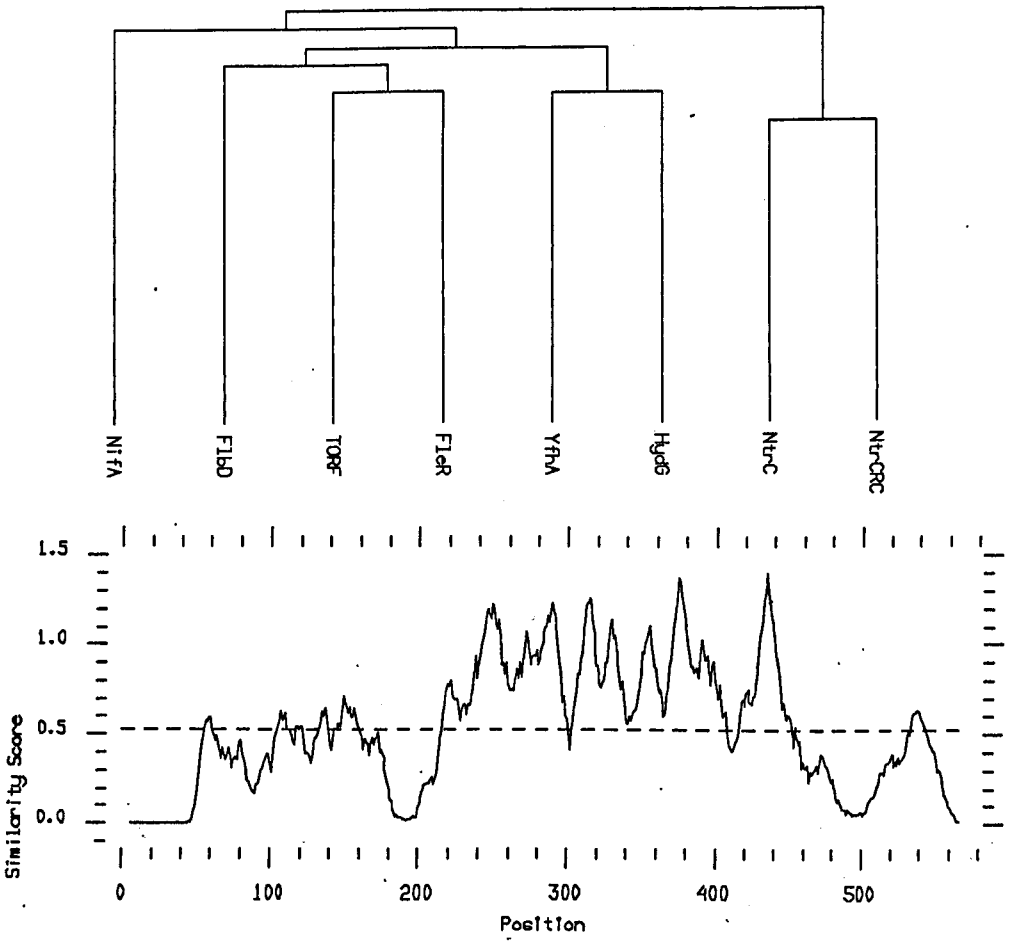

Plurality: 6.50
0
PRETTY of: seq.msf{*} August 9, 1996 17:50
NH
CENTRAL
COOH
H T H
NtrCRC
NtrC
HydG
YfhA
FleR
TORF
FlbD
NifA
Consensus
1
2
3
4
5
6
7

Fig 5.12 Dendrogram showing relationships of members of the two-component sensor regulator family to Torf. Created using the PILEUP program from the GCG package (Deveraux *et al.*, 1984). NtrCRC, *R.capsulatus* NtrC (Jones & Haselkorn, 1989); NtrC, *Rhizobium meliloti* (Szeto *et al.*, 1987); HydG, *E.coli* (Blatiner *et al.*, 1993); YfhA, *E.coli* (Lui & Magasanik, 1993); FleR, *P.aeruginosa* (Richings *et al.*, 1995); FlbD, *C.crescentus* (Ramakrishnan & Newton, 1990); NifA, *K.pneumoniae* (Arnold *et al.*, 1988). Similar proteins are close to each other.

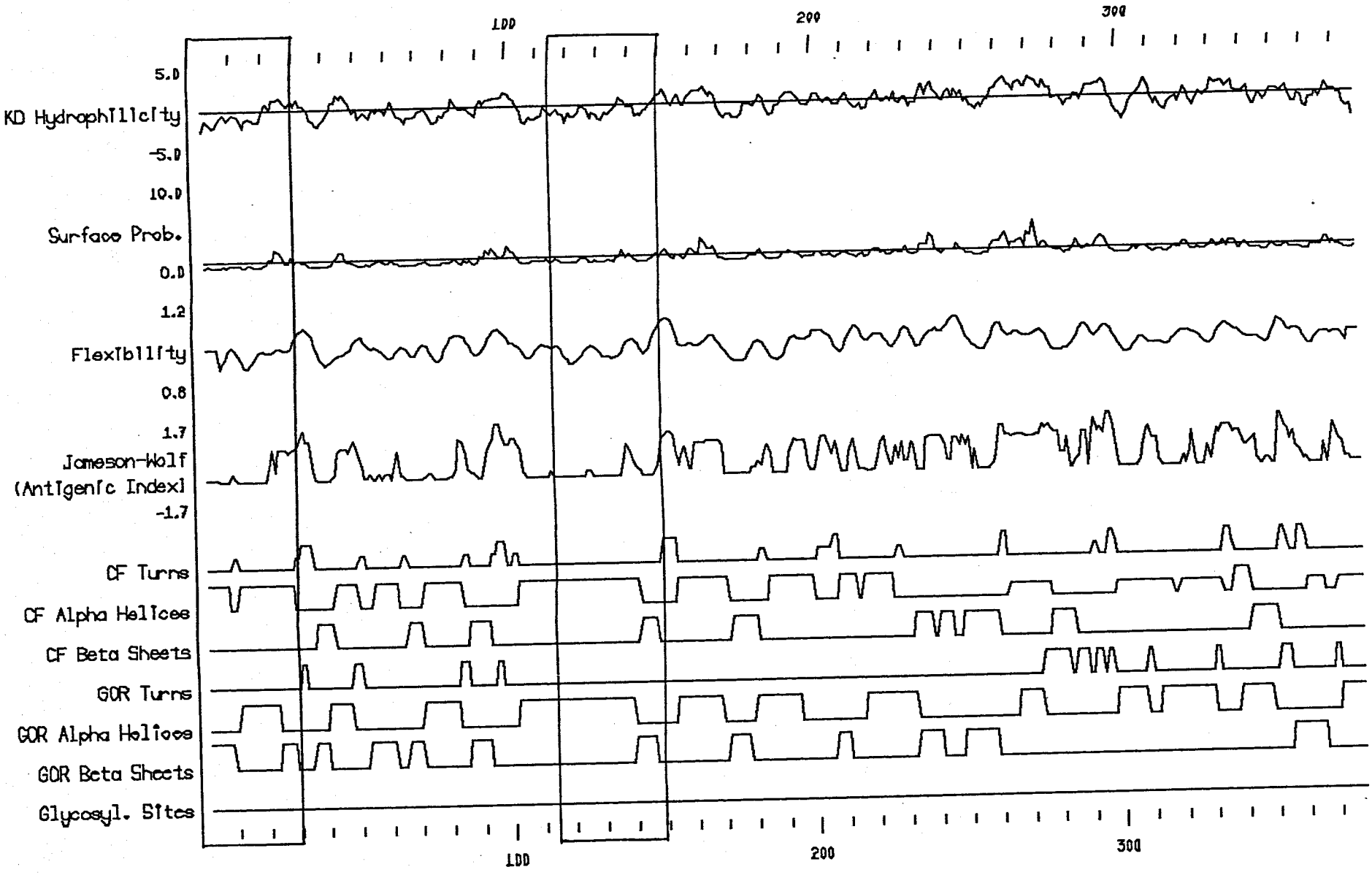
Fig 5.13 Similarity plot of the PILEUP alignment shown in Fig 5.11. Created using the PLOTSIMILARITY program from the GCG package (Deveraux *et al.*, 1984). The graph depicts the level of similarity between the proteins along the alignment (Fig 5.11).

Table 5.1. Proteins with homology to the N-terminal domain of the Torf protein. The alignments were created using the FASTA program from the GCG package (Deveraux *et al.*, 1984). References: HoxV (Chen & Mortenson, 1992 [Smith, 1995 #380); HypF (Colbeau *et al.*, 1993); and RuvB (Robinson, 1994).



Protein	Region homologous in Torf Seq	% Identity	Role of protein	Bacterium
Hox V	45 - 90	31.6	Regulation of Hydrogenase expression	*Azotobacter vinelandii*
Hyp F	56 - 85	44.8	Regulation of Hydrogenase expression	*R. capsulatus*
Pro C	1 - 28	50	Pyrroline - 5-carboxylate reductase	*Mycobacterium leprae*
Ruv B	1 - 139	23	Homologous recombination	*Mycobacterium leprae*

Fig 5.14 Predicted secondary structure of the *R.sphaeroides* Torf protein. The predictions were calculated using the GCG program PEPTIDESTRUCTURE and the data displayed using the PLOTSTRUCTURE program, also from the GCG package (Deveraux *et al.*, 1984). The graphs show the structural information of the residues as a function of; hydrophilicity as determined using the method of Kyte and Doolitle (Kyte & Doolittle, 1982); surface probability according to Emini *et. al.* (Emini *et al.*, 1985); flexibility as determined using the method of Karplus and Shulz (Karplus & Schulz, 1985); turns, alpha helices and beta sheets according to both Chou and Fasman (CF) and Garnier and Robson (GOR) (Chou & Fasman, 1978; Garnier *et al.*, 1978). Glycosylation sites are predicted for the sites where the residues have the composition NXT or NXS. Boxed regions are described in the text.



Residues on protein

concluding that the Torf is involved in flagellar gene regulation, however, given the high degree of homology to FleR and FlbD, it is very likely that it is involved in this regulation.

5.2.4f Characterisation of the *fliE* gene from *R.sphaeroides*

The DNA sequence of the *R.sphaeroides fliE* gene is shown in Fig 5.15. In FliE is a small protein (11.26KDa) with homology to previously isolated FliE proteins (Table 5.2). The direction of transcription of *fliE* in *R.sphaeroides* is also opposite to that in the enteric bacteria i.e. in *R.sphaeroides* it is transcribed in the same direction as *fliF* whereas in *S.typhimurium* it is transcribed opposite to *fliF*. It has no extensive regions of hydrophobicity long enough to span the membrane (data not shown) as determined by the method of Kyte and Doolittle (Kyte & Doolittle, 1982). Muller and co-workers suggest that it forms an adapter between the rod and the MS-ring (Muller *et al.*, 1992).

There is a high level of homology between the C-terminal regions of all isolated FliE proteins as seen in Fig 5.16. The motif LXXVM(X)3QXXS(X)9RNK(X)3AY(X)3MXM is absolutely conserved in all FliE proteins studied to date. No other proteins with this motif were found in the Swissprot protein data base at Seqnet, Daresbury, UK. The RSFLIE protein has statistically significant homology to numerous RecA proteins using the GCG program GAP with the -random feature e.g.. *Thermus aquaticus* subspecies *thermophilus* (Wetmur *et al.*, 1994) with which it has 28.7% identity and 48.1 % similarity. The highest degree of conservation with the RecA proteins is in the region of residue 42 to 58 where the motif LGXXVDDL(X)3QXDXGE is conserved. This domain is lies in the region strongly predicted to form an alpha helix (data not shown). All other FliE proteins do not have a statistically significant degree of conservation with RecA using the GCG program GAP with the -random feature. The significance of this will be discussed later.

5.2.4g Construction of a gene replacement mutant of RS*fliE*

A gene replacement mutant was constructed (as depicted in Fig 5.8) using a similar method to that described in chapters 3 & 4. Exconjugants with the *fliE* gene interruption were isolated and the insertion of the omega cartridge into the chromosome confirmed by Southern blotting (data not shown). Repeated attempts to isolate double recombinants which lacked the presence of pSup202 in the chromosome failed, and only single recombinants containing pSup202 could be isolated. These isolates, known as WS8::*fliE* Ω were found to be non-motile.

5.2.4h Analysis of WS8::*fliE*Ω

Western immunoblot analysis of WS8::*fliE* Ω using anti-flagellin antibody showed that WS8::*fliE* Ω did not contain any extracellular flagellin or intracellular flagellin (data not shown). This suggests that the regulation of flagellin synthesis is dependant on the formation of an intact basal-body structure. This phenotype may have been due to the polar effects of the omega cartridge on expression of the downstream genes *fliF* and *fliG* in WS8::*fliE* Ω.

5.2.4i Complementation analysis of WS8::*fliE* Ω

Cosmids 711 and 753 did not complement WS8::*fliE* Ω back to wild-type levels of motility, moreover it did not complement WS8::*fliE* Ω back to the levels of motility seen in Nm7 complemented by cosmid 711 or 753 (0.1% motility) and again there was no intracellular flagellin (data not shown). This suggests that a truncated product produced in WS8::*fliE* Ω by the insertion of the omega cartridge, resulting in the replacement of the C-terminal 8 residues RDIMNMPV with RPGSGD, and that this interfered with the function of wildtype FliE produced from the cosmids possibly by interacting with it and preventing its interaction with other basal body components.

Fig 5.15 DNA sequence of the *R.sphaeroides fliE* gene. The potential ribosome binding site is shown as SD, upstream of the ATG initiation codon.

Fig 5.16 Protein sequence alignment of RSFliE with previously studied FliE proteins. The alignment was created using the GCG program PILEUP and the consensus displayed using the PRETTY program with a plurality of 4.0 (Deveraux *et al.*, 1984). Absolutely conserved residues are boxed. ▌ represents the position of the Ω cartridge insertion used to create WS8::*fliE*Ω. ECFLIE; *E.coli* FliE (Muller *et al.*, 1992); STFLIE, *S.typhimurium* FliE (Muller *et al.*, 1992); PSFLIE, *P.aeruginosa* FliE (Arora *et al.*, 1996) and BSFLIE, *B.subtilis* FliE (Zuberi *et al.*, 1991).

Table 5.2 Degree of homology between *R.sphaeroides* flagellar proteins studied in this chapter with those from *E.coli, S.typhimurium, B.subtilis* and *P.aeruginosa.* Values are given as a percentage and were calculated using the GCG program GAP (Deveraux *et al.*, 1984). Ident. is the percentage identity and Sim. the percentage similarity. ND, not determined as no homologue cloned as yet.

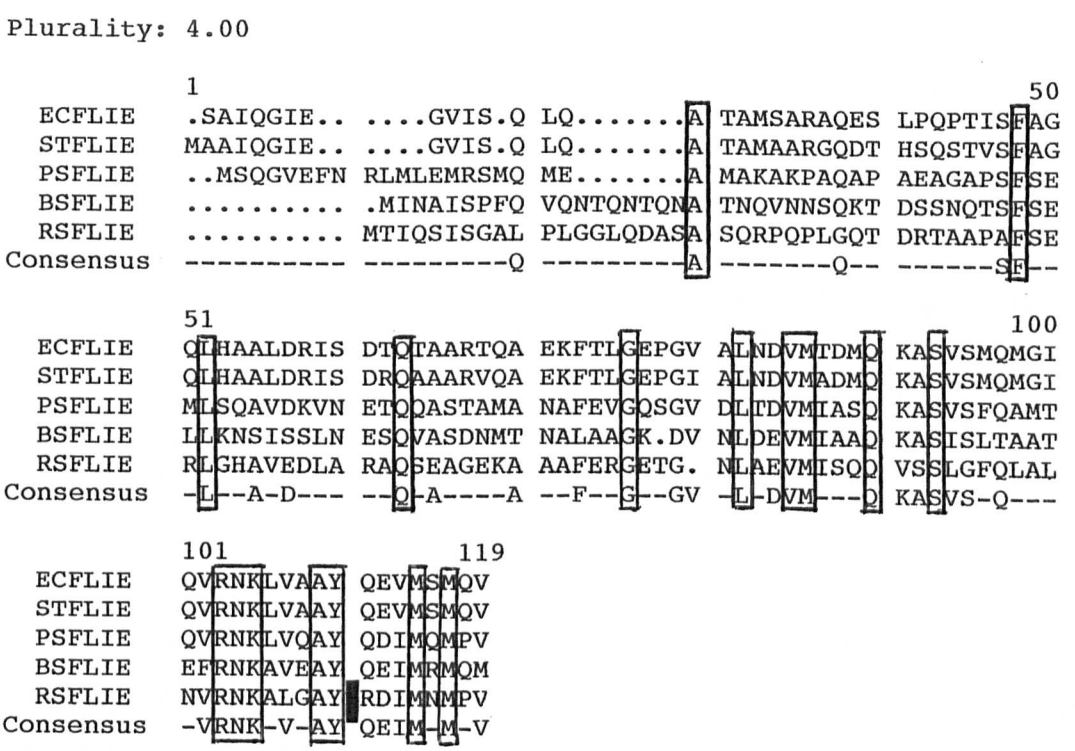
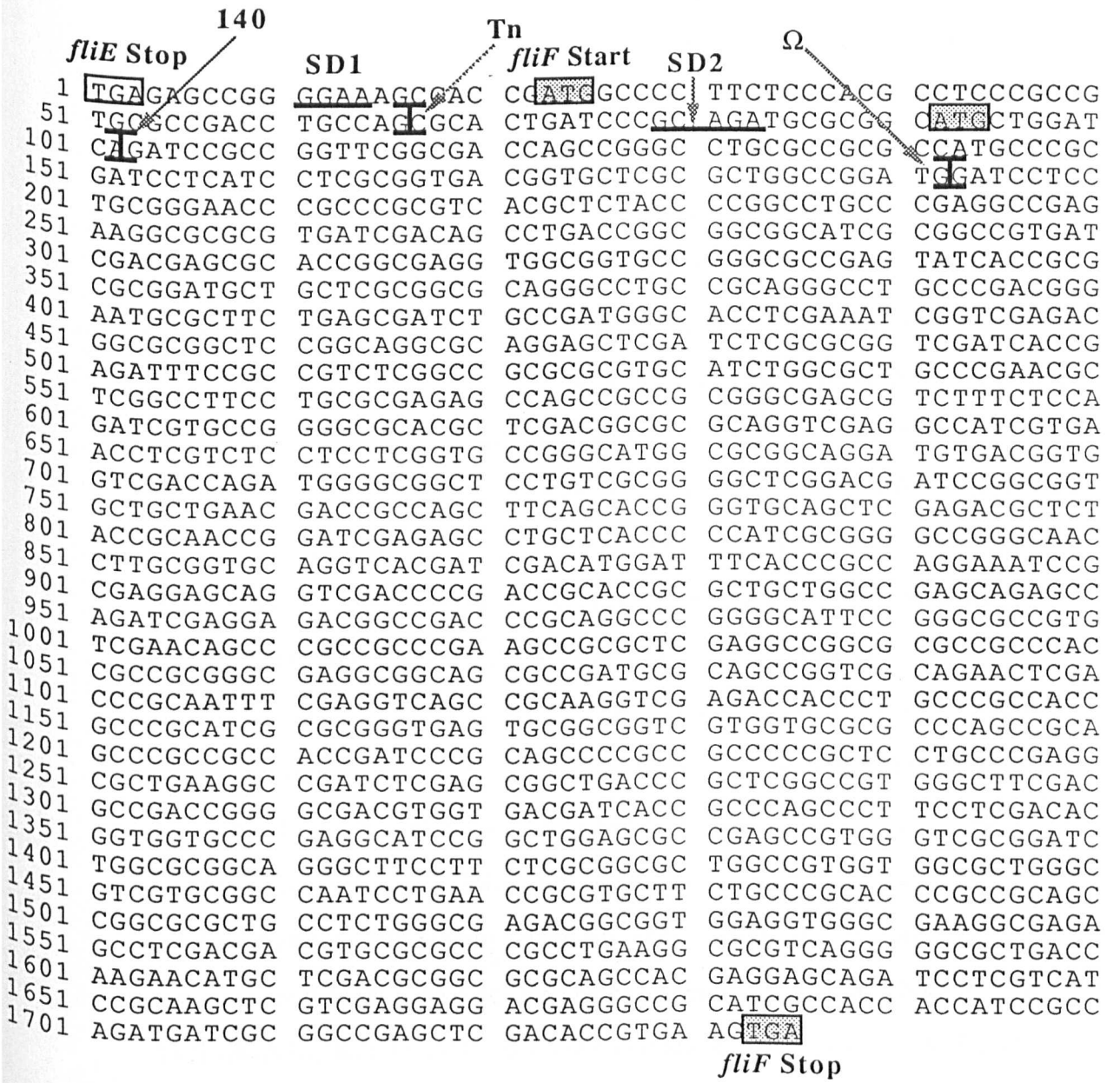


Table 5.2

	FliE		FliF		FliG	
	Ident.	Sim.	Ident.	Sim.	Ident.	Sim.
S.typhimurium	33.3	52	34.8	54.8	41	64.2
E.coli	31.8	54	35.6	54	39	64
C.crescentus	ND	ND	27.3	50.5	31.5	56.1
B.subtilis	30.2	42.5	21	46.2	33.8	60.8
P.aeruginosa	36.8	53.8	31	54.3	41.2	64.5

Fig 5.17 DNA sequence of the *R.sphaeroides fliF* gene. The two potential ATG initiation codons are shown and the ribosome binding sites upstream of these are shown as SD1 and SD2. The site of Tn*phoA* insertion in Nm7 is shown as Tn, with the end of cosmid 140 shown as 140 and the site of the Ω insertion in WS8::*fliF* Ω shown as Ω.


140
Tn
Ω
fliE Stop
SD1
fliF Start
SD2
fliF Stop
1 TGAGAGCCGG GGAAAGCCAC CGATGGCCCC TTCTCCCACG CCTCCCGCCG
51 TGCGCCGACC TGCCAGCGCA CTGATCCCGC AGATGCGCGG CATGCTGGAT
101 CAGATCCGCC GGTTCGGCGA CCAGCCGGGC CTGCGCCGCG CCATGCCCGC
151 GATCCTCATC CTCGCGGTGA CGGTGCTCGC GCTGGCCGGA TGGATCCTCC
201 TGCGGGAACC CGCCCGCGTC ACGCTCTACC CCGGCCTGCC CGAGGCCGAG
251 AAGGCGCGCG TGATCGACAG CCTGACCGGC GGCGGCATCG CGGCCGTGAT
301 CGACGAGCGC ACCGGCGAGG TGGCGGTGCC GGGCGCCGAG TATCACCGCG
351 CGCGGATGCT GCTCGCGGCG CAGGGCCTGC CGCAGGGCCT GCCCGACGGG
401 AATGCGCTTC TGAGCGATCT GCCGATGGGC ACCTCGAAAT CGGTCGAGAC
451 GGCGCGGCTC CGGCAGGCGC AGGAGCTCGA TCTCGCGCGG TCGATCACCG
501 AGATTTCCGC CGTCTCGGCC GCGCGCGTGC ATCTGGCGCT GCCCGAACGC
551 TCGGCCTTCC TGCGCGAGAG CCAGCCGCCG CGGGCGAGCG TCTTTCTCCA
601 GATCGTGCCG GGGCGCACGC TCGACGGCGC GCAGGTCGAG GCCATCGTGA
651 ACCTCGTCTC CTCCTCGGTG CCGGGCATGG CGCGGCAGGA TGTGACGGTG
701 GTCGACCAGA TGGGGCGGCT CCTGTCGCGG GGCTCGGACG ATCCGGCGGT
751 GCTGCTGAAC GACCGCCAGC TTCAGCACCG GGTGCAGCTC GAGACGCTCT
801 ACCGCAACCG GATCGAGAGC CTGCTCACCC CCATCGCGGG GCCGGGCAAC
851 CTTGCGGTGC AGGTCACGAT CGACATGGAT TTCACCCGCC AGGAAATCCG
901 CGAGGAGCAG GTCGACCCCG ACCGCACCGC GCTGCTGGCC GAGCAGAGCC
951 AGATCGAGGA GACGGCCGAC CCGCAGGCCC GGGGCATTCC GGGCGCCGTG
1001 TCGAACAGCC CGCCGCCCGA AGCCGCGCTC GAGGCCGGCG CGCCGCCCAC
1051 CGCCGCGGGC GAGGCGGCAG CGCCGATGCG CAGCCGGTCG CAGAACTCGA
1101 CCCGCAATTT CGAGGTCAGC CGCAAGGTCG AGACCACCCT GCCCGCCACC
1151 GCCCGCATCG CGCGGGTGAG TGCGGCGGTC GTGGTGCGCG CCCAGCCGCA
1201 GCCCGCCGCC ACCGATCCCG CAGCCCCGCC GCCCCCGCTC CTGCCCGAGG
1251 CGCTGAAGGC CGATCTCGAG CGGCTGACCC GCTCGGCCGT GGGCTTCGAC
1301 GCCGACCGGG GCGACGTGGT GACGATCACC GCCCAGCCCT TCCTCGACAC
1351 GGTGGTGCCC GAGGCATCCG GCTGGAGCGC CGAGCCGTGG GTCGCGGATC
1401 TGGCGCGGCA GGGCTTCCTT CTCGCGGCGC TGGCCGTGGT GGCGCTGGGC
1451 GTCGTGCGGC CAATCCTGAA CCGCGTGCTT CTGCCCGCAC CCGCCGCAGC
1501 CGGCGCGCTG CCTCTGGGCG AGACGGCGGT GGAGGTGGGC GAAGGCGAGA
1551 GCCTCGACGA CGTGCGCGCC CGCCTGAAGG CGCGTCAGGG GGCGCTGACC
1601 AAGAACATGC TCGACGCGGC GCGCAGCCAC GAGGAGCAGA TCCTCGTCAT
1651 CCGCAAGCTC GTCGAGGAGG ACGAGGGCCG CATCGCCACC ACCATCCGCC
1701 AGATGATCGC GGCCGAGCTC GACACCGTGA ACTGA

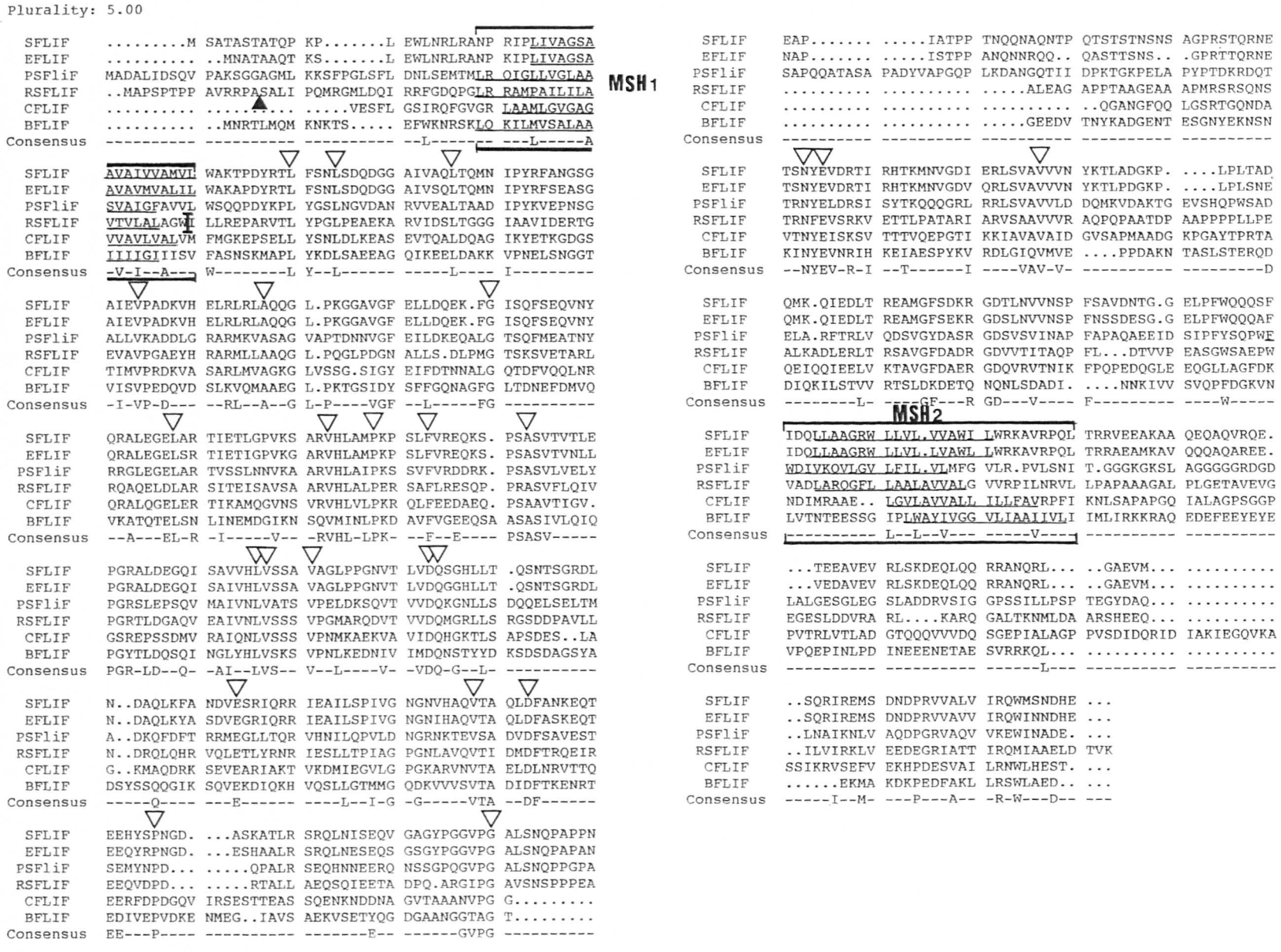
5.2.4j Characterisation of the *fliF* gene from *R.sphaeroides*

The DNA sequence of the *R. sphaeroides fliF* gene is shown in Fig 5.17 with the site of Tn*pho*A insertion and the left hand vector arm junction of cosmid 140 highlighted. The FliF open reading frame is predicted to have two ATG initiation codons (Fig 5.17). I have predicted that the first ATG initiation codon is the correct start point for FliF protein on several facts; 1) The "poor" potential ribosome binding site upstream of the second potential initiation codon (SD2 in Fig 5.17) suggests that it may not function, the first initiation codon (Fig 5.17) possesses a much better potential ribosome binding site (SD1 Fig 5.17) and 2) The rather large intergenic region (88 bp) present between the end of *fliE* and the second initiation codon of *fliF* does not contain any promoter sequences.

RSFliF is predicted to be 60.7 KDa with two predicted membrane spanning helices (Fig 5.18) as determined using the method of Kyte and Doolittle (Kyte & Doolittle, 1982)(data not shown), which is in agreement with other FliF proteins (Zuberi *et al.*, 1991; Ueno *et al.*, 1992; Ramakrishnan *et al.*, 1994; Matsumurra, 1995; Arora *et al.*, 1996). There is good conservation at the N-terminal half of all isolated FliF's (FIG 5.17), with very little conservation elsewhere.

As with other FliF proteins, RSFliF has homology with the YscJ family of virulence exporters, namely YscJ (46% similarity, 23% identity) (Michiels *et al.*, 1991), MxiJ (41.7% similarity, 22.1% identity) (Abdelmounaaim *et al.*, 1992), NolT (45.6% similarity, 26.3% identity) (Meinhardt *et al.*, 1993) and HrpB3 (46.4% similarity, 25% identity) (Fenselau *et al.*, 1992). All such virulence exporters are thought to be outer-membrane proteins based on their homology to YscJ which has been located to the outer membrane (Michiels *et al.*, 1991), but their role in export is not fully understood as extensive analysis of these proteins has not yet been carried out.

Fig 5.18 Protein sequence alignment of RSFliF with previously isolated FliF proteins. The alignment was created using the GCG program PILEUP and the consensus displayed using PRETTY with a plurality of 5.0 (Deveraux *et al.*, 1984). The two predicted membrane spanning regions are shown as MSH1 and MSH2. The site of Tn*pho*A insertion in Nm7 is shown as Δ, the position of the Ω cartridge insertion in WS8::*fliF* Ω as **I** and residues that are absolutely conserved in all 6 sequences shown as ∇. SFLIF, *S.typhimurium* (Jones *et al.*, 1989); EFLIF, *E.coli* (Matsumurra, 1995); PSFliF, *P.aeruginosa* (Arora *et al.*, 1996); CFLIF, *C.crescentus* (Ramakrishnan *et al.*, 1994) and BFLIF, *B.subtillis* (Zuberi *et al.*, 1991).


Plurality: 5.00
MSH1
MSH2

5.2.4k Construction of a gene replacement mutant of RS*fliF*

To determine whether or not the incomplete complementation of Nm7 was due to the effects of an internal Tn*phoA* promoter producing a truncated FliF, which was producing a dominant effect, a null mutant of *fliF* was produced. An omega (Ω) cartridge was cloned into the *Bam*HI site internal to *fliF*, 170bp downstream from the FliF initiation codon (see Fig 5.17 and 5.18). Exconjugants with the *fliF* gene replacement were found and the deleted version of *fliF* in the chromosome of the exconjugants was confirmed by Southern blotting analysis (data not shown) and isolates (known as WS8::*fliF* Ω) were completely non-motile, as for Nm7.

5.2.4l Complementation analysis of WS8::*fliF*Ω

Complementation of WS8::*fliF* Ω by cosmids 711 and 753 gave similar results to that obtained for Nm7 i.e. 1 in 1000 cells were motile. This suggested that the incomplete complemented phenotype of Nm7 was not due to the production of truncated protein causing a dominant effect, but was in fact the characteristics of a polar *fliF* mutation affecting the expression of downstream genes. This lack of complete complementation resembled the characteristics of a *fliI* null mutant we have reported previously in chapter 4 (Goodfellow *et al.*, 1996).

5.2.4m Overexpression, purification of the *R.sphaeroides* FliF protein and the production of a polyclonal antiserum

A GST partial FliF fusion protein, containing FliF from the site of the Ω insertion in WS8::*fliF* Ω (Fig 5.18), was overexpressed in *E.coli* and found to pellet with the membrane/unbroken cells fraction. This may have been due to the formation of inclusion bodies or an association of the FliF portion of the fusion protein with the cytoplasmic membrane. It was purified according to the method of Fragioni and Neel (Frangioni & Neel, 1993) using the detergent sarkosyl to

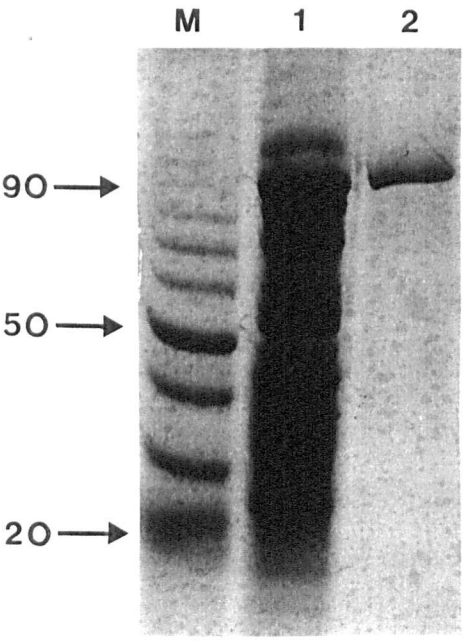
solubilise the protein and Triton X100 to remove the sarkosyl and allow the binding of the GST protein to the glutathione sepharose. Initial small scale purifications yielded very pure FliF-fusion protein (see Fig 5.19 a), however digestion of the fusion protein with Factor X did not yield the predicted fragments (data not shown) suggesting that the purified protein was in fact not FliF i.e. it was predicted that Factor X should digest only at the junction between FliF and GST giving two fragments of 26KDa (GST) and 60 KDa (FliF), however, it was seen to give a fragment of 26KDa and several fragments smaller than 15KDa. The exact size of the fragments produced could not be determined It has been reported that Factor X can cleave at secondary sites which consist of basic residues (Nagai *et al.*, 1985) and it has been suggested that this is due to a partially folded conformation of the target protein (Hall & Riggs, 1996). This may have been occurring, as the partially folded state of the FliF protein was probably due to the method of purification i.e. sarkosyl solubilisation. The fact that the DNA sequence of the construct used to overexpress the fusion protein was correct further strengthens the idea that this incorrect digestion with Factor X was due to a partially folded protein.

A large scale purification of the fusion protein was carried out and the results can be seen in Fig 5.19 b. This yielded fusion protein with less purity than from small scale purifications (Fig 5.19 b, lanes 6-8). Attempts were made to remove the contaminating proteins by passing the purified protein down another glutathione-sepharose column, but a similar product was obtained suggesting that the contaminating proteins were breakdown products of the fusion protein. This was confirmed later using an antibody to the fusion protein (see below). The use of centricon 50 microcentrators (Amincon), dialysis using large molecular weight cut off tubing (50KDa), also failed to remove the contaminating proteins. This suggested that the fusion protein breakdown products formed a complex possibly

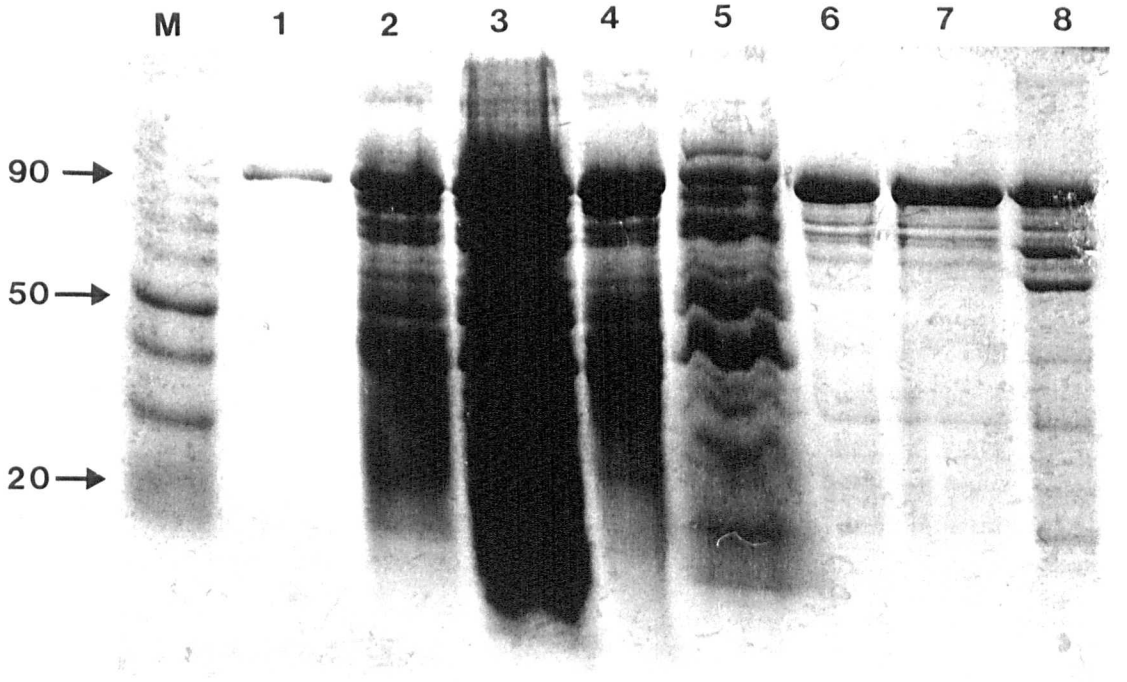
Fig 5.19 A) Small scale purification of the GST-FliF fusion protein. M is 10KDa ladder marker (Gibco BRL); 1 is Sarkosyl solubilised whole cell extract and 2 is purified GST-FliF fusion protein.

B) Large scale purification of the GST-FliF fusion protein. M is as in A. lanes are as follows: 1, small scale purified fusion protein from A; 2, whole cell sample; 3, Sarkosyl solubilised protein; 4, insoluble protein after solubilisation; 5, flowthrough from purification; 6, 7 and 8 three consecutive elutions from the glutathione-sepharose. Sizes are in KDa.

A



B



with the intact product, this was consistent with the observation that, upon the addition of elution buffer in the last step of the purification procedure, a white flocculate precipitate was produced. This precipitate rapidly solubilised upon heating to room-temperature and precipitated again when on ice (data not shown). A similar effect was not seen when the purification of GST alone was carried out using the same procedure and therefore is probably due to the association of the FliF portion of the fusion protein with each other. The self assembly properties of FliF from *S.typhimurium* have been demonstrated previously (Oosawa *et al.*, 1994; Ueno *et al.*, 1994).

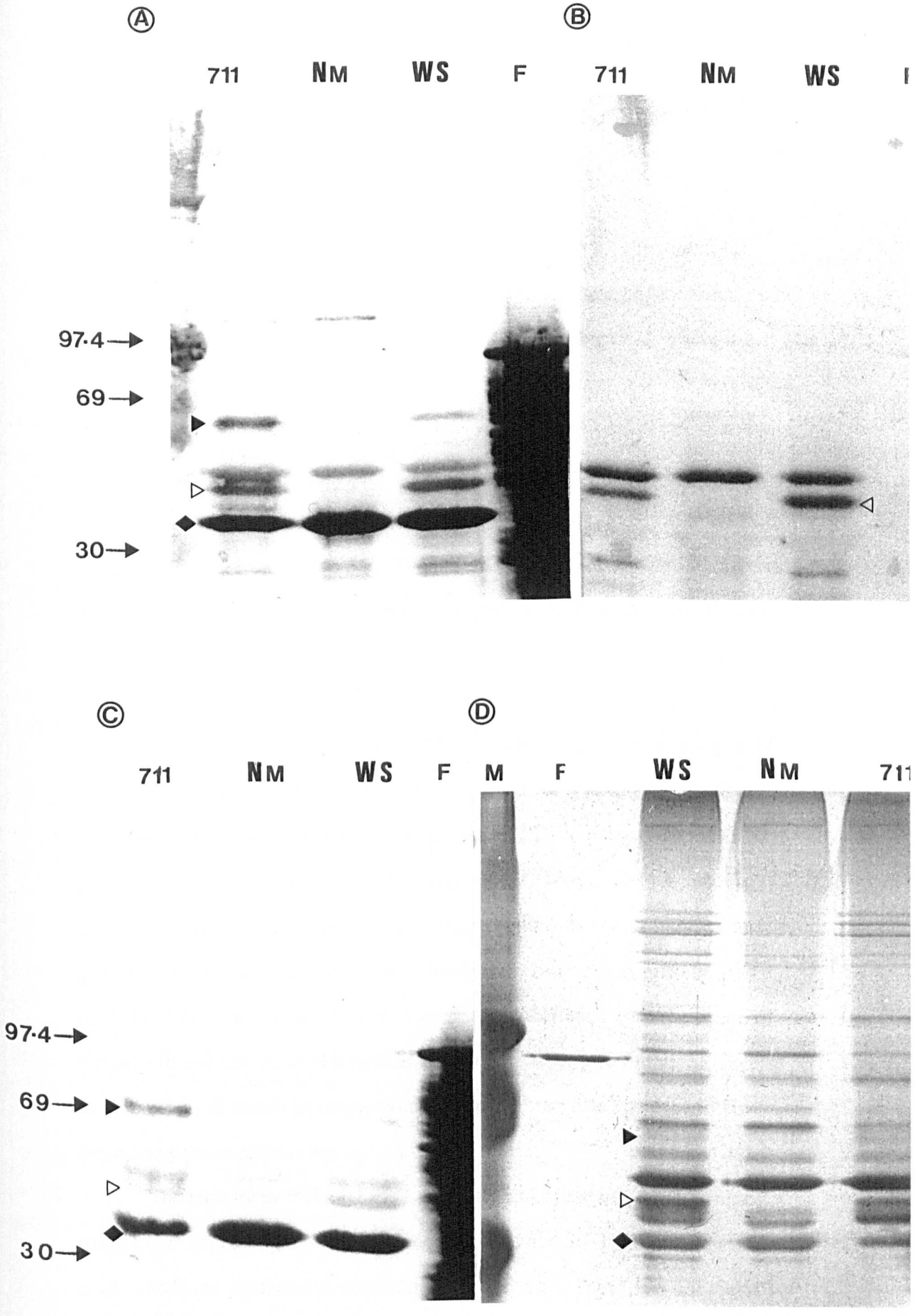
To produce an antiserum against the purified FliF fusion protein, the breakdown products were removed by the electrophoresis of the fusion protein by preparative SDS-PAGE, the gel stained, the protein excised from the gel and ground into a powder. This was used to immunise a single New Zealand White rabbit. Antisera were obtained and used in Western blot analysis of whole cell samples of WS8, Nm7 and Nm7 X cosmid 711. There was a very high degree of cross-reaction of the pre-immune serum and the post-immune serum with other cellular proteins which masked the area of interest on the blots i.e. 60KDa (data not shown). Several approaches were tried to remove the cross-reacting antibodies. The first attempt utilised preadsorption of the serum against an acetone extract of Nm7 cells, this failed to remove the cross-reacting antibodies as the acetone extracted proteins may have been in a different conformation to those recognised by the antiserum. A second approach of using immobilised antigen to purify GST-FliF specific antibodies, surprisingly failed to remove the antibodies that cross-reacted with other cellular proteins at round the 60KDa range.

The use of fractionated cells of Nm7, WS8 and Nm7 X cosmid 711 revealed that the majority of the cross-reacting proteins that were masking the area of interest were in the outer-membrane fraction (data not shown) and that a cross-reacting protein of approximately 67.5KDa was present in WS8 cytoplasmic membranes, absent in Nm7 cytoplasmic membranes and again

present in Nm7 X cosmid 711 cytoplasmic membranes (Fig 5.20 A). This protein only reacted with the post-immune serum and not with the pre-immune serum (Fig 5.20 B) as did the GST-FLIF fusion protein. Allowing for the hydrophobic nature of the FliF protein, this protein approximates to the predicted size of RSFliF (60.7 KDa). A low dilution of the serum (after the second booster injection) was required (1 in 200) in order to obtain reasonable results using the method described in chapter 2. This was suggestive that the antibody was present at a very low concentration and an attempt was made to further purify the antibody by the removal of other serum proteins e.g. albumin, using ammonium sulphate precipitation and DEAE cellulose chromatography. The antibodies produced by this method was shown to be approximately 80% pure by SDS-PAGE (data not shown), but the removal of the contaminating serum proteins did not improve the 'background' cross-reaction seen in Western blot analysis or the affinity of the antibody (Fig 5.20 C).

Other proteins in the cytoplasmic membrane fractions were seen to cross-react specifically with the post-immune serum (Fig 5.20 A & C, lanes 711 & WS). This protein was also seen to be present in the periplasm and at a very high concentration in the outer membrane (data not shown). A protein of approximately 51.5KDa was found to cross-react with pre-immune serum (Fig 5.20 B) and was see to be absent in Nm7 cytoplasmic membranes (lane NM). Silver staining of the membrane fractions confirmed that this protein was in fact missing from Nm7 and not simply masked by another protein that was preventing its cross-reaction with the antibody (Fig 5.20 D). This protein is not flagellin as flagellin was seen to associate with the outer-membrane fraction, and its is also present at the same levels in WS8 as in Nm7X cosmid 711 whereas flagellin is not (data not shown). The protein may well be flagellar associated as it has been previously noted that a

Fig 5.20 Western immunoblot and SDS-PAGE analysis of cytoplasmic membrane fractions of WS8 (WS), Nm7 (NM) and Nm7X cosmid 711 (711) using: A, post-immune serum (1 in 200 dilution); B, pre-immune serum (1 in 200 dilution) and C, ammonium-sulphate precipitated/DEAE chromatography purified antibody (1 in 200 dilution). D is a silver stained SDS-PAGE gel of the same samples as in A- C. F, purified GST-FliF fusion protein; M, Rainbow protein markers (Amersham); closed triangle represents the native FliF protein; open triangle represents an unknown protein missing in Nm7 and the closed diamond probably represents the *R.sphaeroides* porin protein. Approximately 3μg of protein was loaded per lane. Sizes are in KDa


Ⓐ
711
NM
WS
F
97·4
69
30
Ⓑ
711
NM
WS
Ⓒ
711
NM
WS
F
97·4
69
30
Ⓓ
M
F
WS
NM
711

protein of approximately the same size is present in flagellated vesicles purified using anti-flagellin antibodies (Foster, 1991), this will be discussed in more detail later.

5.2.4n Characterisation of the *fliG* gene from *R.sphaeroides*.

The DNA sequence of the *fliG* gene from *R.sphaeroides* is shown in Fig 5.21 and the predicted protein sequence shown in Fig 5.22. There is a good Shine-Dalgarno ribosome binding site 5 bases upstream of the start of the *fliG* gene (Shine & Dalgarno, 1974). This sequence (GAAG) is close to the start of the gene and this may play some role in the regulation of its translation.

The RSFliG protein has good homology with previously studied FliG proteins (Fig 5.22 and Table 5.2) The RSFliG protein is predicted to possess two regions of hydrophobic residues by the method of Kyte and Doolittle (Kyte & Doolittle, 1982) (data not shown), the first spans from residues 11 to 22 and is conserved in most FliG proteins, and the second spans residues 130 to 151 and is less well conserved. The second hydrophobic region spans a region of the protein found to contain a consensus sequence for a lipid attachment site (Hayashi & Wu, 1990), as detected using the MOTIFS program in the GCG package (Deveraux *et al.*, 1984). This sequence is recognised by a specific peptidase which cuts upstream of the cysteine residue to which a glyceride-fatty acid is then attached (Hayashi & Wu, 1990). This motif is not present in any other FliG proteins studied to date, however, as this region is in the centre of the protein and not at the N-terminus, as is required for it to function (Hayashi & Wu, 1990), it would be predicted not to be processed or that the protein would be processed into two halves.

The structures of the RSFliG and other FliG proteins, as determined using the PEPTIDESTRUCTURE program from the GCG package (Deveraux *et al.*, 1984) are predicted to consist mainly of alpha helices (data not shown)

with very few regions predicted to form beta sheets although there is no obvious homology at this structural level.

Domains shown to be important for FliG function appear to be conserved to some degree in all FliG proteins. In Fig 5.22, the domain labelled A represents the region shown to be required for FliF-FliG interaction in the enteric bacteria (Marykwas *et al.*, 1996) and domain B represents a region thought to be involved in torque generation (Lloyd *et al.*, 1996) and FliM/N and MotA interaction (Tang *et al.*, 1996).

The clustering of charged residues in *S.typhimurium* FliG has been noted previously (Kihara *et al.*, 1989) and these clusters of charges are thought to function in gating the proton channel. Several charged residues are highly conserved in all FliG proteins (Fig 5.22), with the majority of them being positioned in the C-terminal domain of FliG, which is consistent with the fact that this region has been implicated in torque generation (Lloyd *et al.*, 1996). The pI of the RSFliG protein differs markedly from other FliG proteins; RSFliG pI is 5.34, *E.coli* FliG pI is 4.64, *S.typhimurium* FliG pI is 4.56, *B.subtillis* FliG pI is 4.48 and the *C.crescentus* FliG pI is 4.69. This has also been noted for the motor proteins MotA and MotB (Shah & Sockett, 1995; Shah & Sockett, 1995).

Analysis of mutations in *S.typhimurium* and *E.coli* has led to the identification of many mutations that affect the direction of rotation, rotation in general and FliG/M interactions. The mutations that are present naturally in the *R.sphaeroides* FliG or where the wild-type residue is conserved amongst most are all FliG proteins are summarised in table 5.3. A greater discussion of these will be presented later in section 5.3 .

Fig 5.21 DNA sequence of the *R.sphaeroides fliG* gene. The potential ribosome binding sites are shown as SD.

fliG Start

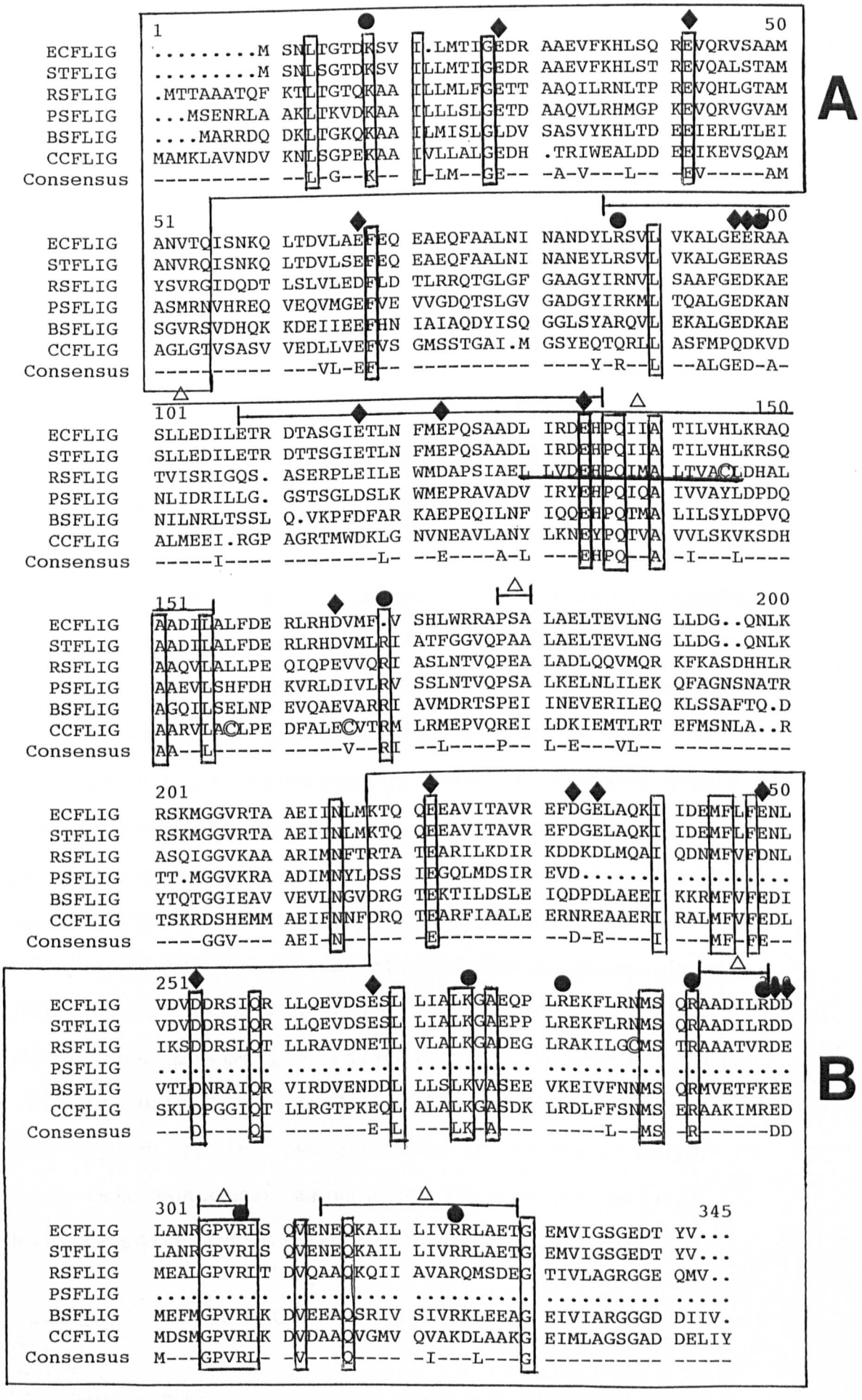
SD

```
   1 GAAGTGACAT GACCACAGCA GCCGCCACCC AGTTCAAGAC GCTCACCGGC
  51 ACGCAGAAGG CCGCCATCCT GCTCATGCTG TTCGGCGAGA CGACGGCCGC
 101 GCAGATCCTG CGCAACCTCA CCCCGCGCGA GGTGCAGCAT CTGGGCACCG
 151 CCATGTACAG CGTGCGCGGC ATCGATCAGG ACACGCTGAG CCTCGTGCTC
 201 GAGGATTTCC TCGACACGCT CCGCCGCCAG ACCGGCCTCG GCTTCGGCGC
 251 GGCGGGCTAT ATCCGCAACG TCCTGTCGGC AGCCTTCGGC GAGGACAAGG
 301 CCGAGACGGT CATCAGCCGC ATCGGCCAGT CCGCCTCCGA ACGCCCGCTC
 351 GAGATCCTCG AATGGATGGA CGCGCCCTCG ATCGCCGAGC TGCTGGTGGA
 401 CGAGCATCCG CAGATCATGG CGCTGACGGT GGCCTGCCTC GATCACGCGC
 451 TGGCAGCCCA GGTGCTGGCG CTGCTGCCCG AGCAGATCCA GCCCGAGGTC
 501 GTCCAGCGCA TCGCCTCGCT GAACACGGTG CAGCCCGAGG CGCTGGCCGA
 551 TCTCCAGCAG GTCATGCAGC GCAAGTTCAA GGCCTCAGAC CACCACTTGC
 601 GCGCGAGCCA GATCGGCGGG GTGAAGGCGG CCGCGCGGAT CATGAACTTC
 651 ACCCGCACCG CGACCGAGGC GCGGATCCTC AAGGACATCC GCAAGGACGA
 701 CAAGGACCTG ATGCAGGCGA TCCAGGACAA CATGTTCGTC TTCGACAATC
 751 TCATCAAGTC CGACGACCGC TCGCTGCAGA CGCTCCTGCG GGCGGTGGAC
 801 AACGAGACGC TGGTGCTGGC GCTGAAGGGC GCGGACGAGG GGCTGCGGGC
 851 GAAGATCCTC GGCTGCATGT CGACGCGGGC CGCGGCCACG GTGCGGGACG
 901 AGATGGAGGC CCTGGGCCCC GTGCGGTTGA CCGACGTCCA GGCCGCGCAG
 951 AAGCAGATCA TCGCGGTGGC CCGGCAGATG TCGGACGAGG GCACGATCGT
1001 GCTCGCGGGC CGCGGCGGCG AGCAGATGGT GTAAGGCATG ACCGCGCAGA
```

fliG Stop

```
1051 CGCCCATCGG ACCCGAGGAT GTGCTGGCCC TGATCCGCGA GACCAACGCG
1101 CGCGGCTCGG CCGCTCGGAC CTGCCCGCGC CCGGGCCGGA AGGCGTTCCG
1151 CCGATGCCGC TCTCGG
```

Fig 5.22 Protein sequence alignment of the RSFliG protein with previously studied FliG proteins from *E.coli* (ECFLIG), *S.typhimurium* (STFLIG), *P.aeruginosa* (PSFLIG), *B.subtillis* (BSFLIG) and *C.crescentus* (CCFLIG) (Kihara *et al.*, 1989; Zuberi *et al.*, 1991; Ramakrishnan *et al.*, 1994; Arora *et al.*, 1996). The alignment was created using the GCG program PILEUP and the consensus displayed using PRETTY with a plurality of 5.0 (Deveraux *et al.*, 1984). Residues that are absolutely conserved in all sequences are boxed. A & B represent domains described in the text. Regions that have been deleted in *S.typhimurium* and *E.coli* and found to cause a Mot$^-$ phenotype are highlighted as |——— | above the sequence with Δ denoting the deletion (Irikura *et al.*, 1993; Lloyd *et al.*, 1996). Open circles highlight cysteine residues possibly involved in di-sulphide bond formation; diamonds represents points at which a positively charge residue is present in at least four of the proteins and the other is not a negatively charged residue; closed circles represent points at which negatively charged residues present in at least four of the proteins and the other is not positively charged. The region underlined in the RSFliG sequence represents the region containing the consensus sequence for a prokaryotic lipid attachment site (Hayashi & Wu, 1990).


Plurality: 5.00
A
B
ECFLIG
STFLIG
RSFLIG
PSFLIG
BSFLIG
CCFLIG
Consensus

5.2.4o Overexpression and purification of the RSFliG protein

The RSFliG protein was overexpressed and purified as a poly-histidine fusion using the Qiaexpessionist system (Qiagen). The cloning of the *fliG* gene into the expression vector pQE30 (Qiagen) was carried out via PCR using a primer (described in chapter 2) that created a *Bgl* II site at the N-terminus of *fliG* and removed the first methionine residue, along with a pUC PCR primer. The template used was the 5.5Kb *Eco*RI fragment illustrated in Fig 5.6. It was found that the addition of DMSO (10%) to the PCR reaction greatly improved the specificity and yield of the product obtained (data not shown). The product was digested with *Bgl* II and *Sma*I and cloned into pQE30 which had been digested with *Bam* HI and *Sma* I to give the construct pQE:HisG. The 5' cloning site (*Bgl* II into *Bam* HI) was sequenced to check that the correct reading frame was being expressed.

Small scale purification revealed that the His-FliG protein was insoluble and it had to be purified under denaturing conditions. A large scale purification was carried out and a very small amount (approx 50μg) of His-FliG could be purified from the soluble protein (Fig 5.23 A, lane 8). Most of the His-FliG bound to the resin was removed during the second wash step suggesting that it was binding with a very low affinity. The soluble fraction was shown to precipitate when eluted from the column (data not shown) and it was decided to use the insoluble proteins as the source for further purification. The urea solubilised His-FliG was readily purified with very few contaminating proteins (Fig 5.23B, lane 8). Again a large amount of His-FliG was found to be removed during the washing of the nickel resin (Fig 5.23, lanes 6 & 7) and this was found to be mainly due to overloading of the column as most of the removed protein could be recovered by passing the wash solutions through another nickel resin column (data not shown). This is consistent with there being a large quantity of His-FliG passing through the column in the

flowthrough fraction (Fig 5.23 B, lane 5). As can be seen from Fig 5.23 B (lane 4) the procedure used for the urea solubilisation did not solubilise all of the His-FliG present and as a consequence the remaining insoluble protein was re-extracted with urea a further 3 times to purify additional His-FliG. The His-FliG protein purified from these additional extractions was of the same purity as that shown in Fig 5.23 B, lane 8. Approximately 10mg of His-FliG was purified from a 225ml culture with approximately one third remaining in the insoluble fraction.

Before use in the FliF binding assay, the His-FliG was renatured by serial dialysis in decreasing concentration of urea until all the urea was removed.

5.2.4p Interaction of RSFliG with RSFliF *in vitro*

After renaturation, a binding assay was carried out to determine if RSFliF and RSFliG interact *in vitro* as they do in the flagellar motor and the results obtained are shown in Fig 5.24. The RSFliF protein seen in the assay (Fig 5.24, lanes 1, 6 & 8) has large amounts of breakdown products present.

The binding of the GST-FliF fusion protein to His-FliG was found to occur *in vitro* (lane 3). This was shown to be specific for the interaction of the FliF portion of the protein with the His-FliG protein because: 1) GST was found not to bind to His-FliG or Talon resin (lanes 4 & 7 respectively) and 2) the GST-FliF fusion protein was found not to bind to the membrane in the Ultrafree filter units or the Talon resin (lanes 6 & 5 respectively).

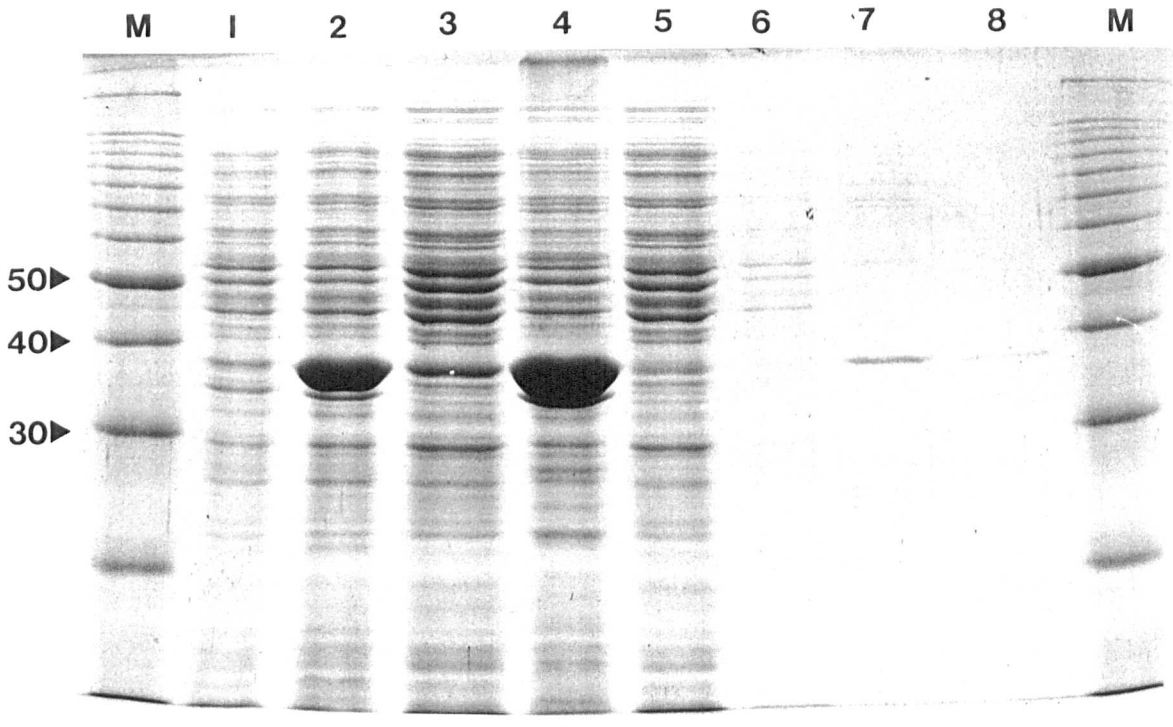
Table 5.3 Mutations isolated in the FliG proteins from *S.typhimurium* and *E.coli* resulting in motility defects. The phenotype describes the characteristics of the mutation on the motility of the cell being either biased CW rotation, biased CCW rotation, not rotation (Mot), and decreasing the interaction between FliG and FliM (G/M interaction). Temp. sens. Mot denotes temperature sensitive phenotypes; Hyd, hydrophobic residue; Pol, polar residue, -ve, negatively charged residue; +ve, positively charged residue; Cons. referred to the fact that the wild-type residue is conserved amongst all/most of the FliG proteins; 'naturally occurs' refers to the fact that the mutant residue is present in the wild-type FliG from either *R.sphaeroides* (RS) or *P.aeruginosa* (PS). Refs: 1 - Irikura *et. al.* 1993 and 2 - Marykwas *et. al.* 1996.

Table 5.3

Mutation	Phenotype	Notes	Ref
E108K	CW bias	-ve to +ve, cons. in all FliG's	1
E125D	CCW bias	Increase in size, cons. in all FliG's	1
I133T	CW bias	Hyd to Pol, present naturally in RS	1
Q168H	CCW bias	Change in size, cons. in most FliG's	1
P169L	CCW bias	Change in size, cons. in most FliG's	1
L172Q	Strong CW bias	Hyd to Pol, cons. in all FliG's	1
E174V	CW bias	-ve to Hyd, cons. in most FliG's	1
G185A	CW bias	G to Hyd, present naturally in RS	1
G195S	CW bias	Cons. in most FliG's	1
E211K	CCW bias	-ve to +ve, cons. in all FliG's	1
E211V	CW bias	-ve to Hyd, cons. in all FliG's	1
E225K	CW bias	-ve to +ve, cons. in all FliG's	1
E237K	CW bias	-ve to +ve, cons. in all FliG's	1
R313H	CW bias	+ve to Pol, cons. in all FliG's	1
R313S	CW bias	+ve to Pol, cons. in all FliG's	1
P127L	Mot	Change in size, cons. in all FliG's	1
Q128H	Mot	Change in size, cons. in all FliG's	1
R160L	Mot	+ve to Hyd, cons in all FliG's	1
R160H	Mot	+ve to Pol, cons. in all FliG's	1
L259P	Mot	Change in size, cons. in all FliG's	1
L259Q	Mot	Hyd to Pol, cons. in all FliG's	1
L259R	Temp. sens. Mot	Hyd to +ve, cons. in all FliG's	1
V135A	G/M interaction	Change in size, present naturally in RS and PS	2
L146Q	G/M interaction	Hyd to Pol, cons. in all FliG's	2
H155P	G/M interaction	Pol to Hyd, present naturally in RS	2
I229T	G/M interaction	Hyd to Pol, cons. in all FliG's	2

Fig 5.23 Large scale purification of His-FliG under native (A) and denaturing (B) conditions). A: 1, un-induced whole cell sample; 2, induced whole cell sample; 3, soluble protein after French press; 4, insoluble protein after French press; 5, flowthrough; 6, wash 1; 7, wash 2 and 8 elution. B: 1, induced whole cell sample; 2, insoluble protein before urea extraction; 3, soluble protein after urea extraction; 4, insoluble protein after urea extraction; 5, flowthrough; 6, wash 1; 7, wash 2 and 8 elution. M, 10KDa ladder (Gibco BRL). Sizes are in KDa.

A



B

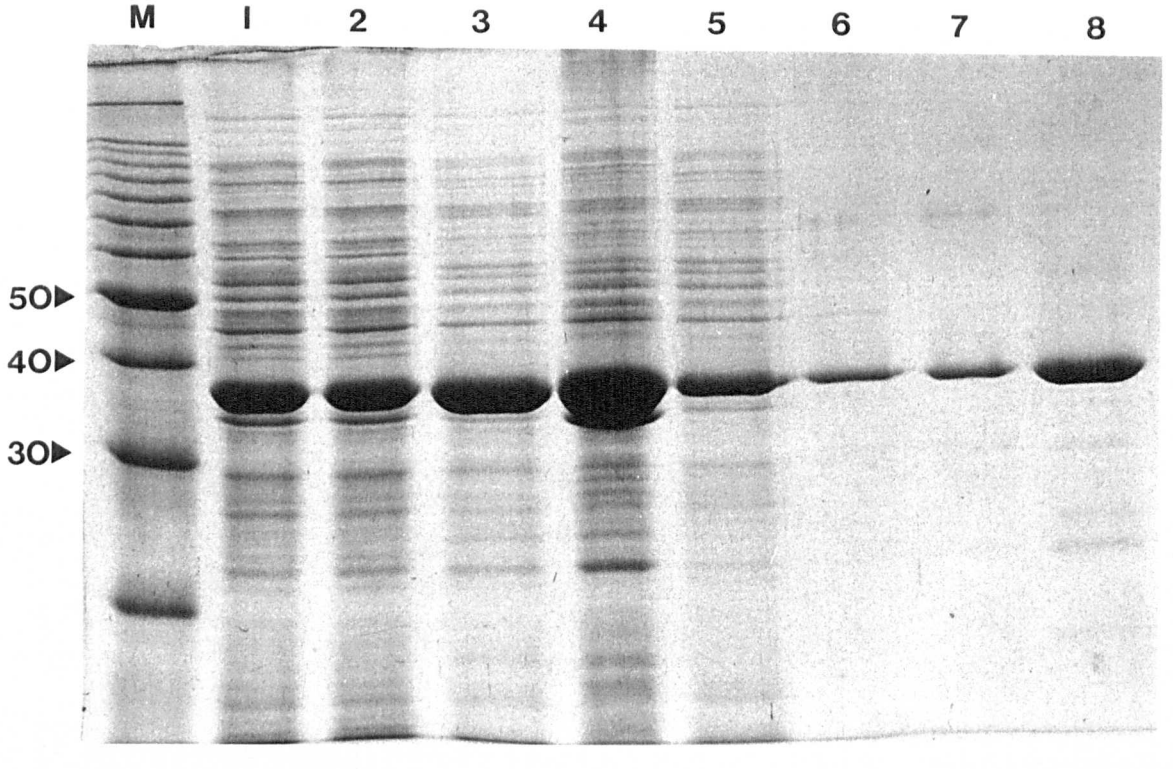
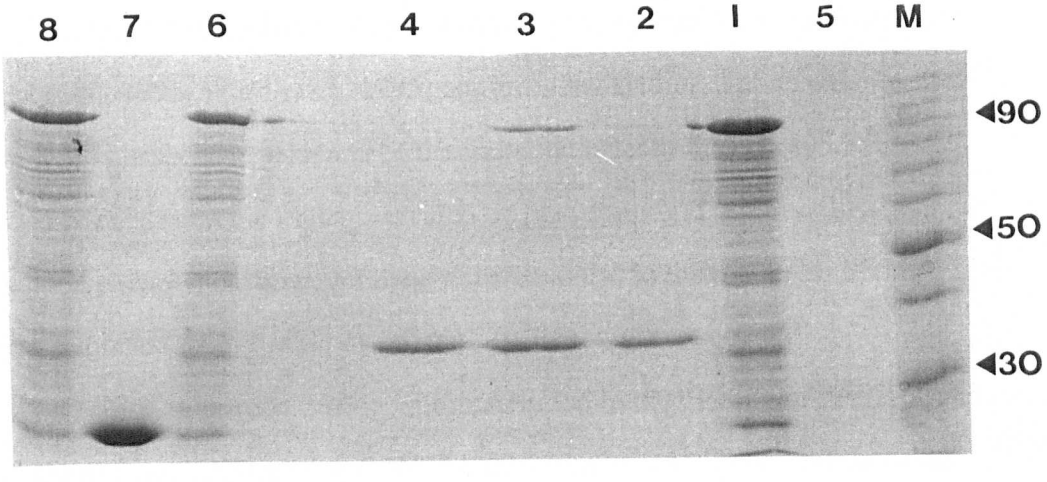


Fig 5.24 *In vitro* binding of His-FliG to the GST-FliF fusion protein. Lanes contained as follows: M, 10KDa ladder (Gibco BRL); 1 & 8, purified GST-FliF fusion protein as used in the assay; 2, purified His-FliG as used in the assay; 3, immidazole elution from His-FliG immobilised on Talon resin to which GST-FliF was added; 4, immidazole elution of immobilised His-FliG to which GST was added; 5, immidazole elution from Talon resin to which GST-FliF was added; 6, flowthrough from Ultrafree filter unit to which GST-FliF was added and 7, flowthrough from Ultrafree filter unit to which GST was added. Sizes are in KDa.


8
7
6
4
3
2
I
5
M
90
50
30

5.3 Discussion

The failure of the initial attempts to clone *fliF* and *fliG* using heterologous hybridisation and functional complementation were obviously due to the differences between the genes/proteins from *R.sphaeroides* and those from *S.typhimurium* and *C.crescentus* as later shown. The fact that the genes were not clustered with the motor genes (*motA* and *motB*) is not surprising as such a situation has not been seen in any other organism. However, the lack of any phenotype for the insertions near the *motA/B* operon (section 5.2.3) is surprising as this suggests that the motor gene operon exists with non-flagellar/essential coding regions either side. In *S.typhimurium* and *E.coli* the motor genes are clustered with the chemotaxis genes (Slocum & Parkinson, 1983) and the identification of the chemotaxis gene operon from *R.sphaeroides* (Ward *et al.*, 1995; Ward *et al.*, 1995) confirms my finding that no chemotactic phenotype is associated with any of the mutation created in section 5.2.3. The possibility exists that the mutants created do possess subtle non-motility associated phenotypes, however they were of no use to the main aim of this project to pursue them further as no motility defect was associated.

The use of 'operon analogy' ultimately led to the cloning of the *fliF* and *fliG* genes from *R.sphaeroides* WS8 confirming that a similar operon exists in *R.sphaeroides* as does in the enteric bacteria, although the direction of transcription of *fliE* differs in *R.sphaeroides* i.e. in enteric bacteria *fliE* transcription proceeds opposite to that of *fliF* whereas in *R.sphaeroides* it is transcribed with *fliF*. Complementation analysis of the *fliF* mutants suggested that the *torf, fliE, fliF* and *fliG* genes are all co-transcribed with the *R.sphaeroides fliI* gene identified in chapter 4. There appears to be a region between *fliG* and *fliI* that would presumably be transcribed. I was unable to complete the sequencing of this region due to time restrictions and could therefore not identify the open reading frame. Ballado and co-workers have shown that upstream of the *fliI* gene from *R.sphaeroides* 241 is a partial open

reading frame with very poor homology to the FliH protein from *S.typhimurium* (Ballado *et al.*, 1996) which is involved in the flagellar specific export pathway (Vogler *et al.*, 1991). A similar case may exist in *R.sphaeroides* WS8 as the intergenic region (approximately 1.2 Kb) is more than large enough to accommodate the *fliH* gene. The size of the *S.typhimurium fliH* gene is 705bp (Vogler *et al.*, 1991) suggesting that an additional gene may lay within this region or that the *R.sphaeroides fliH* gene is larger than its homologues. Evidence for either awaits sequencing of this region.

The identification of a σ^{54} consensus sequence upstream of the *torf* gene and the lack of any other promoter sequences, suggests that the operon starts upstream of the *torf* gene although direct confirmation awaits transcript mapping. The presence of repeated sequences within the region of the consensus sequence is also consistent with the operon being transcribed by σ^{54} (Kutsu *et al.*, 1989). The fact that these repeated sequences are not homologous to any previously identified repeated sequences suggests that the EBP binding to these repeats is novel. As stated in chapter 1, the *motA/B* operon has also been shown to contain a σ^{54} consensus sequence and a set of inverted repeats (Shah & Sockett, 1995) which differ from the repeats found upstream of the *torf* operon. This is consistent with the fact that, in *S.typhimurium* at least, the motor proteins are the last proteins to be expressed (Kubori *et al.*, 1992) i.e. two different EBP's act at the two operons and are probably regulated themselves. The presence of two sets of different inverted repeats within the promoter region may suggest that two different EBP's regulate in conjunction with each other, possibly being regulated with respect to different signals. Again direct proof of this awaits more extensive analysis of the promoter region as it may be that Torf autoregulates expression of itself..

Most σ^{54} promoters contain binding sites for integration host factor (IHF) which is thought to help EBP-σ^{54} interactions by looping of the intervening DNA (Perez-Martin *et al.*, 1994). As no IHF binding site was

present in the *torf* promoter region, the A/T rich region may play a role in this DNA bending as it has been shown that A/T rich regions cause bends in DNA (Koo *et al.*, 1986).

The presence of a region with homology to the *puf* regulatory sequence suggests that premature transcriptional termination may control the expression of the operon under aerobic conditions (Shimada *et al.*, 1993). No further literature is available on the nature of this regulatory protein which prevents me from making any other conclusions. Whether or not this termination occurs awaits transcript mapping experiments but may explain the previously noted effects of aerobic conditions on motility (Foster, 1991).

The first gene identified in the operon, *torf*, encodes a protein with extensive homology to various σ^{54} EBP's (see Fig 5.11) having the greatest degree of homology with the flagellar gene regulators FlbD (Ramakrishnan & Newton, 1990) and FleR (Richings *et al.*, 1995) (Fig 5.12). This is suggestive of the Torf protein being involved flagellar gene regulation but this could not be proved due to problems with the instability of clones containing this region.

What are the features of the Torf protein. EBP's are divided into three domains based on protein sequence alignments (Morett & Segovia, 1993). The N-terminal domain is the most variable (Fig 5.13) and is thought to constitute the signal reception domain (Morett & Segovia, 1993) i.e. the domain that either directly or via another protein senses a signal for activation of the transcriptional activation domain..

Is the activity of the Torf protein modulated? The methods used for signal reception can be separated into three families: 1) regulation by phosphorylation, 2) regulation by protein:protein interaction and 3) direct effector activation (Shingler, 1996). If it is regulated, the Torf protein falls into family 2 or 3 as the conservation of the residues implicated in being important for phosphorylation (see Fig 5.11) (Volz, 1993) is poor whereas FlbD and FleR appear to fall into the first family.

There are no proteins which have homology to the complete activator domain of the Torf protein, however, several proteins were identified that possessed small regions of homology (Table 5.1) but there appears to be no common sequence similarity between these proteins. The regions of homology within these proteins have not been assigned any specific function, consequently no idea as to the significance of the homology to the Torf can be given here and awaits a more extensive analysis of the Torf protein. The homology with RuvB, a protein involved in homologous recombination (Smith & Robinson, 1995), is more extensive (Table 5.1) although the degree of identical residues is low. This is suggestive of structural homology and indeed there are regions of structural homology as shown as the boxed regions in Fig 5.14. These regions are proposed to form similar secondary structures which may suggest a similar function but again extensive analysis of this region in RuvB has not been carried out. A role for this region will be discussed below.

The central region of EBP's is the most conserved region (Fig 5.11) (Morett & Segovia, 1993) and within this region Morett and Segovia have identified 7 conserved regions (labelled 1 - 7 in Fig 5.11). They implicated regions 1 and 5 in ATP binding and hydrolysis based on sequence comparisons with other ATP-bind proteins. As this region is highly conserved in the Torf protein, it could be predicted that it also functions by ATP-hydrolysis during open complex formation. Region 3 has been implicated in contacting σ^{54} and this is also highly conserved in the Torf protein. The other conserved regions have not yet been assigned a function.

Does the Torf protein bind DNA? The C-terminus of all EBP's studied to date possess a DNA-binding helix-turn-helix motif and as is obvious from Fig 5.11, the Torf protein clearly lacks this region. Consequently the interaction of the Torf protein with DNA is debatable i.e. it may interact with DNA via a previously unknown mechanism possibly via the regions which possess homology to RuvB or it may interact with another DNA-binding protein which

then binds to promoter regions. The possibility exists that the Torf protein does not bind to DNA but in fact acts as an anti-sigma factor by binding to σ^{54} and preventing it functioning, as FlgM does for σ^{28} in *S.typhimurium* (Ohnishi *et al.*, 1992). This is unlikely as the ATP-binding domains are very highly conserved and are thought to function during the formation of open complexes during transcription initiation (Weiss *et al.*, 1988; Morett & Segovia, 1993) and presumably would not be required if the function of the Torf protein was simply to repress the activity of σ^{54} by binding to it. I therefore propose that the Torf does bind to DNA, probably via an interaction with another protein which possess a DNA-binding motif which is novel for σ^{54} EBP's. The presence of different enhancer like elements (ELE's) upstream of the *torf* and *motA/B* operons suggest that they are regulated by different EBP's and the possibility exists that the Torf protein is responsible for the regulation of *motA* and *motB* as well as itself. As will be discussed below, it may be that the Torf is a 'universal' flagellar gene regulatory protein and it functions to regulated all σ^{54} dependant flagellar genes via its interaction with other small proteins that confer DNA-binding specificity and activity to the Torf protein. This would minimise the number of large regulatory proteins required to regulate the expression of flagellar genes and would presumably require less metabolic input from the cell.

To understand the role of the Torf protein further studies are needed. It is obvious, due to the uniqueness of this protein that it plays a novel role in flagellar related transcriptional regulation and may represent a new class of transcriptional regulators.

The FliE protein from *R.sphaeroides* is highly conserved at its C-terminus with other FliE proteins (Fig 5.16) and a Ω mutation that removes the last 8 residues was found to be dominant. This is the first dominant mutation to be isolated in any *fliE* gene studied to date, but **what is the role of FliE?** As stated in section 5.2.4f, Muller and co-workers suggested that FliE protein formed an adapter between the rod and FliF, however as they stated their data

does not confirm this conclusively and further work is required. As they stated, a previous study had shown that FliE formed part of the 'rivet' structure consisting of the MS-ring and the rod (Suzuki *et al.*, 1978; Suzuki & Komeda, 1981) but the position in the rivet is unknown. Recently, Aizawa has tentatively demonstrated that FliE is responsible for the integrity of the 'cytoplasmic rod' (C-rod), a structure thought present on the cytoplasmic face of the M-ring, to be part of the flagellar specific export apparatus (see chapter 4 for a greater description) (Aizawa, 1996). It may be that FliE is an adapter between the MS-ring and the other rod proteins, but that part of it projects into the centre of the export channel and plays some role in the export procedure. It is likely that FliE functions only to stabilise the interactions of the export apparatus with the MS-ring as no functional homology between FliE and any export proteins has been noted although this is yet to be proved.

The homology of RSFliE with RecA proteins is intriguing as FliE has not been implicated in DNA-interactions. Whether or not the homology is functionally significant is not clear as the residues that are conserved, excluding the highly conserved region (LGXXVDDL(X)3QXDXGE), consist mainly of I, L and A residues and have no obvious function. Further site-directed mutagenesis of this region may help determine the role of the highly conserved region.

The fact that the *fliE* mutant (WS8::*fliE* Ω) has no internal flagellin, is the first evidence that a regulatory feedback mechanism exists in *R.sphaeroides* flagellum formation whereby morphological checkpoints control expression of distal components. **How does the regulatory feedback mechanism function in *R.sphaeroides?*** How this functions can not be determined from this project, but it is unlikely that it is the result of the export of an anti-sigma factor as seen in the enteric bacteria (Ohnishi *et al.*, 1992) as SDS-PAGE analysis of culture supernatants of *R.sphaeroides* cultures shows only the presence of flagellin and no other secreted anti-sigma factor proteins (data not

shown). It is probable that it functions via the activation/repression of transcriptional regulators possibly by the undetermined mechanism noted in *C.crescentus* (Ramakrishnan *et al.*, 1994). A model could be proposed whereby the Torf activates a variety of different flagellar gene promoters via different specificity proteins (i.e. they confer different DNA-binding specificity's to the Torf protein) and these specificity proteins are themselves regulated by morphological checkpoints i.e. the formation of flagellar substructures. It could be that the specificity-proteins form part of these substructures when they are first expressed, possibly acting as 'scaffold proteins' (or 'macro-chaperones') being involved in stabilising certain substructures of the flagellum until the remaining structures are formed. They would then be released at different points whereby they could interact with the Torf protein or activate other proteins involved in conferring specificity to the Torf which would lead to activation of Torf dependant genes. There is no data to prove or disprove this model but the data presented in this project is consistent with this. A more detailed investigation of the transcripts present in various mutants etc. may help determine how flagellum formation is regulated in this uni-flagellate organism.

The FliF protein from *R.sphaeroides* has all the conserved featured of a typical FliF protein (Fig 5.18) i.e. two membrane spanning helices. As the RSFliF protein reported in this project is the only FliF protein to be rotated unidirectionally, any conserved features with the bi-directionally rotated FliF proteins are significant and may give some clues as to how the proteins function. It is very obvious that the most conserved region is the region thought to constitute the central core of the MS-ring complex (Fig 5.18) and given the homology of this region with export proteins, it is also clear that this region plays some role in the export procedure.

What is the role of FliF in the export procedure? Examining the degree of conservation of all FliF proteins with the export proteins (data not

shown) it is probable that the homology is due to the structure that the proteins form i.e. an export channel. It is of my opinion that the central core of the MS-ring plays a role in flagellar specific export probably creating an environment through which the distal components can easily pass .

Almost all of the absolutely conserved residues in FliF proteins are located in the N-terminal region and are hydrophobic (Fig 5.18). This is consistent with this region being buried in the centre of the MS-ring complex. How the proteins pass through this channel is not known but may involve chaperones (see chapters 1 and 4) which may interact with the proteins being exported possibly creating a hydrophilic protein coat on them. This hydrophilic coat would presumably be extruded from the hydrophobic core of the MS-ring complex causing the partially folded protein to be exported to the distal end of the flagellum where it is assembled. However, it may simply be that the procedure carried out by FliI, i.e. recruitment of flagellar components to the export procedure, 'forces' flagellar components into the export channel at the cytoplasmic side which causes the components already in the channel to be extruded from the distal end where they are assembled. The requirement for energy input in the form of ATP hydrolysis may help the process whereby FliI 'forces' components into the channel. The data from chapter 4 agree with this idea as the absence of FliI would mean that components would not be actively 'forced' into the channel and assembly would be very slow. Combined with the specificity switch required for export of flagellin (see chapter 4), the observed number of filaments formed in a *fliI* mutant is consistent with this idea but direct proof awaits more in depth study of the role of the residues in the channel.

Does FliF play any role in torque generation? It is clear from the absence of any conserved residues in the region thought to constitute the M-ring (MSH2 in Fig 5.18) and consequently to be in close proximity to the motor proteins, that FliF does not play any role in torque generation. This is in agreement with Homma and co-workers (Homma *et al.*, 1987). Given the data

above, the role of FliF is probably threefold: 1) function as a 'mounting plate to which flagellar components are 'bolted'; 2) formation of an export channel and 3) in uni-flagellate organisms, function as the 'site determining' factor in flagellum formation. The ability to function as describe in 3 appears to be sequence related from data obtained in *C.crescentus* (Jenal & Shapiro, 1996) and a similar case may exist in *R.sphaeroides* although the residues playing this role in *C.crescentus* are not conserved in RSFliF. The 'cell cycle' controlled motility may also be sequence determined and with the FliF antibody produced in this project studies could now be carried out to address this and targeting questions:

How is FliF targeted in *R.sphaeroides* ? Examination of the cytoplasmic membrane proteins in Nm7 revealed that a protein of 51.5KDa is absent in Nm7 but present in WS8 and 'psuedo complemented' Nm7 (Fig 5.20). The large amount of this protein present, and the fact that it seems to be present in purified flagellated vesicles (Foster, 1991) may suggest that it does play some role in flagellum formation / function. It may be that this protein makes up the E-ring as described in the introduction, but given the fact that it is present at much higher levels than FliF (Fig 5.20) this is unlikely as the E-ring would have the same stoichiometry as FliF. The possibility therefore exists that this protein represents another novel component of the *R.sphaeroides* flagellum, possibly playing a role in targeting. A more detailed characterisation of this protein e.g. N-terminal sequence determination may reveal the nature of it

Western blot analysis revealed that besides FliF, another protein in the cytoplasmic membrane, periplasm and outer membrane specifically cross-reacts with the FliF antisera (Fig 5.20) The size of this protein (45.8KDa) correlates with the size of the porin protein from *R.sphaeroides* (Baumgardner *et al.*, 1980; Weckesser *et al.*, 1984), which may have isolated with all three fractions due to its location. This may reflect that RSFliF shares some structural homology with the porin or that it is cross-reacting with the antibodies that

recognise the GST portion of the fusion protein but it was impossible to determine which.

The effect of intra-cytoplasmic membrane (ICM) formation, as occurs during the shift from aerobic to photosynthetic growth, on flagellum targeting was not addressed in this project. Previous studies have suggested that the flagellum is assembled in an un-differentiated membrane 'patch' as no photosynthetic apparatus was found in flagellated vesicles (Foster, 1991). With the FliF anti-serum produced as a result of this project, this process of flagellum targeting can now be addressed i.e. is FliF turned over during the transistion to phototrophic growth?

The ability of FliM to complex with FliF (Oosawa *et al.*, 1994) would presumably be conserved in *R.sphaeroides* but **What is the role of this interaction?** It may simply add another point of stabilisation for the switch complex or the export apparatus i.e. stabilising the position of FliN in the export apparatus. It is also possible that the interaction between FliF and FliM is different in each organism which may play some role in torque generation. This is mirrored by the fact that the *R.sphaeroides* FliM protein is markedly different from other FliM proteins suggesting that it plays a major role in determining the direction of rotation. However, it is of my opinion that this interaction is simply to stabilise the switch complex / export apparatus in the correct position for them to function correctly.

The presence of the switch protein FliG in *R.sphaeroides* is as expected given its role in torque generation but - **What is the role of FliG?** As stated previously, a large degree of evidence is present to implicate FliG in torque generation (Garza *et al.*, 1995; Garza *et al.*, 1996; Tang *et al.*, 1996) and it would be predicted to play a similar role in *R.sphaeroides*. Given the fact that FliG is well conserved between a uni-directional flagellum and bi-directional flagella, it could be predicted that it plays only a minor role in switching.

However, given the high number of mutations isolated in FliG that result in switching bias (see Table 5.3) (Irikura *et al.*, 1993) this is unlikely.

Does the RSFliG protein contain the conserved features of FliG proteins? Several Domains of FliG have been implicated in playing certain roles; the domain labelled A in Fig 5.22 has been shown to be responsible for FliF-FliG interactions (Oosawa & Hayashi, 1986; Jenal *et al.*, 1994; Marykwas *et al.*, 1996). There appears to be no common sequence that could be predicted to be the FliF binding site, however, all FliG proteins appear to possess a region of hydrophobicity centred around position 20 in Fig 5.22. The region of FliF implicated in binding to FliG, namely the C-terminus (Oosawa & Hayashi, 1986; Jenal *et al.*, 1994; Marykwas *et al.*, 1996), lacks any obvious sequence homology between FliF proteins but does include part of the second membrane spanning helix. This, together with the conserved region of hydrophobicity in FliG proteins may suggest that the binding occurs in the membrane. The *in vitro* binding assay carried out in this project demonstrates that the RSFliF and the RSFliG proteins do interact. This assay was not quantitative as the FliF protein used contained many breakdown products and this would not allow an accurate estimation of the concentration of complete GST-FliF protein in the sample. Not all of the GST-FliF protein used in the assay possessed the ability to bind to His-FliG (lanes 1 & 3) which may have been due to the methods of purification used to isolate the proteins i.e. both were isolated under denaturing conditions and may be partially folded. Alternatively, the interaction may require a hydrophobic environment, such as would be present in a membranous environment. This assay does provides the first evidence in *R.sphaeroides,* for an interaction between any flagellar proteins and may be useful in determining the exact point of interaction between FliF and FliG as mutant FliF and G proteins could be purified and used in the assay.

There appears to be a high degree of conservation in the C-terminal third of the protein (domain B in Fig 5.22) i.e. 22 out of 126 residues are absolutely

conserved. As these residues, 12 of which are hydrophobic, 5 charged, 3 polar and 2 glycine residues, are highly conserved between the FliG proteins from bidirectional flagella and from the unidirectional flagellum of *R.sphaeroides* (Armitage & Macnab, 1987), they are possibly the more important residues in torque generation. Within this region is the motif GPVRL which is absolutely conserved in all FliG proteins and has been shown to be important for motility as a deletion within this region (see Fig 5.22) has been shown to cause Mot$^-$ phenotype in *S.typhimurium* and a Mot$^-$ / reduced flagellation phenotype in *E.coli* (Irikura *et al.*, 1993; Lloyd *et al.*, 1996). This motif is present in many other proteins in the Swissprot protein database held at Seqnet e.g. *E.coli* molybdenum cofactor biosynthesis protein C (MoaC) (Rivers *et al.*, 1993), *P.aeruginosa* Fe(III)-phyochelin receptor precursor (Ankenbauer & Quan, 1994) and *T.aquaticus* phytoene synthesis gene CrtB (Hoshino *et al.*, 1993), but the significance of this is unclear as they shown no obvious functional similarity. Other deletions shown to cause Mot$^-$ mutations in *S.typhimurium* and Mot$^-$/reduced flagellum formation in *E.coli* are shown in Fig 5.22.

A region recently shown to be responsible for the self association of FliG and its interaction with FliM and N (Tang *et al.*, 1996) also appears to be conserved. This region (Q 227 to R 245 in the *E.coli* sequence) is highly conserved in all FliG proteins (Fig 5.22) but no structural homology is predicted (data not shown). This domain is not sufficient for FliG self association or FliM/N binding (Tang *et al.*, 1996) which is consistent with the finding the a region spanning approximately the central third of the *E.coli* FliG is important for the FliG/M interaction (Marykwas & Berg, 1996).

How does the *R.sphaeroides* FliG differ from that from bi-directional flagella? The differences between FliG from *R.sphaeroides* and those from bi-directional flagella would in principle, given the role of FliG, give a good insight into the mechanisms that cause flagellum rotation and switching. However, in the light of the fact that FliG's are quite similar, whereas the motor

proteins are very different (Shah & Sockett, 1995; Shah & Sockett, 1995) it is possible that it is the motor proteins that govern the unidirectionality. Alternatively, the difference of the motor proteins may simply reflect the different membranous environments that the flagella are position in i.e. the flagellum of *R.sphaeroides* is present in a membrane which possesses the photosynthetic apparatus whereas the others do not. As a consequence of this, *R.sphaeroides* may have developed a novel method of torque generation. RSFliG does still possess some interesting features: it appears that two mutations isolated in the *S.typhimurium* FliG protein are present naturally in the RSFliG protein; namely G185A and I133T which result in a cw biased flagellum (Irikura *et al.*, 1993), and V135A and H155P which cause poor FliG/M interaction (Marykwas & Berg, 1996). The presence of these cw biased mutations could be interpreted to contribute to the cw rotation of the flagellum, however, there are numerous other cases whereby mutation in *S.typhimurium* FliG that result in a cw biased flagellum is present naturally in the FliG proteins from other bidirectionally rotating flagella (data not shown). It is therefore unlikely that these residues exclusively determine the direction of rotation but it would be interesting to determine the flagellar bias of the reverse mutations in RSFliG.

How does FliG function in torque generation? It has previously noted that in *S.typhimurium* FliG there appears to be clustering of charges (Kihara *et al.*, 1989) and the authors suggested that this may play some role in torque generation. However, in the light of the identification of other FliG proteins and the FliG protein from *R.sphaeroides*, this conservation of charges appears not to be as extensive as would be predicted if they played a large role in torque generation. Despite this, there do appear to be some conserved charged residues in the domain important for motility, namely the C-terminal domain (see Fig 5.22) and almost all of the conserved charged residues are predicted to form alpha helical structure (data not shown). Within the C-terminal region there is a

motif whereby a positively charged residue is followed by two negatively charged residues (RDE in *R.sphaeroides*) and deletion of the positively charged residue and 5 upstream residues in *S.typhimurium* FliG (see Fig 5.22) has been shown to give a motility negative phenotype (Irikura *et al.*, 1993). This motif is mirrored in the N-terminus of FliG proteins whereby two negatively charged residues are followed by a positively charged residue. There appears to be no obvious difference between RSFliG and other FliG proteins with respect to charge distribution that might account for the unidirectionality. Whether or not the conserved charged residues play a role in torque generation is not known but would be obvious targets for site directed mutagenesis. As discussed above, the motif GPVRL is absolutely conserved in all FliG proteins and must therefore play a critical role. It is probable, given the conserved nature of the regions, that this C-terminal region interacts with the charged cytoplasmic loop of MotA (Shah & Sockett, 1995). This is in agreement with the recent findings of Tang and co-workers who showed that it is the C-terminus domain that is responsible for the interaction of FliG with MotA (Tang *et al.*, 1996). The low level of interaction that they detected, i.e. MotA had to be overexpressed and could still only be detected by Western blot, may have been due to the fact that the domains need to be charged for their correct interaction. The use of different buffering condition was not examined by the authors but may have resulted in a larger quantity of MotA being isolated.

The presence of a lipid attachment site in RSFliG is interesting as it is obvious that FliG is not an outer membrane protein (based on homology to other FliG proteins) and as a consequence the site would presumed not to function as the signal peptidase responsible for this processing may not have access to the protein. However, this region also appears to be hydrophobic (data not shown) and combined with the fact that it lies in a region shown to be important for FliG/M interactions (Marykwas & Berg, 1996) this may represent a region which transiently interacts with the membrane. It may be the case that

in *R.sphaeroides* the unidirectionality of the motor is controlled by FliM which is consistent with the high degree of difference observed between the *R.sphaeroides* FliM protein and other FliM proteins (Pollitt, 1996;). It may be that FliM alters the conformation of FliG so that the lipid attachment site / hydrophobic region is inserted into the membrane. A more detailed examination of FliM and comparisons with the FliM proteins from bi-directionally rotated flagella may give some idea as to its function and possibly to how flagella rotate.

In conclusion, this part of the project identified and partially characterised a number of flagellar structural proteins and a probable transcriptional regulator. The identification of two motor components in this project, combined with the recent identification of the remaining motor components FliM and FliN , should now enable a detailed investigation of motor function in *R.sphaeroides* WS8. This work will presumably form the basis of a more directed investigation utilising mutagenesis which may reveal the processes that allow flagellar rotation. .

Chapter 6

Concluding discussion

As a result of this project, a motility gene operon was identified and characterised. Unfortunately, due to time constraints a lot of questions could not be answered, some of which include:

How does FliI function in the export pathway?

Are flagellum components in *R.sphaeroides* specifically targeted and turned over during the cell cycle?

What are the roles of the conserved motifs in FliG?

Does RSFliG interact with MotA via the cytoplasmic loop?

From the data generated in this project, most of these questions can now be addressed, but the main question to be addressed is still **how does flagellar rotation occur?** From the data I have obtained, comparing flagellar proteins from *R.sphaeroides* with those from bidirectionally rotated flagella and reviewing the literature, I feel that it is possible to propose a model for rotation. The reader is reminded that this project was undertaken from a molecular standpoint and as a consequence, many of the bio-physical characteristics of the motor as determined by Blair, Berg and Berry etc. are beyond the understanding of this molecular biologist. Therefore the models proposed below, although conforming to current knowledge of the flagellum from a molecular biologists view, may not conform to bio-physical properties of the flagellar motor.

6.1 Unidirectional flagellum rotation - a hypothetical model

A model could be proposed whereby the rotation of the flagellum occurs via the interaction of the C-terminal domain of FliG with the cytoplasmic loop of MotA, with MotB functioning as the elastic linkage as depicted in Fig 6.1. In this model, which could be likened to the models of Berg-Khan and Lauger

depicted in section 1.4, the ligand row would be multiple copies of FliG bound to the base of FliF (Fig 6.1 A). It could be proposed that when a proton enters the channel it causes a conformational change in MotB which then allows the proton to pass into the cytoplasm i.e. the cluster of proton accepting residues on the cytoplasmic side of MotB (as noted in the introduction) channel the protons into a partially formed channel, which is then completed when MotB flexes. This would allow the passage of the protons into the remainder of the channel which would be made up of the membrane spanning helices of MotA (Fig 6.2 B). This might explain the findings of Sharp and co-workers as they predicted that the membrane spanning helix of MotB is tilted during its interaction with the MotA membrane spanning helices (Sharp *et al.*, 1995). This tilting of MotB may only occur transiently to form the channel, but this would not have been detected by Sharp and co-workers as their methodology would not allow it. The flexing of MotB would then bring the charged cytoplasmic loop of MotA into close proximity with the C-terminal domain of a FliG protein. The interaction between the cytoplasmic loop and the conserved residues in the C-terminal domain of FliG would then occur and MotB would contract as the proton passes into the cytoplasm (Fig 6.1 B then A). The continuation of this would result in rotation of FliG and consequently FliF, leading to flagellum rotation. The stopping of the flagellum may be the result of FliM altering the conformation of FliG so that the domain interacting with MotA is 'out of reach' of the cytoplasmic loop (Fig 6.1 C). In this case, proton flux would continue as would the extension/contraction of the MotA/MotB complex (Fig 6.1 D). Alternatively, FliM may interact with MotA so as to cause the opening of the channel, i.e. hold the structure in the extended position where the channel is open. This would then allow proton flux to occur but not extension/contraction of the MotA/MotB complex. This proton flux may be at a reduced level and would explain the 'slippage' or lose coupling of the motor to proton flux (Kaplan *et al.*, 1983).

In this model, the interaction of the cytoplasmic loop with FliG would be transient, but **how is this transient interaction controlled?** It may be that the charge of the MotA cytoplasmic loop is neutralised when the proton passes into the cytoplasm so that it can no longer interacts with FliG. Alternatively, it may be that the interaction is 'broken' by the torque generated by other MotA/MotB complexes i.e. force generating units would presumably be positioned so as to allow continuous rotation of the flagellum and hence when one force generating unit had reached it fully contracted state, there would be several others still contracting and these together may generated enough force to 'break' the interaction. The 'breaking motor's' experiments of Berry and co - workers (Berry *et al.*, 1995) tend to argue for the former. They found that when the flagellum from a cell that can rotate its flagellum only in one direction (chemotaxis deficient *E.coli* cells) was rotated backwards (cw for *E.coli*), it caused it to 'break' i.e. a structural failure within the motor. In the situation whereby the transient interaction is controlled by charge neutralisation, it could be imagined that when the motor was rotated backwards, the interaction between MotA and FliG still remains intact and this may be strong enough to cause the MotA protein to become misalligned with MotB, causing the torque generating unit to effectively stop working. This would not occur if the latter method of transient interaction occurred. Berry and co-workers suggest that it is the MotA/B torque generators that are affected during the 'breaking' as the motors were seen to recover in a series of steps, similar to what is observed when motors lacking either of the motor proteins are supplied with the corresponding protein by induction of expression from a plasmid (Blair & Berg, 1988).

The fact that the *R.sphaeroides* MotB protein lacks the conserved peptidoglycan binding motifs that are thought to attach the periplasmic domain of MotB to the outer membrane suggests that it interacts with the outer membrane in a different manner (Shah & Sockett, 1995). It may be that it

interacts with the P ring, which is thought to remain static whilst the rod rotates within it. Direct evidence for this would presumably come from the co-isolation of the P-ring protein with MotB.

The ability of *R.sphaeroides* to produce a low number of off-spring that rotate their flagella ccw (Packer & Armitage, 1993) is probably due to the mis-incorporation of the MotA/B force generating units into the membrane so as they insert in the opposite orientation i.e. MotB on the right side of MotA. This 'protein phase variation' would explain why the so called 'petite' variants of *R.sphaeroides* that rotate their flagella ccw, do not breed true. This process may occur during the process of photosynthetic membrane differentiation.

6.2 Bidirectional rotation

A similar model could be proposed for the bidirectional flagellum, however, in this case FliM would probably function to alter the conformation of FliG so as to position it closer to the cytoplasmic loop of MotA when it MotB is in the contracted position. This would then result in the 'pushing' of the flagellum in the other direction during switching. The positioning of the force generating units would also be reversed so as to favour the ccw direction of rotation..

As stated above, this model does comply with what is currently known about the molecular nature of the motor but whether or not it would comply with the bio-physical properties of the motor as detected would require extensive modelling which is far beyond the scope of this project.

6.3 Future work

This project has detailed the cloning and partial characterisation of flagellar structural, regulatory, rotor and export proteins from the unidirectional flagellum of *R.sphaeroides*. Due to time constraints, many avenues of

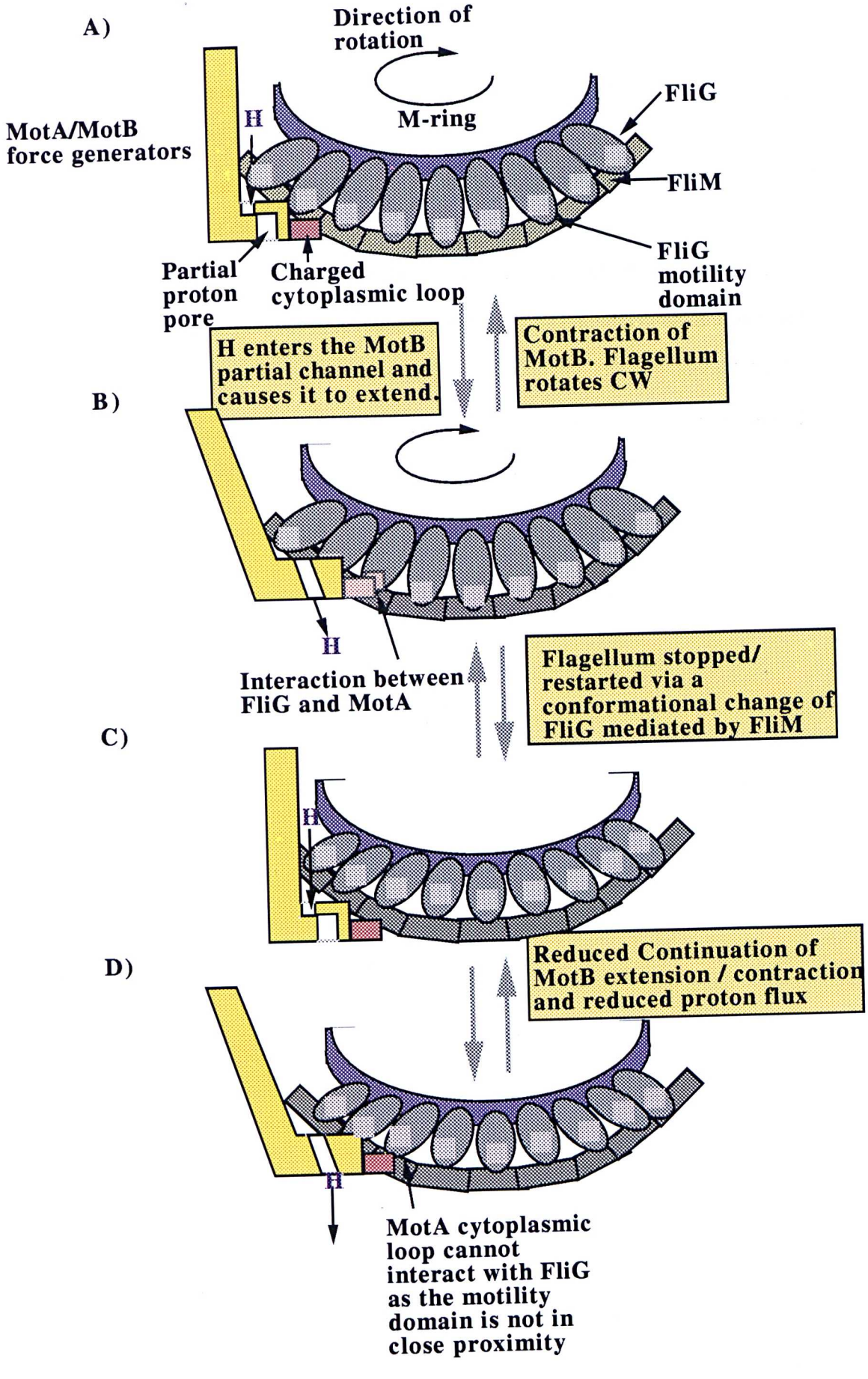
investigation could not be pursued further and what follows is description of what I feel would be the more worthy avenues to pursue.

The elucidation of the method of flagellum rotation was the ultimate goal of this project and others within the research group. The cloning of the rotor proteins FliF and FliG can now allow a more in depth investigation as to the function(s) of these proteins in the unidriectional motor of *R.sphaeroides* i.e. flagellar targeting and rotation. I feel that the most obvious target for further work is the rotor protein FliG. The charged groups within the C-terminal motility domain (discussed in section 5.3) are would be good targets for directed mutagenesis. Testing the effects of the mutations created by complementation is a big problem, as the construction of non-polar mutations in *fliG* will be difficult given its position in the *torf* operon. The use of chemical mutagenesis to create such non-polar mutations was not investigated in this project but offers one possible way of isolating a non-motile strain that could be complemented by the *fliG* gene alone. The isolation of such a strain, in my opinion, is crucial for further investigations to be carried out into the functions of the conserved residues within the motility domain of FliG.

The recent identification of the remaining motor components, FliM and FliN, from *R.sphaeroides*, will allow the interactions between FliG and FliM to be investigated. The *in vitro* binding assay used in this project could be used on mutant FliM and FliG, or FliF, proteins to determine their points of interaction and give some insights into how they function in the motor.

The role of FliF in flagellum targeting can also now be investigated using the FliF antiserum produced in this project. The turnover of FliF during photosynthetic membrane differentiation and the cell cycle are obvious avenues to be pursued. The positioning of the FliF protein in the membrane with respect to the

Fig 6.1 Proposed model for flagellum rotation in *R.sphaeroides*. For simplicity only one MotA/MotB force generating unit, part of the MS-ring complex shown is shown and FliN is also missing although it has been identified in *R.sphaeroides* (Sockett, 1996).


A)
Direction of rotation
M-ring
FliG
FliM
FliG motility domain
MotA/MotB force generators
H
Partial proton pore
Charged cytoplasmic loop
H enters the MotB partial channel and causes it to extend.
Contraction of MotB. Flagellum rotates CW
B)
H
Interaction between FliG and MotA
Flagellum stopped/ restarted via a conformational change of FliG mediated by FliM
C)
H
Reduced Continuation of MotB extension / contraction and reduced proton flux
D)
H
MotA cytoplasmic loop cannot interact with FliG as the motility domain is not in close proximity

photosynthetic apparatus could also be addressed i.e. does FliF only insert into regions devoid of the photosynthetic apparatus?

The flagellar rotation model predicted in sections 6.1 and 6.2 could also be investigated using mathematical models to determine if it confers to what is known about the motor bioenergetically. Again this was beyond the scope of this project.

In conclusion, this project has revealed numerous interesting features of the unidirectional flagellum of *R.sphaeroides*, and the results from the analysis of the export component FliI (detailed in chapter 4) have highlighted the potential of this organism in revealing previously unknown observations that could not be obtained from the enteric.

References

Abdelmounaaim, A., Sansonetti, P. E. & Parsot, C. (1992). "MxiJ, a lipoprotein involved in secretion of *Shigella* Ipa invasins, is homologous to YscJ, a secretion factor of the *Yersinia* Yop proteins." *J. Bacteriol.* **174**(23): 7661 - 7669.

Adler, J. & Templeton, B. (1967). "The effects of environmental conditions on the motility of *Echerichia coli*." *J. Gen. Microbiol.* **46**: 175 - 184.

Allen, L. N. & Hanson, R. S. (1985). "Construction of broad-host-range cosmid cloning vectors: Identification of genes necessary for growth of *Methylobacterium organophilum* on methanol." *J. Bacteriol* **161**: 955-962.

Anderson, D. K., Ohta, N., Wu, J. & Newton, A. (1995). "Regulation of the *Caulobacter crescentus rpoN* gene and function of the purified σ^{54} in flagellar gene transcription." *Mol. Gen. Genet.* **246**: 697 - 706.

Ankenbauer, R. G. & Quan, H. N. (1994). "FptA, the FE(III)-phycolin receptor of *Pseudomonas aeruginosa:* a phrnolate siderophore receptor homologous to hydroxanate siderophore receptors." *J. Bacteriol.* **176**: 307 - 319.

Armitage, J. P. (1992). "Bacterial motility and chemotaxis." *Sci. Progress Oxford* **76**: 451-477.

Armitage, J. P., Ingham, C. & Evans, M. C. W. (1985). "Role of proton motive force in phototactic and aerotactic responses of *R.sphaeroides*." *J. Bacteriol.* **161**: 967.

Armitage, J. P., Kelly, D. J. & Sockett, R. E. (1995). Flagellate motility, behavioural responses and active transport in purple non-sulfur bacteria. Anoxygenic photosynthetic bacteria . Kluwer Academic publishers. 1005 - 1028.

Armitage, J. P. & Macnab, R. M. (1987). "Unidirectional, intermittent rotation of the flagellum of *Rhodobacter sphaeroides*." *J. Bacteriol* **169**: 514-518.

Arnold, W., Rump, A., Klipp, W., Priefer, U. B. & Puhler, A. (1988). "Nucleotide sequence of a 24206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumonia*." *J. Mol. Biol.* **203**: 715 - 738.

Arora, S. K., Ritchings, B. W., Elmira, E. C., Lory, S. & R., R. (1996). "Cloning and characterisation of the *Pseudomonas aruginosa fliF* gene involved in flagellar biosynthesis and bacterial adherence to mucin." *Infect. Immun.* **64**(6): 2130 - 2136.

Ballado, T., Campos, A., Camarena, L. & Dreyfus, G. (1996). "Flagellar genes from *Rhodobacter sphaeroides* are homologous to genes of the *fliF* operon of *Salmonella typhimurium* and to the type-III secretion system." *Gene* **170**: 69 - 72.

Baumgardner, D., Deal, C. & Kaplan, S. (1980). "Protein composition of *Rhodopseudomonas sphaeroides* outer membrane." *J. Bacteriol.* **143**(1): 265 - 273.

Benson, A. K. & Newton, A. (1994). "The role of FlbD in regulation of flagellar gene transcription in *Caulobacter crescentus*." *12th Forum in Microbiqlogy*: 420 -430.

Benson, A. K., Ramakrishnan, G., Ohta, N., Feng, J., Ninfa, A. & Newton, A. (1994). "The *Caulobacter crescentus* FlbD protein acts at *ftr* sequence elements both to activate and repress transcription of cell cycle regulated flagellar genes." *Proc. Natl. Acad. Sci. USA* **91**: 4989 - 4993.

Berg, H. C. (1974). "Dynamic properties of bacterial flagellar motors." *Nature* **249**: 77 - 79.

Berry, R. M. (1993). "Torque and switching in the bacterial flagellar motor - an electrostatic model." *Biophysical Journal* **64**: 961 - 973.

Berry, R. M., Turner, L. & Berg, H. C. (1995). "Mechanical limits of bacterial flagellar motors by electrorotation." *Biophys. J.* **69**: 280 - 286.

Bertin, P., Terao, E., Lee, E. H., Lejeune, P., Colson, C., Danchin, A. & Collatz, E. (1994). "The H-NS protein is involved in the biogenesis of flagella in *Escherichia coli*." *J. Bacteriol.* **176**: 5537 - 5540.

Birnboim, H. C. & Doly, J. (1979). "A rapid alkaline extraction procedure for the screening of recombinant plasmid DNA." *Nuc. Acids Res* **7**: 1513-1523.

Blair, D. F. (1995). "How bacteria sense and swim." *Ann. Rev. Microbiol.* **49**: 489 - 522.

Blair, D. F. & Berg, H. C. (1988). "Restoration of torque in defective flagellar motors." *Science* **242**: 1678 - 1681.

Blair, D. F. & Berg, H. C. (1990). "The MotA protein of *E.coli* is a proton-conducting component of the flagellar motor." *Cell* **60**: 439 - 449.

Blair, D. F., Kim, D. Y. & Berg, H. C. (1991). "Mutant MotB proteins in *Escherichia coli*." *J. Bacteriol.* **173**: 4049 - 4055.

Blair, D. F., Kim, D. Y. & Berg, H. C. (1991). "Mutations in the MotA protein of *Escherichia coli* reveal domains critical for proton conductance." *J. Mol. Biol.* **221**: 1433 - 1442.

Blattner, F. R., Burland, V. D., Plunkett, G., Sofia, H. J. & Daniels, D. L. (1993). "Analysis of the *Escherichia coli* genome IV. DNA sequence analysis of the region from 89.2 to 92.8 minutes." *Nuc. Acid. Res.* **21**: 5408 - 5417.

Bourret, R. B., Borkovich, K. A. & Simon, M. I. (1991). "Signal transduction pathways involving protein phosphorylation in prokaryotes." *Annu. Rev. Biochem.* **60**: 401 - 441.

Brun, Y. V., Marczynski, G. & Shapiro, L. (1994). "The expression of asymetry during *Caulobacter*cell differentiation." *Annu. Rev. Biochem.* **63**: 419 - 450.

Buck, M. & Cannon, W. (1992). "Specific binding of the transcription factor sigma-54 to promoter DNA." *Nature* **358**: 422 - 424.

Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). "XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase." *Biotechniques* **5**: 376.

Burchard, R. P. (1981). "Gliding motility of prokaryotes: ultrastructure, physiology and genetics." *Ann. Rev. Microbiol.* **35**: 497.

Cannon, W., Claveriemartin, F., Austin, S. & Buck, M. (1993). "Core RNA polymerase assists binding of the transcription factor sigma(54) to promoter DNA." *Mol. Microbiol.* **8**: 287 - 298.

Caplan, S. R. & Kara-Ivanov, M. (1993). "The Bacterial flagellar motor." *Int Rev. Cytol.* **147**: 97 - 164.

Chang, A. C. Y. & Cohen, S. N. (1978). "Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic plasmid." *J. Bacteriol.* **134**: 1141 - 1156.

Chen, J. C. & Mortenson, L. E. (1992). "Identification of six open reading frames from a region of the *Azotobacter vinelandii* genome likely involved in dihydrogen metabolism." *Biochim. Biophys. Acta.* **1131**: 119 - 202.

Chen, L.-S., Mullin, D. & Newton, A. (1986). "Identification, nucleotide sequence, and control of developementally regulated promoters in the hook operon region of *Caulobacter crescentus.*" *Proc. Natl. Acad. Sci. USA* **83**: 2860 - 2864.

Chien, C.-T., Bartel, P. L., Sternglanz, R. & Fields, S. (1991). "The Two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest." *Proc. Natl. Acad. Sci. USA* **88**: 9578 - 9582.

Chou & Fasman (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. in Enzymol.* **47**: 45 - 148.

Chun, S. Y. & Parkinson, J. S. (1987). "Bacterial motility: Membrane topology of the *Escherichia coli* MotB protein." *Science* **239**: 276 - 278.

Cohen-Bazire, G., Sistrom, W. R. & Stanier, R. Y. (1957). "Kinetic studies of pigment synthesis by non-sulphur purple bacteria." *J. Cell. Comp. Physiol.* **49**: 25 - 68.

Colbeau, A., Richaud, P., Toussaint, B., Caballero, F. J., Elster, C., Delphin, C., Smith, R. L., Chabert, J. & Vignais, P. M. (1993). "Organisation of the genes necessary for hydrogenase expression in *Rhodobacter capsulatus* - sequence analysis and identification of 2 Hyp regulatory mutants." *Mol. Microbiol.* **8**: 15 - 29.

Conley, M. P. & Berg, H. C. (1984). "Chemical modification of *Streptococcus* flagellar motors." *J. Bacteriol.* **158**: 832 - 843.

Cyr, D., Langer, T. & Douglas, M. G. (1994). "DnaJ-like proteins: molecular chaperones and specific reulators of Hsp70." *TIBS* **19**: 176 -181.

De Mot, R. & Vanderleyden, J. (1994). "The C-terminal sequnce conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both Gram-positive and Gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan." *Mol. Micro.* **12**(2): 333 - 334.

De Pamphilis, M. L. & Adler, J. (1971). "Fine structure and isolation of the hook-basal body complex of flagella from *Escherichia coli* and *Bacillus subtilis*." *J. Bacteriol.* **105**: 384 - 395.

Dean, G. E., Macnab, R. M., Stader, J., Matsumura, P. & Burks, C. (1984). "Gene sequence and predicted amino acid sequence of the *motA* protein, a membrane-associated protein required for flagellar rotation in *Escherichia coli*." *J. Bacteriol.* **159**(3): 991 - 999.

Deveraux, J., Haeberli, P. & Smithies, O. (1984). "A comprehensive set of sequence analysis programmes for the VAX." *Nucleic Acids Res.* **12**: 387 - 395.

Dingwall, A., Garman, J. D. & Shapiro, L. (1992). "Organization and ordered expression of *Caulobacter* genes encoding flagellar basal body rod and ring proteins." *J. Mol. Biol.* **228**: 1147 - 1162.

Dingwall, A., Zhuang, W. Y., Quon, K. & Shapiro, L. (1992). "Expression of an earley gene in the flagellar hierarchy is sensitive to an interruption in DNA replication." *J. Bacetriol.* **174**: 1760 - 1768.

Doll, L. & Frankel, G. (1993). "*fliU* and *fliV*, two flagellar genes essential for biosynthesis of *Salmonella typhimurium* and *Escherichia* coli flagella." *J. Gen. Microbiol.* **139**: 2415 - 2422.

Dreyfus, G., Williams, A. W., Kawagishi, I. & Macnab, R. M. (1993). "Genetic and biochemical analysis of Salmonella typhimurium Fli I, a flagellar protein related to the catalytic subunit of the F_0F_1 ATPase and to virulence proteins of mammalian and plant pathogens." *J. Bacteriol* **175**: 3131-3138.

Driks, A. & DeRosier, D. J. (1991). "Additionla structures associated with bacterial flagellar basal body." *J. Mol. Biol.* **221**: 669 - 672.

Eisenbach, M. (1990). "Functions of the flagellar modes of rotation in bacterial motility and chemotaxis." *Mol. Microbiol.* **4**: 161 - 167.

Eisenbach, M. (1996). "Control of bacterial chemotaxis." *Mol. Micro.* **20**(5): 903 - 910.

Emerson, S. U., Tokuyasu, K. & Simon, M. I. (1970). "Bacterial flagella: polarity of elongation." *Science* **169**: 190 - 192.

Emini, E. A., Hughes, J. V., Perlow, D. S. & Boger, J. (1985). "Induction of Hepatitis-A virus-neutralizing antibody by a virus specific synthetic peptide." *J. Virol.* **55**(3): 105 - 132.

Fenselau, S., Balbo, I. & Bonas, U. (1992). "Determinants of pathogenicity in *Xanthomonas campestris* pv. *vesicatoria* are related to proteins

involved in secretion in bacterial pathogens of animals." *Mol. Plant-Microbe Interact.* 5(5): 390 - 396.

Fiedler, W. & Rotering, H. (1988). "Properties of *E.coli* mutant lacking membrane-derived oligosaccharides." *J. Biol. Chem.* **263**: 14684 - 14689.

Fields, S. & Song, O.-K. (1989). "A novel genetic system to detect protein-protein interactions." *Nature* **340**: 245 - 246.

Foster, J. C. A. (1991). "Biosynthesis of the flagellum of *Rhodobacter sphaeroides*." DPhil thesis, Oxford University.

Foster-Hartnett, D. & Kranz, G. (1992). "Analysis of the promoters and upstream sequences of *nifA1* and *nifA2* in *Rhodobacter capsulatus*; activation requires *ntrC* but not *rpoN*." *Mol. Microbiol.* **6**(8): 1049 - 1060.

Francis, N. R., Irikura, V. M., Yamaguchi, S. & DeRosier, D. J. (1992). "Localisation of the *Salmonella typhimurium* flagellar switch protein FliG to the cytoplasmic M-ring face of the basal body." *Proc. Natl. Acad. Sci. USA* **89**: 6304 - 6308.

Francis, N. R., Sosinsky, G. E., Thomas, D. & DeRosier, D. J. (1994). "Isolation, characterisation and structure of bacterial flagellar motors containing the switch complex." *J. Mol. Biol.* **235**: 1261 - 1270.

Frangioni, J. V. & Neel, B. G. (1993). "Solubilization and purification of enzymatically active Glutathione S-Transferase (pGEX) fusion proteins." *Anal.Biochem.* **210**: 179-187.

Fuhr, G. & Hagedorn, R. (1989). "Dielectric motors - a new hypothesis for the bacterial flagellum." *J. Theor. Biol.* **139**: 39 - 59.

Fung, D. C. & Berg, H. C. (1995). "Powering the flagellar motor of *Escherichia coli* with an external voltage source." *Nature* **375**: 809 - 812.

Garber, A. T., Retief, J. D. & Dixon, G. H. (1989). "Isolation of dynein heavy chain cDNAa from trout testis which predict an extensive carboxyl-terminal alpha-helical coliled-coil domain." *EMBO J.* **8**: 1727 - 1734.

Garnier, J., Oguthorpe, D. J. & Robson, B. (1978). "Ananlysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins." *J. Mol. Biol.* **120**: 97 - 120.

Garza, A. G., Harris-Haller, L. W., Stoebner, R. A. & Manson, M. D. (1995). "Motility protein interactions in the bacterial flagellar motor." *Proc. Natl. Acad. Sci. USA* **92**: 1970 - 1974.

Garza, G., Biran, R., Wohlsclegel, J. A. & Manson, M. D. (1996). "Mutations in *motB* suppressible by changes in stator and rotor components of the bacterial flagellar motor." *J. Mol. Biol.* **258**: 270 - 285.

Gething, M. J. & Sambrook, J. (1992). "Protein folding in the cell." *Nature* **355**: 33 - 44.

Gibson, J. L. & Tabita, F. R. (1988). "Localisation and mapping of CO_2 fixation genes within two gene clusters in *Rhodobacter sphaeroides*." *J. Bacteriol.* **170**: 2153 - 2158.

Giuliano, G., Pollock, D., Stapp, H. & Scolnik, P. A. (1988). "A genetic-physical map of the *Rhodobacter capsulatas* carotenoid biosynthesis gene cluster." *Mol. Gen. Genet.* **213**: 78 - 83.

Glagalin (1978). "The proton pump is a molecular engine of motile bacteria. *Nature* 272: 280 - 282

Gober, J. W. & Marques, M. V. (1995). "Regulation of cellular differentiation in *Caulobacter crescentus*." *Microbiol. Rev.* **59**: 31 - 47.

Godchaux, W., Gorski, L. & Leadbetter, E. R. (1990). "Outer membrane polysaccharide deficiency in two nongliding mutants of *Cytophaga johnsonae*." *J. Bacteriol.* **172**: 1250.

Gomelsky, M. & Kaplan, S. (1995). "Genetic eveidence that PpsR from *Rhodobacter sphaeroides* 2.4.1. functions as a repressor of *puc* and *bchf* expression." *J. Bacteriol.* **177**(6): 1634 - 1637.

Goodfellow, I. G., Pollitt, C. E. & Sockett, R. E. (1996). "Cloning of the *fliI* gene from *Rhodobacter sphaeroides* WS8 by analysis of a transposon mutant with impaired motility." *FEMS* **142**: 111 - 116

Hall, I. & Riggs, P. (1996). Unpublised results. NEB. Catalogue, 96/97

Harayama, S. & Iino, T. (1976). "Phototactic response of aerobically grown *Rhodospirillum rubrum*" *J. Gen. Microbiol.* **94**: 173.

Harshey, R. M. & Toguchi, A. (1996). "Spinning tails: homologies among bacterial flagellar systems." *Trends in Microbiology* 4(6): 226 - 231.

Hayashi, S. & Wu, H. C. (1990). "Lipoproteins in bacteria." *J. Bioenerg. Biomembr.* **22**(3): 451 - 471.

Helmann, J. D. (1991). "Alternative sigma factors and the regulation of flagellar gene expression." *Mol. Micro.* **5**(12): 2875 - 2882.

Hirota, N. & Imae, Y. (1983). "Na^+-driven motors of an alkalophilic *Bacillus* strain YN-1." *J. Biol. Chem.* **258**: 10577 - 10581.

Homma, M., Aizawa, S., Dean, D. E. & R.M., M. (1987). "Identification of the M-ring protein of the flagellar motor of *Salmonella typhimurium*." *Proc. Natl. Acad. Sci. USA* **84**: 7483 - 7487.

Homma, M., DeRosier, D. J. & Macnab, R. M. (1990). "Flagellar hook and hook-associated proteins of *Salmonella typhimurium* and their relationship to other axial components of the flagellum." *J. Mol. Biol.* **213**: 819 - 832.

Homma, M., DeRosier, D. J. & Macnab, R. M. (1990). "FlgB, FlgC, FlgF and Flg G. A family of structurally related proteins in the flagellar basal body of *Salmonella typhimurium*." *J. Mol. Biol.* **211**: 465 - 477.

Homma, M., Iino, T. & Macnab, R. M. (1988). "Identification and characterisation of the products of six region III flagellar genes (*flaAII.3* through *flaQII*) of *Salmonella typhimurium*." *J. Bacteriol.* **170**(5): 2221 - 2228.

Homma, M., Komeda, Y., Iino, T. & Macnab, R. M. (1987). "The *flaFIX* gene product of *Slamonella typhimurium* is a flagellar basal body component with a signal peptide fro export." *J. Bacteriol.* **169**(4): 1493 - 1498.

Hoshino, T., Fujii, R. & Nakahara, T. (1993). "Molecular cloning and sequence-analysis of the *crtB* gene of *Thermus-Thermophilus* HB27, an extreme thermophile producing carotenoid-pigments." *Appl. Environ. Microbiol.* **59**: 3150 - 3153.

Hughes, K. T., Gillen, K. L., Semon, M. J. & Karlinsey, J. E. (1993). "Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator." *Science* **262**: 1277 - 1280.

Iino, T. (1969). "Polarity of flagellar growth in *Salmonella*." *J. Gen. Microbiol.* **56**: 227 - 239.

Iino, T. (1974). "Assembly of Salmonella flagellin *in vitro* and *in vivo*." *J. Supramol. Struc.* **2**: 372.

Imae, Y. (1985). Sensing and response in microorganisms. Amsterdam, Elsevier. 73 - 81.

Irikura, V. M., Kihara, M., Yamaguchi, S., Sockett, H. & Macnab, R. M. (1993). "*Salmonella typhimurium fliG* and *fliN* mutations causing defects in assembly, rotation, and switching of the flagellar motor." *J. Bacteriol.* **175**(3): 802 - 810.

Iwazawa, J., Imae, Y. & Kobayashi, S. (1993). "Study of the torque of the bacterial flagellar motor using a rotating electric field." *Biophys. J.* **64**: 925 - 933.

Jenal, U. & Shapiro, L. (1996). "Cell cycle-controlled proteolysis of a flagellar motor protein that is asymmetrically distributed in the *Caulobacter* predivisional cell." *EMBO* **15**(10): 2393 - 2406.

Jenal, U., White, J. & Shapiro, L. (1994). "*Caulobacter* flagellar function, but not assembly, requires FliL, a non-polarly localised membrane protein present in all cell types." *J. Mol. Biol.* **243**: 227 - 244.

Jones, C. J., Homma, M. & Macnab, R. M. (1989). "L-, P-, and M-ring proteins of the flagellar basal body of *Salmonella typhimurium*: Gene sequences and deduced protein sequences." *J. Bacteriol.* **171**(7): 3890 - 3900.

Jones, C. J. & Macnab, R. M. (1990). "Flagellar assembly in Salmonella typhimurium: analysis with temperature-sensitive mutants." *J.Bacteriol.* **172**: 1327-1339.

Jones, C. J., Macnab, R. M., Okino, H. & Aizawa, S.-I. (1990). "Stoichiometric analysis of the flagellar hook-(basal-body) complex of *Salmonella typhimurium*." *J. Mol. Biol.* **212**: 377 - 387.

Jones, R. & Haselkorn, R. (1989). "The DNA sequence of the *Rhodobacter capsulatus ntrA, ntrB* and *ntrC* gene analogues required for nitrogen fixation." *Mol. Gen. Genet.* **215**: 507 - 516.

Kagawa, H., Owaribe, K., Asakura, S. & Takahashi, N. (1976). "Flagellar hook protein from *Salmonella* SJ25." *J. Bacteriol.* **125**: 68 - 73.

Kami-Ike, N., Kudo, S. & Hotani, H. (1991). "Rapid changes in flagellar rotation induced by external electric pulses." *Biophysical Journal* **60**: 1350 - 1355.

Kaplan, S., Cain, B. D., Donohue, T. J., Shepherd, W. D. & Yen, G. S. L. (1983). "Biosynthesis of the photosynthetic membranes of *Rhodopseudomonas sphaeroides*." *J. Cell. Biochem.* **22**: 15 - 29.

Karplus, P. A. & Schulz, G. E. (1985). "Prediction of Chain felxibility in proteins - a tool for the selection of peptide antigens." *Naturwissenschaften* **72**(4): 212 - 213.

Katayama, E., Shiraishi, T., Oosawa, K., Baba, N. & Aizawa, S.-I. (1996). "Geometry of the flagellar motor in the cytoplasmic membrane of *Salmonella typhimurium* as determined by stereo-photogrammetry of quick-freeze deep-etch replica images." *J. Mol. Biol.* **255**: 458 - 475.

Kawagishi, I., Homma, M., Williams, A. W. & Macnab, R. M. (1996). "Characterisation of the flagellar hook length control protein FliK of *Salmonella typhimurium* and *Escherichia coli*." *J. Bacteriol.* **178**(10): 2954 - 2959.

Kawagishi, I., Muller, V., Williams, A. W., Irikura, V. M. & Macnab, R. M. (1992). "Subdivision of flagellar region III of the *Escerichia coli* and *Salmonella typhimurium* chromosomes and identification of two additional flagellar genes." *J. Gen. Microbiol.* **138**: 1051 - 1065.

Keen, N. T., Tamaki, S., Kobayashi, D. & Trollinger, D. (1988). "Improved broad-host range plasmids for DNA cloning in Gram-negative bacteria." *Gene* **70**: 191-197.

Khan, I. H., Reese, T. S. & Khan, S. (1992). "The cytoplasmic component of the bacterial flagellar motor." *Proc. Natl. Acad. Sci. USA* **89**: 5956 - 5960.

Khan, S. & Berg, H. C. (1983). "Isotope and temperature effects in chemiosmotic coupling to the flagellar motor of *Streptococcus*." *Cell* **32**: 913 - 919.

Khan, S., Dapice, M. & Reese, T. S. (1988). "Effects of *mot* gene expession on the structure of the flagellar motor." *J. Mol. Biol.* **202**: 575 - 584.

Kihara, M., Homma, M., Kutsukake, K. & Macnab, R. M. (1989). "Flagellar switch of *Salmonella typhimurium*: gene sequences and deduced protein sequences." *J. Bacteriol* **171**(6): 3247 - 3257.

Kohler, T., Harayama, S., Ramos, J.-l. & Timmis, K. (1989). "Involvement of *Pseudomonas putida* RpoN σ factor in regulation of various metabolic functions." *J. Bacteriol.* **171**(8): 4326 - 4333.

Komeda, Y., Icho, T. & Iino, T. (1977). "Effects of *galU* mutation on flagellar formation in *Escheherichia coli*." *J. Bacteriol.* **129**: 908 - 915.

Komeda, Y., Kutsukake, K. & Iino, T. (1980). "Definition of additional flagellar genes in *Escherichia coli* K12." *Genetics* **94**: 277 - 290.

Koo, H.-S., Wu, H. M. & Crothers, D. M. (1986). "DNA bending at adenine/thymine tracts." *Nature.* **320**: 501 - 506.

Kornacker, M. G. & Newton, A. (1994). "Information essential for cell-cycle-dependant secretion of the 591-residue *Caulobacter* hook protein is confined to a 21-amino-acid sequence near the N-terminus." *Mol. Microbiol* **14**: 73-85.

Kubori, T., Shimamoto, S., Yamaguchi, K., Namba, K. & Aizawa, S.-I. (1992). "Morphological pathway of flagellar assembly in *Salmonella typhimurium.*" *J. Mol. Biol.* **226**(2): 433 - 446.

Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H. & Fischer, H.-M. (1991). "*Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the σ^{54} gene (*rpoN*)." *J. Bacteriol.* **173**(3): 1125 - 1138.

Kutsu, S., Santero, E., Keener, J., Popham, D. & Weiss, D. (1989). "Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism." *Micro. Rev.* **53**(1): 367 - 376.

Kutsukake, K. & Iino, T. (1994). "Role of the FliA-FlgM regulatory sytem on the transcriptional control of the flagellar regulon and flagellar formation in *Salmonella typhimurium.*" *J. Bacteriol.* **176**(12): 3598 - 3605.

Kutsukake, K., Minamino, T. & Yokoseki, T. (1994). "Isolation and characterisation of FliK-independent flagellation mutants from *Salmonella typhimurium.*" *J. Bacteriol.* **176**: 7625 - 7629.

Kutsukake, K., Ohya, Y. & Iino, T. (1988). "Operon structure of flagellar genes in *Salmonella typhimurium.*" *Mol. Gen. Genet.* **214**: 11 - 15.

Kutsukake, K., Ohya, Y. & Iino, T. (1990). "Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium.*" *J. Bacteriol.* **172**: 741 - 747.

Kuwajima, G., Kawagishi, I., Homma, M., Asaka, J.-I., Kondo, E. & Macnab, R. M. (1989). "Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway." *Proc. Natl. Acad. Sci. USA* **86**: 4953-4957.

Kyte, J. & Doolittle, R. F. (1982). " A simple method for displaying the hrdophobic characater of a protein." *J. Mol. Biol.* **157**: 105 - 132.

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

Lapidus, I. R. & Berg, H. C. (1982). "Gliding motility of *Cytophaga* sp. strain U67." *J. Bacteriol.* **151**: 384.

Larsen, S. H., Adler, J., Gargus, J. J. & Hogg, R. W. (1974). "Chemomechanical coupling without ATP: the source of energy for motility and chemotaxis in bacteria." *Proc. Natl. Acad. Sci. USA* **71**: 1239 - 1243.

Lauger (1974)."Ion transport and rotation of bacterial flagella: *Nature* **268**: 360 - 362.

Lauger, P. (1988). "Torque and rotational rate of the bacterial flagellar motor." *Biophys. J.* **53**: 53 - 65.

Lederberg, J. & Iino, T. (1956). "Phase variation in *Salmonella.*" *Genetics* **41**: 744 - 747.

Lloyd, S. A., Tang, H. T., Wang, X., Billings, S. & Blair, D. F. (1996). "Torque generation in the flagellar motor of *Escherichia coli*: evidence of a ditrect role for FliG but not FliM or FliN." *J. Bacteriol.* **178**(1): 223 - 231.

Lui, J. & Magasanik, B. (1993). "The *glnB* region of the *Echerichia coli* genome." *J. Bacteriol.* **175**: 7441 - 7449.

Macnab, R. M. (1992). "Genetics and biogenesis of bacterial flagella." *Annu. Rev. Genet.* **26**: 131-158.

Macnab, R. M. (1996). "The bacterial flagellum". Society of General microbiology., London.

Macnab, R. M. & DeRosier, D. J. (1988). "Bacterial flagellar structure and function." *Can. J. Microbiol* **34**: 442-451.

Macnab, R. M. & Ornston, M. K. (1977). "Normal-to-curly flagellar transitions and their role in bacterial tumbling. Stabilisation of an alternative quaternary structure by mechanical force." *J. Mol. Biol.* **112**: 1 - 30.

Malakooti, J., Ely, B. & Matsumurra, P. (1994). "Molecular characterisation, nucleotide sequence, and expression of the *fliO*, *fliQ* and *fliR* genes of *Escherichia coli.*" *J. Bacteriol.* **176**: 189 - 197.

Maniatas, T., Fritsch, E. F. & Sambrook, J. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Manson, M. D., Tedesco, H. C., Berg, H. C., Harold, F. M. & van der Drift, C. (1977). "A protonmotive force drives bacterial flagella." *Proc. Natl. Acad. Sci. USA* **74**: 3060 - 3064.

Manson, M. D., Tedesco, P. M. & Berg, H. C. (1980). "Energetics of flagellar rotation in bacteria." *J. Mol. Biol.* **138**: 541 - 561.

Maquat, L. E. & Reznikoff, W. S. (1978). "*In vitro* analysis of the *Escherichia coli* RNA polymerase interaction with wild-type and mutant lactose promoters." *J. Mol. Biol.* **125**: 467 - 490.

Marczynski, G. T., Lentine, K. & Shapiro, L. (1995). "A developmentally regulated chromosomal origion of replication uses essential transcription elements." *Genes Dev.* **9**: 1543 - 1557.

Marykwas, D. L. & Berg, H. C. (1996). "A mutational analysis of the interaction between FliG and FliM, two components of the flagellar motor if *Escherichia coli.*" *J. Bacteriol.* **178**(5): 1289 - 1294.

Marykwas, D. L., Schmidt, S. A. & Berg, H. (1996). "Interacting components of the flagellar motor of *Escherichia coli* revealed by the two-hybrid system in yeast." *J. Mol. Biol.* **256**: 564 - 576.

Meijer, W. G. & Tabita, F. R. (1992). "Isolation and characterisation of the *nifUSVW-rpoN* gene cluster from *Rhodobacter sphaeroides*." *J. Bacteriol.* **174**(12): 3855 - 3866.

Meinhardt, K. W., Krishnan, H. B., Balatti, P. A. & Pueppke, S. G. (1993). "Molecular cloning and characterisation of a sym plasmid locus that regulates cultivar-specific nodulation of soybean by *Rhizobium fredii* USDA257." *Mol. Microbiol.* 9(1): 17 - 29.

Meister, M., Caplan, S. R. & Berg, H. C. (1989). "Dynamics of a tightly coupled mechanism for flagellar rotation." *Biophys. J.* **55**: 905 - 914.

Merrick, M. J. (1993). "In a class of its own - the RNA polymerase sigma factor σ^{54} (σ^{N})." *Mol. Micro.* **10**(5): 903 - 909.

Meyer, R. J., Figurski, D. & Helinski, D. R. (1977). "Physical and genetic studies with restriction endonucleases on the broad host range plasmid RK2." *Mol. Gen. Genet.* **152**: 129 - 135.

Michiels, T., Vanooteghem, J., Lambert de Rouvroit, C., China, B., Gustin, A., Boudry, P. & Cornelis, G. R. (1991). "Analysis of *virC*, an operon involved in the secretion of Yop proteins by *Yersinia enterocolitica*." *J. Bacteriol.* **173**(16): 4994 - 5009.

Minamino, T., Iino, T. & Kutsukake, K. (1994). "Molecular characterisation of the *Salmonella typhimurium flhB* operon and its protein products." *J. Bacteriol.* **176**: 7630 - 7637.

Minnich, S. A. & Newton, A. (1987). "Promoter mapping and cell cycle regulation of flagellin gene transcription in *Caulobacter crescentus*." *Proc. Natl. Acad. Sci. USA* **84**: 1142 - 1146.

Mirel, D. B., Lustre, V. M. & Chamberlin, M. J. (1992). "An operon of *Bacillus subtillis* motility genes transcribed by the sigma-D form of RNA polymerase." *J. Bacteriol.* **174**: 4197 - 4204.

Mitchell, P. (1984). "Bacterial flagellar motors and osmoelectric rotation by an exially transmembrane well and turnstile mechanism." *FEBS* **176**(2): 287 - 294.

Moore, M. D. & Kaplan, S. (1989). "Construction of Tn*phoA* gene fusions in *Rhodobacter sphaeroides*: isolation and characterisation of a respiratory mutant unable to utilise dimethyl sulfoxide as a terminal electron acceptor during anaerobic growth in the dark on glucose." *J. BActeriol.* **171**: 4385 - 4394.

Morett, E. & Buck, M. (1989). "In *vivo* studies on the interaction of RNA polymerase-σ^{54} with the *Klebsiella pneumiae* and *Rhizobium meliloti nifH* promoters: the role of NifA in the formation of an open promoter complex." *J. Mol. Biol.* **210**: 65 - 77.

Morett, E. & Segovia, L. (1993). "The sigma-54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains." *J. Bacteriol.* **175**(19): 6067 - 6074.

Muller, V., Jones, C. J., Kawagishi, I., Aizawa, S.-I. & Macnab, R. M. (1992). "Characterization of the *fliE* gene of *Escherichia*

coli and *Salmonella typhimurium*and identification of the FliE protein as a component of the flagellar hook-basal body complex." *J. Bacteriol.* **174**(7): 2298 - 2304.

Mullin, D., Minnich, S., Chen, L.-S. & Newton, A. (1987). "A set of positively regulated flagellar gene promoters in *Caulobacter crescentus* with sequence homology to *nif* gene promoters of *Klebsiella pneumoniae*." *J. Mol. Biol.* **195**: 939 - 943.

Mullin, D. A. & Newton, A. (1989). "Ntr-like promoters and upstream regulatory sequence *ftr* are required for transcription of a developmentally regulated *Caulobacter crescentus* flagellar gene." *J. Bacteriol.* **171**(6): 3218 - 3227.

Mullin, D. A. & Newton, A. (1993). "*Cis*- and *Trans*-acting elements required for regulation of flagellar gene transcription in the bacterium *Caulobacter crescentus*." *Cell. and Mol. Biol. Res.* **39**: 361 - 369.

Mullin, D. A., Van Way, S. M., Blakenship, C. A. & Mullin, A. H. (1994). "FlbD has a DNA-binding activity nesr its carboxy terminus that recognises *ftr* sequences involved in positive and negative regulation of flagellar gene transcription in *Caulobacter crescentus*." *J. Bacteriol.* **176**: 5971 - 5981.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). "Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction." *Cold Spring Harbour Symp. Quant. Biol.* **51**: 263 - 273.

Murata, T., Shimizu, H. & Yano, M. (1989). "A model for bacterial flagellar motor - free energy transduction and self organisation of rotational motion." *J. Theor. Biol.* **139**: 531 - 559.

Nagai, K., Perutz, M. F. & Poyart, C. (1985). "Oxygen binding properties of human mutant hemoglobin synthesised in *Escherichia coli*." *Proc. Natl. Acad. Sci. USA.* **82**: 7252 - 7255.

Nagel, J. F., Mille, M. & Morowitz, H. L. (1980). "Theory of hydrogen bonded chains in bioenergetics." *J. Biol. Chem.* **72**: 3959 - 3971.

Namba, K., Yamashita, I. & Vondervistz (1989). "Structure of the core and central channel of bacterial flagella." *Nature* **342**: 648 - 655.

Newton, A., Ohta, N., Ramakrishnan, G., Mullin, D. & Raymond, G. (1989). "Genetic switching in the flagellar hierarchy of *Caulobacter crescentus* requires negative as well as positive regulation of transcription." *Proc. NAtl. Acad. Sci. USA* **86**: 6651 - 6655.

Nishino, T., Kitamura, E., Matsuzaki, H., Nishijima, S., Matsumoto, K. & Shibuya, I. (1993). "Flagellar formation depends on membrane acidic phospholipids in *Escherichia coli*." *Biosci. Biotechnol. Biochem.* **57**: 1805 - 1808.

Ohnishi, K., Kutsukake, K., Suzuki, H. & Iino, T. (1990). "Gene *fliA* encode an alternative sigma factor specific for flagellar operons in *Salmonella typhimurium*." *Mol. Gen. Genet.* **221**: 139 - 147.

Ohnishi, K., Kutsukake, K., Suzuki, H. & Iino, T. (1992). "A novel transcriptioal regulation mechanism in the flagellar regulon of *Salmonella*

typhimurium: an anti-sigma factor inhibits the activity of the flagellum specific sigma factor σ^F" *Mol. Microbiol.* **6**: 3149 - 3157.

Ohta, N., Chen, L. S., Swanson, E. & Newton, A. (1985). "Transcriptional regulation of a periodically controlled flagellar gene operon in *Caulobacter crescentus*." *J. Mol. Biol.* **186**: 107 - 115.

Okino, H., Isomura, M., Yamaguchi, S., Magariyama, Y., Kudo, S. & Aizawa, S. (1989). "Release of flagellar filament-hook-rod complex by a *Salmonella typhimurium* mutant defective in the M ring of the basal body." *J. Bacteriol* **171**(4): 2075 - 2082.

Oosawa, F. & Hayashi, S. (1986). "The loose-coupling mechanism in molceular machines." *Adv. Biophys.* **22**: 151 - 183.

Oosawa, K., Ueno, T. & Aizawa, S.-I. (1994). "Overproduction of the bacterial flagellar switch proteins and their interactions with the MS ring complex in vitro." *J. Bacteriol.* **176**(12): 3683 - 3691.

Packer, H. & Armitage, J. P. (1993). "The unidirectional flagellar motor of *Rhodobacter sphaeroides* WS8 can rotate either clockwise or counterclockwise: characterization of the flagellum under both conditions by antibody decoration." *J. Bacteriol* **175**: 6041-6045.

Park, D. (1990). "Construction of mobilisation vectors derived from plasmids RP4, pUC18 and pUC19." *Gene* **93**: 135 - 137.

Perez-Martin, J., Rojo, F. & de Lorenzo, V. (1994). "Promoters responsive to DNA bending: a common theme in prokaryotic gene expression." *Microbiol. Rev.* **58**: 268 - 290.

Pierce, N. F., Cray, W. C., Kaper, J. B. & Mekalanos, J. L. (1988). "Determinants of immunogenicity and mechanisms of protection by virulent and mutant *Vibrio cholerae* 01 in rabbits." *Infect. Immun.* **56**: 142.

Piire, N. W. (1964). "The Leeuwenhoek lecture, 1963." *Proc. Roy. Soc.* **160**: 149 - 166.

Pollitt, C. E. (1996). Dphil Thesis: "Analysis of genes involved in flagellar biosynthesis and switching in *Rhodobacter sphaeroides*.". Nottingham.

Poole, P. S., Sinclair, D. R. & J.P, A. (1988). "Real time computer imaging of free swimming tethered rotating cells." *Analytical Biochem* **175**: 52 - 58.

Popham, D. L., Szeto, D., Keener, J. & Kutsu, S. (1989). "Function of a bacterial activator protein that binds to transcriptional enhancers." *Science* **243**: 629 - 635.

Prentki, P. & Krisch, H. M. (1984). "*In vitro* insertional mutagenesis with a selectable DNA fragment." *Gene* **29**: 303 - 313.

Pugsley, A. P. (1993). "The complete general secretory pathway in Gram-negative bacteria." *Microbiol. Revs* **57**: 50-108.

Quon, K. C., Marczynski, G. T. & Shapiro, L. (1996). "Cell cycle control by an essential bacterial two-component signal transduction protein." *Cell* **84**: 83 - 93.

Ramakrishnan, G. & Newton, A. (1990). "FlbD of *Caulobacter crescentus* is a homologue of the NtrC protein and activated sigma-54 dependent flagellar gene promoters." *Proc. Natl. Acad. Sci. USA* **87**: 2369 - 2373.

Ramakrishnan, G., Zhao, J.-L. & Newton, A. (1994). "Multiple structural proteins are required for both transcriptional activator and negative autoregulation of *Caulobacter crescentus* flagellar genes." *J. Bacteriol.* **176**(24): 7587 - 7600.

Richings, B. W., Almira, E. C., Lory, S. & Ramphal, R. (1995). "Cloning and phenotypic characterisation of *fleS* and *fleR*, new response regulators of *Pseudomonas aeruginosa*which regulate motility and adhesion to mucin." *Infect. Immun.* **63**(12): 4868 - 4876.

Rivers, S. L., McNairn, E. & Blasco, F. (1993). "Molecular genetic analysis of the *moa* operon of *Escherichia coli* K-12 required for molybdenum cofactor biosynthesis." *Mol. Microbiol.* **8**: 1071 - 1081.

Robinson, K. (1994). Unpublished, EMBL accession number: ML0011.

Roman, S. J., Frantz, B. B. & Matsumurra, P. (1993). "Gene sequence, overproduction, purification and determination of the wild-type level of the *Escherichia coli* flagellar switch protein FliG." *Gene* **113**: 103 - 108.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). "DNA sequencing with chain terminating inhibitors." *Proc Natl Acad Sci* **74**: 5463-.

Sasakawa, C. K., Komatsu, K., Tobe, T., Fukuda, I., Suzuki, T. & Yoshikawa, M. (1993). "Eight genes in region 5 that form an operon are essential for invasion of epithelial cell by *Shigella flexerni* 2a." *J. Bacteriol.* **175**: 2334 - 2346.

Sasse-Dwight, S. & Gralla, J. D. (1988). "Probing the *glnALG* upstream activation mechanism *in vivo*." *Proc. Natl. Acad. Sci. USA* **85**: 8934 - 8938.

Saunders, L. (1993). "Molecular genetic analyses and partial purification of polyhooks from a *Rhodobacter sphaeroides* WS8N presumptive *fliK* mutant." Undergraduate project dissertation, Nottingham University.

Shah, D. S., Perehenic, T., Stevens, S. & Sockett, R. E. (1996). "The flagellin gene from *Rhodobacter sphaeroides* WS8." Manuscript in preperation.

Shah, D. S. & Sockett, R. E. (1995). "Analysis of the *motA* flagellar motor gene from *Rhodobacter sphaeroides*, a bacterium with a unidirectional, stop-start flagellum." *Mol. Micro.* **17**(5): 961 - 969.

Shah, D. S., Armitage, J.P. & Sockett, R. E. (1995). "*Rhodobacter sphaeroides* WS8 encodes a polypeptide that is similair to MotB of *E.coli*." *J. Bacteriol.* **177**: 2929 - 2932.

Shapiro, L. & Maizel, J. (1973). "Synthesis and structure of *Caulobacter crescentus* flagella." *J. Bacteriol.* **113**: 478 - 485.

Sharp, L. L., Zhou, J. & Blair, D. F. (1995). "Feautures of the MotA proton channel structure revealed by tryptophan-scanning mutagenesis." *Proc. Natl. Acad. Sci. USA* **92**: 7946 - 7950.

Sharp, L. L., Zhou, J. & Blair, D. F. (1995). "Tryptophan-Scanning Mutagenesis of MotB, an integral membrane protein essential for flagellar rotation in *Escherichia coli.*" *Biochemistry* **34**: 9166 - 9171.

Shi, W., Bogdanov, M., Dowham, W. & Zushman, D. R. (1993). "The *pss* and *psd* genes are required for motility and chemotaxis in *Escherchia coli.*" *J. Bacteriol.* **175**: 7711 - 7714.

Shi, W., Zhou, Y., Wild, J., Adler, J. & Gross, C. A. (1992). "DnaK, DnaJ, and GrpE are required for flagellum synthesis in *Eschericgia coli.*" *J. Bacteriol.* **174**: 2236 - 2240.

Shimada, H., Ohta, H., Masuda, T., Shioi, Y. & Takamiya, K.-I. (1993). "A putative transcription factor binding to the upstream region of the *puf* operon in *R.sphaeroides.*" *FEBS* **328**(1): 41 - 44.

Shine, J. & Dalgarno, L. (1974). "The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites." *Proc. Natl. Acad. Sci. USA* **71**: 1342 - 1346.

Shingler, V. (1996). "Signal sensing by σ^{54}-dependent regulators: derepression as a control mechanism." *Mol. Microbiol.* **19**(3): 409 - 416.

Silva-Herzog, E., Macnab, R. M. & Dreyfus, G. (1995). The study of FliI a member of the flagellar export apparatus. BLAST III, Austin, Texas.,

Silverman, M. & Simon, M. (1974). "Characterisation of *Escherichia coli* flagellar mutants that are insensitive to catabolite repression." *J. Bacteriol.* **120**: 1196 - 1203.

Simon, R., Priefer, U. & Puhler, A. (1983). "A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagensis in Gram negative bacteria." *Bio/Technology* **1**: 37-45.

Sistrom, W.R. (1960). " A requirement for sodium in the growth of *Rhodopseudomonas sphaeroides.*" *J. Gen. Microbiol,* **22**: 778 - 785

Slocum, M. K. & Parkinson, J. S. (1983). "Genetics of methyl-accepting chemotaxis proteins in *Escherichia coli*:: organisation of the *tar* region." *J. Bacteriol.* **155**: 565 - 577.

Smith, D. B. & Johnson, K. S. (1988). "Single step purification of polypeptides expressed in *Escherichia coli* as fusions to glutathione S-tranferase." *Gene* **67**: 31 - 40.

Smith, D. R. & Robinson, K. (1995). Unpublished. Accession number P40833.

Sockett, H., Yamaguchi, S., Kihara, M., Irikura, V. M. & Macnab, R. M. (1992). "Molecular analysis of the flagellar switch protein FliM of *Salmonella typhimurium.*" *J. Bacteriol* **174**(3): 793 - 806.

Sockett, R. E. (1986). Dphil Thesis: "Biochemistry of motility and chemotaxis in photosynthetic bacteria." University of London.

Sockett, R. E. & Armitage, J. (1991). "Isolation, characterization, and complementation of a paralysed flagellar mutant of *Rhodobacter sphaeroides* WS8." *J.Bacteriol* **173**: 2786-2790.

Sosinsky, G. E., Francis, N. R., DeRosier, D. J., Wall, J. S., Simon, M. N. & Hainfield, J. (1992). "Mass determination and estimation of subunit stoichiometry of the bacterial hook-basal body flagellar complex of *Salmonella typhimurium*by scanning transmission electron microscopy." *Proc. Natl. Acad. Sci. USA* **89**: 4801 - 4805.

Southern, E. M. (1975). "Detection of specific sequences among DNA fragments seperated by gel electrophoresis." *J. Mol. Biol.* **98**: 503 - 518.

Stader, J., Matsumura, P., Vacante, D., Dean, D. & Macnab, R. M. (1986). "Nucleotide sequence of the *EScherichia coli motB* gene and site-limited incorporation of its product into the cytoplasmic membrane." *J. Bacteriol.* **166**(1): 244 - 252.

Stallmeyer, M. J., Hahneberger, K., Sosinsky, G. E., Shapiro, L. & DeRosier, D. (1989). "Image reconstruction of the flagellar basal body of *Caulobacter crescentus*." *J. Mol. Biol.* **205**: 511 - 518.

Stallmeyer, M. J. B., Aizawa, S.-I., Macnab, R. M. & DeRosier, D. J. (1988). "Image reconstruction of the basal body of *Caulobacter crescentus*." *J. Mol. Biol.* **205**: 519.

Stephens, C. M. & Shapiro, L. (1993). "An unusual promoter controls cell-cycle regulation and dependence on DNA replication of the *Caulobacter fliLM* early flagellar operon." *Mol. Micro.* **9**(6): 1169 - 1179.

Stephens, C. M., Zweiger, G. & Shapiro, L. (1995). "Coordinate cell cycle control of a *Caulobacter* DNA methyltransferase and the flagellar genetic hierarchy." *J. Bacteriol* **177**: 1662-1669.

Stock, J. B., Ninfa, A. J. & Stock, A. M. (1989). "Protein phosphorylation and regulation of adaptive responses in bacteria." *Micro. Rev.* **53**: 450 - 490.

Stolz, B. & Berg, H. C. (1991). "Evidence for the interaction between MotA and MotB, torque-generating elements of the flagellar motor of *Eschericia coli*." *J. Bacteriol.* **173**: 7033 - 7037.

Suzuki, T., Iino, T., Horiguchi, T. & Yamaguchi, S. (1978). "Incomplete flagellar structures in nonflagellate mutants of *Salmonella typhimurium*." *J. Bacteriol.* **133**: 904 - 915.

Suzuki, T. & Komeda, Y. (1981). "Incomplete flagellar structures in *Escherichia coli* mutants." *J. Bacteriol.* **145**: 1036 - 1041.

Szeto, W. W., Nixon, B. T., Ronson, C. W. & Ausbel, F. M. (1987). "Identification and characterisation of the *Rhizobium meliloti ntrC* gene - *R.meliloti* has seperate regulatory pathways for activation of nitrogen-fixation genes in free-living and symbiotic cells." *J. Bacteriol.* **169**(4): 1423 - 1432.

Tai & Kaplan, S. (1985). "Intracellular localisation of phospholipid transfer activity in *Rhodopseudomonas sphaeroides* and a possible role in membrane biosynthesis." *J. Bacteriol.* **164**: 181 - 186.

Tang, H., Billings, S., Wang, X., Sharp, L. & Blair, D. F. (1995). "Regulated underexpression and overexpression of the FliN protein of *Escherichia coli* and evidence for an interaction between FliN and FliM in the flagellar motor." *J. Bacteriol.* **177**(12): 3496 - 3503.

Tang, H. & Blair, D. F. (1995). "Regulated underexpression of the FliM protein of *Escherichia coli* and evidence for its location in the flagellar motor distinct from the MotA/MotB torque generators." *J. Bacteriol.* **177**(12): 3485 - 3495.

Tang, H., Braun, T. F. & Blair, D. F. (1996). "Motility protein complexes in the bacterial flagellar motor." *J. Mol. Biol.* **261**: 209 - 221.

Tartof, K. D. & Hobbs, C. A. (1987). "Improved media for grwoing plasmid and cosmid clones." *BRL Focus* 9(2): 12.

Tauschel, H. D. (1987). "ATPase activity of the polar organelle demonstrated by cytochemical reaction in whole unstained cells of *Rhodopseudomonas palustris*." *Arch. Microbiol* **148**: 159-161.

Taylor, B. L. (1983). "Role of proton motive force in sensory transduction in bacteria." *Ann. Rev. Microbiol.* **37**: 551.

Tso, W. W. & Alder, J. (1974). "Negative chemotaxis in *Esherichia coli*." *J. Bacteriol.* **118**: 560-576.

Ueno, T., Oosawa, K. & Aizawa, S. (1992). "M Ring, S ring and proximal rod of the flagellar basal body of *Salmonella typhimurium* are composed of subunits of a single protein, FliF." *J. Mol. Biol.* (227): 672 - 677.

Ueno, T., Oosawa, K. & Aizawa, S. (1994). "Domain structures of the MS ring component protein (FliF) of the flagellar basal body of *Salmonella typhimurium*." *J. Mol. Biol.* **236**: 546-555.

Van Gijsegem, F., Gough, C., Zischek, C., Niquex, E., Arlat, M., Genin, S., Barberis, P., German, S., Castello, P. & Boucher, C. (1995). "The hrp gene locus of *Pseudomonas solanacearum*, which controls the production of a type III secretion system, encodes eight proteins related to components of the bacterial flagellar biogenesis complex." *Mol. Microbiol* **15**(6): 1095 -1114.

Van Way, S. M., Newton, S., Mullin, A. H. & Mullin, D. A. (1993). "Identification of the promoter and negative regulatory element, ftr4, that is needed for cell cycle timing of *fliF* operon expression in *Caulobacter crescentus*." *J. Bacteriol.* **175**: 367 - 376.

Venkatesan, M. M., Buysse, J. M. & Oaks, E. V. (1992). "Surface presentation of the *Shigella flexerni* invasion plasmid antigens requires the products of the *spa* locus." *J. Bacteriol.* **174**: 1990 - 2001.

Vieira, J. & Messing, J. (1982). "The pUC plasmids, an M13mp7-derived system of insertion mutagenesis ans sequencing with synthetic universal primers." *Gene* **19**: 259 - 268.

Vogler, A. P., Homma, M., Irikura, V. M. & Macnab, R. M. (1991). "*Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of Fli I to F_0F_1, vacuolar and archaebacterial ATPase subunits." *J. Bacteriol* **173**: 3564-3572.

Volz, K. (1993). "Structural conservation in the CheY superfamily." *Biochemistry* **32**(44): 11741 - 11753.

Wagenknecht, T. (1986). "A plausible mechanism for flagella rotation in bacteria." *FEBS Lett.* **196**: 193 - 197.

Walker, J. E., Saraste, M., Runswick, M. J. & 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 J.* **1**: 945 - 951.

Ward, M. J., Bell, A. W., Hamblin, P. A., Pavker, H. L. & Armitage, J. P. (1995). "Identification of a chemotaxis operon with two *che*Y genes in *Rhodobacter sphaeroides*." *Mol. Microbiol.* **17**: 357 - 366.

Ward, M. J., Harrison, D. M., Ebner, M. J. & Armitage, J. P. (1995). "Identification of a methyl-accepting chemotaxis protein in *Rhodobacter sphaeroides*." *Mol. Microbiol.* **18**: 115 - 121.

Wattiau, P., Woestyn, S. & Cornelis, G. R. (1996). "Customized secretion chaperones in pathogenic bacteria." *Mol. Microbiol.* **20**(2): 255 - 262.

Weckesser, J., Zalman, L. S. & Nikaido, H. (1984). "Porin from *Rhodopseudomonas sphaeroides*." *J. Bacteriol.* **159**(1): 199 - 205.

Weiss, V., Batut, J., Klose, E., Keener, J. & Kutsu, S. (1988). "The phosphorylated form of the enhancer-binding protein NTRC has ATPase activity that is essential for activation of transcription." *Cell* **67**: 155 - 167.

Welch, M., Oosawa, K., Aizawa, S. & Eisenbach, M. (1993). "Phosphorylation-dependant binding of a signal molecule to the flagellar switch of bacteria." *Proc. Natl. Acad. Sci. USA* **90**: 8787 - 8791.

Wetmur, J. G., Wong, D. M., Oritz, B., Reichert, F. & Gelfand, D. H. (1994). "Cloning, sequencing, and expression of RecA proteins from 3 distntly related thermophilic eubacteria." *J. Biol. Chem.* **269**: 25928 - 25935.

Williams, A. W., Yamaguchi, S., Togashi, F., Aizawa, S.-I., Kawagishi, I. & Macnab, R. M. (1996). "Mutations in *fliK* and *flhB* affecting flagellar hook and filament assembly in *Salmonella typhimurium*.." *J. Bacteriol.* **178**(10): 2960 - 2970.

Wilson, M. L. & Macnab, R. M. (1988). "Overproduction of the MotA protein of *Escherichia coli*and estimation of its wild-type level." *J. Bacteriol.* **170**(2): 588 - 597.

Wilson, M. L. & Macnab, R. M. (1990). "Co-overproduction and localization of the *Escherichia coli* motility proteins MotA and MotB." *J. Bacteriol.* **172**(7): 3932 - 3939.

Wingrove, J. A. & Gober, J. W. (1994). "A σ54 transcriptional activator also functions as a pole specific repressor in *Caulobacter.*" *Genes Dev.* **7**: 1979 - 1992.

Woestyn, S., Allaoui, A., Wattiau, P. & Cornelis, G. R. (1994). "YscN, the putative energizer of the *Yersinia* Yop secretion machinery." *J. Bacteriol* **176**: 1561-1569.

Wolkin, R. H. & Pate, J. L. (1984). "Translocation of motile cells of the gliding bacterium *Cytophaga johnsonae* depends on a surface component that may be modofoed by sugars." *J. Gen. Microbiol.* **130**: 2651 - 2669.

Wolkin, R. H. & Pate, J. L. (1986). "Phage adsorption and cell adherence are motility-dependent characteristics of the gliding bacterium *Cytophaga johnsonae.*" *J. Gen. Microbiol.* **132**: 355 - 367.

Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989). "Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants." *Nucl Acids Res* **17**: 3469-3478.

Wu, J., Benson, A. K. & Newton, A. (1995). "Global regulation of a σ^{54}-dependent flagellar gene family in *Caulobacter crescentus* by the transcriptional activator FlbD." *J. BActeriol.* **177**(11): 3241 - 320.

Yamaguchi, S., Aizawa, S., Kihara, K., Isomura, M., Jones, C. & Macnab, R. M. (1986). "Genetic eveidence for a switching and energy-transducing complex in the flagellar motor of *Salmonella typhimurium.*" *J. Bacteriol* **168**(3): 1172 - 1179.

Yamaguchi, S., Fujita, H., Ishihara, A., Aizawa, S.-I. & Macnab, R. M. (1986). "Subdivision of flagellar genes of *Salmonella typhimurium* into regions responsible for assembly, rotation, and switching." *J. Bacteriol.* **166**: 187 - 193.

Yanish-Perron, C., Viera, J. & Messing, J. (1985). "Improved M13 phage cloning vectors and host strains:nucleotide sequences of the M13mp18 and pUC19 vectors." *Gene* **33**: 103-119.

Yokoseki, T., Iino, T. & Kutsukake, K. (1996). "Negative regulation by FliD, FliS, and FliT of the export of the flagellum-specific anti-sigma factor, FlgM, in *Salmonella typhimurium..*" *J. Bacteriol.* **178**: 899 - 901.

Zhuang, W. Y. & Shapiro, L. (1995). "*Caulobacter* FliQ and FliR membrane proteins, required for flagellar biogenesis and cell division, belong to a family of virulence factor export proteins." *J. Bacteriol.* **177**: 343 - 356.

Zuberi, A. R., Bischoff, D. S. & Ordal, G. O. (1991). "Nucleotide sequence and characterization of a *Bacillus subtilis* gene encoding a flagellar switch protein." *J. Bacteriol* **173**(2): 710 - 719.

Zuberi, A. R., Ying, C., Bischoff, D. S. & Ordal, G. O. (1991). "Gene-protein relationships in the flagellar hook-basal body complex of *Bacillus subtilis*: sequences of the *flgB, flgC, flgG, fliE and fliF* genes." *Gene* **191**: 23 - 31.

Appendix 1

GST protein purification solutions:

STE: 10mM Tris-HCl pH 8.0, 150mM NaCl and 1mM EDTA.

1 X PBS: 140mM NaCl, 2.7mM Kcl, 10mM Na_2HPO_4 and 1.8mM KH_2PO_4 pH 7.3.

PBS-Tween 20: 1 X PBS containing 0.3 % (v/v) Tween-20

Histidine Tagged protein purification solutions:

8 X Binding buffer: 40mM imidazole, 4M NaCl and 160mM Tris-HCl pH 7.9

8 X Charge buffer: 400mM NiSo4

8 X Wash buffer: 480mM imidazole, 4M NaCl, 160mM Tris-HCl pH 7.9.

20mM Imidazole buffer: combine 15ml of 1 X binding buffer with 4.1ml of 1 X Wash buffer.

4 X Elution buffer: 4M imidazole, 2M NaCl and 80mM Tris-HCl pH 7.9.

Growth Media:

Luria Bertani (LB): 1 litre contained 5g yeast extract, 10g Tryptone and 5g NaCl. Agar was added to 1.5% in solid media

Terrific broth: A 900 ml\solution was made containing 24g yeast extract, 12g Tryptone and 4ml of glycerol. After autoclaving 100ml of a solution containing 0.17 M KH_2PO_4, 0.72M K_2HPO_4.

2 X YT: One litre contained 16g yeast extract, 16g Tryptone, 5g NaCl and 2.5g K_2HPO4.

Succinate media: One litre contained 20ml od Solution C (see below), 9g KH_2PO_4, 1.5g NH_4SO_4, 1.5g NaCl, 3ml of vitamin solution (see below), adjusted to pH 7.2.

Solution C: 400 ml contained 2.4g Nitriloacetic acid, 25mg EDTA, 110mg $ZnSO_4 \cdot 7H_2O$, 70mg $FeSO_4 \cdot 7H_2O$, 15µl 3M H_2SO4, 20mg $MnSO_4 \cdot 4H_2O$,

4mg $CuSO_4 \cdot 5H_2O$, 1.2mg H_3BO_3, 2mg $CoCl_2 \cdot 6H_2O$, 5.8g $MgSO_4 \cdot 7H_2O$, 1.84mg $MoNH_4$ and 0.67g $CaCl_2 \cdot 2H_2O$.

Vitamin solution: 100ml contained 1g Nicotinic acid, 0.5g Thiamine•HCl and 0.01g Biotin.

Miscellaneous:

10 X TBE: One litre contained 108g Tris base, 55g Boric acid and 7.4g EDTA.

TE: 10mM Tris-Hcl pH 8.0 and 1mM EDTA pH 8.0

Lysis Buffer (DNA mini preps): 50mM glucose, 10mM EDTA and 25mM Tris-Hcl pH 8.0.

Chromoprep buffer: 10mM Tris-HCl pH 8.0, 1mM EDTA and 1% SDS.

1 X TNE: 100mM Tris-HCl pH8.0, 50mM NaCl and 5mM EDTA.

Alkaline SDS (DNA mini preps): An euqal volume of 2%SDS and 0.4M NaoH were mixed and used immediately.

AP buffer (Western Blots): 100 mM Tris-HCl, pH 9.5, 100mM NaCl and 5mM $MgCl_2$.

Phosphate buffer pH 6.4: 100ml contained 27.8ml of 1M K_2HPO_4 and 72.2ml of 1M KH_2PO_4.

SDS-PAGE solutions:

10 X Tris-Glycine buffer: 0.25M Tris base, 1.92 M Glycine and 1% SDS.

2 X SDS-PAGE sample buffer: 250µl 0.25M Tris-Hcl pH 6.8, 100µl of 20% SDS, 100µl 0.1% Bromophenol blue, 50µl βMercaptoethanol, 280µl glycerol and 20µl of SDW.

Coomassie stain: 10% methanol, 10% acetic acid and 0.25% Coomassie blue.

Sequenase Version 2.0 reagents (DNA sequencing):

Enzyme dilution buffer: 10mM Tris-HCl pH 7.5; 5mM DTT; 0.5mg/ml BSA

Mn Buffer: 0.15M Sodium isocitrate; 0.1M $MnCl_2$

Sequenase reaction buffer 5 X: 200mM Tris-Hcl pH 7.5; 100mM $MgCl_2$; 250mM NaCl

Stop Solution: 95% formamide; 20mM EDTA; 0.05% bromophenol blue; 0.05% xylene cyanol FF.

Termination mixes: Termination mixes contained 80μM of each dNTP, except for the nucleotide with which terminations were created, which was present at 8μM along with 80μM of the ddNTP and 50mM NaCl

Southern Hybridisation solutions (Southern Blots):

Hybridisation solution: 100ml contained 5 X SSC, 5 X Denharts solution, 0.5% SDS and 100μg/ml of salmon sperm DNA.

50 X Denhardts solution: 500ml contained 5g Ficol (Type 400), 5g PVP and 5g BSA (fraction V).

Southern Denaturing solution: One litre contained 0.5M sodium hydroxide and 1.5M NaCl.

Southern Neutralising solution: One litre contained 1.5M NaCl and 1M Tris base, pH 7.5.

20 X SSC: One litre contained 3M NaCl and 0.3 M sodium citrate, pH 7.0.

TBS-Tween20: One litred contained 100mM Tris base, 150mM NaCl and 0.05% (v/v) Tween20, pH 7.5

Southern blocking Solution: 100ml contained 3g of Bovine serum albumin in TBS-Tween20.

Final Wash buffer: One litre contained 100mM Tris base, 100mM NaCl and 50mM $MgCl_2 \cdot 6H_20$, pH 9.5.

Appendix 2

R.sphaeroides codon usage chart:

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	27.00	10.48	0.13
Gly	GGA	9.00	3.49	0.04
Gly	GGT	9.00	3.49	0.04
Gly	GGC	167.00	64.83	0.79
Glu	GAG	101.00	39.21	0.71
Glu	GAA	41.00	15.92	0.29
Asp	GAT	31.00	12.03	0.28
Asp	GAC	81.00	31.44	0.72
Val	GTG	70.00	27.17	0.43
Val	GTA	1.00	0.39	0.01
Val	GTT	1.00	0.39	0.01
Val	GTC	92.00	35.71	0.56
Ala	GCG	148.00	57.45	0.42
Ala	GCA	17.00	6.60	0.05
Ala	GCT	17.00	6.60	0.05
Ala	GCC	169.00	65.61	0.48
Arg	AGG	1.00	0.39	0.01
Arg	AGA	0.00	0.00	0.00
Ser	AGT	2.00	0.78	0.02
Ser	AGC	36.00	13.98	0.29
Lys	AAG	64.00	24.84	0.86
Lys	AAA	10.00	3.88	0.14
Asn	AAT	8.00	3.11	0.12
Asn	AAC	59.00	22.90	0.88
Met	ATG	74.00	28.73	1.00
Ile	ATA	0.00	0.00	0.00
Ile	ATT	7.00	2.72	0.05
Ile	ATC	126.00	48.91	0.95
Thr	ACG	60.00	23.29	0.42
Thr	ACA	3.00	1.16	0.02
Thr	ACT	3.00	1.16	0.02
Thr	ACC	77.00	29.89	0.54
Trp	TGG	56.00	21.74	1.00
End	TGA	9.00	3.49	0.75
Cys	TGT	0.00	0.00	0.00
Cys	TGC	13.00	5.05	1.00
End	TAG	0.00	0.00	0.00
End	TAA	3.00	1.16	0.25
Tyr	TAT	18.00	6.99	0.38
Tyr	TAC	30.00	11.65	0.62

Leu	TTG	4.00	1.55	0.01
Leu	TTA	0.00	0.00	0.00
Phe	TTT	4.00	1.55	0.04
Phe	TTC	100.00	38.82	0.96
Ser	TCG	59.00	22.90	0.48
Ser	TCA	1.00	0.39	0.01
Ser	TCT	3.00	1.16	0.02
Ser	TCC	22.00	8.54	0.18
Arg	CGG	41.00	15.92	0.26
Arg	CGA	2.00	0.78	0.01
Arg	CGT	7.00	2.72	0.05
Arg	CGC	104.00	40.37	0.67
Gln	CAG	99.00	38.43	0.95
Gln	CAA	5.00	1.94	0.05
His	CAT	13.00	5.05	0.27
His	CAC	35.00	13.59	0.73
Leu	CTG	158.00	61.34	0.54
Leu	CTA	0.00	0.00	0.00
Leu	CTT	12.00	4.66	0.04
Leu	CTC	120.00	46.58	0.41
Pro	CCG	82.00	31.83	0.56
Pro	CCA	1.00	0.39	0.01
Pro	CCT	6.00	2.33	0.04
Pro	CCC	58.00	22.52	0.39