Regulation of the expression and positioning of chemotaxis and motor proteins in *Rhodobacter sphaeroides*



David Arthur Wilkinson

Merton College
University of Oxford

Thesis submitted for the Degree of Doctor of Philosophy

Trinity Term 2010

Abstract

Bacteria achieve directed motion through their environments by integrating propulsion with chemical detection in the process of chemotaxis. Central to this process are the macromolecular protein structures of the flagellar motor and the chemoreceptor arrays, which are responsible for motility and chemical sensing, respectively. These protein complexes localise to different discrete subcellular positions in different bacterial species, and their correct subcellular localisation is often essential to their function.

In the monotrichous α-proteobacterium *Rhodobacter sphaeroides*, the flagellum is subpolar and two distinct sets of chemotaxis proteins localise to discrete polar and cytoplasmic positions within the cell.

In this study, the development of software for the analysis of fluorescent microscopy images allowed cellular morphologies and the localisation and distribution of the chemoreceptor arrays of *R.sphaeroides* to be characterised in detail, showing that protein partitioning at cell division results in an asymmetric separation of both cytoplasmic and membrane-bound protein components between daughter cells.

The design of a fluorescence-based assay for the analysis of gene expression assisted in demonstrating that expression of both the chemotaxis and motor genes of *R.sphaeroides* is regulated by the sigma factor, FliA, and its inhibitor, FlgM. FliA was then used to achieve varying expression of the chemotaxis genes, and the concentration dependence of array clustering was explored in microscopy images, revealing important differences between cluster formation in *R.sphaeroides* and other species.

Additionally, FliA was identified as a regulator of flagellar number in *R.sphaeroides*, controlling a negative feedback-loop in the hierarchy of flagellar assembly that represses flagellar formation upon secretion of FlgM. The complex regulatory pathway controlling *R.sphaeroides* flagellar assembly is the first identified system where completion of a single flagellum directly inhibits the production of a second, a mechanism that may be important to many monotrichous bacterial species.

Declaration

I, David Wilkinson, declare that unless otherwise stated, all presented observations within this thesis are the product of my own work, performed within the Armitage lab in the Department of Biochemistry, University of Oxford, between 1st October 2006 and 30th September 2010.

Acknowledgements

There are a number of people who I would like to thank for helping to make this work possible.

Dr Mark Roberts, Dr Kathryn Scott and Dr Steven Porter have helped to direct me in the lab, and have been fountains of knowledge throughout.

Dr George Wadhams has helped in countless ways, and I could not have achieved any of this work without his continuing support and assistance.

Prof. Judy Armitage has allowed me to explore my own ideas and supported my decisions along the way.

Most importantly, my wife and daughter have been a constant source of encouragement, and have been there for me when I needed them. For this I am eternally grateful.

Abbreviations

All abbreviations used in this thesis are listed below (Table 1).

Abbreviation	Meaning
1D	One Dimensional
3D	Three Dimensional
Å	Angstrom
ATP	Adenosine Tri Phosphate
bEBP	Bacterial Enhancer Binding Protein
bp	Base Pairs
BSA	Bovine Serum Albumin
CCD	Charge-Coupled Device
CCW	Counter ClockWise
CFP	Cyan Fluorescent Protein
CW	ClockWise
Da	Dalton
DAPI	4',6-DiAmidino-2-PhenylIndole
DASs	Downstream Activation Sequence
DIC	Differential Interference Contrast
DIG	DIoxyGenin
DNA	Deoxyribo Nucleic Acid
dNTPs	Deoxynucleoside 5'-triphosphate
EM	Electron Microscopy
FITC	Fluorescein IsoThioCyanate
FRAP	Fluorescence Recovery After Photobleaching
GFP	Green Fluorescent Protein
HAPs	Hook Accessory Proteins
HBB	Hook Basal Body
His	Histidine
HK	Histidine Kinase
IM	Inner Membrane
LB	Luria Broth
Lrp	Leucine-responsive Regulatory Protein
MCP	Methyl-accepting Chemotaxis Proteins
OD	Optical Density
OM	Outer Membrane
PALM	Photo Activatable Light Microscopy
PBD	Peptidoglycan Binding Domain
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
RNA	Ribo Nucleic Acid
RNAP	RNA-Polymerase
SDS	Sodium Dodecyl Sulphate
SNM	Stochastic Nucleation Model
T3S	Type III Secretion
T3SS	Type III Secretion System
TIRF	Total-Internal Reflection Fluorescence
TLP	Transducer-Like Protein
UASs	Upstream Activation Sequence
UV	Ultra-Violet
YFP	Yellow Fluorescent Protein

Table 1**:** List of Abbreviations.

Contents

1) Introduction 12

1.1) Bacterial Motility and Chemotaxis 12

1.2) Flagella-Mediated Taxis 13

1.3) The Flagellum 15

1.3.1) The Filament 16

1.3.2) The Hook 18

1.3.3) The Basal Body 18

1.4) Flagellar Assembly 21

1.4.1) Protein Targeting and Secretion. 21

1.4.2) Sigma Factors and the Regulation of Gene Expression 23

1.4.3) The Flagellar Assembly Hierarchy 25

1.4.4) Regulating Flagellar Number 28

1.4.5) FliA and FlgM 29

1.5) Chemotaxis 32

1.6) Chemoreceptor Arrays 35

1.6.1) Architecture 35

1.6.2) Dynamics 36

1.6.3) Positioning 37

1.7) Subcellular Protein Localisation in Bacteria 42

1.7.1) Cell Division 43

1.7.2) Cytoskeletons 43

1.7.3) DNA Partitioning 43

1.7.4) Whole genome analysis 45

1.8) The model organism *Rhodobacter sphaeroides* 45

1.8.1) Motility 46

1.8.2) Flagellar Assembly 47

1.8.3) Chemotaxis 49

1.8.4) Chemoreceptor positioning 52

1.9) Aims 55

2) Materials and Methods 56

2.1) Strains and Plasmids: 56

2.2) Antibiotic concentrations 60

2.3) Growth Conditions 60
2.3.1) *E.coli* liquid cultures 60
2.3.2) *E.coli* agar plates 60
2.3.3) *R.sphaeroides* liquid cultures 60
2.3.4) *R.sphaeroides* agar plates 60
2.3.5) Bacterial preservation 61
2.4) Genetic Techniques 61
2.4.1) Polymerase chain reaction (PCR) 61
2.4.2) Agarose gel electrophoresis 61
2.4.3) Gel Extraction 61
2.4.4) Mini/Midiprep 62
2.4.5) *R.sphaeroides* genomic DNA extraction 62
2.4.6) Restriction digests 62
2.4.7) Dephosphorylation 62
2.4.8) DNA Ligation 63
2.4.9) Competent cells 63
2.4.10) Chemical transformation 63
2.4.11) DNA sequencing 63
2.4.12) Conjugation 63
2.4.13) *R.sphaeroides* chromosomal replacement 64
2.4.14) Southern blotting 64
2.4.15) DIG labelling 65
2.5) Protein Techniques 65
2.5.1) Protein purification 65
2.5.2) Sonication 65
2.5.3) Bradford assay 65
2.5.4) Acrylamide electrophoresis 65
2.5.5) Coomassie staining 66
2.5.6) Western Blotting 66
2.6) Phenotype Analysis 66
2.6.1) Growth curves 66
2.6.2) Soft-agar swim plates 67
2.6.3) Fluorimetry 67
2.6.4) Biofilm formation 67

2.6.5) Negative staining 68
2.7) Microscopy 68
2.7.1) Tethered cell analysis 68
2.7.2) Phototaxis assay 68
2.7.3) Epifluorescence microscopy 69
2.7.4) Confocal microscopy 69
2.7.5) Electron microscopy 69
2.8) Image analysis 69
2.8.1) Epifluorescence image analysis 69
2.8.2) Tethered cell analysis 70
2.9) Bioinformatics: 70
2.9.1) DNA sequencing fragment alignment 70
2.9.2) Multiple protein sequence alignment 70
3) The analysis of epifluorescence images of *R.sphaeroides* in MATLAB 71
3.1) Epifluorescent images of *R.sphaeroides* 71
3.2) Mean grey level analysis 72
3.3) Lengths, widths and positions analysis 73
3.4) Why was it done in this way? 74
3.5) Existing image analysis software 74
3.6) Phase Contrast 76
3.7) Differential interference contrast 77
3.8) Defining cells within images 79
3.8.1) Stage 1 - Whole-image object detection 80
3.8.2) Stage 2a – Cell edge refinement 82
3.8.3) Stage 2b – Fluorescent cluster position assignment 83
3.9) Storing data for fast analysis 86
4) Clusters and morphology in the average cell 87
4.1) Morphology Analysis 87
4.2) Population analysis of chemotaxis protein cluster positions 89
4.2.1) Polar chemotaxis proteins 89
4.2.2) Numbers of polar and lateral clusters in a wild-type population 90
4.2.3) Cells with one cluster 93
4.2.4) Cells with two clusters 93
4.2.5) Cells with three clusters 94

4.2.6) Cells with more than three clusters 95
4.3) Cytoplasmic chemotaxis proteins 97
4.3.1) Numbers of cytoplasmic clusters in a wild-type population 97
4.3.2) Cells with one cluster 100
4.3.3) Cells with two clusters 100
4.3.4) Cells with more than two clusters 101
4.4) Cluster separation in growing cells: 102
4.5) Asymmetry in polar cluster separation 103
4.6) Asymmetry in cytoplasmic cluster partitioning 107
4.7) The effect of asymmetry in cluster separation 109
4.8) Chapter discussion 112
5) Regulation of expression of the chemotaxis genes by FliA and FlgM 116
5.1) Deletion of *flgM* 116
5.2) FlgM and FliA regulate *che*Op2 expression 118
5.3) FlgM and FliA regulate *che*Op3 expression 120
5.4) Discussion: 121
5.5) Studying gene expression using fluorescence 122
5.6) Development of a fluorimetry-based protocol for measuring copy number 123
5.7) The effect of FliA and FlgM activity on *che*Op2 and *che*Op3 expression: 127
5.7.1) Complementing the deletion of *fliA* 128
5.7.2) Complementing the *flgM* deletion 132
5.7.3) The effect of FlgM on the activity of FliA: 138
5.8) Chapter discussion 140
6) Protein clustering in non-wild-type strains 142
6.1) The effect of concentration on polar protein clusters 142
6.2) The effect of concentration on cytoplasmic protein clusters 148
6.3) Patterns of localisation in filamentous cells 154
6.3.1) Dual-tagged strains for simultaneous study of polar and cytoplasmic clusters 154
6.3.2) Cell morphology analysis of filamentous cells 155
6.3.3) Polar/Lateral clusters in filamentous cells 157
6.3.4) Cytoplasmic clusters in filamentous cells: 161
6.4) Observation and isolation of long-cell spontaneous mutants: 165
6.5) Phenotypic characterisation of long-cell Mutants: 168
6.6) Cytoplasmic cluster positioning in long-cell mutants 171

6.7) Polar/Lateral cluster positioning in long-cell mutants 173
6.8) Chapter Discussion 174
7) FliA and FlgM regulate motility and flagellar genes 177
7.1) Motility and Chemotaxis 177
7.1.1) Chemotaxis with varying FliA activity 177
7.1.2) Chemotaxis with varying FlgM activity 180
7.1.3) FliA and FlgM regulate flagellar number in *R.sphaeroides* 183
7.1.4) FliA and FlgM are required for biofilm formation 187
7.1.5) Discussion: 189
7.2) FlgM 190
7.2.1) Making the YFP-tagged FlgM reporters 190
7.2.2) Phenotypic analysis of YFP-tagged FlgM mutants: 192
7.2.3) FliA regulates FlgM expression 193
7.2.4) FlgM is secreted 194
7.2.5) Observation of FlgM via epifluorescence 195
7.2.6) Discussion 196
7.3) MotB 197
7.3.1) Introduction 197
7.3.2) Making the YFP-tagged MotB reporter 198
7.3.3) Phenotypic Analysis of the YFP-MotB Strain: 200
7.3.4) YFP-MotB cannot be seen in the motor 201
7.3.5) FliA/FlgM regulation of MotB expression 201
7.3.6) Quantitative Western Blot Analysis MotB expression levels: 202
7.3.7) Discussion 206
7.4) FliM 206
7.4.1) Introduction 206
7.4.2) YFP-FliM is detectable within the motor 208
7.4.3) FliA/FlgM regulation of FliM expression 209
7.4.4) Discussion 209
7.5) Chapter Discussion 210
8) General Discussion 212
9) Appendices 218
9.1) List of Primers 218
9.2) Solutions and Chemicals 219

9.2.1) Luria broth 219
9.2.2) Succinate medium 219
9.2.3) 5 x DNA loading dye 220
9.2.4) ½ x TBE 220
9.2.5) Chromoprep lysis buffer 220
9.2.6) TFB1 221
9.2.7) TFB2 221
9.2.8) M22 medium 221
9.2.9) 2YT medium 221
9.2.10) Protein wash buffer 222
9.2.11) Protein elution buffer 222
9.2.12) 5 x protein loading dye 222
9.2.13) SDS running buffer 222
9.3) MATLAB code 223
10) References 224
11) Publications 233

1) Introduction

1.1) Bacterial Motility and Chemotaxis

Bacteria will often encounter changes in their local environments, and have the ability to sense and adapt to many of these changes by modifying their behaviour or by altering their gene expression.

Most bacteria have the ability to move within their environments towards conditions that are more favourable for their growth, and away from conditions that may inhibit their growth. The process by which cellular motion is directed towards or away from these stimuli is referred to as chemotaxis when the stimulus is a chemical, aerotaxis when the stimulus is oxygen, thermotaxis when the stimulus is temperature and so on. Bacteria can move in a variety of different ways (for a review see (68)); *Myxococcus xanthus* can move across solid surfaces using either gliding motility (143) or by twitching motility, which involves the extension and retraction of type IV pili (19), whereas *Flavobacterium johnsoniae* achieves rapid surface gliding motility via a mechanism which involves focal adhesion and the hydrolysis of ATP (110). In aqueous environments most bacteria swim by the rotation of semi-rigid, helical flagellar filaments which protrude from the cell bodies (6). In spirochetes such as *Treponema pallidum*, which causes syphilis, flagella are found in the periplasm and rotate to generate a corkscrew-like motion of the cell body which allows swimming through high viscosity fluids (24). The bacterial parasite *Spiroplasma melliferum* achieves swimming without flagella by the contraction of three intracellular, helical fibril filaments which alter cellular morphology in a way which generates motion (151). The best characterised motility systems are the flagella-swimming apparatus of the enteric γ-proteobacteria *Escherichia coli* and *Salmonella*, which provide the paradigm for flagella-mediated motion and chemotaxis (153).

Being able to move towards better environmental conditions confers an obvious, selective survival advantage upon motile bacteria. When the advantage provided by motility is removed, however, the selection process favours bacteria which are not motile. This was demonstrated in *Salmonella,* where spontaneous mutation reduced proportions of flagellated cells to zero after only ten days of growth in stirred liquid culture (93). This shows that bacterial motility is an energetically expensive venture, but is ultimately a rewarding investment under natural conditions. This investment also allows for some of the more complex survival-linked bacterial behaviours such as host infection and virulence, symbiosis or biofilm formation. For example, mutations which impair chemotaxis and motility have been shown to reduce pathogenicity in *Helicobacter pylori*, which

infects gastric tissue causing peptic ulcers, and *Vibrio cholerae*, which infects the small intestine causing cholera (20, 161). Formation of the symbiotic relationship between the nitrogen-fixing soil bacterium *Azospirillum brasilense* and the roots of wheat has been shown to require chemotaxis (152). Biofilms, complex multicellular structures that form on solid surfaces, are important factors in various industrial, medical and environmental situations (66). Their formation requires active motility in many species, such as *V.cholerae* and *Agrobacterium tumefaciens* (102, 106).

The importance of directed motion in so many different bacterial lifestyles highlights why an improved understanding of chemotaxis is essential. Bacteria can cause problems such as medically or agriculturally associated infections but can also be beneficial to the growth of crops, in the synthesis of chemicals or in our diet. The study of chemotaxis and motility may assist in the exploitation of, or the development of treatments against, bacteria. Some of the principles of signal transduction that have come from the study of the chemotaxis pathway are even being developed within the emerging field of synthetic biology, which aims to engineer biological systems to act as machines to perform particular tasks or as sensors to indicate the presence of chemical pollutants or poisons (111).

1.2) Flagella-Mediated Taxis

Bacteria that swim using flagella display a wide variety of phenotypes. For example, the α-proteobacteria *Rhodobacter sphaeroides* and *Caulobacter crescentus* are monotrichous (possess a single flagellum) but have flagella that form at different cellular positions (sub-polar and polar, respectively) (5, 120). *Esherichia coli* and *Salmonella* cells are peritrichous (possess several flagella) and have, on average, between five and eight flagella which protrude from the sides of the cell and coil together to form a flagellar bundle when swimming (6). *Vibrio fischeri* is also peritrichous, but forms a “tuft” of flagella at one pole of its cell (163). The γ-proteobacterium *Proteus mirabilis*, which causes urinary tract infections, can have hundreds of flagella which are used in a unique form of swarming motility (115) (Figure 1).

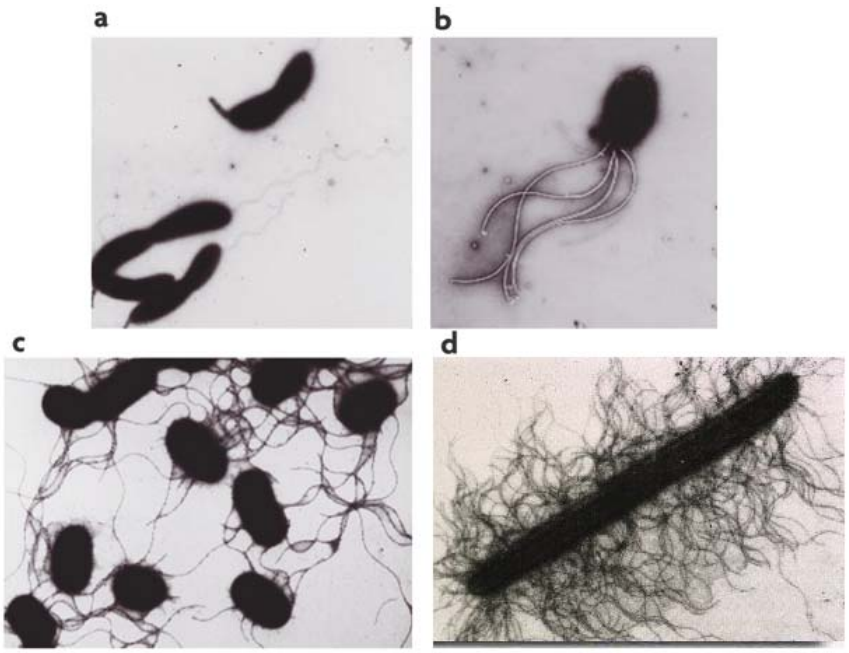


Figure 1: Electron micrograph images of the flagella phenotypes for **a)** monotrichous *Caulobacter cresentus*, **b)** lophotrichous *Vibrio fischeri,* **c)** peritrichous *Salmonella* and **d)** *Proteus mirabilis* (adapted from (25) and http://www.bhrgroup.co.uk/press/prm290102.htm).

The paradigm for flagella-mediated taxis states that counter-clockwise (CCW) rotation of the flagellar motor results in smooth swimming, whereas clockwise (CW) rotation results in a "tumble" which allows for a change in swimming direction. Tumbles interspacing swims in taxis result in a "random walk" pattern of swimming with no net motion. In chemotaxis, the CCW/CW bias is altered in response to a chemical stimulus resulting in a "biased random walk" and generating net movement towards attractants and away from repellents (36). However, with the wide variety of morphologies and flagella phenotypes observed across different bacterial species, it is unsurprising that different species swim in different ways. In monotrichous *R.sphaeroides*, a tumble event occurs when CCW rotation is stalled causing the flagellum to stop rotating and allowing the cell to re-orientate via Brownian motion (124). In peritrichous *E.coli,* "tumbles" result from one or many flagella switching from CCW to CW rotation; the filaments of these flagella uncouple from the flagellar bundle and "kick-out" from the cell causing an average change in swimming direction of 62 ° (12). In other bacteria, such as *Chromatium minus*, flagella "tumbles" result in a direct reversal (180 °) of swimming direction (105). These differences in swimming behaviour are thought to be optimised for the environments encountered by each species (90).

1.3) The Flagellum

The flagellar motor is one of nature's most complex macromolecular machines. Assembly of a bacterial flagellum requires a minimum of 13 genes which encode the flagellar structural components, and 25 further genes which encode components controlling its expression and assembly (94). In Gram negative bacteria, the flagellar motor spans two lipid membranes, the periplasm and the peptidoglycan layer, a distance of approximately 50 nm. The flagellar filament can contain tens of thousands of copies of a single protein (FliC) arranged as protofilaments around a 30 Å channel and can reach lengths of over ten times that of the bacterial cell. The complexity of the flagellar motor is such that some anti-evolutionary theorists have cited it as evidence for intelligent design (Figure 2) (28).

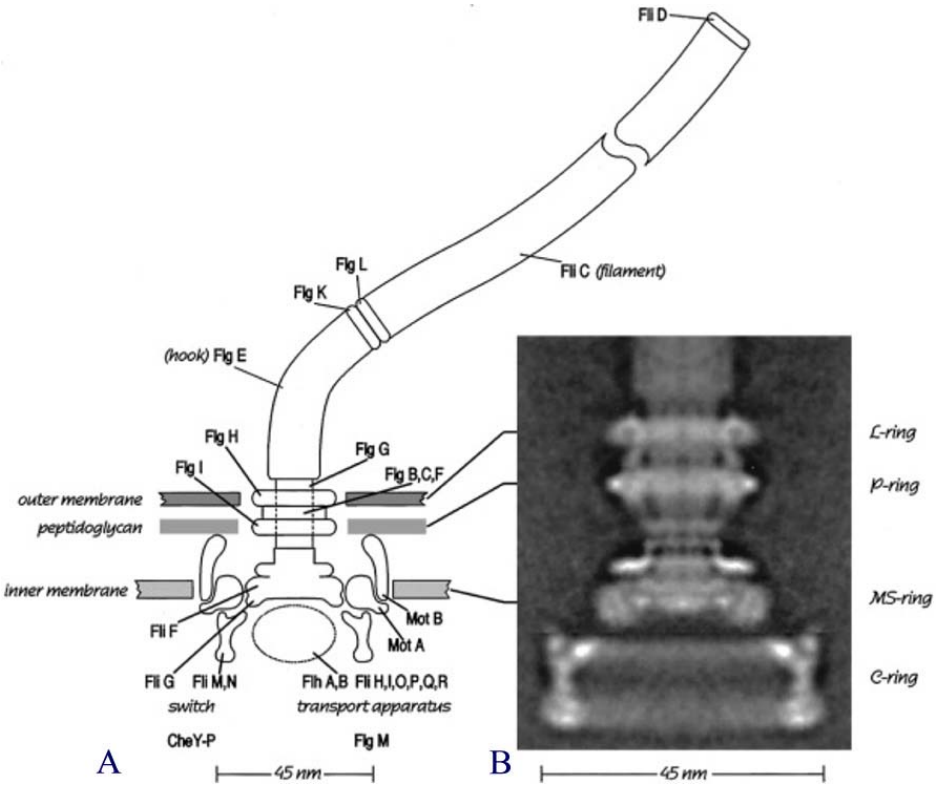


Figure 2: The Structure of the flagellar motor – **A)** Schematic and **B)** Electron micrograph reconstruction representations of the complex structure of the flagellar motor (taken from (11)). The protein components that make up the individual sections of the flagellar motor are indicated.

The subtleties of flagellar quaternary structures differ from species to species, as observed between *Salmonella* and the spirochete *Borrelia burgdorferi* (Figure 3). However the core structural assembly of all bacterial flagella is the same – consisting of three main sections; the basal body, the hook and the filament (11).

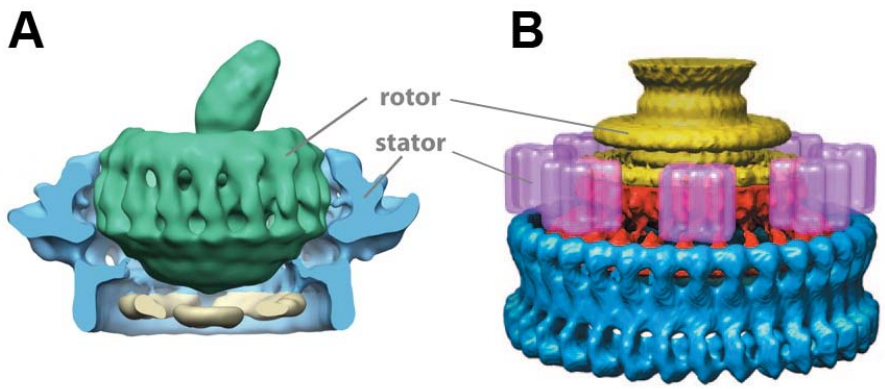


Figure 3: Models of the flagellar basal bodies constructed using cryo-EM tomography for **A)** *Borrelia burgdorferi* (adapted from (87)) and **B)** *Salmonella enteric sv. typhimurium*, The two main parts of the motor are displayed: the stator units, MotA and MotB (in purple), and the rotor (yellow: L- and P-rings, red: MS-ring and blue: C-ring). (adapted from (149)).

1.3.1) The Filament

In *E.coli*, FliC (or flagellin) is a 51.3 kDa protein which polymerises to form the 11 protofilaments of the flagellar filament. Each protofilament stretches the length of the filament, wrapped around a hollow central cavity which has a diameter of approximately 30 Å. The filament itself has a diameter of approximately 220 Å and can grow to a length of 15 µm (13). Divergence in the structure of the flagellar filament has been demonstrated between different species of bacteria (Figure 4) (49). However, the primary role of all filaments is the same – a propeller that drives cellular motion by rotation, powered by the flagellar basal body.

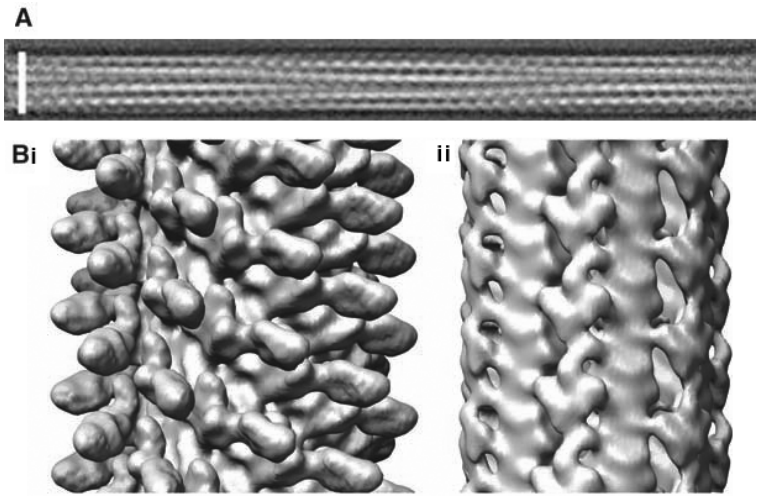


Figure 4: Electron micrograph reconstructions of flagella filaments from **A) and Bii)** *Campylobacter jejuni* and **Bi)** *Salmonella*. *C.jejuni* filaments are ~180 Å in diameter. *Salmonella* filaments are ~220 Å in diameter (adapted from (49)).

Filaments assemble at the distal end from unfolded FliC monomers that are exported through the flagella apparatus (otherwise referred to as the type-three secretion system (T3SS)) and assembled via the action of the cap protein FliD (165). The FliC protein forms a stable, semi-rigid polymer that can adopt one of two structural conformations (R-type or L-type) based upon the nature of the torque transmitted to the filament via the hook and basal body (95). Combinations of protofilaments with R and L conformations are responsible for the variety of filament morphologies that are observed in free-swimming bacteria (Figure 5). Changes in filament morphology, in turn, result in the changes in swimming direction that are essential to chemotaxis.

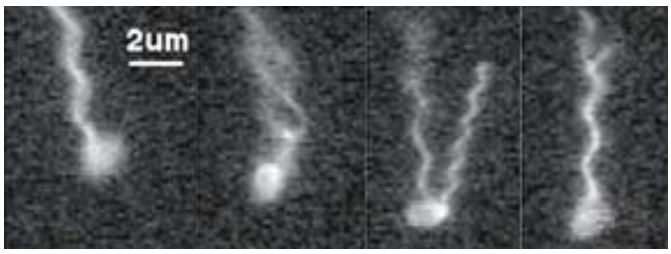


Figure 5: Images of a fluorescently-labelled swimming *E.coli*, with filaments that adopt different conformations (adapted from (13)).

1.3.2) The Hook

The flagellar hook is the joint between the basal body and the filament. Similarly to FliC, the 42 kDa protein FlgE is exported through the flagellar basal body in an unfolded state and incorporated into the hook by FliD. The filament and hook are connected by the Hook Accessory Proteins (HAPs) FlgK and FlgL. The length of the hook is known to be controlled by the regulatory protein FliK (104), and hook sizes and structures are divergent across bacterial species (41). In peritrichous bacteria, such as *E.coli*, flagellar filaments form a bundle which increase filament rigidity and ensure that all filaments push in the same direction. The hook of *E.coli* is therefore a flexible structure, which acts like a hinge between the cell body and the filament and appears curved in electron micrographs (135) (Figure 6B). In the monotrichous bacterium *R.sphaeroides*, the hook appears straight and is thought to have a more rigid structure (120, 160) (Figure 6A).

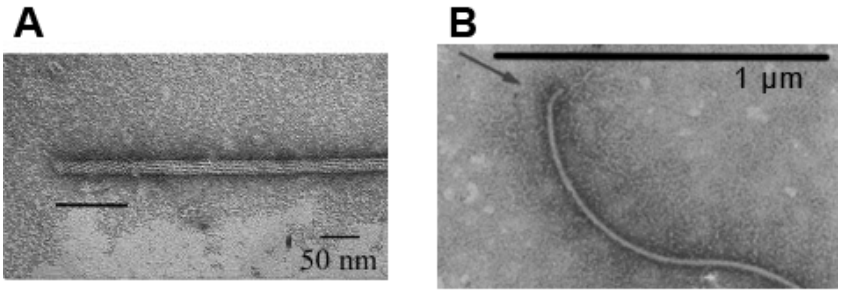


Figure 6: Electron microscopy images of flagellar hook structures from A) *Rhodobacter sphaeroides* (adapted from (120)) and **B)** *Escherichia coli* (adapted from (73)). Hook and filament regions are visible in both images; the hook regions and scales are indicated.

1.3.3) The Basal Body

The flagellar rotor, or basal body, consists of a series of ring-shaped protein polymers that are located within different layers of the cellular membrane. The hook is connected to the rod, composed of FlgB, FlgC and FlgF, which passes through the central cavities of these ring structures. The L-ring, composed of FlgL, is the outermost of these rings and is located in the outer membrane. The P-ring, composed of FlgH, is similar to the L-ring and is found within the peptidoglycan layer. Neither L nor P-rings rotate, but act as cylindrical linings that allow the rod to traverse the cell membrane, analogous to the bushings of mechanical devices that house axles and reduce the friction of rotation.

The rod is connected to the MS-ring which spans the inner membrane of the cell. The MS-ring is composed of FliF, rotates and has 28-fold rotational symmetry (46). It joins to the C-ring, composed of FliG, FliM and FliN, which is the section of the rotor that is exposed to the cytoplasm of the cell. Until recently, investigations into the structure of the C-ring had postulated that FliG formed a 1:1 interaction with the MS-ring, therefore having 28-fold symmetry, whereas FliM had approximately 34-fold symmetry. However, a recent crystal structure of full-length FliG from *Thermotoga maritima* has provided evidence that FliG may have 34-fold symmetry and form a 1:1 interaction with FliM (Figure 7) (82). Additionally, each monomer of FliM is known to interact with a stable tetramer of FliN, which has more than 100 copies in each rotor (113).

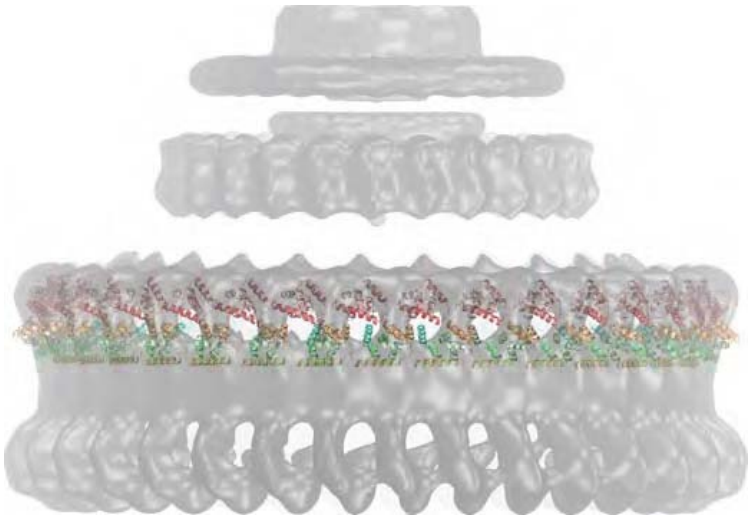


Figure 7: Proposed C-ring structure. Overlaid EM density map of the C and MS-rings of the basal body with the crystal structure of the FliG. 16 FliG monomers are displayed, which are arranged with 32-fold rotational symmetry, FliG alpha helices are coloured in random colours (taken from (82), supplementary information).

Protein interactions with the C-ring in the cytoplasm are responsible for switching from CCW to CW rotation in *E.coli* (164), and the C-ring is often referred to as "the switch complex". The N-terminal region of FliM has been shown to bind the phosphorylated form of the protein CheY, which results in switching in chemotaxis (40).

Rotation of the flagellar motor is achieved through the rotor's interaction with the torque-generating units that comprise the stator. The stator transforms the energy from the flux of positively-charged ions across the inner membrane of the cell into a structural change. The integral membrane proteins MotA and MotB harness this energy from the flow of protons (H^+), whereas

PomA and PomB harness energy from the flow of sodium ions (Na^+). In *Shewanella oneidensis*, MotAB and PomAB complexes are functionally interchangeable and can both act simultaneously on a single rotor, allowing efficient use of Na^+ ions under low and high salt environmental conditions (114). No high-resolution structures of the stator complex are available; however, extensive analyses via site-directed mutagenesis and chemical cross-linking provide information on the structure and mechanism of the torque-generating complex (26, 34, 35). The stator complex consists of four A and two B subunits. MotA (or PomA) has four transmembrane α-helices and its tetramer encloses a dimer of MotB (or PomB), which has only one transmembrane α-helix. The transmembrane region of the B subunit contains an essential aspartate residue (Asp32 in MotB, Asp24 in PomB) which is involved in the mechanism of ion transduction (Figure 8). The C-terminus of the B subunit contains a peptidoglycan binding domain (PBD) which is thought to anchor the stator complex to the peptidoglycan layer (34). The accepted model for the transmission of torque to the rotor complex involves a conformational change that is induced by the translocation of ions through the stator, and passed on to the switch complex via interaction with FliG (14).

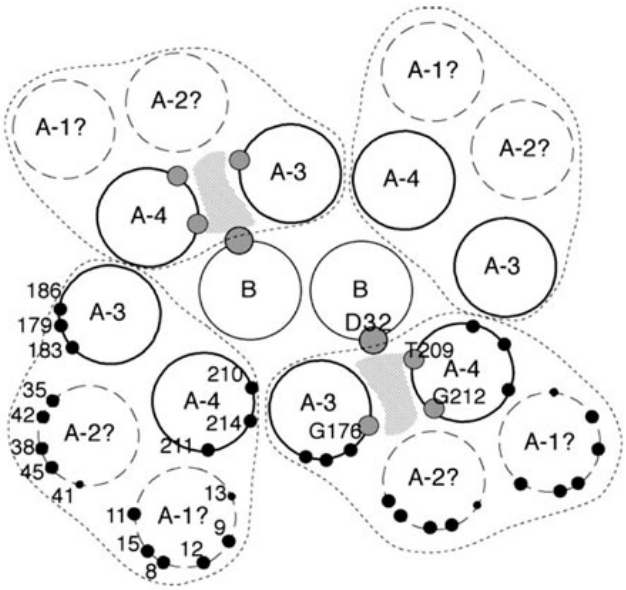


Figure 8: The Stator Complex – The $MotA_4B_2$ complex. Transmembrane helices are represented by circles for the 4-transmembrane protein MotA (A-1, A-2, A-3 and A-4) and the 1-transmembrane protein MotB (B). Black dots represent the positions of point mutations that were used to assess the mechanism of proton flux. Proton flux is proposed to occur through the shaded channel of this complex (taken from (13), reprinted from (14, 77)).

1.4) Flagellar Assembly

There are a number of logistical problems that must be overcome for the successful and efficient formation of flagella. Flagellar proteins are targeted and delivered to specific cytoplasmic, membranous, periplasmic and extracellular locations, and must come together in the correct order to synthesise what is a large and intricate machine. In addition, the number of each protein required at each location will vary dramatically; each flagellum may require only a handful of rotor, stator or hook proteins but may need in excess of 20,000 flagellin proteins, meaning that strict control over gene expression must be achieved. Furthermore, whether a flagellum is made or not must be accomplished as a function of both the environment that a cell is in and of how many flagella the cell has already made and this must be regulated differently between monotrichous and peritrichous species.

1.4.1) Protein Targeting and Secretion.

It is important that proteins, which are synthesised in the cytoplasm of bacterial cells, can be targeted to different locations both inside and outside of the cell (15). For example, bacteria produce many digestive enzymes which must be secreted from the cell in order to digest complex molecules into more simple organic compounds which can then be absorbed into the cell and metabolised. Alternatively, enzymes responsible for important cytoplasmic events such as DNA replication, transcription and translation must be retained within the cytoplasm, and the channels and pores responsible for transduction of metabolites into the cell must be incorporated into the cellular membrane.

Proteins are retained in the cytoplasm unless they contain a signal peptide sequence, which signals the final intended destination of that protein to the cell. Signal peptides are usually N-terminally encoded, but can also be more complex structural properties of fully folded proteins (75). Various chaperone proteins recognise these signal peptides and assist in the delivery of specific proteins to their final location (39).

Integral membrane proteins of the flagellar rotor are all targeted via the Sec pathway, whereas rod, hook and filament proteins are secreted through the central cavity of the Hook Basal Body (HBB) via the Type III Secretion System.

1.4.1.1) Type Three Secretion

T3SSs are used by many bacterial pathogens to inject harmful chemicals into eukaryotic host cells. The T3SS employed by the flagellar motor is controlled by six integral membrane proteins (FlhA, FlhB, FliO, FliP, FliQ and FliR) and three soluble proteins (FliH, FliI and FliJ), which have strong sequence homology with the injectisome proteins of *Pseudomonas*, *Salmonella*, *E.coli* and *Shigella* (among others) and the two systems are suggested to have co-evolved (108). The soluble components of the T3SS have been shown to have similar structures to those of the ATP-synthase complex, and a mechanism is proposed whereby a transient attachment of FliH, FliI and FliJ allows for unwinding and secretion of proteins in a process that is known to require the energy from ATP hydrolysis (44). Chaperone proteins interact with the membrane proteins of the T3S machinery, which form a complex in association with the flagellar rotor, passing proteins to the HBB for secretion. Recent analysis of the interactions formed between chaperone and T3S proteins in *Bacillus subtilis* has shown that as well as facilitating secretion of specific proteins, chaperone interactions may play a role in regulating the rate at which individual proteins are secreted via this complex (Figure 9) (9).

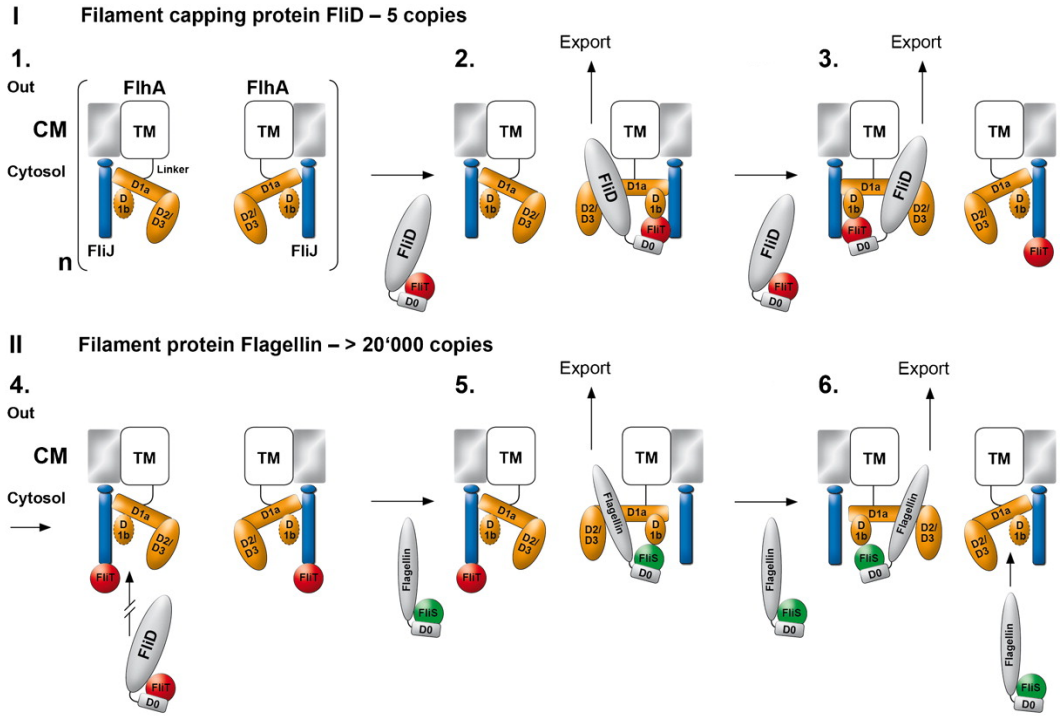


Figure 9: Type III secretion of FliD and FliC via two different chaperones – The base of the HBB, located within the cytoplasmic membrane (CM) is represented in grey. FlhA forms a complex at the base of the HBB, with transmembrane domains (TM) and cytoplasmic domains D1a, D1b and D2/D3 (yellow) which are involved in forming interaction with the chaperone proteins **i, 1-4)** The interaction with the FliD (grey) chaperone protein, FliT (red), allows secretion of one FliD molecule, but inhibits further secretion. **ii, 5-6)** Sequential rounds of FliC (Flagellin, grey) secretion are not inhibited by the FliC-specific chaperone, FliS (green).

1.4.2) Sigma Factors and the Regulation of Gene Expression

Regulation of the amount of each gene product made is not only essential to flagellar formation, but is universally encountered with the expression of all genes. The amount of a protein within each cell is determined by the relative rates of its synthesis and its degradation. Synthesis of each protein requires the transcription of its encoding gene from DNA to mRNA, and the translation of its mRNA to a protein sequence. Additionally, mRNA and proteins are constantly broken down and proteins are often exported from the cell, reducing cellular copy number (Figure 10).

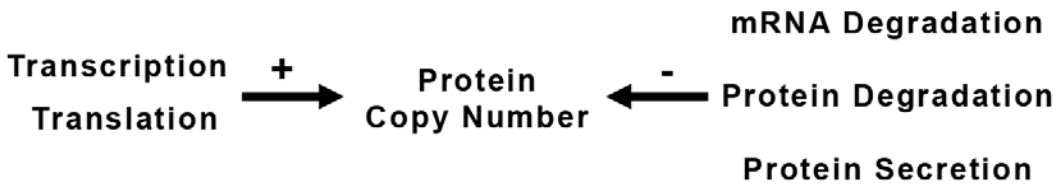


Figure 10: Factors that affect cellular protein copy numbers.

In bacteria, the processes of transcription and translation are often coupled, meaning that regulation of the rate of synthesis of a protein is predominantly controlled by the rate of activation of transcription.

The initiation of transcription involves the binding of the RNA Polymerase (RNAP) holoenzyme complex to a promoter, a non-coding sequence of DNA found upstream of the START codon of each gene. Specificity of RNAP to a particular promoter sequence is achieved by the σ subunit of the holoenzyme complex, and different σ factors compete for binding to RNAP in the cytoplasm. Binding to the promoter allows DNA "melting", separation of the two strands of the DNA molecule, and formation of the "transcription bubble" (101). RNAP then polymerises RNA nucleotides in the 5' to 3' direction, using the 3' to 5' DNA strand as a template. After the polymerisation of approximately ten RNA nucleotides, the σ factor dissociates from RNAP, which enters the elongation phase of transcription. Elongation continues until transcription is terminated in either a ρ-dependent or ρ-independent manner. ρ-dependent termination involves the recognition of a specific termination sequence on the RNA transcript, which is bound by a hexameric complex of ρ-proteins, destabilising the mRNA/DNA interaction. ρ-independent termination is achieved by transcription of a palindrome-containing sequence of RNA, which forms an RNA hairpin, preventing further elongation (119).

There are two main, structurally dissimilar, families of bacterial σ factor; σ^{70} (alternatively RpoD) and σ^{54} (alternatively RpoN). The majority of σ factors belong to the σ^{70} class and have four main protein domains of which two, σ^{2} and σ^{4}, are involved in the recognition of promoter sequences, forming interactions with DNA -10 and -35 bp from the start-site of transcription, respectively (97).

In *E.coli*¸ σ^{70} serves to transcribe the genes required for cell viability and is referred to as the "house-keeping" σ factor. Alternative members of both the σ^{70} and σ^{54} families act to transcribe genes and gene clusters (operons) in response to specific environmental stimuli, and their activity is regulated both at the level of transcription and by protein-protein interactions with effector proteins such as the anti-σ factors and bacterial Enhancer Binding Proteins (bEBPs) (128). For example, σ^{F}, which is one of the σ^{70} family members of the spore-forming bacterium *Bacillus subtilis*, is responsible for the compartment-specific transcription of genes involved in spore formation (103). The anti-σ^{F} protein SpoIIAB, which is also a protein kinase, binds and inhibits σ^{F} but is in turn regulated by the anti-anti-σ^{F} protein SpoIIAA. The switch between SpoIIAB binding σ^{F} and binding SpoIIAA, triggered by formation of the asymmetric septum, is responsible for successful sporulation (37).

The σ^{54} family is structurally and functionally dissimilar to the σ^{70} family, forming interactions with DNA -12 and -24 bp from the start-site of transcription. Binding of the σ^{54} holoenzyme complex is not sufficient to initiate transcription, and relies on DNA-bending and the energy of ATP-hydrolysis performed by bEBPs that bind specific DNA Activation Sequences that can be Upstream (UASs) or Downstream (DASs) of the promoter. ATP hydrolysis results in a local restructuring of the DNA molecule, which facilitates the initiation of transcription (Figure 11) (116, 128).

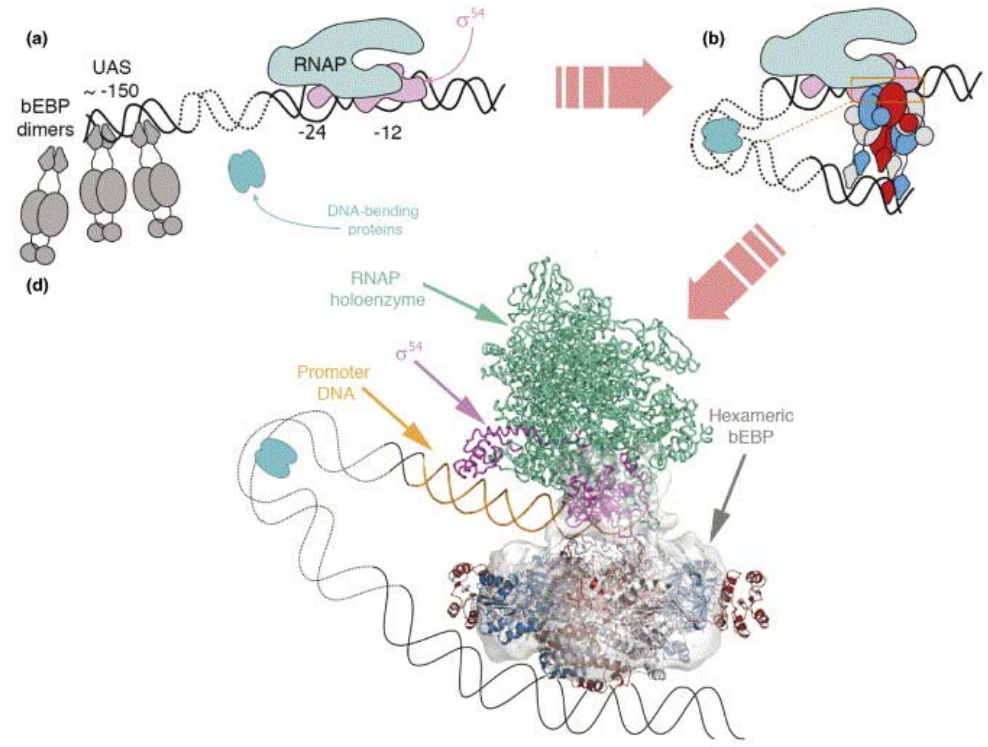


Figure 11: Representation of the initiation of transcription at a σ^{54} family promoter. **A)** The holoenzyme complex (RNAP, in blue, and σ^{54}, in purple) binds the promoter, bEBPs (in grey) bind to the UAS, DNA bending proteins (in blue) associated in-between these two loci **B)** DNA melting and initiation of transcription is facilitated by DNA bending and association of bEBPs (in red and blue) with the holoenzyme complex. The crystal structure diagram below provides a more accurate representation of what activation at a σ^{54} family promoter may look like at the molecular level (adapted from (128)).

E.coli, which grows mainly in the nutrient-rich environment of the gut, has only seven σ factors whereas other bacteria possess many more, allowing for tighter control of gene expression over the whole chromosome. For example, *R.sphaeroides* has 17 (92) and *Streptomyces coelicolor* has 65 (10). The sometimes complex regulation of gene expression allows cells to efficiently control the amounts of protein that are made in response to their requirements, which is essential to the conservation of energy and in adapting to ever-changing environmental conditions.

1.4.3) The Flagellar Assembly Hierarchy

The macromolecular structure of the flagellar motor is assembled from the inside out, and its ordered assembly is achieved using one of the more complex systems of gene regulation that are found in the bacterial kingdom. The paradigm for flagella assembly comes from *E.coli* and *Salmonella*, and provides a model for a multi-tiered hierarchical cascade of gene expression. The

position of each gene product within this hierarchy is referred to as its class, with *E.coli* and *Salmonella* having three classes of flagellar genes (Figure 12).

The class I *flhDC* master operon encodes two proteins, FlhD and FlhC. These proteins form a hetero-oligomeric complex consisting of four FlhD and two FlhC proteins. The $FlhD_4FlhC_2$ complex is a trans-acting transcriptional activator which is responsible for class II gene expression (158) and has been shown to bind to the ~-40 and ~-80 regions of three class II operons of *E.coli*: *fliA*, *flhB* and *fliL* (88).

Class II flagellar genes are required for the formation of the HBB structures which make up the T3SS. Also, FliA (σ^{28}) and FlgM, which regulate Class III gene expression, are class II genes. FliA and FlgM are a σ factor/ anti-σ factor pair that generate a checkpoint in flagellar assembly, FlgM is secreted upon completion of the HBB, allowing FliA to upregulate class III gene expression (discussed later). The class III flagellar genes consist of those that are needed to complete formation of the flagella, such as flagellin (see chapter 1.3.1) and the stator components (see chapter 1.3.3).

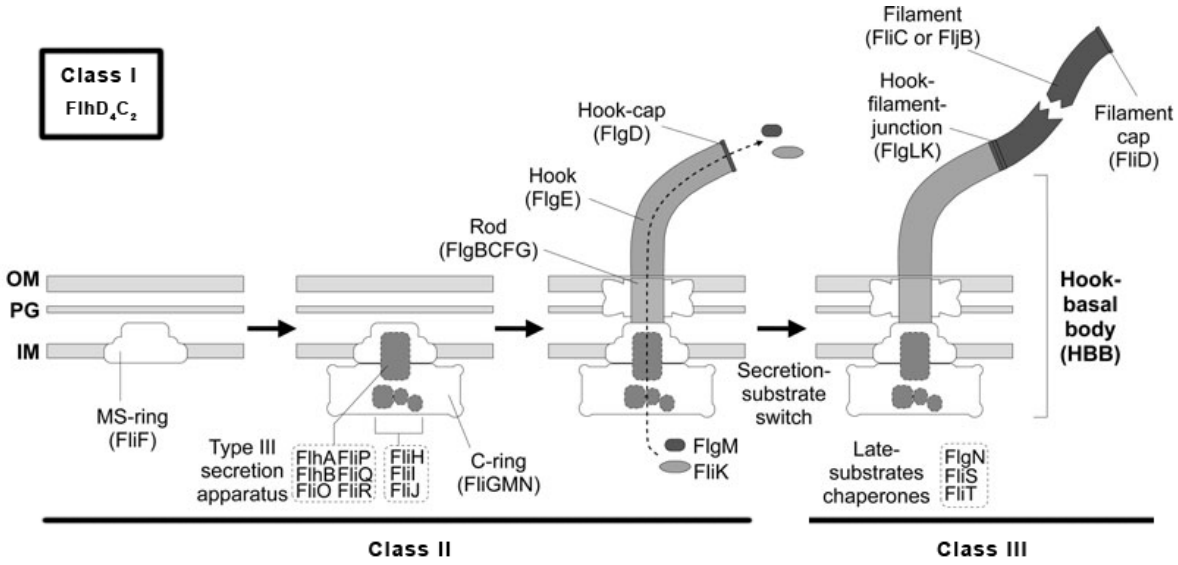


Figure 12: Flagellar assembly hierarchy of *Salmonella*: The ordered synthesis of a flagellum from the inside out. The expression class of each component within the hierarchy of flagellar assembly is indicated (adapted from (25)). The stator components are not shown in this diagram, but are class III genes in *Salmonella*.

Other bacterial species, such as *R.sphaeroides* (122), *Helicobacter pylori* (23), *Pseudomonas aeruginosa* (100) and *V.cholerae* (126) have a four-tiered hierarchy of flagellar gene expression.

In *P.aeruginosa*, the class I master regulator of flagellar gene expression is FleQ, a transcriptional activator of σ^{54} mediated gene expression which, interestingly, lacks a consensus sequence in the UASs that it binds (72). FleQ activity promotes the expression of the class II flagellar genes which include genes involved in the synthesis of the flagellar basal body, FlgM which inactivates a constitutively expressed FliA (47), FleN which inhibits FleQ and is involved in maintaining the monotrichous nature of the cell (30) and FleS and FleR which form a two-component signalling system responsible for the expression of the class III flagellar genes (132). Activation of class III gene expression by the FleSR signalling system allows completion of the HBB and the subsequent secretion of FlgM. Secretion of FlgM results in an increase in the concentration of free FliA, in turn resulting in class IV gene expression and the completion of the flagellar motor (Figure 13) (33).

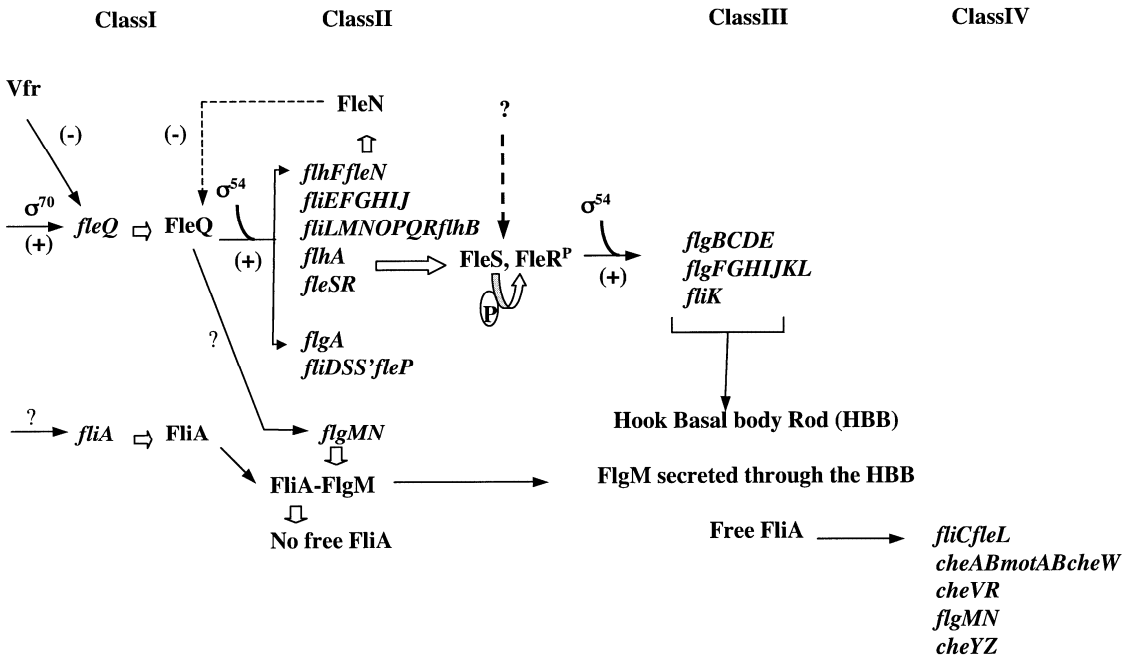


Figure 13: The four classes of flagellar gene expression in *P.aeruginosa* (taken from (33)).

Although in essence similar to the three-tiered hierarchy of *E.coli* and *Salmonella*, this four-tiered hierarchy allows for increased control over flagellar assembly with additional regulation of class III genes by FleS and FleR, as well as a delicate interplay between the activity of FleQ and its inhibition by FleN.

The spirochetes and the bacterium *B.subtilis* have very different methods of regulating flagellar assembly. Spirochete flagella are expressed with the house-keeping genes of the cell, and flagellar

assembly is regulated post-translationally (107). In *B.subtilis*, only two classes of flagellar genes exist with the majority of flagellar and chemotaxis genes in one long operon, also encoding σ^D which promotes transcription of the remaining flagellar genes (3).

1.4.4) Regulating Flagellar Number

As some species of bacteria are monotrichous whereas others are peritrichous, different bacteria must possess different mechanisms for sensing whether or not they have made a flagellum (or flagella) and altering the "decision" to produce a new flagellum accordingly. Various different factors have been shown to affect the number of flagella made by different bacteria (reviewed in (100)). For example, in *P.aeruginosa* (see above) Vfr has been shown to down-regulate expression of the master regulator, FleQ, by competitively binding to its DNA promoter sequence. Vfr therefore down-regulates flagellar number (31). As previously mentioned, FleN also regulates the master regulator of *P.aeruginosa* by directly binding to FleQ, preventing binding to its UAS. Deletion of FleN was shown to be sufficient to change *P.aeruginosa* from being monotrichous to being peritrichous (Figure 14) (30, 32).

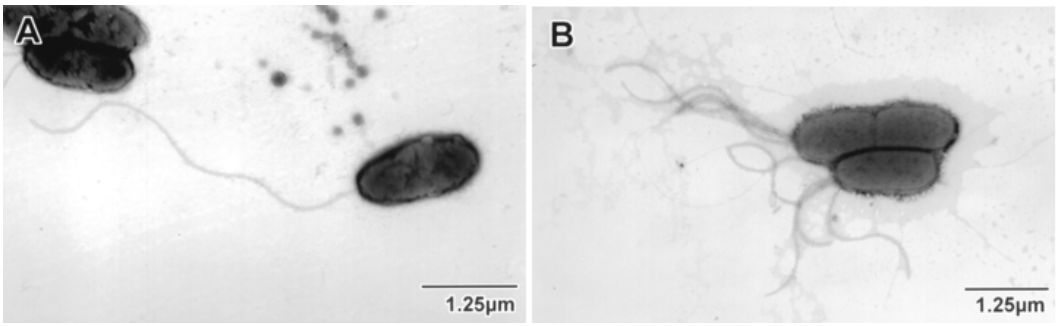


Figure 14: The effect of deleting FleN in *P.aeruginosa*. Electron micrographs of **A)** Monotrichous, wild-type strain PAK and **B)** Peritrichous *fleN* knock-in strain PAK-N (adapted from (30)).

In monotrichous *C.crescentus*, the formation of a polar flagellum is tightly linked to the distinctive cell cycle of this species. *C.crescentus* cells grow asymmetrically, with flagellated "swarmer cells" budding-off from surface-attached "stalk cells". The genetic and morphological characteristics of swarmer-cell progeny are different from the stalk cells, and the asymmetric activation of flagellar gene expression is controlled by the oscillatory action of the proteins CtrA and GcrA, which are also linked to DNA replication in these cells (29). CtrA and GcrA are known to trigger flagellar assembly, but altering their activity has not been shown to increase flagellar number.

The Leucine-responsive Regulatory Protein (Lrp) has been shown to be essential for hyperflagellation in *P.mirabilis*, and is thought to upregulate FlhDC activity (58). Lrp is a metabolic sensor that responds to the presence of L-leucine and L-alanine within cells, its function as both an activator and an inhibitor has been shown to be conserved across different bacterial species, and is likely to regulate flagella number in response to the availability of nutrients (86).

In *Salmonella*, FliT and FliZ were shown to have antagonistic effects on the number of flagella produced by the cell (79). Both FliT and FliZ are simultaneously class II and class III flagellar genes, with their expression being predominantly regulated by the σ factor FliA. FliZ enhances the activity of the master regulator, $FlhD_4C_2$, by increasing its stability and thus promoting the synthesis of more flagella (134). FliT is a flagellar-specific chaperone protein, which promotes the secretion of the cap protein FliD via the T3SS. FliT has been shown to have a second role as a repressor of $FlhD_4C_2$, mediated by direct interaction with the $FlhD_4C_2$ complex and therefore repressing flagellar synthesis (64). The fact that FliT and FliZ act antagonistically but are regulated in the same way has not been fully explained, but it has been suggested that repression of flagellar assembly by FliT could be temporally regulated by the secretion of FliD in a similar way to the regulation of FliA by FlgM. Repression when FliD levels are low may assist in slowing down flagellar synthesis when the existing number of flagella is high, whereas FliZ mediated activation (which is independent of FliT) may increase the rate of flagellar synthesis when the existing number of flagella is low (48).

1.4.5) FliA and FlgM

One important aspect of flagellar assembly, and part of the focus of this study, is the regulation of the final class of flagellar genes that is achieved through the activity of the σ factor FliA (σ^{28}) and its anti-σ factor, FlgM.

In *E.coli*, the *fliA* and *fliZ* (see above) genes are expressed from overlapping promoters at the *fliAZY* locus (134). *fliA* has a class II promoter, and its expression is controlled by the activity of the flagellar master regulator. Free FliA, which belongs to the σ^{70} class of σ factors, binds to the class III promoter and activates transcription of the class III flagellar genes. Class III genes include those genes responsible for completion of the flagellar motor, such as flagellin and the stator complex, but also the genes necessary for chemotaxis. *E.coli* chemotaxis genes are encoded at different chromosomal positions, and the expression of each of these loci has been linked to the activity of FliA (148).

FlgM is expressed from both class II and class III promoters (53), it is a 10 kDa protein that has been shown to bind FliA and destabilises the FliA/RNAP holoenzyme complex, therefore repressing class III gene expression (21, 22). Western blot analysis of the intracellular and extracellular fractions of cell cultures has shown that FlgM is secreted upon completion of the HBB, and mutants defective in HBB formation have decreased efficiency in FlgM secretion – suggesting that FlgM is secreted via the T3SS (62). The inhibition of FliA by FlgM has been demonstrated in many different species; for example, T3S deficient mutants of *Salmonella* produce fewer class III genes (52) and *P.aeruginosa* strains with mutations in *flgM* have been shown to upregulate flagellin expression (47).

The interaction between FliA and FlgM has been best demonstrated in the co-crystal structure obtained from *Aquifex aeolicus*, where FliA was shown to be held in an inactive conformation by the binding of FlgM (Figure 15) (144).

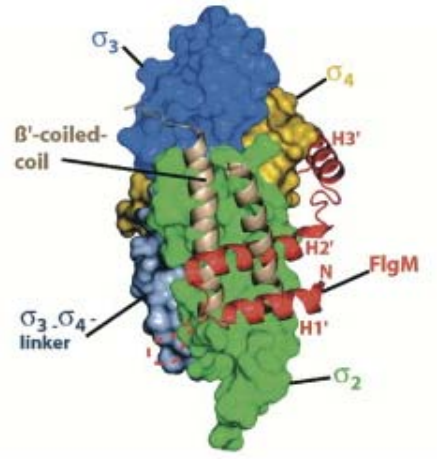


Figure 15: Co-crystal structure of the FliA/FlgM complex from *Aquifex aeolicus*. The region of FliA that does not interact with FlgM is displayed as a mesh and the individual domains of the molecule are coloured differently, the σ3-σ4 linker is light blue, the σ3 domain is dark blue and the σ4 domain is yellow. The FlgM/FliA interaction is displayed as a cartoon, FlgM is displayed in red and FliA in brown (adapted from (144)).

Study of the crystal structure, and sequence alignments of FlgM molecules from different species, allowed the identification of three C-terminally located conserved residues (I85, A86 and L89) which have been demonstrated to be essential to the binding of FlgM to FliA in *Yersinia pseudotuberculosis* (Figure 16) (38).

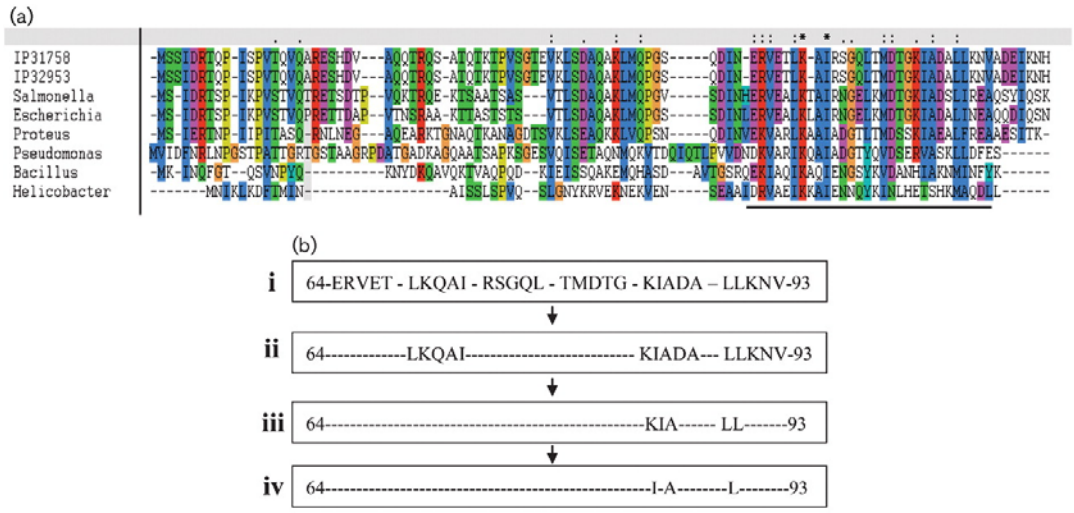


Figure 16: Identification of the key residues of FlgM. **A)** Alignment of the sequences of FlgM from several different species of bacteria. The different colours indicate different conserved residues across the different species. **B)** Identification of the 3 essential amino acids of FlgM. **i)** The conserved regions of FlgM were subdivided into short sequences that were mutated and functionality was assessed. **ii, iii and iv)** Successive rounds of mutagenesis and phenotype analysis were used to isolate the three residues which are essential to FlgM functionality in *Y.pseudotuberculosis*. (taken from (38)).

T3S-dependent activation of class III gene expression has commonly been referred to as a "checkpoint" in the hierarchy of flagellar synthesis. However, the relevance of this checkpoint is limited to the formation of a single flagellum since in peritrichous species T3S will occur constantly as long as at least one flagellum is present. Recent work suggests that in *Salmonella*, it is the rate of secretion of FlgM that fine-tunes the dynamics of FliA activity, thus not generating a "checkpoint", but a stable equilibrium where the correct number of class III gene products are generated relative to the flagellated state of the cell (18).

Additionally, filament shearing, which occurs commonly in liquid environments due to shear-flow of the liquid medium, may influence FlgM secretion and FliA activity. The evidence for this is contradictory, as FliA activity is independent of shearing in *Salmonella* (133), but has been shown to increase expression of FliA regulated virulence factors in *V.cholerae* (89). This interesting distinction may be explained by the fact that *Salmonella* is peritrichous, whereas *V.cholerae* is monotrichous as it seems likely that the "checkpoint" of HBB formation plays a more stochastic role in monotrichous bacteria (although this has not been shown).

It has also been observed that FliA activity can control flagellar number in peritrichous bacteria, but this has been shown to be due to the class III regulation of FliZ, which together with FliA forms a

positive feedback loop within the flagellar assembly pathway, enforcing the "decision" to become motile by promoting expression of a second flagellum after completion of the first (134).

1.5) Chemotaxis

The bacterial chemotaxis signalling pathway is one of the best characterised in biology, with predictive models having been successfully developed for the reduced-complexity signalling system of *E.coli*, which provides the paradigm for bacterial chemotaxis (7).

Excluding the 1 mm bacterial giants of some colourless sulphur bacteria, such as *Thiomargarita namibiensis* (71), bacterial cells are too small to sense chemical gradients along the length of their cells, which can be as small as 0.2 µm across in some species of *Mycoplasma*.

As discussed previously, flagellar-based bacterial swimming results in a 3D random walk through a liquid environment. In *E.coli*, the flagellar motor switches between CCW and CW rotation resulting in smooth swimming and tumbles, respectively. Instead of sensing absolute concentrations of chemical attractants or repellents, the chemotaxis system detects changes in local concentrations of these chemicals and adjusts the frequency of switching to produce net motion in a favourable direction. For example, a reduction in the concentration of a chemoattractant will increase the CW bias of the flagellar motor, resulting in an increased frequency of tumbling until the cell is reoriented at random into a more favourable direction (Figure 17).

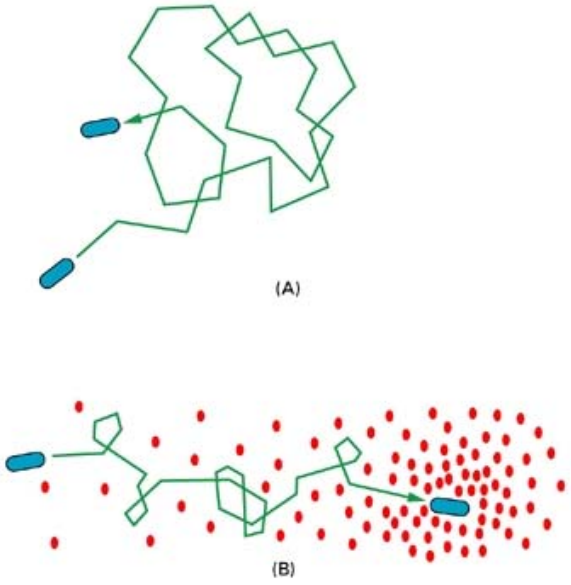


Figure 17: Net movement by a biased random walk. Schematic representations of bacteria moving using **A)** a random walk and **B)** a biased random walk, influenced by the concentration of chemoattractant, which is indicated by red dots (taken from http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=cell&part=A3946).

The signalling pathway used to achieve this temporally regulated switch is an example of a two-component signal transduction system. Two-component signalling is a common form of signal transduction in bacteria and involves the detection of an environmental stimulus by a histidine kinase and the phosphorylation of a response regulator protein by this kinase. Once present in sufficient quantities, the phosphorylated form of the response regulator goes on to trigger a change in behaviour, or in gene expression such as those linked with quorum sensing, nitrogen sensing and osmosensing (reviewed in (16)).

In chemotaxis, the histidine kinase (CheA) does not have a dedicated sensory domain, but instead interacts with the trans-membrane sensory Methyl-accepting Chemotaxis Proteins (MCPs), and this interaction is dependent on the structural adaptor protein, CheW. The MCPs form trimers of dimers that sense chemicals in the periplasm and this conformation has been shown to be essential to the function of the MCPs (Figure 18) (59). *E.coli* has five distinct MCPs each of which senses a range of different amino acids, sugars and dipeptides as well as pH, temperature and redox potential (146). Changes in periplasmic chemical concentration induce conformational changes in the structure of the MCP/CheW/CheA complex that alter the rate at which CheA dimers can autophosphorylate (59, 60). This, in turn, alters the rate at which the two response regulators, CheY and CheB, accept phosphoryl groups from CheA.

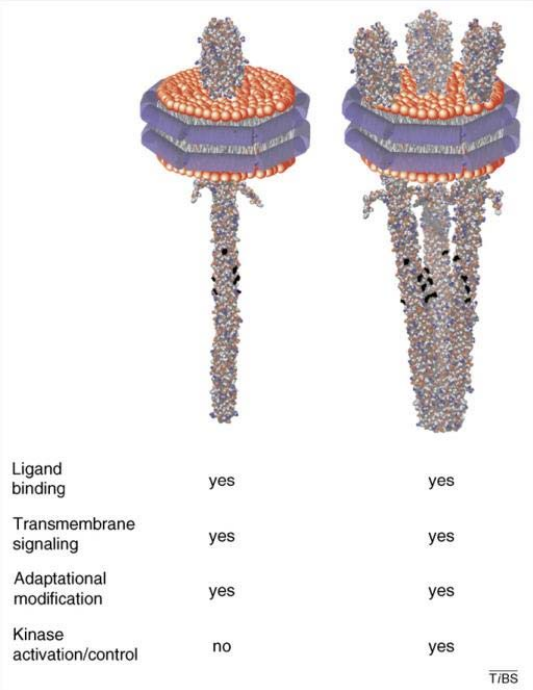


Figure 18: MCP trimers of dimers. Molecular model diagrams of a single MCP dimer and the trimer of dimers arrangement of MCPs. These MCPs are displayed assembled within a 10 nm diameter Nanodisk, a small protein assembly that stabilises a hydrophobic lipid bilayer core. The functional dependence on the trimer conformation is indicated below each diagram

The two response regulator proteins of *E.coli* chemotaxis have distinct roles. The phosphorylated form of CheY, CheY-P, binds to FliM in the flagellar rotor and induces switching, whereas CheB phosphorylation is involved in adaptation.

The adaptation pathway involves the antagonistic action of the methyltransferase, CheR, and the methylesterase, CheB, on the MCP proteins. The methylation state of the MCPs determines their sensitivity to different concentrations of attractant and repellent (43). CheB-P has a 100-fold increased rate of de-methylation over CheB, whereas CheR activity remains constant (4). Therefore, increased autophosphorylation of CheA results in a decreased methylation state of the MCPs. The adaptation of methylation state allows the chemotactic response to detect relative changes in concentration over large ranges of absolute concentrations.

Finally, the flagellar switching signal is terminated by the action of the CheY-P specific phosphatase, CheZ, which has also been shown to localise to the polar regions of the cell in *E.coli* where it interacts with a truncated version of CheA - $CheA_{short}$ (57). Although a relatively complex variation of the canonical two-component signalling pathway (Figure 19), the chemotaxis pathway is exquisitely simple (it can be achieved using as few as 7 protein components, (127)) and remarkably robust (it withstands variations in the concentrations of its components,(2)).

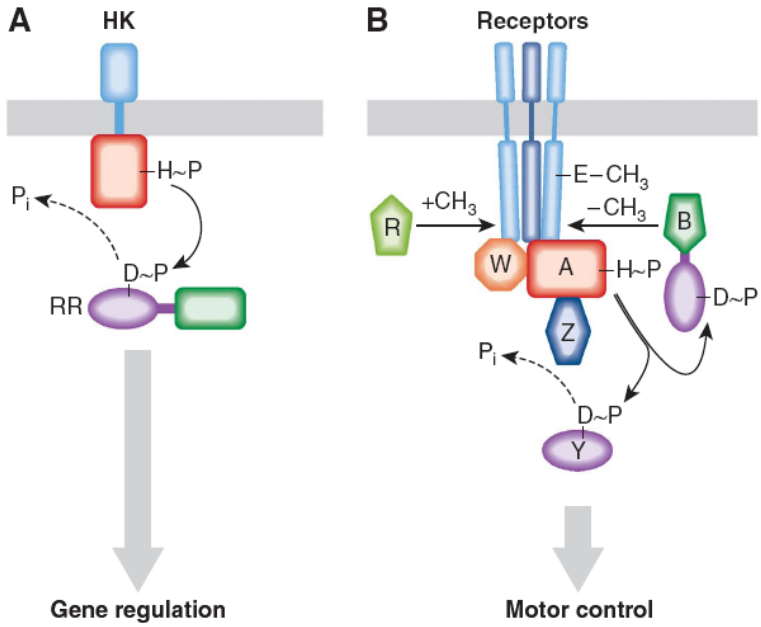


Figure 19: Two-component and chemotaxis signalling in bacteria. Schematic representations of **A)** the canonical two-component system, composed of a sensory Histidine Kinase (HK) and response regulator (RR), **B)** the chemotaxis pathway of *E.coli*, composed of receptor trimers of dimers, the histidine kinase CheA, structural adaptor CheW, response regulators CheY and CheB, methyltransferase CheR and phosphatase CheZ (taken from (145)).

The chemotaxis proteins found at the poles of the cell are arranged into large protein arrays that can contain tens of thousands of protein molecules.

1.6) Chemoreceptor Arrays

Many different approaches have been used to study the localisation, dynamics, structure and function of the large chemoreceptor arrays that are present in many bacterial species (reviewed in (145)). The ability to visualise these, and other, subcellular protein structures has improved vastly in recent years, with the development of fluorescent imaging techniques including epifluorescence, Total Internal Reflection Fluorescence (TIRF) and super-resolution Photo Activated Localisation Microscopy (PALM) which have all been used to study chemoreceptor arrays. Additionally, improvements in electron cryo-tomography have enabled 3D reconstructions of whole cells and subcellular structures to be assembled, which have also revealed valuable information about the structure of chemoreceptor arrays.

1.6.1) Architecture

A recent cross-species analysis of the structures of chemoreceptor arrays performed by electron cryo-tomography identified a universal architecture to the arrangement of these structures (17). When observed side-on, the periplasmic receptor domains, cytoplasmic receptor domains and CheA/CheW base plate of the arrays were all visible (Figure 20).

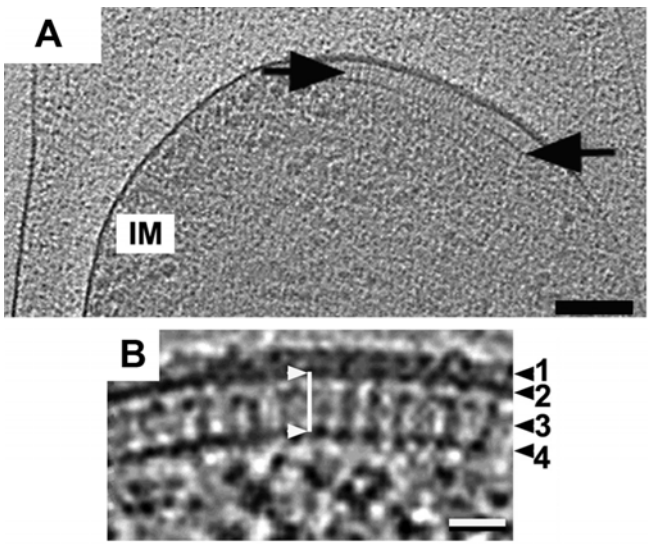


Figure 20: The characteristic appearance of a chemoreceptor array. **A)** Cross section through a *T.maritima* cell. Arrows mark the chemoreceptor array. The position of the inner-membrane is also indicated (IM). Scale bar represents 100 nm. **B)** Enlarged view of a cross-section through a *T.maritima* cell. The positions of the periplasmic receptor domains (1), inner-membrane (2), cytoplasmic receptor domains (3) and CheA/CheW base plate are all indicated. Scale bar represents 25 nm (adapted from (17)).

3D reconstruction and observation of the same structures as if from on-top revealed a hexagonal packing to the chemoreceptor array complex. Pleasingly, this hexagonally packed structure agreed exactly with a proposed model for the arrangement of cytoplasmic components within this array, which had been put together (literally) from 3D-printed models of the crystal structures of CheA, CheW and the cytoplasmic domains of the MCPS (Figure 21) (139).

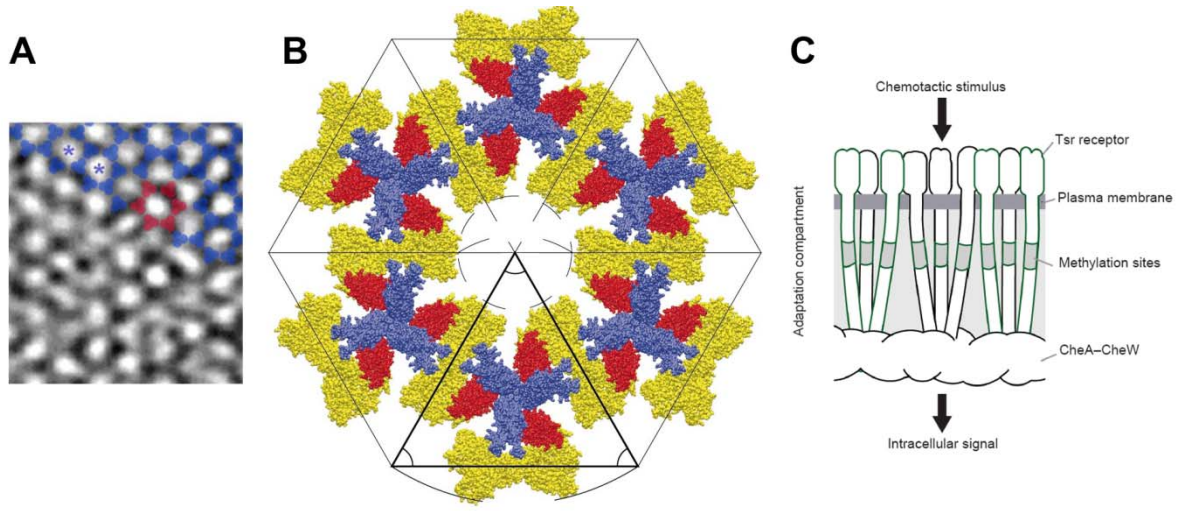


Figure 21: The hexagonal packing of the chemoreceptor array. **A)** Top-down view of an electron cryo-tomography 3D reconstruction of a chemoreceptor lattice, hexagonal packing of trimers of dimers is indicated in blue and red (adapted from (17)). **B)** Bottom-up view of a model of CheA dimers (yellow), CheW (red) and MCP cytoplasmic domains (blue) constructed from NMR and crystal structures. **C)** Side-on representation of the model proposed in **(B)** (**(B)** and **(C)** adapted from (139)).

1.6.2) Dynamics

Derivatives of the naturally occurring Green Fluorescent Protein (GFP) are widely used in the study of subcellular morphology. Genetic manipulation allows a fluorescent protein to be fused to another non-fluorescent protein, which allows the position of that protein to be visualised via fluorescence microscopy. Fluorescent proteins are susceptible to bleaching by prolonged exposure to light, and a recent study employed this property to study the dynamics of the chemotaxis proteins in the chemoreceptor array.

Directed photo-bleaching of fluorescent-fusion variants of the chemotaxis proteins at the poles of *E.coli* cells was achieved using a laser. Time-lapse fluorescence imaging of the cell (visualised using TIRF microscopy) allowed the rate of diffusion of the remaining fluorescent proteins into the bleached pole to be calculated (76, 137). This experimental procedure is known as Fluorescence

Recovery After Photobleaching (FRAP) and provides a method of estimating the rate of protein exchange in a particular subcellular region (Figure 22).

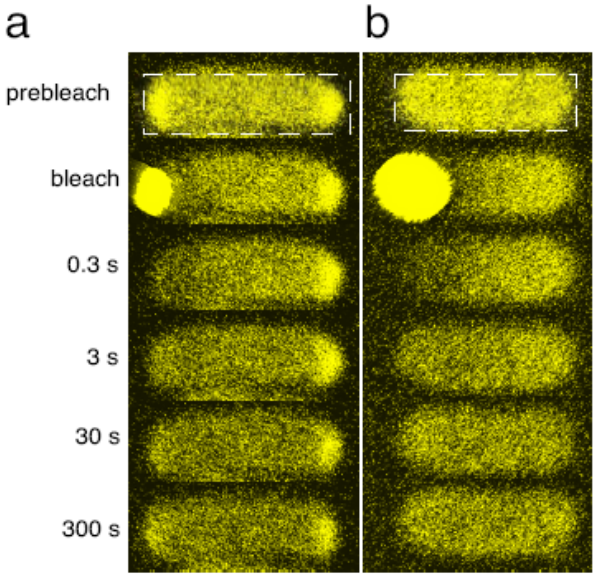


Figure 22: Studying the dynamics of the chemotaxis proteins of *E.coli*. Diagram showing the process of FRAP. Fluorescent-fusion proteins (in this case **a)** CheZ-YFP and **b)** delocalised CheZ-YFP), are visualised using fluorescence microscopy techniques at time 0, a laser is used to bleach the polar region of the cell. Sequential images at times 3 s, 30 s and 300 s allow the rate of exchange of fluorescent and bleached regions of the cell to be calculated (taken from (137)).

These experiments showed that the core elements of the chemoreceptor arrays are very stable structures. MCP proteins did not exchange, and CheW and CheA proteins exchanged slowly over time scales approaching that of cell division (~12 minutes). The phosphatase, CheZ, also exchanged slowly (~8 minutes), whereas adaptation enzymes CheR and CheB exchanged more rapidly (~15 seconds). Unsurprisingly CheY, which must be phosphorylated and diffuse to the flagellar rotor in order to induce switching within the time scale of the chemotaxis response (~0.1 seconds), showed rapid exchange dynamics.

1.6.3) Positioning

The use of epifluorescent imaging of fluorescent-fusion chemotaxis proteins in *E.coli* has also allowed for the study of positioning of the chemoreceptor arrays. In addition to polar clusters of chemotaxis proteins, clusters have been observed to form on the lateral edge of rod-shaped *E.coli* cells. Furthermore, these lateral clusters were seen to form at periodically positioned loci within the cell, corresponding to the 1/8, 1/4, 1/2 , 3/4 and 7/8 positions of the cell's long axis (Figure 23) (147).

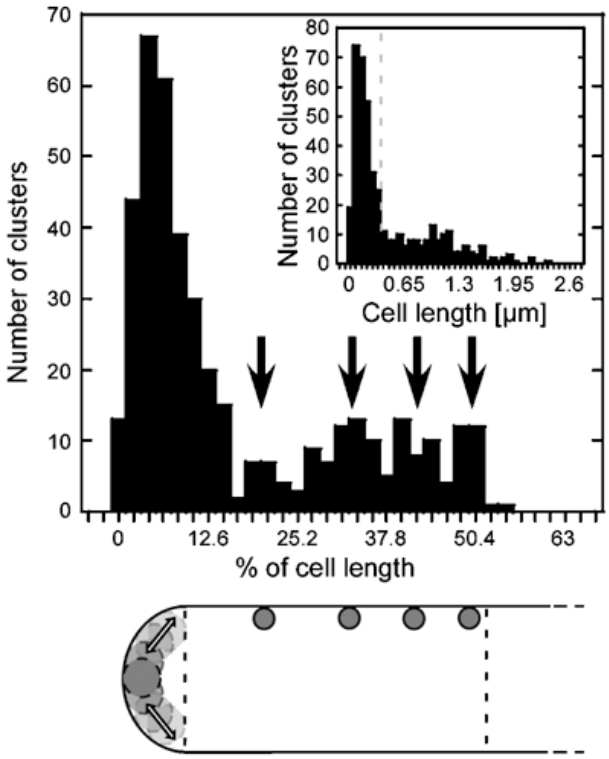


Figure 23: Periodical positioning of lateral chemoreceptor clusters. The positions of chemoreceptor clusters as observed by epifluorescence are shown relative to the cell length and as absolute positions (inset). Distances are measured to the nearest pole. A schematic of the positions of these clusters within a rod-shaped cell is presented underneath (adapted from (147)).

Timecourse fluorescence microscopy whilst growing the same cells on a surface revealed that clusters that form at lateral positions move to the poles of the cell after sequential rounds of cell division. This led to the hypothesis that future sites of cell division are pre-defined by an unknown cytoskeletal element within the cell, and that chemotaxis protein clusters localise with this cytoskeletal element. In support of this theory, the origins of replication of DNA (which also localise to discrete positions within the cell, see later) were suggested to co-localise to these future sites of cell division (Figure 24) (147).

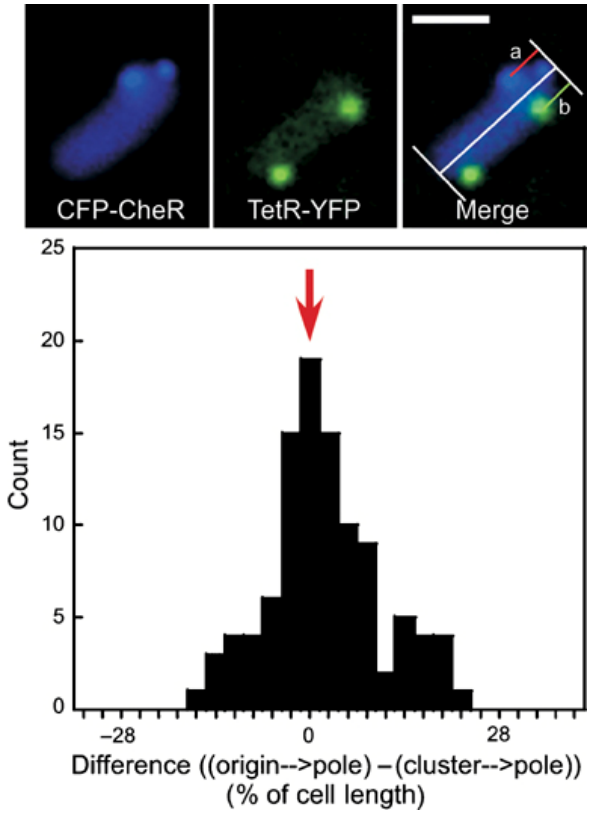


Figure 24: Colocalisation of lateral chemoreceptor clusters and origins of replication. Images show the localisation patterns of CFP-CheR (blue), and TetR-YFP which marks the origin of DNA replication (green). Measurement of the distance between these clusters shows that they are found in similar positions within the cell (adapted from (147)).

Additionally, timecourse fluorescence microscopy of the same cells, but immobilised, provided evidence that lateral clusters are static, as they were not observed to diffuse with time, whereas clusters at the poles of the cell were seen to diffuse within the polar region (indicated in Figure 23, see above) (147).

The receptors of chemotaxis protein clusters have been shown to insert into the cell membrane along the whole length of the cell (140). Receptors will either join to existing clusters via 2D diffusion within the membrane, or nucleate a new cluster. The nucleation of new clusters has been shown to be dependent on the concentration of chemotaxis proteins in the cell, as increasing the expression of the class III chemotaxis genes by deletion of the anti-σ factor, FlgM, resulted in an increase in the number of lateral clusters within cells (148).

Lateral cluster nucleation is proposed to be followed by association with the “pre-division site” of the cell, which is suggested to cause the reduced diffusion rates that are observed in lateral clusters and the periodicity in cluster positioning. This model for cluster assembly fits well with the

suggested universal mechanism of bacterial protein localisation which is coined “diffusion and capture” in a recent review (Figure 25) (138).

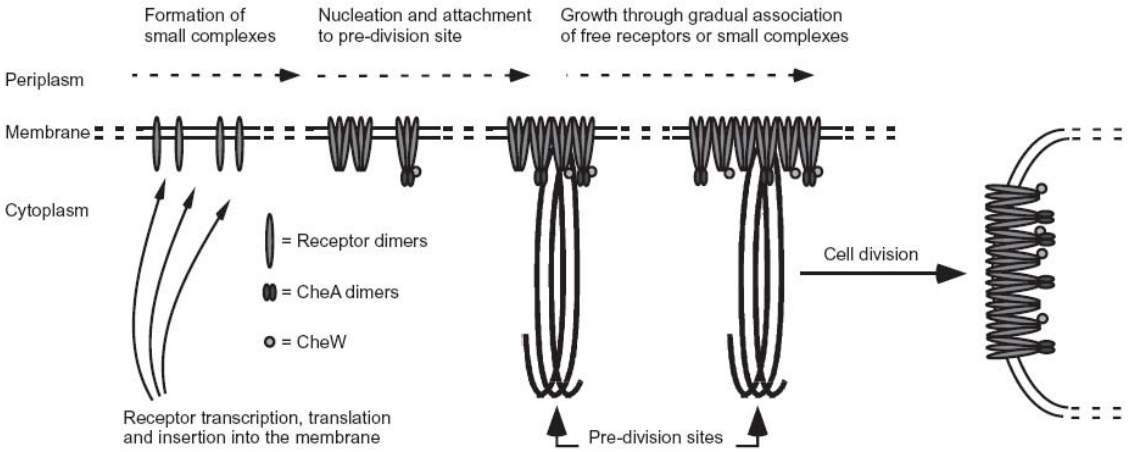


Figure 25: Proposed mechanism of chemoreceptor positioning. Chemoreceptors are inserted into the cytoplasmic membrane and are free to diffuse; nucleation of new lateral clusters occurs and is followed by association with pre-division sites. Cell division results in clusters being positioned at the new pole (taken from (147))

However, computer modelling of the interactions formed by chemotaxis receptors that was performed in response to the observation of periodicity in localisation suggests that these patterns may be explained not by “diffusion and capture”, but “diffusion and self-association”. Simulations of diffusing particles which had the ability to associate with each other into large complexes produced similar localisation patterns to those that were observed in *E.coli*, without the additional cytoskeletal element (Figure 26). This is called the Stochastic Nucleation Model (SNM) (157).



Figure 26: Simulated localisation patterns of chemoreceptor clusters. The results of simulating particle diffusion and association within an elongating 2D environment are shown. Cluster formation occurs as rapidly diffusing particles grow upon collision with others, reducing their diffusing rate due to an increase in mass. A comparison to observed localisation patterns via epifluorescence is inlayed (taken from (157)).

Without the identification of the “pre-division site element” it is nearly impossible to distinguish between the two models of cluster positioning. Further evidence has been obtained in support of stochastic nucleation by the use of the super-resolution imaging technique, PALM.

In PALM, a derivative of GFP is fused to a protein which is not fluorescent, but becomes fluorescent upon exposure to light of a particular wavelength; fluorophores with this property are termed “photo-activatable fluorophores”. Cells are immobilised and fixed to reduce diffusion of subcellular particles, then small numbers of fluorophores are photo-activated at a time. Visualisation of single molecules is achieved using a high signal-to-noise setup such as TIRF microscopy. By visualising single molecules with high-resolution imaging techniques, the exact position of each fluorophore can be determined to a precision of 5 nm. In comparison, epifluorescence techniques are limited in their resolution by the wavelength of light used; here, two fluorescent particles cannot be distinguished unless they are separated by a distance which is approximately half the wavelength of the fluorescent light that they emit. A PALM image is a composite image from many successive rounds of photo-activation and image acquisition. Images for each fluorophore are captured until that fluorophore bleaches, avoiding the detection of the same fluorophore more than once. Sequential rounds of imaging mean that a composite image can take hours to generate, and therefore this technique is not suitable for living-cell analysis.

A study of chemoreceptor positioning that used PALM was able to detect single molecules and small clusters that would not be visible via epifluorescence techniques (Figure 27). The probability of seeing a small cluster was evaluated relative to the total number of proteins detected and the presence of one or two preformed polar clusters. In agreement with the SNM, small lateral clusters were seen to be more likely to form away from existing clusters and were more prevalent in cells with two occupied poles. This in turn allowed for an extension of the SNM (55).

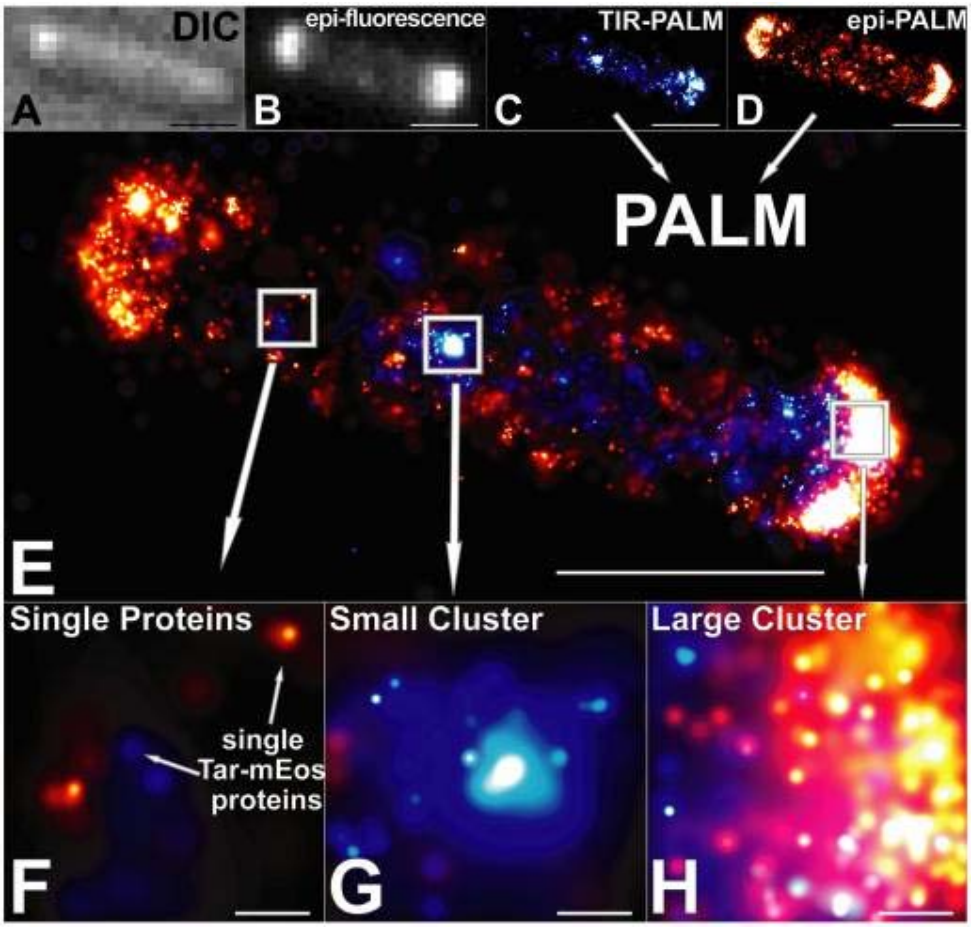


Figure 27: Imaging chemoreceptor clusters using PALM. **A)** Light image, **B)** epifluorescence image, **C)** TIRF-acquired PALM image, **D)** epi-acquired PALM image and **E)** composite PALM image of the same cell. Cells contained a photo-activatable fusion of the chemoreceptor Tar. Scale bars represent 1 µm. **F)** Single proteins, **G)** small lateral clusters and **H)** large polar clusters were all visible using this technique. Scale bars represent 50 nm (taken from (55)).

The patterns of chemotaxis cluster positioning that have been extensively observed in *E.coli* are thought to play important roles in cluster segregation and in chemotactic signalling (145).

1.7) Subcellular Protein Localisation in Bacteria

The distinctive pattern of polar and lateral positions of chemoreceptor arrays is just one example of complex subcellular localisation within bacteria. Whereas once bacteria were thought of as "amorphous reaction vessels", it is now known that the subcellular structures formed by bacteria are as diverse and intricate as some of those that can be observed in eukaryotic cells. In large part, our understanding of the subcellular organisation of bacteria can be credited to the discovery and genetic manipulation of the naturally occurring protein fluorophores, such as GFP (reviewed in (138)).

1.7.1) Cell Division

Cell division is one process that intuitively must require some form of subcellular localisation. Many bacterial cells divide symmetrically about their middle, and the identification of the mid-cell position of the cell is a precursor to cellular septation. This is achieved by the inhibition of polymerisation of the tubulin-like protein FtsZ, which forms the Z-ring at cell septation loci. Gradients of protein inhibitors of FtsZ are formed relative to the poles of the cell, using different mechanisms, which result in the minimum concentration of these inhibitors being present at the mid-cell, and subsequently Z-ring formation occurs at this position (reviewed in (91)). In *E.coli*, the FtsZ inhibitor protein, MinC spatially oscillates between the poles of the cell in association with the cytoplasmic membrane via a mechanism that involves the MinD and MinE proteins and the hydrolysis of ATP. Oscillation means the MinC is present in the lowest concentration at the mid-cell, allowing polymerisation of FtsZ (61, 129).

FtsZ polymerisation is inhibited by naturally occurring proteins *in vivo*, but can also be inhibited chemically by the antibiotic cephalexin. Cephalexin-treated cells are elongated due to the inhibition of cell division and are commonly used to study subcellular protein organisation (96, 141).

1.7.2) Cytoskeletons

The regulation of where cell division occurs is one way of thinking about how bacteria regulate cell length. Control of the diameter of a bacterial cell is less well understood, but is known to be controlled by cytoskeletal proteins such as MreB and RodA. MreB is a structural homologue of the eukaryotic cytoskeletal protein, actin, and has been observed to form helical filaments along the length of bacterial cells. Localisation of proteins associated with the growth of the peptidoglycan layer to periplasmic positions determined by the cytoplasmic location of the MreB filament is thought to result in the preferential growth of a rod-shaped cell wall which maintains the diameter of the cell. Therefore, the pitch of the MreB cytoskeletal filament determines diameter. Mutation or chemical-inhibition of MreB filaments result in cell shape defects (reviewed in (166)).

1.7.3) DNA Partitioning

The segregation of replicated DNA into daughter cells is one of the most fundamental processes in biology. There are three known mechanisms of DNA segregation in bacteria (recently reviewed in (51)): The first involves actin-like ATPases such as ParM, which push plasmids apart by forming dynamically unstable filaments (69). The second involves tubulin-like GTPases, such as TubZ, which

move DNA by the rapid treadmilling of cytoskeletal filaments (80). The third involves cytoskeletal P-Loop ATPases (ParAs).

ParA proteins are defined by a variant Walker A box ATPase motif (1). ParAs are commonly found on plasmid DNA and are transcribed with their partner proteins, ParBs. ParAs form dimers that bind to ParBs, which in turn bind to specific centromeric DNA sequences, such as *pars* (109), which are found on plasmids and in multiple positions of several bacterial genomes. ParAs form filaments in the presence of non-specific sequences of DNA, and contact with bound ParBs promotes ATP hydrolysis by ParAs and depolymerisation of the ParA filaments along DNA molecules. Rapid detachment and reattachment of ParB-bound DNA to the depolymerising end of a ParA filament results in a "pulling" force which drives DNA segregation (Figure 28). The rapid depolymerisation of ParAs has been observed to result in oscillation of ParA protein from one end of the cell to the other, and this shuttling is thought to be responsible for even segregation of plasmid and DNA components (131).

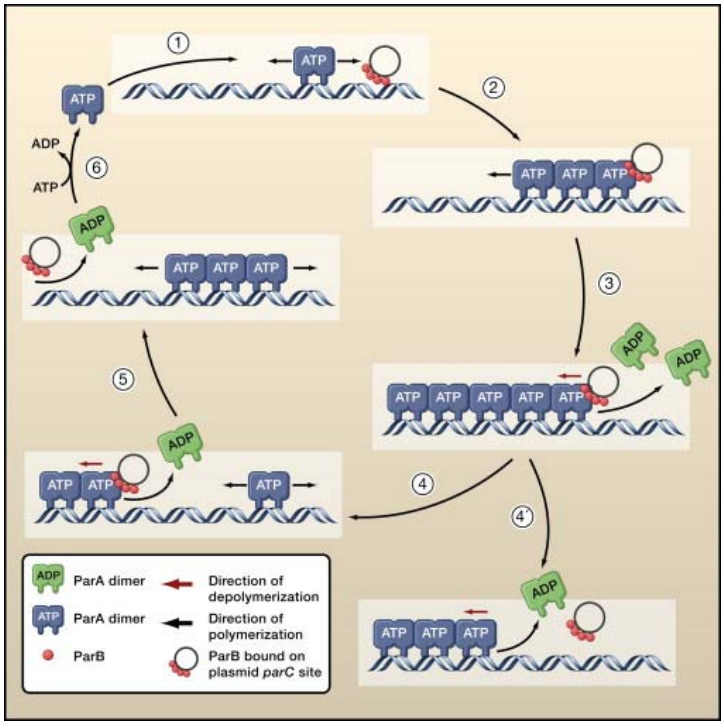


Figure 28: The ParA Pulling Mechanism. ParA-ATP dimers bind cooperatively to nucleoid DNA, leading to the formation of ParA filaments. **1)** Formation of filaments begins with a nucleating core. **2)** Subsequently, a growing filament contacts a plasmid via ParB. **3)** ParBs stimulate the ATPase activity of ParA-ATP at the end of the filament. **4)** ParA-ATP is converted to its ADP form and released from the DNA, leaving a new ParA-ATP filament end accessible. For each depolymerisation event, **4')** the plasmid can either detach or **4)** remain attached to the end of the depolymerising ParA filament. **5)** The plasmid is moved. **6)** The ParA-ATP subunit is released by ParB and free ParA-ADP is rejuvenated to ParA-ATP and the cycle repeats (taken from (51)) .

1.7.4) Whole genome analysis

Given that our knowledge of the organised nature of the subcellular environment is increasing rapidly with improving fluorescence microscopy techniques, one interesting question is; how many bacterial proteins display subcellular localisation patterns? A recent study in *C.crescentus* used high-throughput mutagenesis techniques to generate fluorescent fusions for every protein encoded in its genome. This was coupled with high-throughput microscopy and image analysis which allowed the subcellular positioning of each protein to be determined. Remarkably, 10 % of all proteins expressed from the genome of *C.crescentus* were seen to form at discrete subcellular positions, and analysis of the positions at which protein clusters could be seen showed that the poles and the mid-cell were "hot-spots" for protein localisation (Figure 29) (159).

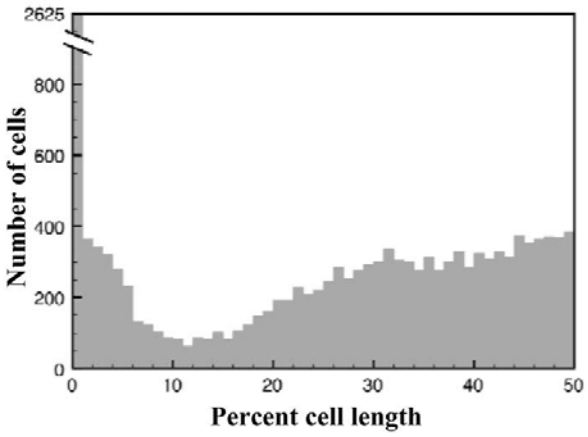


Figure 29: Quantitative analysis of fluorescent protein localization in *C.crescentus*. A histogram of the location of the fluorescent peak in individual cells, broken into 1% bins, shows enrichment for polar and mid-cell localisation and a region, between ≈5% and 20% of the cell length, where protein clusters are is relatively rarely found (taken from (159)).

1.8) The model organism *Rhodobacter sphaeroides*

Rhodobacter sphaeroides belongs to the α-subgroup of proteobacteria, it is a metabolically diverse bacterium which can grow photosynthetically as well as in the dark, can fix nitrogen and carbon dioxide and is capable of emitting hydrogen (84). It is a rod-shaped bacterium with similar size and morphology to *E.coli*, but divides more slowly than *E.coli* with an optimal doubling time of approximately one hour (personal communication, Dr M.Kojadinovic). The inner membrane of *R.sphaeroides* is known to form complex invaginated structures when grown photosynthetically,

which accommodate large numbers of photosynthetic complexes (78). *R.sphaeroides* is a model organism for the study of bacterial photosynthesis, motility and chemotaxis and is being developed in Korea for its potential in producing photosynthetic biofuels (84).

1.8.1) Motility

R.sphaeroides is monotrichous and swims with a single subpolar flagellum. The *R.sphaeroides* flagellum can only rotate in one direction, CCW, to produce smooth swimming. Tumbles are achieved in *R.sphaeroides* by arrest of the motor. Motor braking results in the same conformational changes of flagellin that are observed in other species, and *R.sphaeroides* filaments can be observed in both semi-rigid helical arrangements and relaxed, coiled arrangements when visualised using fluorescence staining (Figure 30). When the filament is relaxed, *R.sphaeroides* cells reorient via Brownian motion (112).

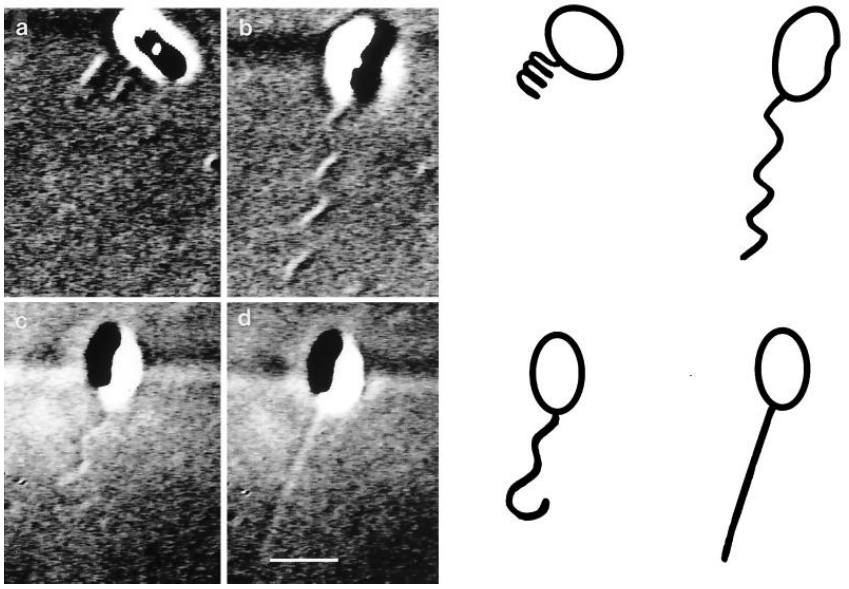


Figure 30: DIC images of the flagellar filament of *R.sphaeroides*. Various flagellar conformations: **(a)** a coiled form, **(b)** a functional helix, **(c)** a helix relaxing into a coil, and **(d)** an apparently straight conformation. The cartoon shows the likely cell size and the flagellar shape (taken from (142)).

Interestingly, *R.sphaeroides* possesses two full sets of flagellar genes that are spread across its two circular chromosomes. Only one set of flagellar genes is known to be expressed under laboratory and physiological conditions. However, variants which express the second set of flagellar genes were isolated via the forced evolution of motility knock-out strains. Characterisation of these isolates revealed a polar, tufted flagellar which is similar to that of *Vibrio fischeri* (Figure 31) (120).

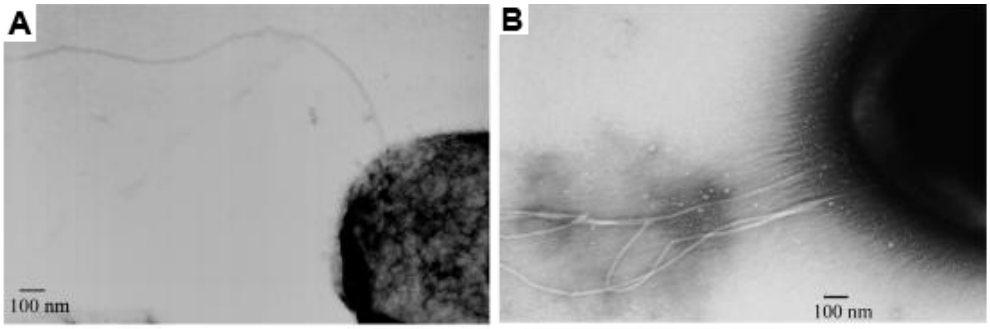


Figure 31: Electron microscopy images of the two different types of flagella assembled by *R.sphaeroides*. **A)** Representative WS8N cell, expressing the first, subpolar set of flagellar genes. **B)** Representative isolated mutant, expressing the second polar, tufted set of flagellar genes (taken from (120)).

1.8.2) Flagellar Assembly

R.sphaeroides has a four-tier hierarchy of flagellar assembly, which is similar to that of *P.aeruginosa*. The class I master regulator of *R.sphaeroides* is FleQ, which by itself is only capable of binding the promoter of a single class II operon, which contains FleT and six other genes that are involved in formation of the HBB (122). FleT is a bEBP which lacks the helix-turn-helix motif of other bEBPs; it has been shown to bind FleQ resulting in the formation of the FleQ/FleT complex. The FleQ/FleT complex promotes the expression of class III genes, which encode the remainder of the components of the HBB, and the stator proteins. Computational analysis of known class III promoters allowed identification of the FleQ/FleT consensus sequence, which was successfully used to predict other class III regulated gene loci (Figure 32) (116).



Figure 32: The FleQ/FleT consensus sequence. **A)** The FleQ/FleT consensus sequence determined by computational analysis of the known promoter sequences of the genes listed in **(Bi)**. **Bi)** The first genes of gene clusters known to be class III regulated, *flgB*, *fliO* and *flgG* operons contain flagellar HBB genes. The *motA* operon encodes the stator complex. The *flgA* operon encodes genes involved in the T3SS and includes the *flgM* gene. **Bii)** Gene clusters that were predicted, then confirmed to be class III regulated by the determination of the consensus sequence in **(A)** (adapted from (116)).

Class IV flagellar gene expression is controlled by the activity of FliA, which is expressed from the weak class III promoter of the *flhA* operon (116). Interestingly, the σ factor FliA and its anti-σ factor, FlgM are encoded in juxtaposed divergent operons, which are regulated in a manner which is thought to involve DNA bending and the preferential alignment of UASs and DASs that have been identified for the individual operons (Figure 33), although the mechanism and the reason for this rare form of bacterial transcriptional regulation is not fully understood.

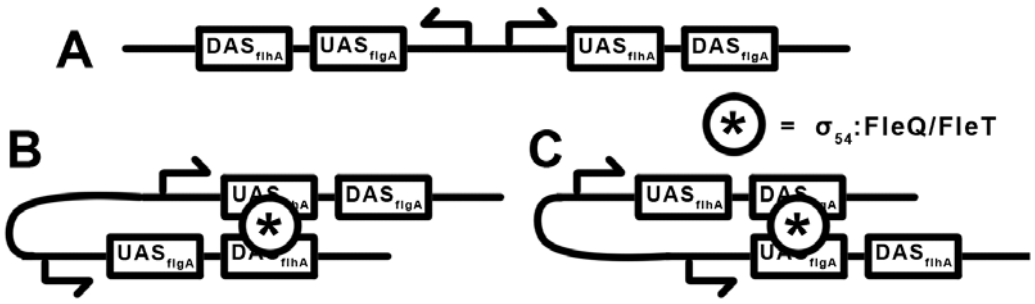


Figure 33: The divergent *flgA* and *flhA* operons. **A)** UASs and DASs specific to the *flhA* and *flgA* operons encoding FliA and FlgM, respectively, are encoded either side of the class III promoter sequences of each operon. **B)** DNA-bending arrangement thought to promote *flhA* transcription. C) DNA-bending arrangement thought to promote *flgA* transcription.

To date, experimental analyses of the activity of FliA promoters of *R.sphaeroides* has demonstrated that the *fliC*, *fliS*, *fliD* and *flgK* operons (122) as well as the second chemotaxis-gene operon (*che*Op2) (98) are regulated by class IV promoters. The *flgM* gene is suggested to possess its own class IV promoter sequence (8), but the experimental evidence for this is unpublished. No experimental analysis of the inhibitory affect of FlgM on FliA has been presented in *R.sphaeroides*.

A summary of the flagellar hierarchy of *R.sphaeroides* is presented in Figure 34.

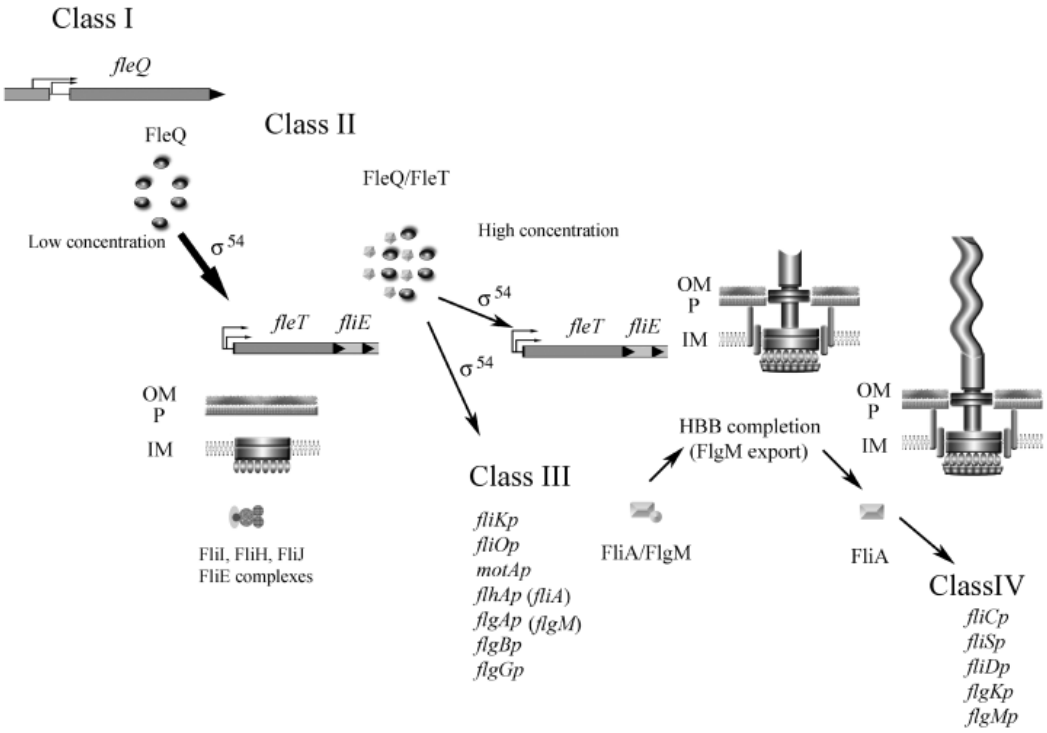


Figure 34: Schematic representation of the regulatory steps in the biogenesis of the *R.sphaeroides* flagellum. IM, inner membrane; OM, outer membrane; P, peptidoglycan layer (taken from (122)).

1.8.3) Chemotaxis

The chemotaxis pathway of *R.sphaeroides* is a step up in complexity from that of *E.coli*. Three main operons and several other genomic loci encode homologues of the *E.coli* chemotaxis genes (Figure 35) (56). Only two of the three main operons (*che*Op2 and *che*Op3) are essential for chemotaxis under laboratory conditions, *che*Op1 components have been shown to interact with components of the second flagellar system when expressed *in vitro* (120).

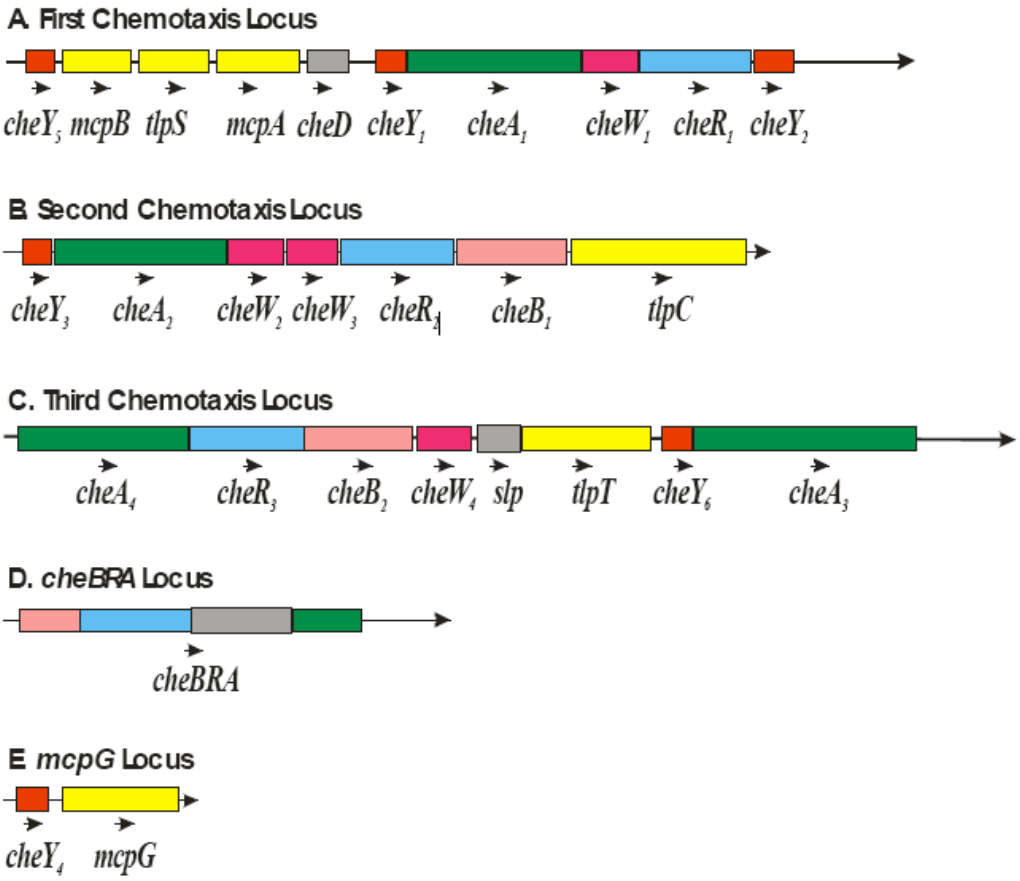


Figure 35: The chemotaxis loci of *R.sphaeroides*. Genes encoded in **A)** cheOp1, **B)** cheOp2, **C)** cheOp3 (*slp* has since been renamed *ppfA*), **D)** cheBRA and **E)** the mcpG locus. In addition to the receptors encoded by these loci, there are an additional seven chemoreceptors at unlinked loci.

The components of the two major chemotaxis operons have been shown to localise to distinct subcellular regions; *che*Op2 components are mainly polar, whereas *che*Op3 components are found at a discrete position in the cytoplasm of the cell (154, 155) (discussed later).

The chemotaxis signalling pathway of *R.sphaeroides* is also complex. Polar *che*Op2-encoded components form a similar two-component system to that of *E.coli*, transmembrane receptors sense chemical concentrations in the cytoplasm, and transmit this information to $CheA_2$ via the adaptor proteins $CheW_2$ and $CheW_3$. Adaptation is thought to occur at the poles via *che*Op2 encoded proteins $CheR_2$ and $CheB_1$. The polar histidine kinase, $CheA_2$, is known to transfer phosphoryl groups to $CheB_1$ and three of the six CheY homologues that are present in *R.sphaeroides*, $CheY_3$, $CheY_4$ and $CheY_6$. $CheY_6$ is known to be essential for motor stopping. While the exact roles of $CheY_3$ and $CheY_4$ are unknown, one or both of these proteins are required for chemotaxis.

The cytoplasmic clusters formed by *che*Op3-encoded components have homologues of all of these proteins. Transducer Like Proteins (TLPs) are cytoplasmic homologues of the polar receptor

proteins and lack the transmembrane domain of the MCPs. TLPs are presumed to have the same function as MCPS, but it is not known what ligands they bind (if any), nor how this signal is communicated to the cytoplasmic histidine kinases, $CheA_3$ and $CheA_4$. $CheA_4$ forms a homodimer in the cytoplasmic cluster, and how it interacts with $CheA_3$ is unclear. Together the cytoplasmic CheA proteins function as both a kinase and a phosphatase, replacing the role of *E.coli* CheZ – which is absent from the *R.sphaeroides* chemotaxis proteins (123). $CheA_3$ only transfers phosphoryl groups to components from *che*Op3, $CheB_2$ and $CheY_6$. The role of the cytoplasmic cluster is not fully understood, but it is hypothesised that it acts to sense the metabolic state of the cells, and somehow alters chemotaxis in response to the availability of different nutrients (Figure 36) (124, 145).

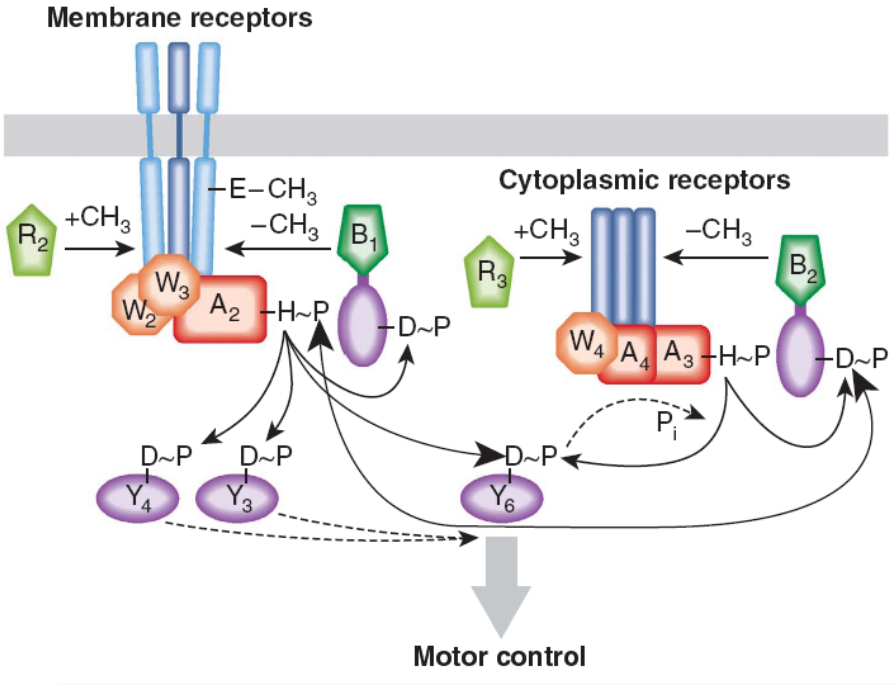


Figure 36: Chemotactic complexity in *R.sphaeroides*. The model of signalling in *R.sphaeroides*. *che*Op2 encoded components localise to the poles and detect extracellular chemical concentrations, whereas *che*Op3 encoded components localise to the cytoplasm and detect the metabolic state of the cell (taken from (145)).

The deconstruction of the chemotaxis pathway of *R.sphaeroides* has proved a difficult task, as deletion of any individual component of this pathway results in a loss of chemotaxis. Also, mutations that result in changes in chemotaxis gene expression, or in disruption of cluster structure, abolish chemotaxis, showing that the structure and function of these components are intimately linked and that the dynamics of phosphotransfer must be tightly controlled to generate successful chemotactic signalling.

1.8.4) Chemoreceptor positioning

As previously mentioned, the chemotaxis proteins of *R.sphaeroides* form clusters at discrete cytoplasmic and polar regions within the cell, and the positioning of these proteins is a requirement for chemotaxis (Figure 37) (154-156).

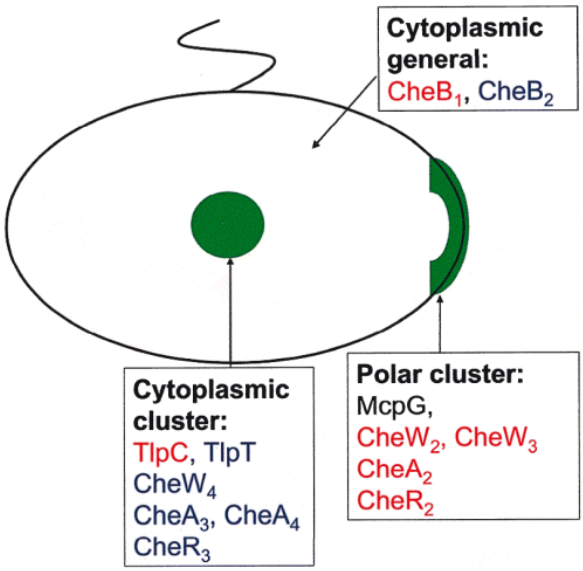


Figure 37. Known localization of chemotaxis proteins in *R.sphaeroides*. Proteins from *che*Op_2 are shown in red and those from *che*Op_3 in blue.

The polar chemoreceptor arrays of *R.sphaeroides* have not been studied in the same detail as those of *E.coli* (see section 1.6); however, epifluorescence microscopy of different fluorescent-fusion proteins has shown which protein components display polar localisation (154, 156): McpG, McpH, McpJ, $CheA_2$, $CheW_2$, $CheW_3$ and $CheR_2$. Also, electron cryo-tomography of the polar chemoreceptor cluster of *R.sphaeroides* has shown the same hexagonal packing that is observed in other bacteria (see section 1.6) and predicts that there will be approximately 2,200 receptor molecules in the average polar cluster (17). Laterally positioned clusters have not previously been observed in *R.sphaeroides*.

Epifluorescent studies of the cytoplasmic chemotaxis cluster have shown that it contains a number of chemotaxis proteins (156): TlpC, TlpT, $CheW_4$, $CheA_3$, $CheA_4$ and $CheR_3$. The cytoplasmic chemotaxis clusters have been shown to demonstrate interesting localisation patterns. In cells with a single cytoplasmic cluster, this cluster is located at the mid-point of the cell and in cells with two cytoplasmic clusters, they are positioned at the ¼ and ¾ positions of the cell (Figure 38) (150).

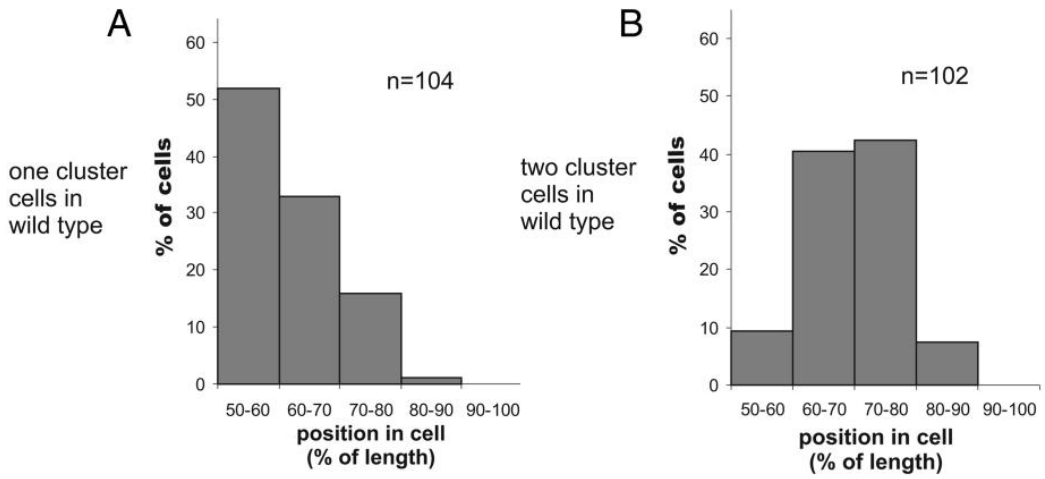


Figure 38: Positioning of the cytoplasmic protein clusters. Histograms show the distance of clusters from the farthest cell pole relative to the cell length. Data are symmetrically distributed around mid-cell. ***A)*** Wild-type cells with one cluster found around mid-cell. ***B)*** Wild-type cells with two clusters localized to the 1/4 and 3/4 positions.

The partitioning and positioning of the cytoplasmic clusters has been shown to be dependent on the activity of a ParA homologue, Protein Partitioning Factor A (PpfA), which is encoded with the chemotaxis genes of *che*Op3

1.8.4.1) Protein Partitioning Factor A

PpfA is an ATP driven enzyme that interacts with the cytoplasmic cluster and facilitates cluster partitioning and localisation in *R.sphaeroides*. PpfA shows strong sequence homology to ParA (see section 1.7.3), and recent advances in our understanding of PpfA show that it most likely forms DNA-binding interactions at its N-terminus (Dr Mark Roberts, personal communication). Study of the YFP-fused cytoplasmic chemotaxis proteins TlpT and $CheW_4$ by epifluorescence have shown that deletion mutants lacking PpfA only ever form a single cytoplasmic cluster, and that this cluster is positioned randomly within the cell (150).The same phenomenon was observed in filamentous cells, generated by treatment with cephalexin, where cytoplasmic chemotaxis clusters were seen to form at seemingly regular loci along the cell length in the presence of PpfA, but most cells formed only one cluster after PpfA deletion (Figure 39).

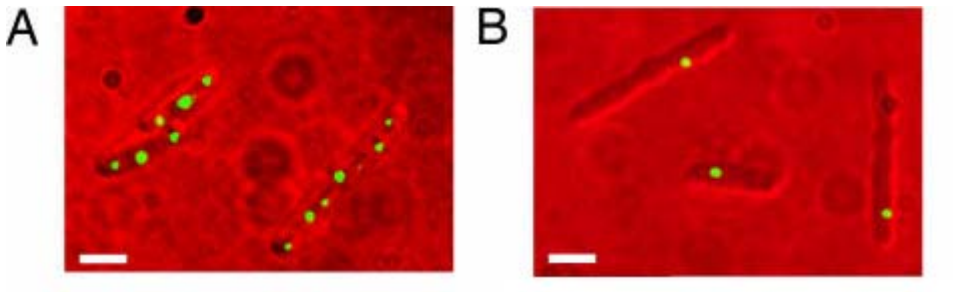


Figure 39: Fluorescence images of the cytoplasmic cluster position in filamentous cells. **A)** Wild-type background, clusters appear at "regular" positions along the cell length. **B)** Δ*ppfA,* only single clusters are observed in cells. Scale bars represent 2 µm.

The exact mechanism that PpfA uses to achieve cluster partitioning is unknown, and the attempts to purify active PpfA in order to obtain a better understanding of its function have been unsuccessful.

1.9) Aims

This project aims to study components of the chemotaxis and motility systems of *R.sphaeroides*, and how they are regulated by the activity of FliA and FlgM.

Previous work in other species has shown that altering FliA and FlgM expression levels can achieve varying expression levels of chemotaxis proteins. It will be investigated here whether or not the same is possible in *R.sphaeroides*, and what can be deduced about chemotaxis protein clustering with varying expression levels. Further analysis will investigate the roles of FliA and FlgM in the regulation of motility, to deduce whether the FlgM secretion checkpoint exists in *R.sphaeroides*, and, if so, how it ties into the flagellar assembly hierarchy.

2) Materials and Methods

2.1) Strains and Plasmids:

All strains used in this study are listed below (Table 2).

Strain	Species	Genotype/Description	Source
DH5α	*E.coli*	General purpose cloning strain. Allows for blue/white selection.	Gibco BRL
XL1-BLUE	*E.coli*	General purpose cloning strain. Tetracyclin resistant.	Stratagene
M15[pREP4]	*E.coli*	Expression strain, allows tight expression from IPTG inducible plasmids.	QIAGEN
S17-1 λpir	*E.coli*	pK18*mobsacB* mobilising strain used for conjugation. Streptomycin resistant.	(117)
WS8N	*R.sphaeroides*	Spontaneous naladixic acid resistant strain of the wild-type WS8	(142)
JPA467	*R.sphaeroides*	*ΔfliA*	(98)
JPA500	*R.sphaeroides*	*mcpG-gfp*	(156)
JPA527	*R.sphaeroides*	*Δ$cheW_3$*	(155)
JPA1215	*R.sphaeroides*	Δ$cheY_{1\text{-}5}$ $CheY_6$D57N Non-motile control strain.	(155)
JPA1301	*R.sphaeroides*	*Δche*Op3 Non-chemotactic control strain.	(155)
JPA1328	*R.sphaeroides*	*ΔppfA*	(150)
JPA1399	*R.sphaeroides*	*ΔcheW4*	(155)
JPA1418	*R.sphaeroides*	*yfp-$cheW_3$*	(156)
JPA1444	*R.sphaeroides*	*yfp-$cheB_2$*	(156)
JPA1447	*R.sphaeroides*	*cfp-$cheW_3$, $cheW_4$-yfp*	(156)
JPA1457	*R.sphaeroides*	*$cheW_4$-yfp*	(156)

JPA1629	*R.sphaeroides*	*yfp-cheY$_3$*	(156)
JPA1638	*R.sphaeroides*	*yfp-fliM*	Roberts and Wadhams Unpublished.
JPA2001	*R.sphaeroides*	*ΔflgM*	This Study
JPA2002	*R.sphaeroides*	*cfp-cheW$_3$, cheW$_4$-yfp, ΔflgM*	This Study
JPA2003	*R.sphaeroides*	*cfp-cheW$_3$, cheW$_4$-yfp, ΔppfA*	This Study
JPA2004	*R.sphaeroides*	*cfp-cheW$_3$, cheW$_4$-yfp, ΔflgM, ΔppfA*	This Study
JPA2006	*R.sphaeroides*	*yfp-motB1*	This Study
JPA2018	*R.sphaeroides*	*yfp-cheW$_3$, ΔflgM*	This Study
JPA2019	*R.sphaeroides*	*cheW$_4$-yfp, ΔflgM*	This Study
JPA2022	*R.sphaeroides*	*yfp-cheW$_3$, ΔfliA*	This Study
JPA2023	*R.sphaeroides*	*cheW$_4$-yfp, ΔfliA*	This Study
JPA2032	*R.sphaeroides*	*yfp-fliM, ΔfliA*	This Study
JPA2033	*R.sphaeroides*	*yfp-fliM, ΔflgM*	This Study
JPA2034	*R.sphaeroides*	*yfp-flgM*	This Study
JPA2035	*R.sphaeroides*	*flgM-yfp*	This Study
JPA2038	*R.sphaeroides*	*yfp-motB1, ΔflgM*	This Study
JPA2039	*R.sphaeroides*	*yfp-motB1, ΔfliA*	This Study
JPA2050	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA467 Derivative.	This Study
JPA2051	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA1638 Derivative.	This Study
JPA2052	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA2032 Derivative.	This Study
JPA2053	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA2004 Derivative.	This Study
JPA2054	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA1418 Derivative.	This Study

JPA2055	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA2019 Derivative.	This Study
JPA2056	*R.sphaeroides*	Spontaneous long-cell isolated mutant. JPA500 Derivative.	This Study

Table 2: Strains list.

All plasmids used in this study are listed below (Table 3).

Name	Description	Source
p*eyfp*	An ampicillin resistance plasmid that contains the *eyfp* gene sequence.	ClonTech
pREP4	A tetramycin resistance plasmid that contains the *lacI* and *lacQ* genes, and improves leaky expression in *E.coli* strains.	QIAGEN
pInd4	A kanamycin resistance *R.sphaeroides* IPTG inducible expression plasmid with MCS to allow the creation of C-terminally 6 x his-tagged proteins	(65)
pInd4:*fliA*	pInd4 containing the in-frame *fliA* gene for expression analysis.	This Study
pInd4:*flgM*	pInd4 containing the in-frame *flgM* gene for expression analysis.	This Study
pInd4:*yfp*	pInd4 containing the in-frame *yfp* gene for expression analysis and fluorescence marking.	Dr M. Roberts (unpublished)
pInd4:*cfp*	pInd4 containing the in-frame *cfp* gene for expression analysis and fluorescence marking.	Dr M. Roberts (unpublished)
pQE60	An ampicillin resistance plasmid with MCS to allow the creation of C-terminally 6 x his-tagged proteins	QIAGEN
pQE60:*cheW*$_3$	pQE60 containing the in-frame 6 x his-tagged *cheW*$_3$ gene for protein purification.	(125)

pQE60:*cheW*$_4$	pQE60 containing the in-frame 6 x his-tagged *cheW*$_4$ gene for protein purification.	(125)
pQE60:*motB*$_C$	pQE60 containing the in-frame 6 x his-tagged cytoplasmic fragment of the *motB* gene for protein purification.	This Study
pQE80	An ampicillin resistance plasmid with MCS to allow the creation of N-terminally 6 x his-tagged proteins	QIAGEN
pQE80:*fliA*	pQE60 containing the in-frame 6 x his-tagged *fliA* gene for protein purification.	This Study
pQE80:*flgM*	pQE60 containing the in-frame 6 x his-tagged *flgM* gene for protein purification.	This Study
pK18*mobsacB*	A kanamycin resistance suicide plasmid that allows for blue/white selection and sucrose sensitive negative selection.	(117)
pK18*mobsacB:ΔfliA*	pK18*mobsacB* containing the flanking regions of the *fliA* gene, with *fliA* deleted.	This Study
pK18*mobsacB:ΔflgM*	pK18*mobsacB* containing the flanking regions of the *flgM* gene, with *flgM* deleted.	This Study
pK18*mobsacB:yfp-flgM*	pK18*mobsacB* containing the flanking regions of the *flgM* gene, with the N-terminal *yfp-flgM* fusion.	This Study
pK18*mobsacB:flgM-yfp*	pK18*mobsacB* containing the flanking regions of the *flgM* gene, with the C-terminal *flgM-yfp* fusion.	This Study
pK18*mobsacB:yfp-motB*	pK18*mobsacB* containing the flanking regions of the *motB* gene, with the N-terminal *yfp-motB* fusion.	This Study
pK18*mobsacB:ΔppfA*	pK18*mobsacB* containing the flanking regions of the *ppfA* gene, with *ppfA* deleted.	(150)

Table 3: Plasmids list

2.2) Antibiotic concentrations

Antibiotic	Concentration for *R.sphaeroides* (μg ml^{-1})	Concentration for *E.coli* (μg ml^{-1})
Naladixic Acid	25	-
Kanamycin	25	25
Ampicillin	-	100

Table 4: Antibiotic concentrations used for different bacterial species.

2.3) Growth Conditions

2.3.1) *E.coli* liquid cultures

Strains of *E.coli* were grown in Luria Broth medium (appendix 9.2.1) with shaking, at 37 °C. Antibiotics were added when appropriate.

2.3.2) *E.coli* agar plates

E.coli colonies were grown on solidified Luria Broth containing 2 % agar, containing appropriate antibiotics. Inoculated plates were incubated overnight at 37 °C.

2.3.3) *R.sphaeroides* liquid cultures

Strains of *R.sphaeroides* were grown in Succinate medium (appendix 9.2.2). Stationary-phase starter cultures were grown at 30 °C, photoheterotrophically in a light cabinet under white, non-fluorescent light of an intensity 50 μM m^{-2} s^{-1}. Microaerobic cultures were grown at 30 °C with shaking. Antibiotics were added when appropriate; however, when cells were grown for microscopy, naladixic acid was not added as it induces cellular autoagglutination.

2.3.4) *R.sphaeroides* agar plates

R.sphaeroides colonies were grown on solidified Luria Broth containing 2 % agar and naladixic acid. Kanamycin was also added when appropriate. Inoculated plates were incubated for at least 48 hours at 30 °C.

2.3.5) Bacterial preservation

Stationary-phase cultures of *E.coli* and *R.sphaeroides* strains were preserved in a 3:2 mixture of succinate medium and sterile glycerol (50 % v/v). Cryotubes (Starstedt) containing 1.5 ml of this mixture were flash frozen in liquid nitrogen and placed at -80 °C.

2.4) Genetic Techniques

2.4.1) Polymerase chain reaction (PCR)

PCR reactions for amplification of a section of genomic of plasmid DNA were carried out in a final volume of 50 µl of liquid. Reaction mixtures contained 1 µl of the appropriate DNA template, as well as appropriate primers at a final concentration of 100 µM, 5 µl of 2.5 mM dNTPs and the appropriate quantity of 5 x *Pfu* polymerase buffer (Promega) and Pfu polymerase enzyme (Promega), as specified by the manufacturers' instructions. For overlap extension PCR, 1 µl of each overlapping DNA fragment was added in the place of the template DNA. The remaining volume was made up with sterile Milli-Q water and PCR conditions were varied by the addition of up to 10 % (v/v) DMSO.

A thermal-cycler was used to vary the temperature of the PCR reaction mixture in order to achieve denaturation (98 °C), primer annealing (different temperatures) and polymerase elongation (72 °C) for each PCR cycle. An initial denaturation of 10 minutes was followed by 30 PCR cycles and a final elongation of 5 minutes. Elongation times were varied based on the base-pair length of the expected product of the PCR reaction.

2.4.2) Agarose gel electrophoresis

0.8 % to 2 % Agarose (Roche) gels were prepared in ½ x TBE (9.2.4). 5 x DNA loading Dye containing RNAse (appendix 9.2.3) was added to each DNA sample before loading it into the gel. Electrophoresis was performed at a constant voltage between 80 V and 150 V. DNA within gels was stained with ethidium bromide solution (1 ng ml^{-1}) for 20 minutes and visualised on a long wavelength UV transilluminator.

2.4.3) Gel Extraction

DNA bands were excised from agarose gels using a razor blade. The QIAquick Gel Extraction Kit (QIAGEN) was used according to the manufacturer's protocol to obtain purified DNA, which was eluted in 20-50 µl of the provided elution buffer, or in sterile Milli-Q water.

2.4.4) Mini/Midiprep

Plasmid DNA was extracted from overnight stationary-phase cultures of E.coli (DH5α) cells using the QIAprep Spin Miniprep kit (QIAGEN), or the HiSpeed Plasmid Midi Kit (QIAGEN) according to the manufacturer's protocol. Purified plasmid DNA was eluted in appropriate quantities of the provided elution buffer, or in sterile Milli-Q water.

2.4.5) *R.sphaeroides* genomic DNA extraction

Genomic DNA was extracted from stationary-phase starter cultures of *R.sphaeroides* using the technique detailed in (54). Cells were pelleted via centrifugation at 13,000 x g for one minute and the growth medium discarded. The cell pellet was flash frozen in liquid nitrogen and immediately resuspended in 0.5 ml of chromoprep lysis buffer (appendix 9.2.5) which had been preheated to 65 °C. 0.1 mg of proteinase K (Sigma) was added, and the mixture was incubated at 42 °C for a minimum of 2 hours. DNA was then extracted twice in 0.5 ml phenol:chloroform:isoamyl-alcohol (25:24:1), and precipitated in 1 ml ethanol (100 %) at -20 °C. The precipitated DNA was pelleted by centrifugation at 13,000 x g for 30 minutes, washed in 1 ml ethanol (70 %) and re-pelleted. The wash solution was discarded and the pellet was dried overnight at 37 °C before being resuspended in 50 µl of sterile Milli-Q water. Extracted DNA was stored at 4 °C.

2.4.6) Restriction digests

Purified DNA samples were digested in volumes between 20 µl and 50 µl, the appropriate 10x buffer and enzyme quantity were used, as specified by the manufacturer's protocols. Reaction mixtures were incubated at 37 °C for a minimum of 2 hours for plasmid DNA, a minimum of 4 hours for PCR fragment DNA and overnight for genomic DNA.

2.4.7) Dephosphorylation

1 µl of Calf Intestinal alkaline Phosphatase (New England Biolabs) enzyme was added to each DNA sample in the appropriate buffer, as specified by the manufacturer's protocol. The mixture was incubated for a minimum of 1 hour at 37 °C.

2.4.8) DNA Ligation

Digested and purified DNA fragments and plasmids were mixed in a 3:1 ratio. Appropriate quantities of 5 x DNA ligase buffer (New England Biolabs), T4 DNA ligase enzyme and purified MilliQ water were added in the mix, as specified by the manufacturer's protocol. Ligation mixtures were incubated overnight at between 16 °C and 20 °C.

2.4.9) Competent cells

E.coli strains were grown to OD_{600} 0.4 - 0.6, and harvested by gentle centrifugation at 4 °C. Cells were resuspended in 15 ml of TFB1 (appendix 9.2.6) and left on ice for 30 minutes. Cells were then spun at 2,000 x g for 15 minutes at 4 °C, and resuspended in 2 ml of TFB2 (appendix 9.2.7). These cells were then flash frozen and stored at -80 °C.

2.4.10) Chemical transformation

Competent *E.coli* cells were defrosted on ice for 20 minutes. 1 µl of purified plasmid DNA, or 25 µl ligation mixture was mixed with the defrosted cells, and transformation of the DNA was induces by heat shock at 42 °C for 1 minute. 1 ml of Luria Broth was added to heat-shocked cells, and antibiotic resistance genes were left to express for 1 hour at 37 °C before cells were spread onto agar plates containing appropriate antibiotics.

2.4.11) DNA sequencing

Plasmid DNA was sequenced at the in-house GeneService DNA sequencing facility on an AB1377 Sequencer (Applied Biosystems) with the use of BigDye™ dye terminators (PE Biosystems). DNA sequences were aligned using the Staden PreGap Linux plug-in and inspected in the Clone Manager 9 software package.

2.4.12) Conjugation

Plasmid DNA cannot be transformed into *R.sphaeroides*, as *R.sphaeroides* cannot be made competent. Introduction of plasmid encoded genetic information is achieved via conjugation, which involves the direct transfer of DNA from one bacterial species to another via pili (117). Plasmids pK18*mobsacB* and pInd4 have been designed to allow for bacterial conjugation. *E.coli* S17-1λpir cells that contained the plasmid to be conjugated were grown to an early exponential phase of growth and mixed gently in a ratio of 1:10 with stationary-phase *R.sphaeroides* cells. Cell mixtures were pipetted onto a 0.2 µm nitrocellulose filter that had been placed on a 2 % LB agar plate. Cells were

incubated overnight, resuspended and spread onto LB agar plates containing kanamycin and naladixic acid.

2.4.13) *R.sphaeroides* chromosomal replacement

Chromosomal replacement is achieved in *R.sphaeroides* by the use of the pK18*mobsacB* plasmid. This plasmid is known as a "suicide plasmid", and lacks the origin of replication required for plasmid DNA replication. Thus, in order for the kanamycin antibiotic resistance of pK18*mobsacB* to be conferred onto colonies of *R.sphaeroides* cells, the plasmid DNA must be incorporated into the genome via recombination. Recombination events occur between homologous DNA sequences, therefore incorporation of repeats of sections of the genome into pK18*mobsacB* results in the directed incorporation of the plasmid DNA into the genome after a single round of selection by growth in the presence of kanamycin. Any mutations within the pK18*mobsacB*-inserted DNA sequence, such as insertions, deletions or point mutations are incorporated into the genome in this fashion.

After the first round of selection and recombination, plasmid-encoded genes and two copies of the target DNA sequence are found within the genome – one wild-type and one mutant. A second round of selection allows for removal of the plasmid DNA and the wild-type DNA sequence by growth in the presence of sucrose, which is metabolised into a toxic product by the *sacB*-encoded gene from pK18*mobsacB*.

Cells were grown overnight in the absence of antibiotics and spread onto M22 (appendix 9.2.8), 10 % sucrose, 2 % agar plates. Cells that grew on sucrose plates had undergone the second round of recombination. Single colonies were replica-plated onto naladixic acid-containing and kanamycin and naladixic-acid containing plates in order to test for the loss of the plasmid-encoded kanamycin resistance gene. The resulting colonies could either contain wild-type chromosomal DNA or mutant DNA, and mutant strains were distinguished from wild-type via chromosomal DNA extraction (chapter (2.4.5) and southern blot (chapter (2.4.14)).

2.4.14) Southern blotting

10 µg of extracted chromosomal DNA was digested overnight with an appropriate restriction enzyme and run by electrophoresis on an agarose gel. DNA within gels was stained with ethidium bromide solution (1 ng ml^{-1}) for 20 minutes and visualised on a long wavelength UV transilluminator. DNA was denatured and neutralised before being capillary blotted onto a Hybond-N membrane as

detailed in (136). Blots were probed with an appropriate DNA probe using the dioxygenin (DIG) Nucleic Acid Detection Kit (Roche) according to the manufacturers' instructions.

2.4.15) DIG labelling

Southern Blot probes were prepared from suitable fragments of DNA extracted from midiprep preparations of plasmid DNA (chapter (2.4.4)). The DIG DNA labelling kit was used to randomly prime these DNA sequences as per the manufacturers' instructions.

2.5) Protein Techniques

2.5.1) Protein purification

E.coli cells containing expression plasmids were grown in 2YT media (appendix 9.2.9) containing appropriate antibiotics to an OD_{600} of 0.8 before induction of by the addition of 100 µM IPTG. After a variable period of induction of a number of hours, cells were harvested by centrifugation at 6,000 xg for 15 minutes. Cell pellets were resuspended in protein wash buffer (appendix 9.2.10) and lysed by sonication (chapter (2.5.2)). Cell debris was removed from cell lysates by a 20 minute centrifugation at 35,000 xg and filtration through a 0.4 µm filter. Cleared cell lysates were added to a Nickel-NTA agarose drip column (QIAGEN) which had been equilibrated with protein wash buffer. Following column-binding nickel columns were washed with a minimum of 60 column volumes of protein wash buffer. Protein was eluted with protein elution buffer (appendix 9.2.11).

2.5.2) Sonication

Resuspended cell pellets were sonicated on ice using a Vibracell sonicator (Sonics and Materials Incorporated). Six 20 second full power bursts, separated by 20 second pauses, were used to achieve efficient cell lysis.

2.5.3) Bradford assay

Protein concentrations were calculated by a Bradford-based assay (BioRad). Standard curves were generated using BSA protein standards (BioRad).

2.5.4) Acrylamide electrophoresis

12 to 15 % acrylamide SDS-PAGE gels were prepared using standard acrylamide gel mixtures and polymerised by the addition of TEMED and Ammonium-per-Sulphate (APS). 5 x protein loading dye

(appendix 9.2.12) was added to each protein sample before loading it into the gel. Electrophoresis was performed in SDS-running buffer (appendix 9.2.13) at a constant voltage of 140 – 200 mV.

2.5.5) Coomassie staining

The protein content of acrylamide gels was stained for visualisation using Coomassie blue stain as detailed in (136).

2.5.6) Western Blotting

Cells were grown microaerobically to an OD_{700} of 0.5 – 0.7. The fraction 1 ml corresponding to 1 ml of cells at OD_{700} 0.6 was harvested via centrifugation at 13,000 xg. Any excess liquid was removed by careful pipetting. Cells were vigorously resuspended in appropriate volumes of 5x protein loading dye and all samples were boiled before being loaded and run on acrylamide gels.

Protein from acrylamide gels was transferred onto a 0.2 µm nitrocellulose PVDF membrane (BioRad) using the BioRad Mini-Protean II transfer kit. Membranes were blocked overnight at room temperature in 5 % dried milk dissolved in PBS, then washed for 2 x 10 minutes in PBS. Membranes were then incubated in 1 % dried milk in PBS containing a suitable dilution of primary antibody before washing in 0.2 % Tween in PBS for 2 x 10 minutes and PBS for 2 x 10 minutes. Membranes were then incubated in 1 % dried milk in PBS containing a suitable dilution of secondary horse-radish peroxidise (HRP) conjugated antibody. Millipore Immobilon reagents were used to visualise bands in a light-emitting reaction that was observed using a Syngene G:Box gel-documentation system. For quantitative western-blot analysis, an appropriate range of purified protein standards was run on each gel. Band intensities were quantified using the Syngene GeneTools software package.

2.6) Phenotype Analysis

2.6.1) Growth curves

Unless otherwise stated, *R.sphaeroides* growth rate was assessed by the growth of strains in a Fluostar Optima plate reader. Stationary-phase photoheterotrophic cultures were diluted by a factor of 100 in succinate medium. Cells were incubated at 30 °C with intermittent shaking. Absorption at wavelength of 700 nm was measured every 20 minutes over 30 hours using a combination of appropriate optical paths and filters.

2.6.2) Soft-agar swim plates

Soft-agar allows bacteria to swim using flagella. As the bacteria grow on a soft-agar plate, they metabolise the carbon sources in the surrounding medium, generating a local gradient of chemoattractant concentration. This gradient is sensed by the chemotaxis machinery of the bacteria, and cells that have better chemotactic efficiency will swim further along this gradient. The distance travelled by cells is, thus, a measure of chemotactic efficiency. However, the soft-agar swim plate assay is also sensitive to changes in growth rate, metabolism and motility.

Chemotactic efficiency was assessed on soft-agar swim plates. 0.2 % BioTek agar in M22 media was autoclaved and left to set in 140 mm petri dishes at 30 °C overnight. Propionate was added at a final concentration of 100 µM as the only carbon source in the medium. Antibiotics and IPTG were also added where appropriate. 5 µl of overnight stationary-phase photoheterotrophic culture was inoculated gently onto the surface of the soft-agar plate, and plates were incubated at 30 °C for 48 hours. Swim diameters were measured using a ruler and back-lighting to aid visualisation.

2.6.3) Fluorimetry

R.sphaeroides cells were grown to an OD_{700} of 0.6-0.7. For each sample, the fraction of cells corresponding to 200 µl, 1 ml, 2 ml and 4 ml of cells at OD_{700} of 0.6 was collected by bench-top centrifugation at 13,000 xg. Excess medium was removed via careful pipetting, and the resulting pellet was resuspended in 200 µl of fresh succinate medium. Resuspended cells were transferred to a black, clear-bottomed Corning Costar 96-well plate and a fluorescence reading was taken using appropriate YFP filters on a FLUOstar Optima fluorescence plate reader with a gain setting of 2200. A cellular autofluorescence baseline was calculated using WS8N cells that did not contain YFP protein, and subtracted from the measured values. For protein copy number quantification, fluorescence signal was compared to that obtained from known quantities of purified YFP (provided by George Wadhams), resuspended in the same number of mid-log WS8N cells (further discussed in chapter (5.6)).

2.6.4) Biofilm formation

Long-edged glass coverslips were placed into 15ml of succinate medium in a sterile 50 ml falcon tube, leaving approximately 0.5 cm of glass exposed to the air, while the remaining surface was submerged. The growth medium was inoculated with equal numbers of stationary-phase cells containing YFP and CFP reporter plasmids and incubated at approximately 30 °C in a light cabinet under white, non-fluorescent light of an intensity 50 µM m^{-2} s^{-1} for 48 hours. Coverslips were

carefully removed from the medium and one side was cleaned with an ethanol-soaked tissue before incorporating the coverslip into a tunnel-slide using double-sided sticky tape. M22 minimal media was added to the tunnel slide, and the submerged, unclean surface was imaged using a Nikon A1R confocal microscope.

2.6.5) Negative staining

Mid-log phase cell culture of OD_{700} ~0.6 was gently pelleted at 3,000 xg in a desktop centrifuge and resuspended in fresh succinate medium. The washed cell culture was applied to a copper grid coated with a thin carbon film. Grids were blotted with filter paper and negatively stained with a 2 % solution of uranyl acetate, then washed with two drops of water to remove excess stain and air-dried.

2.7) Microscopy

2.7.1) Tethered cell analysis

10 µl of mid-log phase cells were immobilised on a round glass coverslip using the α-FliC antibody (142). This coverslip was incorporated into a custom-made flow cell and minimal M22 medium containing 100µM propionate, a known chemoattractant, was passed over the sample. Cells were visualised under a NIKON Optiphot-150 phase contrast microscope and images were recorded using a DALSA Genie HM640 black and white camera at 100 frames per second. To test the response to chemoattractant, medium with no propionate was passed through the flow cell and the change in speed of rotation of the cells was observed. A stop in rotation upon removal of propionate corresponds to a wild-type response, whereas no response is associated with cells that cannot undergo wild-type chemotaxis.

Tethered cell videos were analysed using custom in-house software that had been developed in Python by Dr Mila Kojadinovic.

2.7.2) Phototaxis assay

Mid-log phase photoheterotrophically grown cells were placed under a coverslip on a standard microscope slide, and visualised under a NIKON Optiphot-150 phase contrast microscope. The light-path of the microscope was obscured briefly, and upon restoration of the image the behaviour of swimming cells was noted. Wild-type populations that have been grown photoheterotrophically pause in swimming in response to the removal of light, whereas mutants with defects in chemotaxis do not necessarily stop.

2.7.3) Epifluorescence microscopy

The system used for the acquisition of epifluorescence images of *R.sphaeroides* cells consisted of an inverted Nikon Eclipse TE200 microscope, with Differential Interference Contrast (DIC) enhancement, a mercury lamp with multiple filter combinations optimised for YFP, GFP, CFP, DAPI and FITC fluorescent imaging (among others). Images are acquired using a Hamamatsu ORCA-ER CCD camera, and the Simple PCI 6 software package.

2.7.4) Confocal microscopy

The system used for the acquisition of confocal images of *R.sphaeroides* biofilms consisted of an inverted Nikon A1R confocal microscope, with DIC objectives and a laser system for the illumination of the sample. The most common microscope setup used a 40 x DIC objective, and YFP and CFP lasers to sequentially acquire images of the different fluorophores in the sample. Suitable light-filters were used to exclude any non-fluorescent light, and light signal was detected using a photon-multiplier tube, therefore images were composite scans. The microscope setup was controlled by the Nikon NIS-Elements software package, and image files were stored in the Nikon .nd2 file format.

2.7.5) Electron microscopy

Images were acquired on a CCD camera using a Philips CM120 electron microscope (courtesy of Dr Catherine Venien-Bryan) equipped with a LaB6 filament (FEI) operating at 120 kV under low dose conditions with different magnifications.

2.8) Image analysis

2.8.1) Epifluorescence image analysis

Data on cellular morphologies and fluorescent protein clustering was extracted from epifluorescence images using custom in-house MATLAB-encoded software that was developed as part of this study. The software is discussed in detail in chapter (3), and the MATLAB code for this software is presented in appendix (9.3).

2.8.2) Tethered cell analysis

In-house software developed in Python by Dr Mila Kojadinovic was used to obtain a measure of cell rotation within digital videos of tethered cells (chapter (2.7.1)). A one-step user-defined threshold was used to define cell objects within each video frame, and regions of interest containing tethered cells were also user-defined. The centroid coordinate of the cell was determined for each frame of the video, and rotation speed was calculated relative to the mean position of the cell over all frames.

2.9) Bioinformatics:

2.9.1) DNA sequencing fragment alignment

DNA sequences were aligned using the Staden PreGap Linux plug-in which also generated a consensus sequence, which was used to inspect sequenced DNA for point mutations.

2.9.2) Multiple protein sequence alignment

Multiple protein sequence alignments were achieved and rendered using the free online tools ClustalW and UniProt.

3) The analysis of epifluorescence images of *R.sphaeroides* in MATLAB

3.1) Epifluorescent images of *R.sphaeroides*

In order to measure fluorescent protein cluster patterns in *R.sphaeroides*, custom software was developed that assisted in the analysis of the sorts of images that could be obtained using the available optical setup. The optical setup used for the acquisition of epifluorescence images is described previously (chapter 2.7.3). *R.sphaeroides* cellular imaging is routinely carried out with either 2 or 3 separate light channels, which correspond to different filter and lighting combinations. The images captured on a 12-bit monochrome ORCA –ER camera have their values linearly scaled (multiplied by 16) to 16-bit data sets that are saved in standard RGB uncompressed '.tif' format, with each channel of the image containing separate light-channel data. For example, a standard image may contain a DIC light image, a YFP-fluorescence image and a CFP-fluorescence image, which will be saved into the Red, Green and Blue channels of the .tif image, respectively (Figure 40).

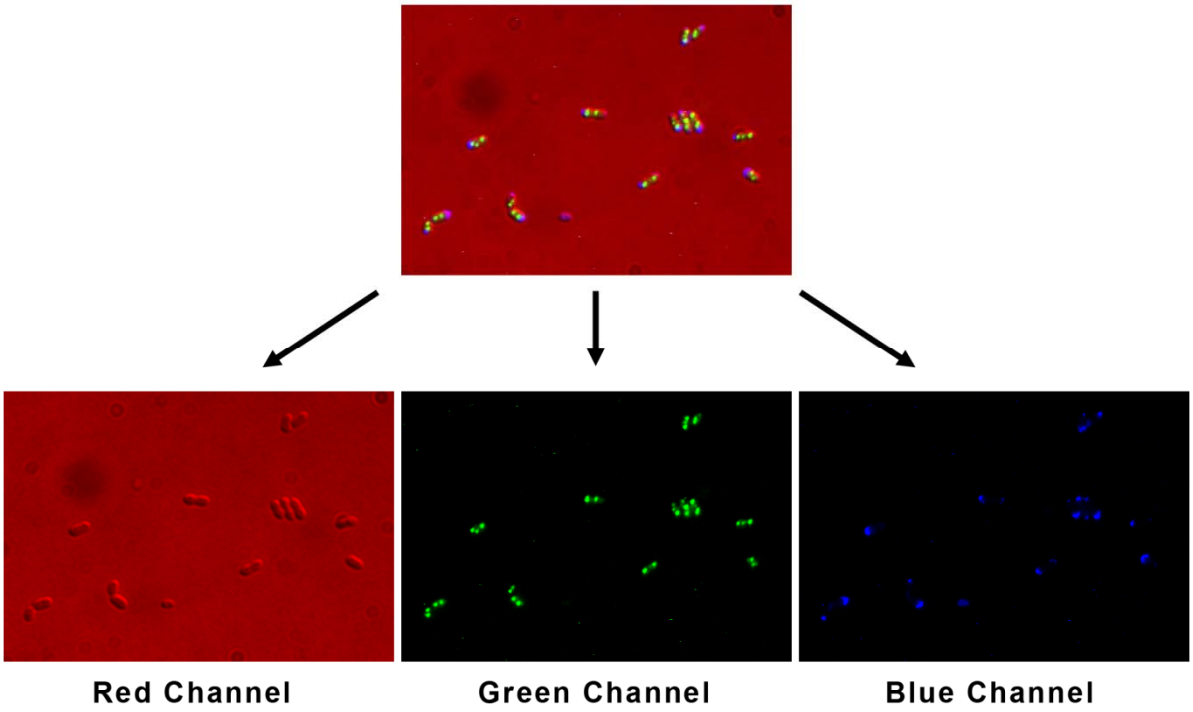


Figure 40: Example RGB image – An example image of *R.sphaeroides* cells expressing both YFP-labelled and CFP-labelled proteins. A DIC light image is saved in the red channel of the image, the YFP data is saved in the green channel of the image and the CFP data is saved in the blue channel of the image. Red, green and blue are false-colour representations of what is actually 16-bit greyscale data.

Although *R.sphaeroides* cells demonstrate interesting and complex subcellular protein localisation patterns, to date the analysis of images acquired in this manner has been achieved in a fairly rudimentary way, by one of two methods:

3.2) Mean grey level analysis

In order to address the question of whether or not protein within bacterial cells is clustered or diffuse, and to compare differences in clustering between different mutant strains, mean grey level analysis has been used (141, 155, 156). Images acquired on the same day, using consistent microscope settings are subjected to thresholding of the colour channel containing the data acquired for the fluorophore of interest. A threshold value is chosen based upon the maximum value obtained for images of wild-type, non fluorescent cells; the value is chosen by the user within the Simple PCI software package with the assistance of a mask-overlay which displays the areas of the image which fall above and below the chosen value. Fluorescence values that fall above the threshold value within each subsequent image are deemed to be the product of emitted fluorescence from a fluorophore of interest, and "objects" are defined within each image as adjacent pixels that fall above this threshold. For each object, a mean value is calculated. In order for this analysis to produce interpretable results, the threshold value chosen must be set deliberately high so that only bright clusters are defined as objects in cells containing clustered protein, and "whole cell" regions are defined in cells containing diffuse protein (Figure 41). In this way, cells that have obvious clusters give a high mean value for fluorescence, and cells with no obvious clustering give a lower mean value (assuming that the concentration of protein in these cells remains constant).

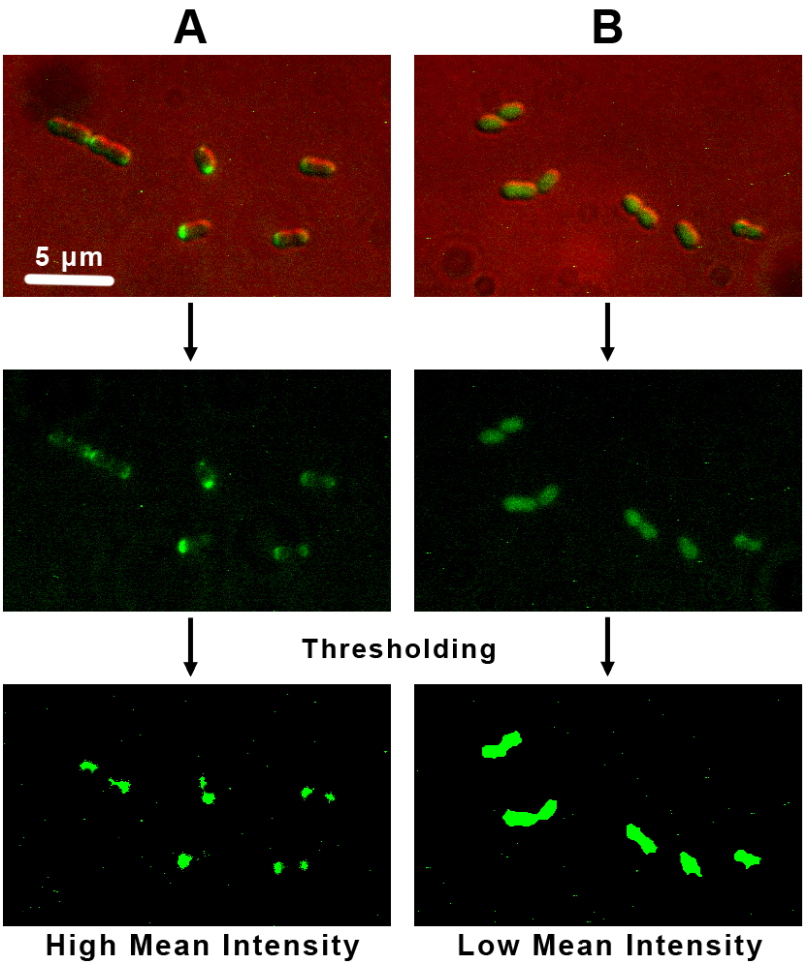


Figure 41: Example of mean grey analysis. Fluorescence intensity data is extracted from each image, in this instance the green channel of the image. A single thresholding step defines "objects" within these images and the mean intensities of each object are calculated. For **A)** clustered protein the mean intensity obtained will be greater than for **B)** diffuse protein.

The mean values obtained via this analysis do not represent the mean values of fluorescence for the cells within these images, as this would leave no way of distinguishing between diffuse and clustered protein for cells with the same concentration of protein. The value chosen for thresholding, although consistent from image to image is therefore relatively arbitrary and can be "subjectively optimised" by the user to give the best results for each data set.

3.3) Lengths, widths and positions analysis

Questions concerning cell morphology, cell lengths, widths and the positions of objects such as fluorescent protein clusters and cellular septation sites have been addressed in *R.sphaeroides* using the Simple PCI measurement tool (141, 150). The Simple PCI software package offers a tool for

measurement between two user defined positions within an image; a line is drawn between the two positions and the distance between the objects is returned as a value in pixels, which is then converted to a measurement in microns using a scale factor which is calculated from an image of a 100 µm graticule. When a user wants to measure more than one distance within the same cell, the measurement process must be repeated, relying on an accurate second definition of the same reference site for measurement. As *R.sphaeroides* cells are rod-shaped, the position of objects within cells is most commonly measured relative to the long axis of the cell; the positions measured rely on accurate by-eye assignment of this long axis which, again, must be repeated from object to object. In addition, when an object lies off-centre within a cell, the position at which a line drawn orthogonal to the long axis would intercept with the long axis is estimated by-eye by the user.

As well as this technique being subject to user-error, it is also particularly time-consuming, meaning that the size of data sets analysed has been limited to not much more than about 100 cells within a population and for a small number of different strains (141, 150). Also, the simultaneous assignment of positions for multiple objects (i.e./ septation sites, cytoplasmic clusters and polar clusters), has only been attempted in a small number of cases (141).

3.4) Why was it done in this way?

The reasons behind why the analysis of *R.sphaeroides* images has been undertaken in this manner are mainly classical; the Hamamatsu ORCA-ER camera, which is used for the image acquisition on the Nikon Eclipse TE200 microscope, is designed to integrate with the Simple PCI software, which also offers simple analysis tools (described above). For convenience, this software has been used for both image acquisition and analysis.

Also, to date, the questions being asked about protein clusters and clustering within *R.sphaeroides* have been mainly based on positioning, and the intensity information of the image has been ignored meaning that no more complex system of analysis was required.

3.5) Existing image analysis software

The ability to extract data from images is a frequent problem faced in many scientific disciplines, many different approaches have been used to analyse fluorescent cluster positions within bacteria. The variety of previous work has led to the development of different software packages for image analysis, each designed specifically for the measurement of a particular trait (Table 5).

Software Package	Description	Reference
ImageJ	Free, online Java-based image analysis software, with various open-source plug-in functions.	http://rsbweb.nih.gov/ij/
MATLAB	Mathworks software designed for mathematical simulation and data handling. The image processing Toolkit provides many functions which are specifically for image processing.	http://www.mathworks.co.uk/
CellProfiler	Freely available online software designed for the analysis and tracking of biological images. Object definition is exclusively achieved from thresholding of phase contrast or fluorescence data sets.	http://www.cellprofiler.org/
Other options	Any programming language can be used for the processing of image data, as images are simply 2D matrices of numbers. For example, the analysis of tethered cell videos in this project is achieved using custom software programmed in Python (see chapter 2.8.2).	

Example	Software Package	Reference
Z-stack analysis of chemotaxis protein cluster positions in *E.coli* and the integrated values of fluorescence within each cluster. Cell positions are defined from phase contrast images.	ImageJ	(147, 148)
High-throughput cell segmentation and cluster positioning in *C.crescentus* cells, cell outlines are defined from phase contrast images and peak-finding algorithms position clusters within each cell.	MATLAB	(159)
Cluster positioning analysis of asymmetrically distributed protein clusters In *E.coli*. Cells are defined from phase contrast images.	ImageJ	(162)
Tracking of the position of protein aggregates within lineages of growing cells within surface-growing monolayers of *E.coli*, and the effect of protein aggregation on the rate of cell growth. Cell outlines are defined from phase contrast images.	MATLAB	(85)

Table 5: List of software packages that were available for image analysis at the start of this project. **Top)** Commonly used software packages for image analysis. **Bottom)** Four examples of image analysis used in the investigation of protein clustering in bacteria.

The first main step in the analysis of cells within an image, and thus the first step in analysis using existing software, is the separation of objects within the image from background. Of the existing programs that were available for use, those that were suitable for the analysis of bacterial cells were based around cell definition of a phase-contrast image. More generalised cell-defining software packages, such as those found in the plug-in functions of the general image analysis tool ImageJ, were not easily integrated for simultaneous cell definition and cluster position analysis (as

described below). Therefore, none of the available software at the start of this project was able to be integrated with the analysis of our image data sets.

The main reason that existing software packages were unable to be used for image analysis was the use of DIC contrast enhancement to image the cells themselves. In order to reliably define cell boundaries within images of biological samples which are predominantly water based and hence optically similar to the surrounding aqueous media, optical contrast enhancement techniques are used which allow visualisation of the cells without the use of stains, which are potentially harmful to a living cell preparation. Phase contrast and DIC are the two most commonly used methods for enhancing the contrast of the light image, they work based on different optical principles and have mutually exclusive optical set ups.

3.6) Phase Contrast

Light that passes through any material other than a vacuum has its amplitude and phase altered by interaction with that material. The use of light-excluding rings within the condenser and objective of a microscope allows light that has an altered phase relative to that of light passing through an empty sample to have its path blocked. For a biological sample, light that passes through a cell undergoes phase alterations that are different to those of the light that does not pass through a cell, this light is blocked from the camera and the cell appears dark on the image, around the edges of the cell (where the sample is thinner), not all light is obscured by the phase-rings and a light halo is seen around each object (Figure 42).

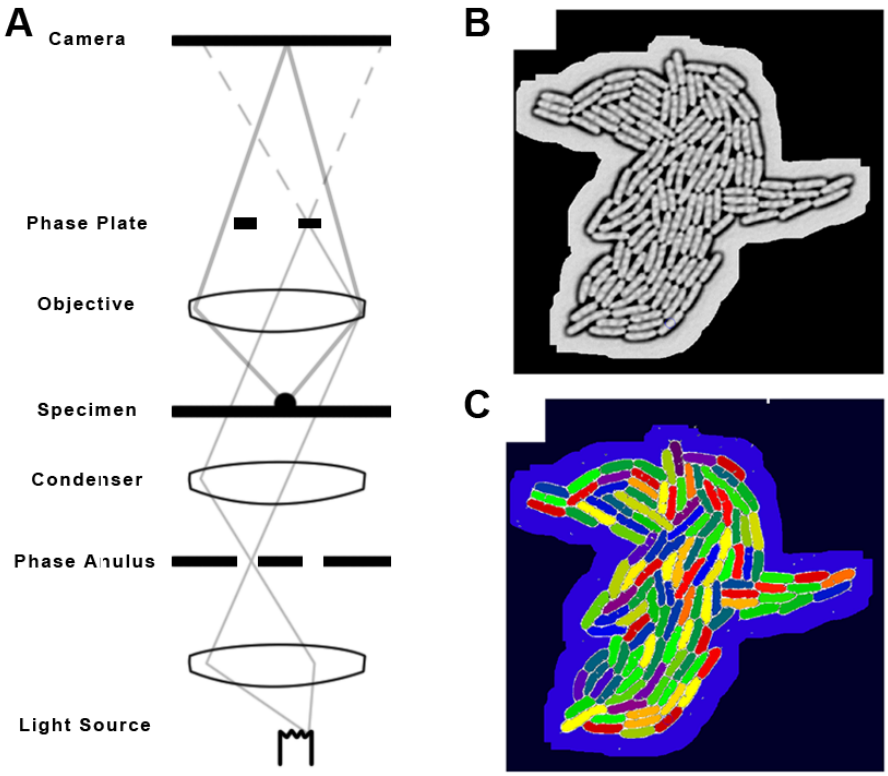


Figure 42: Phase contrast setup and object separation. **A)** Schematic diagram of the light path through a phase contrast enhanced microscope. Light passes through a phase annulus before passing through the specimen, after being focused by the objective, the phase plate blocks the path of light that has an unaltered phase. **B)** Example phase contrast image, a monolayer of *E.coli* cells is imaged here. NB. This image has been background-adjusted by the subtraction of a neutral-field image (a second image containing no cells), and therefore cells appear light with dark halos. **C)** Cell segmentation of the image from **(B)** achieved by a one-step threshold algorithm, random colours are used to represent individual cells (**B** and **C** adapted from (85)).

As cells appear dark on a light background, the computational procedure for the identification of those cells within an image is straight forward. A simple one-step threshold-based identification of objects is sufficient to obtain crisp outlines of cells within an image, with careful parameter selection this is even achievable for touching cells within a monolayer – as has been demonstrated (Figure 42 and (85)).

3.7) Differential interference contrast

With DIC, polarised light is spatially separated using a Quartz-based Wollaston prism, effectively splitting a single light beam into parallel beams separated in space (Figure 43A). This light is passed through the liquid sample, and subjected to phase and amplitude changes determined by what material the light passes through. Before being focused onto the image plane, the light is recombined using a second Wollaston prism, which has the opposite effect to the first Wollaston prism. At an object, where one edge of the split light beam passes through the sample and the other

does not, reintegration of the light beams will lead to the constructive interference of light, whereas at the opposite edge reintegration causes destructive interference. A second polariser set at 90 ° to the first blocks all light which has not had its polarisation altered by passage through the sample. The resultant image contains a pseudo-3D object where one side is light, and the other side is dark. The light-dark orientation is determined by the orientation of the Wollaston prism, and edges that lie parallel to this orientation appear grey in the resultant image, with no enhanced contrast (Figure 43B).

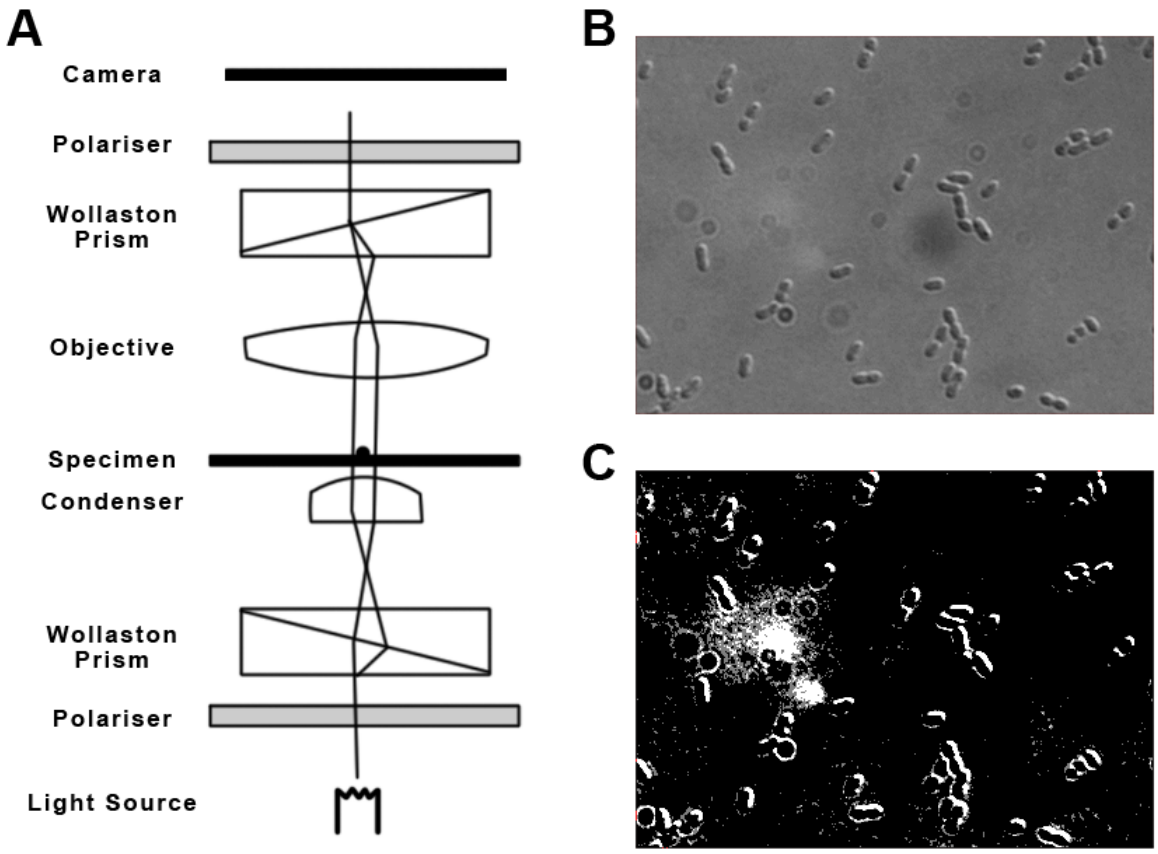


Figure 43: DIC setup and object separation. . **A)** Schematic diagram of the light path through a DIC enhanced microscope. Light passes through a polariser and Wollaston prism before passing through the specimen. Light is then focused by the objective and passed through a second Wollaston prism and polariser before reaching the camera. **B)** Example DIC image of a typical field of view of *R.sphaeroides* cells. Cells have a pseudo-3D appearance. **C)** Cell segmentation of the image from **(B)** achieved by a one-step threshold algorithm. One-step thresholding is insufficient to obtain accurate definitions of cells, and imaging imperfections are picked-out as objects.

The challenge for computational analysis of DIC images is a result of the uneven, pseudo-three dimensional appearance of the each object within the image as well as the appearance of objects of

the same shape and size differing with their relative orientations. For these images the simple single-step thresholding approach used on phase contrast images does not work. Various methods have been employed to automatically define DIC-enhanced objects within images, with varying levels of complexity (70). However, as only the edges of each object have enhanced contrast, each method relies on inferring the position of the object based on its edges. The results of object definition by these methods show varying levels of accuracy, and are very sensitive to image imperfections caused by factors such as dust on optical surfaces (Figure 43C).

The reason for choosing DIC over phase contrast for a microscope system comes from the principle of light-exclusion in phase contrast. The phase ring of a phase contrast objective acts as a physical barrier to light, meaning that less of the emitted light from a fluorophore will be detected at the camera. Although the image-analysis of phase contrast images is easier, the sensitivity of detection is reduced and therefore the minimum number of fluorophores needed to produce a measurable signal above background will be higher for phase contrast. All high resolution biological microscopy techniques rely on the detection of weak signals, and simultaneous contrast enhancement and detection of weak signals (down to single fluorescent molecules) becomes achievable when DIC is used to enhance contrast.

No existing software packages were found that offered a quick or simple method for performing analysis of fluorescent cluster numbers, positions and intensities for cells defined from a DIC image. Therefore, custom software was designed using the MATLAB programming package with a license for the Image Analysis Toolkit.

3.8) Defining cells within images

Cell sizes, division site positions, fluorescent protein cluster positions and fluorescence intensity measurements were all deemed relevant to the analysis of cluster positioning within *R.sphaeroides*. Groups of cells and overlapping cells were excluded from the analysis, as where cells touch each other. Cluster marking, cell edge definition and the assignment of fluorescent signal to a particular cell were not possible. Cells that were included in the analysis were spatially separate from other cells within the field of view and had a single, straight long axis where the overall shape of the cell could be described by two semi-circular ends with a straight central section, or “rod-like” (Figure 44). The septation and division of cells is an important part of the bacterial cell-cycle and a variety of proteins have been shown to localise to existing and future sites of cell division (91, 147). Thus, for the majority of the analysis included in this work it was decided to include all cells that had a maximum of one visible septation site and that maintained a straight rod shape.

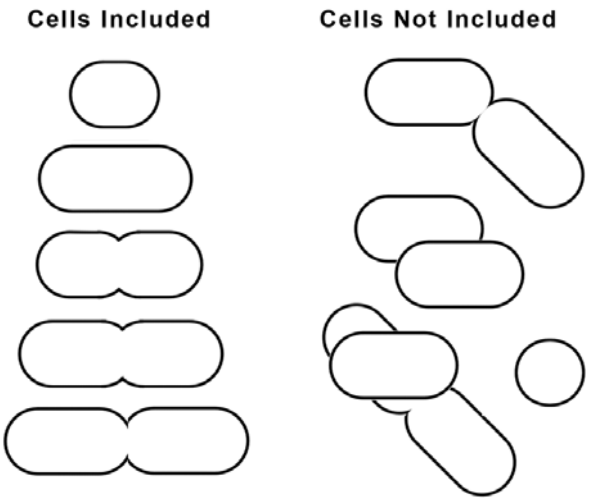


Figure 44: Cells included in the analysis – A schematic representation of the types of cells which were included, or not included in the image analysis. All cells that were included in the analysis could be described as "rod-like", having semicircular ends and a straight central section. Overlapping, circular or post-division cells were not included.

The inaccuracies of DIC-based cell detection combined with the added complication of wanting to accurately position the septation sites meant that the process of cell definition was not able to be fully automated. In order for the analysis software to remain effective, manual input from the user was designed to be kept to a minimum and to remain as simple as possible. To achieve this, definition of the outline of each cell was performed in two separate stages:

3.8.1) Stage 1 - Whole-image object detection

An interface allowed object detection from the DIC image, the green-channel fluorescence image or the blue-channel fluorescence image. As with previous mean grey level analysis, a single user defined threshold could be applied to the green or blue channels of the image, which generated a black and white image mask. Where the whole area of the cell gave detectable signal above background, relatively accurate cell outlines could be determined using this single step. However, where subcellular protein localisation was uneven, or when fluorescence signal was absent or of comparable values to background noise, the outlines determined by this method were inaccurate. In these instances, a simple edge-finding algorithm could be applied to the DIC (Red) channel of the image; this detected cell positions but gave numerous false-positives and relatively inaccurate estimates of cell position. The masks generated from any of the three channels could be summed in different combinations in order to give the best first approximation of cell position (Figure 45).

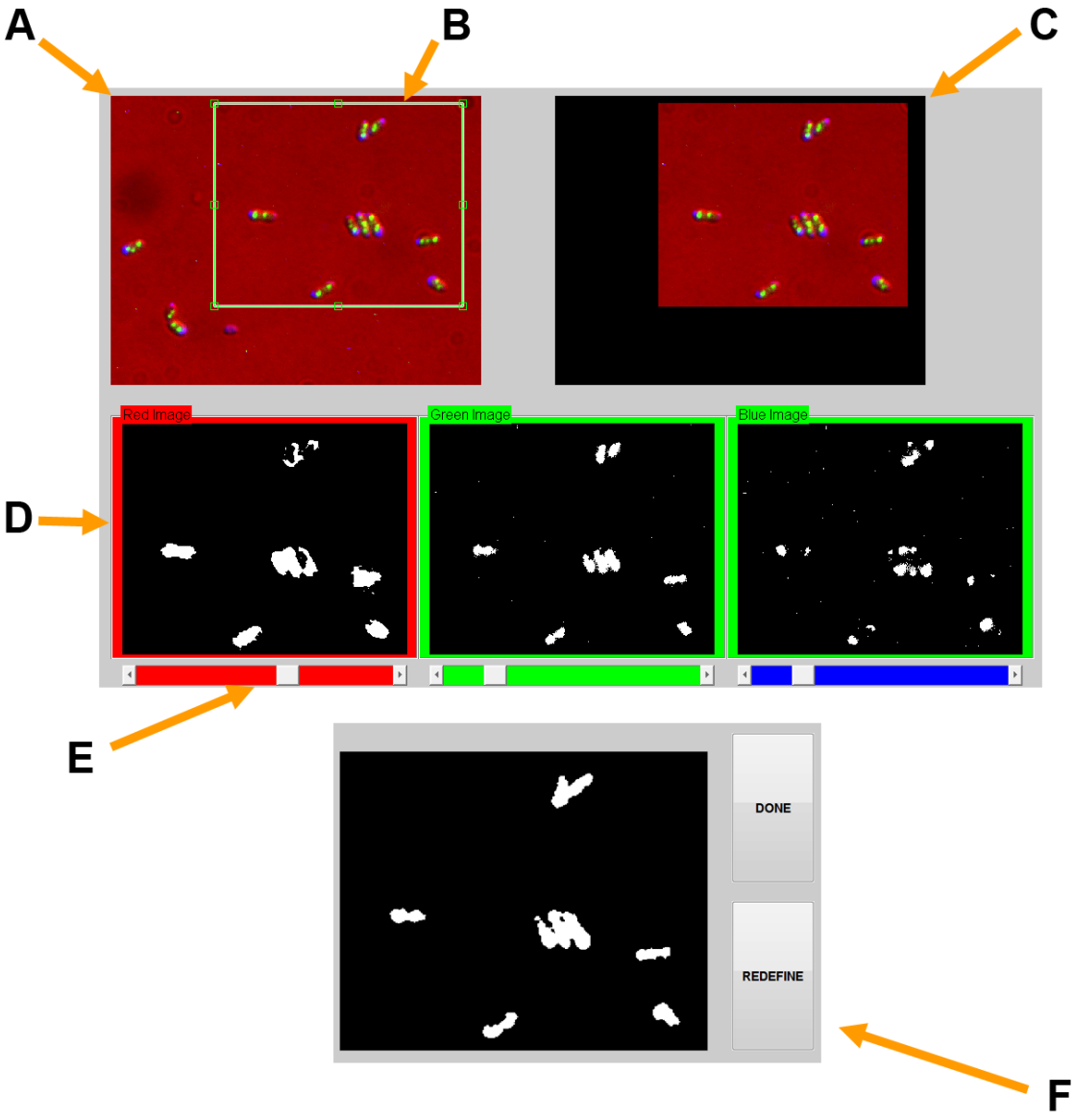


Figure 45: Stage one user interface – Screen shots of the user interface for whole-image object definition, or "stage one" of image processing. **A)** The original image is displayed here. **B)** A green box allows the image to be cropped in order to exclude regions of the image that are visibly out of focus, or contain no useful data. **C)** The cropped region of the image is displayed here. **D)** Black and white mask images generated using simple object definition algorithms are displayed for the red, green and blue channel images. The masks generated can be selected or deselected for use in generating the final mask by clicking on them. **E)** Slider bars for adjustments of the parameters used to generate the red, green and blue masked images; slider values correspond to threshold values for the green and blue channels, but represent a contrast threshold for the red channel. **F)** The final black and white mask generated by summation of the masks in **(D)**.

3.8.2) Stage 2a – Cell edge refinement

For each object defined within an image at stage 1, a second interface allowed an accurate redefinition of the outline of that object, or where the detected object was not to be included in the analysis the object could be discarded. An initial estimate of the central position of the object was made based on the mask defined in stage 1, the user then clicked at one end of the cell, and a rod-shaped object was positioned over the cell based on these two positions. The defined position of the centre of the cell could be easily adjusted by a key-based method until the outline of the cell was deemed satisfactory by the user. Cell width could also be adjusted, and for cells with visible septation sites, the site of division could also be defined (Figure 46). The rate at which cells could be positioned using this approach was determined by the accuracy of the initial estimate of cell position, and therefore the use of fluorescence data (where available) alongside the DIC image to obtain better estimates of cell shape and position at stage 1 increased the rate at which images could be processed.

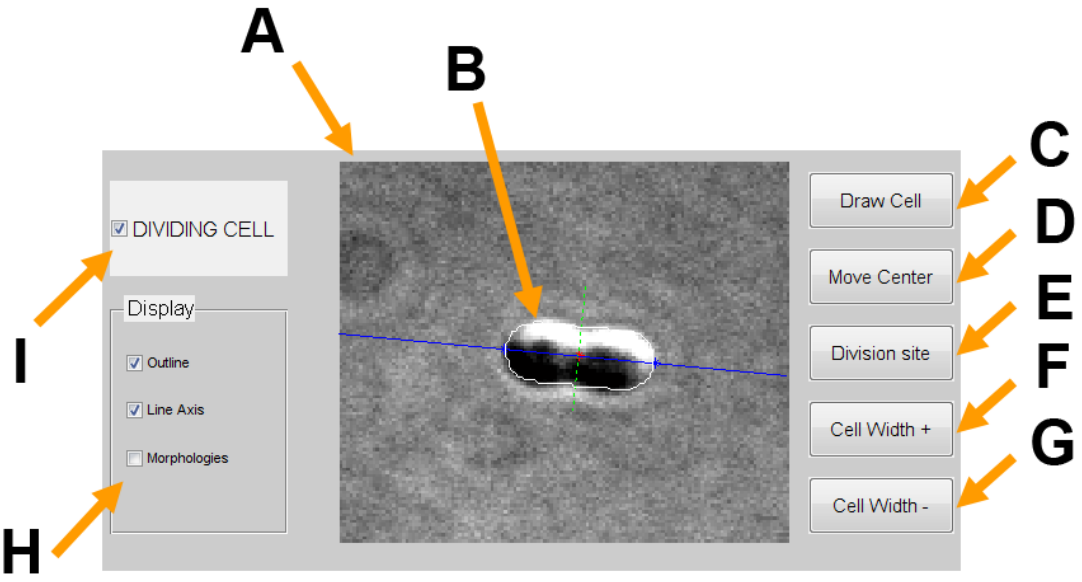


Figure 46: Stage 2a user Interface – Screen shot of the user interface used for cell edge refinement at stage 2a. **A)** The DIC (red) channel image is displayed in greyscale within this box, cropped around the cell of interest. Clicking within this region will alter the shape of the defined mask. **B)** Visual aids for placing the rod-shaped object over the cell are displayed here, these include blue stars to mark either end of the cell, a red star to mark the centre of the cell, a blue line to mark the long axis of the cell, a dotted green line to mark (if present) the division site of the cell and a white line which marks the edge of the rod-shaped object defined as the cell. **C)** A button which allows the cell end to be defined upon clicking on the image. **D)** A button which allows the cell centre to be redefined upon clicking on the image. **E)** A button which allows the cell division site to be moved upon clicking on the image. **F & G)** Buttons which adjust the width of the defined rod-shaped object. **H)** Check-boxes that allow the appearance of the visual aids in **(B)** to be altered. **I)** A check box to define whether or not a cell has a visible site of septation.

3.8.3) Stage 2b – Fluorescent cluster position assignment

As part of the interface in stage 2a, a second window allowed the positions of fluorescent clusters from the blue and green channels of the image to be defined. To remove the element of manual error from the method of definition of cluster positions and produce consistent results from different datasets, a three-step algorithm was designed. The centre of each cluster was defined as the area with the local maximum intensity within the cell; therefore regional-maximum detection was used, isolating areas with pixel values greater than their surrounding pixels (Figure 47A). Due to the total recorded intensity for each pixel being a sum of the emitted fluorescence from the sample and Gaussian noise from the camera, peak-detection generated some false positives for cluster assignment, therefore a single-stage 3x3 median filter (equivalent to an image blur, or for 1D data – a 3 point moving average) was applied to the image before peak detection to remove the effect of noise (Figure 47B). Peak detection was improved, but peaks were also found in regions of low fluorescence which would not be assigned as clusters in a by-eye analysis. Therefore, a threshold value was applied to the image dataset, with a minimum value of at least 10 % greater than the average value of pixels found in the areas defined as “non-cell” in stage 1. This threshold could then be manually increased to exclude cellular background-fluorescence, leaving only the positions of clusters (Figure 47C). The cluster position was defined as the X-Y mean position of the pixels that remained. To aid in assignment of clusters, the green and blue images from around the cell were displayed with 2 x linearly scaled contrast to one side of the light-channel image, and the original 3-colour image was displayed above (Figure 48).

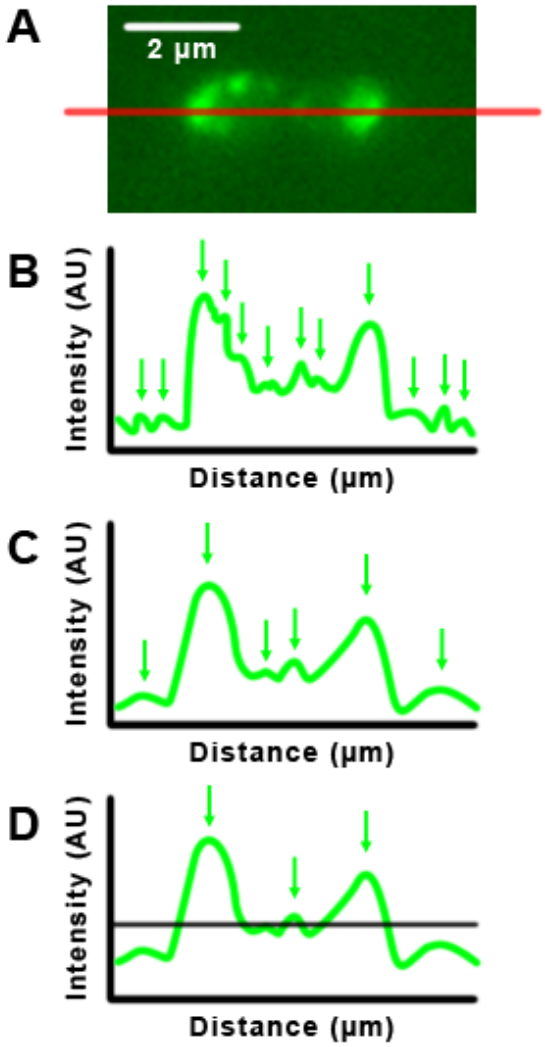


Figure 47: Fluorescence peak assignment. Schematic representation of the three-step algorithm used to define the positions of fluorescent peaks within images. **A)** An example *R.sphaeroides* cell containing YFP-tagged fluorescent protein clusters. **B**, **C** and **D)** schematic 1D fluorescence intensity profiles as if plotted along the red line in **(A)**, arrows mark positions that are defined as peaks at each stage. **B)** Raw image data, multiple false-positive peaks are seen. **C)** Median-filtered image data, the intensity profile is smoothed, and fewer false-positive peaks are observed. **D)** A threshold value is chosen (marked by the black line), and all peaks that fall below this value are excluded, if false positives still remain in the assignment after the selection of an optimum threshold, the cell is discarded from the analysis.

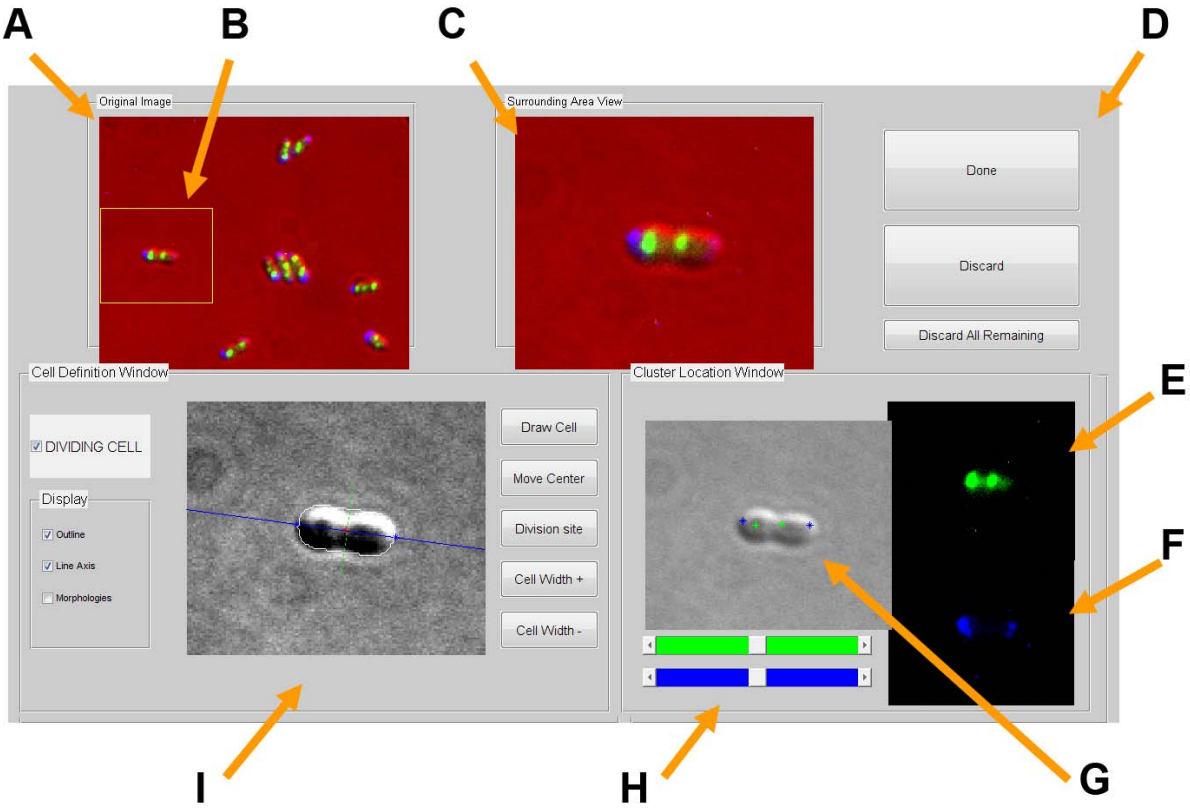


Figure 48: Stage 2b user Interface – Screen shot of the user interface used for fluorescent cluster positioning at stage 2b. **A)** The original Image is displayed here. **B)** A yellow box marks the cropped region around the current object of interest. **C)** The RGB image of the current object of interest is displayed here. **D)** Buttons which allow the cell morphology and cluster positions defined by the user to be accepted, or for that object to be discarded. **E)** The green-channel image is displayed here, pixel intensities are multiplied by 2 to increase contrast to aid in the visualisation of dim clusters. **F)** The blue-channel image is displayed here, pixel intensities are multiplied by 2 to increase contrast to aid in the visualisation of dim clusters. **G)** Defined cluster positions are displayed as green and blue stars, and overlaid with the greyscale DIC (red) channel image. **H)** Slider bars that allow the thresholds of cluster detection to be adjusted for the green and blue channels. **I)** The user interface of Stage 2a.

Although this method of definition works well for the majority of images, low signal:noise images, or cells with high fluorescence backgrounds generated slight inaccuracies in cluster assignment, the user was therefore required to judge whether the achieved definition was satisfactory, and either accept or discard the data.

Due to the inability to fully automate the method of cell and cluster definition, a certain degree of user subjectivity was retained for each stage of the assignment of cell lengths, widths and positions. And therefore for consistency, it is recommended that a single user process a single dataset. For this work, each cell from each image was processed in this manner for a total of approximately 10,000 included cells and many more during the development of the software.

Despite the subjectivity, the provided framework did improve the restrictions placed on definitions of clusters, and the visual-interface assisted in the accurate measurement of cellular morphologies and was deliberately designed to make consistent analysis achievable.

3.9) Storing data for fast analysis

As the values of the fluorescence background varied from day to day and with individual microscope settings, images for comparable datasets were all captured on the same day, with the same settings and non-fluorescent cell samples were imaged for the estimation of cellular background auto-fluorescence intensity. The calculated background-intensity values were therefore subtracted from the image after processing of the full data set, and intensity data was extracted during this post-processing step. To enable re-visiting of the original image data and to perform this level of post-processing (which is a computationally demanding process), the user defined data was saved for each image with a selection of other metadata parameters including several masks of the potential regions of interest within each cell (whole-cell regions, polar regions, positions of individual daughter cells and cluster positions) and a cropped version of the original image. Storing the data in this way allowed for fast processing of large data sets after the initial labour-intensive process of object definition, with the data from approximately 4000 cells being compiled in less than one minute.

4) Clusters and morphology in the average cell

Using the image analysis software developed in chapter (3), cell sizes, shapes and patterns of fluorescent protein cluster number and positioning were examined in otherwise wild-type populations of *R.sphaeroides* cells.

4.1) Morphology Analysis

The development of the MATLAB image analysis program discussed in chapter (3) allowed for a large increase in the number of cells that could have their lengths, widths, sites of cell division, and protein cluster locations defined and measured for any individual data set. Although the full automation of this process was not possible and a large amount of manual input was still required in order to obtain data, dataset sizes obtained in this way could regularly include more than 300 cells with no upper limit to a sample size, and, importantly, contained all information about each cell which allows data to be revisited and probed for different trends.

In order to verify that the data thus generated were comparable with previous data on *R.sphaeroides*, cellular morphologies were investigated for 3812 cells from mid-log phase cultures of cells grown on different days from different samples and different strains which had similar growth rates. Cell lengths, widths, and division site positions were determined for the entire population (Figure 49).

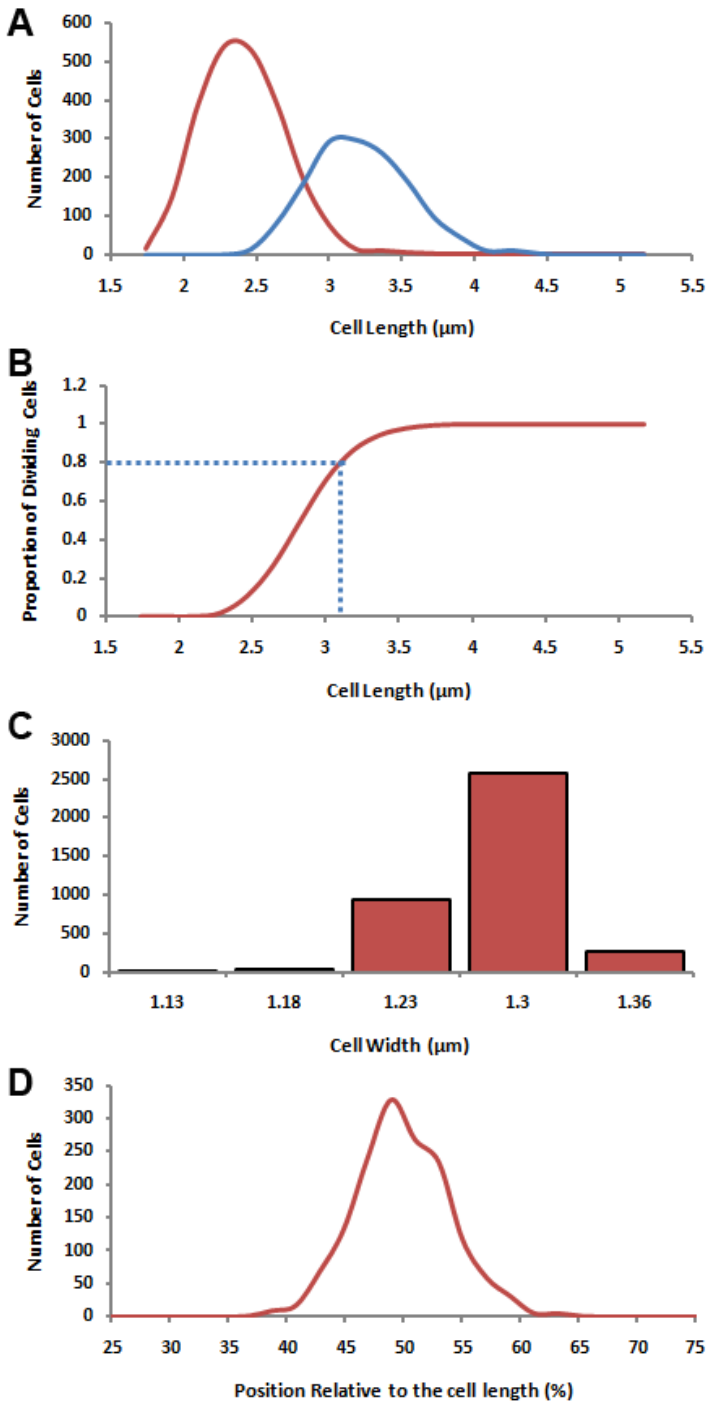


Figure 49: Morphology Analysis of 3812 cells. **A)** Cell Lengths**:** Smoothed histogram of the number of cells with no visible septation sites (red) and number of cells with visible septation sites (blue). **B)** Proportion of Dividing Cells with Cell Length**:** Best-fit sigmoid curve (data was fitted using MATLAB) of the proportion of dividing cells with visible septation sites with cell length. The position at which 80 % of the population were seen to be septate is marked (dotted blue line). **C)** Cell Widths**:** Bar representation of the number of cells defined with different widths. NB. The method of defining cells widths in the MATLAB analysis program discussed in chapter 3 (3.8.2) produced a discrete number-set. **D)** Positions of Division Sites: Smoothed histogram of septation-site positions measured as a percentage of the whole cell length.

In good agreement with data obtained by P.Slovak (141), cell widths weres on average 1.29 µm, with a standard deviation of 0.04 µm. The minimum length of a cell within a mid-log phase population was seen to be 1.74 µm, with cells growing to a maximum length of 5.17 µm before budding off into individual daughter cells. Division sites were visible in cells of 2.5 µm in length and above, and above 3.19 µm in length more than 80 % of the cells measured had visible septation sites. On average, division sites were found at the exact mid-cell position with a mean position of 49.97 +/- 0.04 % when measured as the relative position along the cell's length. These data, and a comparison to previous data is summarised in Table 6.

	This Study	P. Slovak (2005)
Minimum Cell length (µm)	1.74	1.73
Average Cell Width (µm)	1.29 ± 0.04	1.17 ± 0.01
Septation Initiation (µm)	2.5 – 3.19	2.67 ± 0.03

Table 6: Comparison of Morphology Analysis with previous data: The above data is summarised for this study, and compared to data acquired by P.Slovak in 2005. Numbers represent the means and standard deviations of all measured cells.

4.2) Population analysis of chemotaxis protein cluster positions

4.2.1) Polar chemotaxis proteins

Of the fluorescent protein-fusion reporter strains that were available in *R.sphaeroides*, the protein marker YFP-CheW$_3$ was chosen as the reporter of polar-localised proteins. Previous work on this strain, JPA1418, and wild-type WS8N strains provides several reasons for this choice:

1) Growth rate in this mutant is unaffected (99).

   Cellular morphologies and gene expression patterns that are dependent on the stage and rate of growth will be comparable between wild-type and JPA1418.

2) Chemotactic efficiency (measured on soft-agar swim plates) is as wild-type.

   Fusion of the CheW$_3$ protein to YFP leaves chemotaxis unaffected. As protein clustering is also important to chemotaxis signalling (99), this implies that the morphologies, numbers and positions of clusters observed using CheW$_3$ will be comparable to that of wild-type cells.

3) The copy-number of $CheW_3$ is relatively high.

Previous quantitative western blot analyses estimate the copy number of $CheW_3$ to be approximately 10,000 copies per cell when grown under microaerobic laboratory conditions (98, 99). A high copy number of fluorescent proteins will give a high signal to noise ratio in images, making the image processing stage easier.

Three cultures of JPA1418 and three cultures of WS8N cells were grown microaerobically from 1/30 dilutions of independently grown overnight cultures. Cells of OD_{700} = 0.55 - 0.65 were immobilised on agarose pads, and were visualised on a Nikon TE200 microscope, with a YFP exposure time of 1.2 seconds, which was optimised by eye to give high signal to noise, with minimum saturation of the camera. All images were taken on the same day, using consistent microscope settings in order for the intensity values to be comparable across images. Captured images were analysed using the custom-built MATLAB analysis tool discussed in chapter (3).

4.2.2) Numbers of polar and lateral clusters in a wild-type population

In over 600 cells from a mid-log phase culture, the numbers of distinguishable YFP-$CheW_3$ clusters observed within a population of individual and dividing rod-shaped cells varied between zero and 6 clusters, with 95 % of these cells containing between 1 and 3 clusters (Table 7).

Number of Visible Clusters	Frequency
0	3
1	108
2	368
3	117
4	26
5	1
6	1
TOTAL	624

Table 7: The number of polar and lateral YFP-$CheW_3$ clusters observed in a population of JPA1418 cells.

In order to determine the effects of cell growth on the number of clusters in a cell, cells were grouped according to their lengths. Twenty groups of equal bin width were defined between the minimum and the maximum values obtained for cell length, and the proportion of cells with 0, 1, 2 and more than 2 distinguishable clusters was determined for cell lengths that fell within each bin (Figure 50).

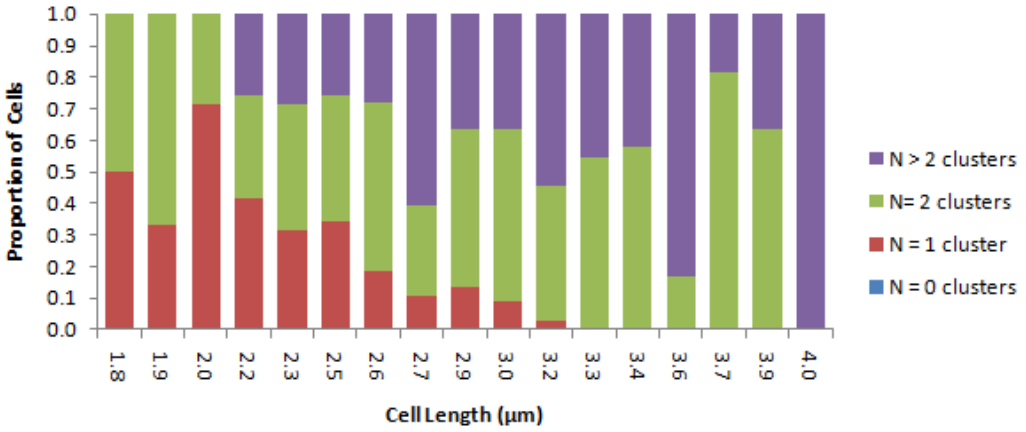


Figure 50**: Cluster Numbers With Cell Length:** The proportion of cells found in each bin with zero (blue), one (red), two (green) and more than two (purple) YFP-CheW_3 clusters. Results are plotted for cell lengths between 2 and 4 µm.

From this analysis, it was observed that newly divided cells had a higher tendency to possess a single polar cluster; cells with 2 clusters were common across a large range of cell lengths, whereas the proportion of cells with more than 2 clusters rose gradually with cell length. The cluster number therefore increases with cell length, and when average cluster number is plotted for the same bin sizes (Figure 51), an upwards trend between 2 and 3 µm is followed by a plateau in the line between 3 and 4 µm, corresponding to the range of cell lengths which fall between the beginning and completion of cell division. Cells with lengths greater than 4 µm represented a minority of the population (~1 %), and were excluded from the data set. The completion of cell division may therefore lead to new cluster formation, presumably at the new pole, and indeed the fact that many newly divided cells have only one polar cluster but the majority of growing cells possess two supports this hypothesis.

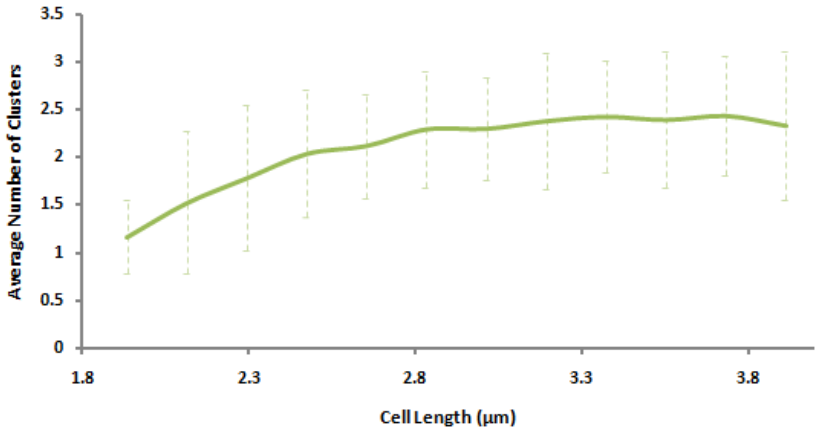


Figure 51**: Average number of polar clusters with cell length:** Error bars represent plus or minus one standard deviation from the mean.

To examine the positioning of YFP-CheW_3 clusters in more detail, the data obtained for cells with different numbers of clusters was examined individually. For definitions of cluster positions, "polar" was defined as any cluster that fell within the end semi-circle of the rod shaped cell (i.e. within half the width of the cell from the end of the cell), and "lateral" was any position outside the polar regions (Figure 52). NB. As the cell grows, the proportion of the total volume of the cell that falls within the "polar" region decreases, and the "lateral" volume increases until formation of the new pole.

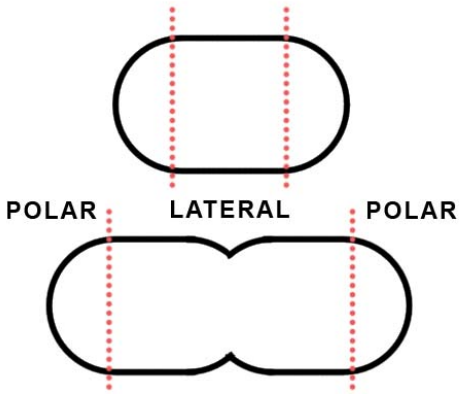


Figure 52**: Schematic diagram of the definitions of lateral and polar positions:** Any position inside of the two red dotted lines was defined as "lateral"; any position in the end semi-circles of the cell (hemispheres in reality) was defined as polar. Cell length, or the presence of septation sites did not alter these definitions.

4.2.3) Cells with one cluster

As expected, in 98 % of 108 cells which had only one visible cluster, that cluster was observed in a polar position (Figure 53). Polar localised protein was seen to collect at the 12.5 % and 87.5 % positions within cells; the spacing from the far end of the pole represents the distance between the visible outer shell of the cell in a DIC image and the cytoplasmic side of the inner membrane.

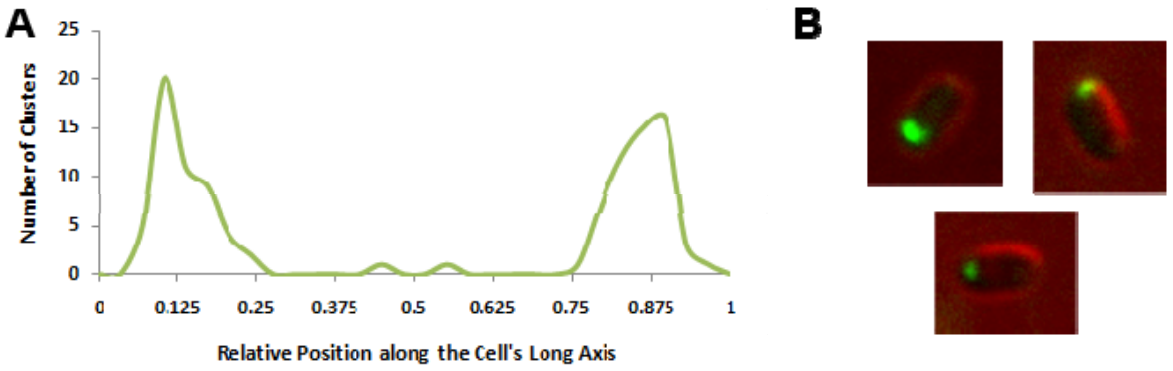


Figure 53**: A)** One Cluster Positions: Positions of clusters in cells with only one cluster, positions are calculated relative to the whole cell length **B)** Example Images: 3 example cells with only one cluster, contrast has been adjusted for clarity.

4.2.4) Cells with two clusters

Here, again, the vast majority of observed clusters were found at polar positions, with 92 % of the observed cells having a polar cluster at each pole of the cell. Interestingly, of the remaining 8 % of cells in this population approximately half were seen to have one polar and one laterally positioned cluster, whereas the other half contained two distinguishable clusters at only one pole of the cell (Table 8).

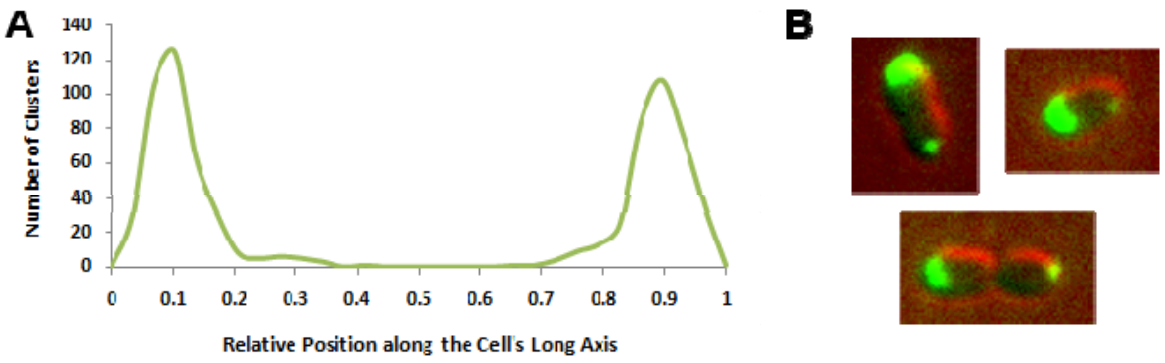


Figure 54**: A)** Two Cluster Positions: Positions of clusters in cells with two clusters, positions are calculated relative to the whole cell length **B)** Example Images: 3 example cells with two clusters, contrast has been adjusted for clarity.

Polar proteins were found at the 10 % and 90 % positions within cells, which is slightly different from one-cluster cells because two-cluster cells are on average longer than one-cluster cells.

4.2.5) Cells with three clusters

99 % of cells with three clusters had at least one cluster at either pole of the cell. The third cluster could either be lateral (45 %), or polar (55 %) showing that cell poles regularly accommodate multiple spatially distinct clusters (Figure 55).

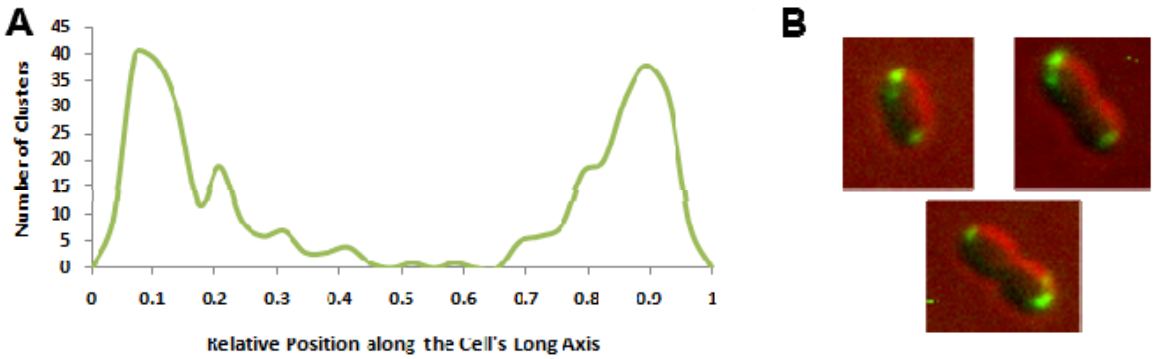


Figure 55**: A)** Three Cluster Positions: Positions of clusters in cells with three clusters, positions are calculated relative to the whole cell length **B**) Example Images**:** 3 example cells with three clusters, contrast has been adjusted for clarity.

4.2.6) Cells with more than three clusters

When there were more than 3 clusters present in a cell, the majority of cells had both multiple clusters at one pole and a laterally positioned cluster (Figure 56), no cells were seen which had poles with no cluster.

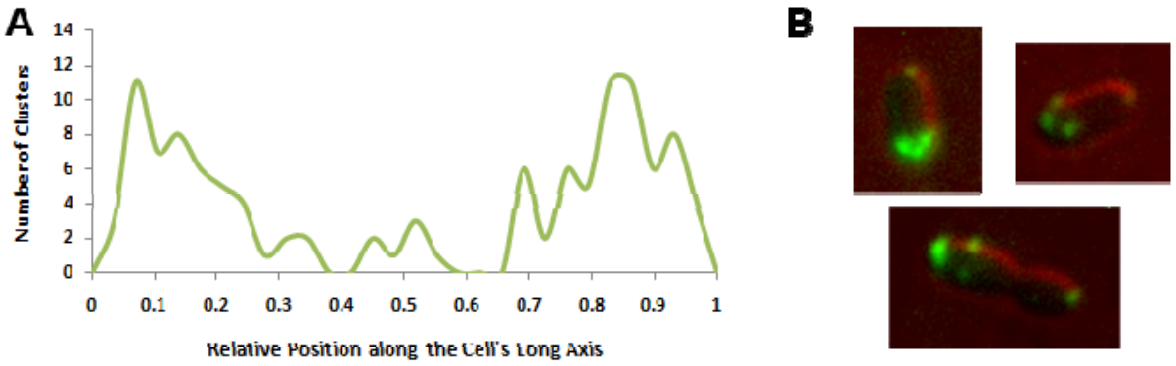


Figure 56**: A)** Greater than Three Cluster Positions**:** Positions of clusters in cells with more than three clusters, positions are calculated relative to the whole cell length **B)** Example Images**:** 3 example cells with more than three clusters, contrast has been adjusted for clarity.

The above data are summarised below in Table 8.

Number of Clusters	Total Number of Cells	Number of Cells with lateral clusters	Number of Cells with an empty pole	Number of Cells with poles containing >1 cluster
0	3	-	-	-
1	108	2	108	-
2	368	15	27	12
3	117	52	1	65
>3	28	19	0	24
TOTAL	624	88	136	101

Table 8**: Polar Cluster Positions Summary:** Summary of the data presented in sections (4.2.3), (4.2.4), (4.2.5) and (4.2.6).

Intriguingly, from these data (Table 8) the formation of laterally positioned clusters appears relatively rare within a wild-type population of cells, with only 14 % of the total number of cells

observed containing lateral clusters. The presence of more than one cluster at a single pole is more frequent, occurring in 16 % of observed cells.

As the positioning of lateral clusters has been suggested to be linked to future sites of cell division in *E.coli* (147), the positions of all lateral clusters were plotted for dividing and non-dividing cells within this population. Positions were calculated to the nearest pole and relative to the total cell length. For dividing cells, positions were also calculated from the site of cell division and relative to cell length (Figure 57).

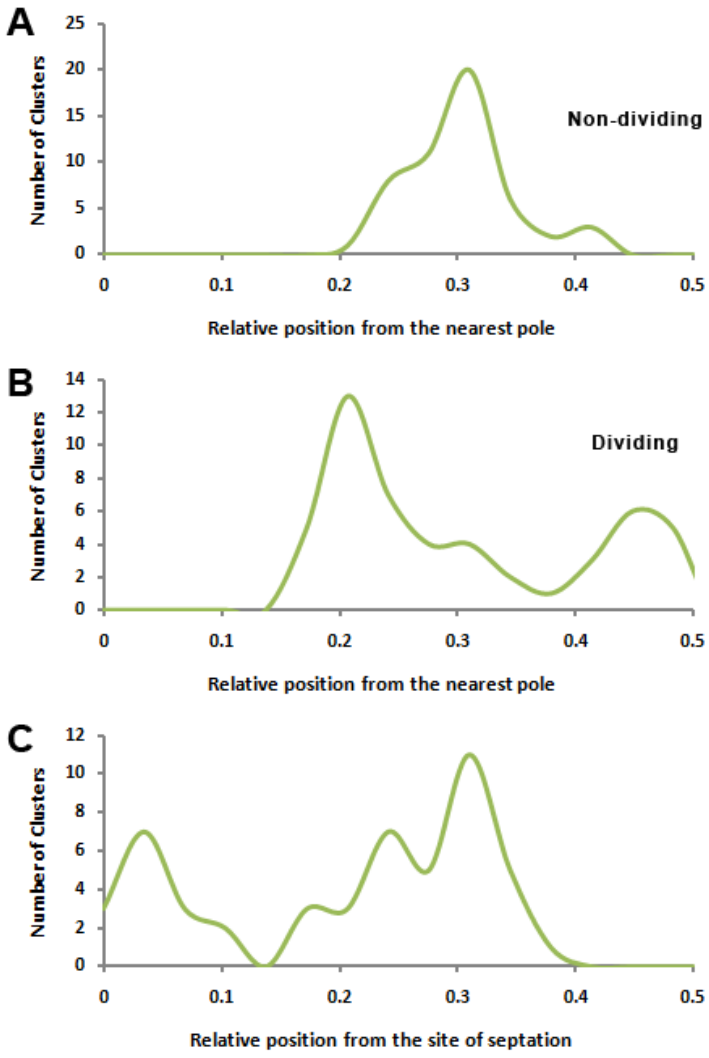


Figure 57**: Lateral Cluster Positions. A)** Positions of lateral clusters from cells with no visible septation sites, measured to the nearest pole and relative to the whole cell length. **B)** Positions of lateral clusters from cells with a visible septation site, measured to the nearest pole and relative to whole cell length. **C)** The same data as **(B)**, but measured to the site of septation and relative to whole cell length.

Lateral clusters were no more common in dividing cells than in non-dividing cells; 41 of 88 cells with lateral clusters had visible septation sites and 47 did not, suggesting that septation does not encourage the formation of a new cluster. However, the distribution of positions observed in dividing cells was clearly different from those in non-dividing cells (Figure 57).

In non-dividing cells, if lateral cluster formation and positioning were entirely random, a uniform distribution of lateral clusters at positions spanning the whole length of the cell would be expected, however the observed pattern of localisation suggests that clusters are more likely to form closer to an already existing pole than near the mid-cell. In dividing cells, this pattern is also observed, however there is a second group of lateral clusters which is found to localise near the future site of cell division.

4.3) Cytoplasmic chemotaxis proteins

As with the YFP-$CheW_3$ mutant, the homologous fluorescent fusion protein from *cheOp3*, $CheW_4$-YFP has previously been shown to exhibit wild-type growth rates, chemotactic efficiency and to have a high copy number when grown under microaerobic conditions (98, 99). For these reasons, strain JPA1457 was chosen as the reporter strain for the study of cytoplasmic chemotaxis protein localisation.

4.3.1) Numbers of cytoplasmic clusters in a wild-type population

As with JPA1418, three cultures of JPA1457 and three cultures of WS8N cells were grown microaerobically from 1/30 dilutions of independently grown overnight cultures. Cells of OD_{700} = 0.55 - 0.65 were immobilised on agarose pads, and were visualised as for JPA1418, with a YFP exposure time of 1.1 seconds. All images were taken on the same day, using consistent microscope settings in order for the intensity values to be comparable across images. Captured images were analysed using the custom-built MATLAB analysis tool discussed in chapter (3).

In 589 cells from a mid-log phase culture, the numbers of distinguishable $CheW_4$-YFP clusters observed within a population of individual and dividing rod-shaped cells varied between zero and 4 clusters, with 94 % of these cells containing between 1 and 3 clusters (Table 9).

Number of Visible Clusters	Frequency
0	27
1	168
2	299
3	88
4	7
TOTAL	589

Table 9: The number of polar and lateral $CheW_4$-YFP clusters observed in a population of JPA1457 cells.

As with YFP-$CheW_3$, the effect of cell growth on the number of clusters was investigated: Twenty groups of equal bin width were defined between the minimum and the maximum values obtained for cell length, and the proportion of cells with 0, 1, 2 and more than 2 distinguishable clusters was determined for cell lengths that fell within each bin (Figure 58).

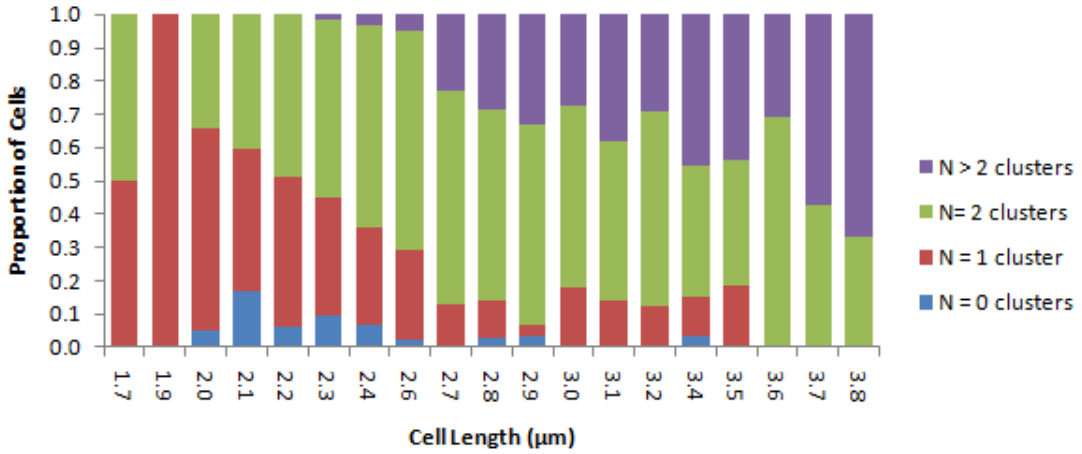


Figure 58**: Cluster Numbers With Cell Length:** The proportion of cells found in each bin with zero (blue), one (red), two (green) and more than two (purple) $CheW_4$-YFP clusters. Results are plotted for cell lengths between 2 and 4 µm.

Similar to the observations for $CheW_3$ cluster numbers, a large proportion of newly divided cells contained only one protein cluster and the proportion of cells with only one cluster decreased with cell length. A small proportion of newly divided cells were also seen to have no clusters. Cells with two clusters were the most common, and represented a large proportion of cells of all lengths. The proportion of the population that contained more than two clusters was seen to increase with cell

length. As a general trend, cluster numbers are seen to increase with cell length, and plotting average cluster number with cell length reveals an upwards trend with cell length (Figure 59).

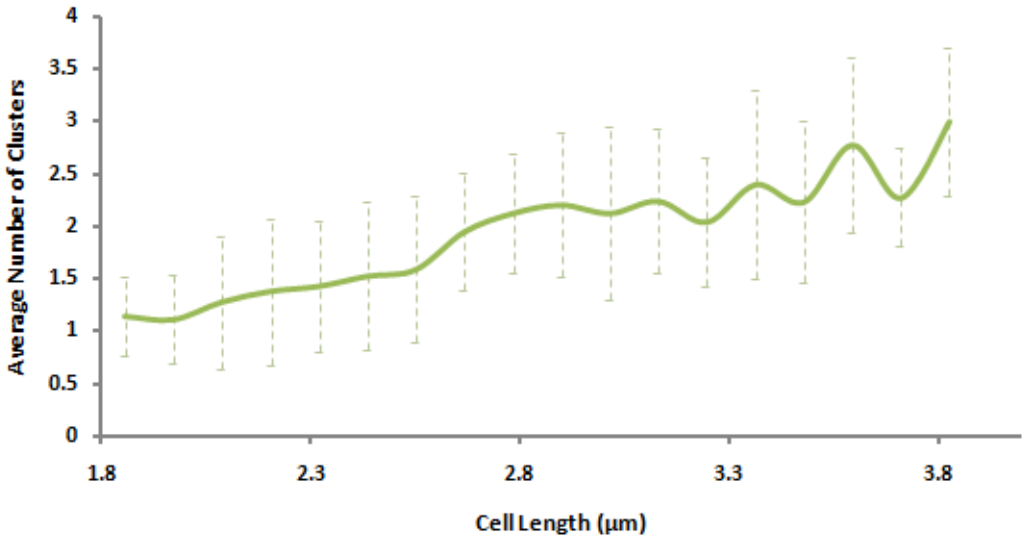


Figure 59**: Average number of cytoplasmic clusters with cell length:** Error bars represent plus or minus one standard deviation from the mean.

To examine the positioning of $CheW_4$-YFP clusters in more detail, the data obtained for cells with different numbers of clusters was examined individually using the same definitions of “polar” and “lateral” as defined in section (4.2.2).

4.3.2) Cells with one cluster

As has been observed previously (150), the average position of a single cluster is the centre of the cell, however, the distribution of cluster positions ranged between 25 % and 75 % of the whole cell length with a broad spread between these two extremes (Figure 60). In 18 out of 168 cells (11 %), the cluster's centre fell within the polar region of the cell.

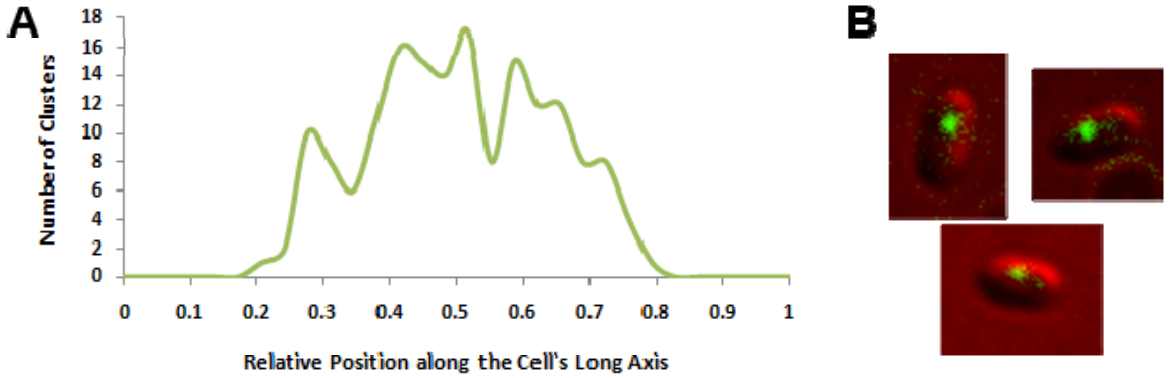


Figure 60**: A)** One Cluster Positions: Positions of clusters in cells with only one cluster, positions are calculated relative to the whole cell length **B)** Example Images: 3 example cells with only one cluster, contrast has been adjusted for clarity.

4.3.3) Cells with two clusters

In cells with two clusters, two clear populations of cluster position were observed which, on average, were seen around the 30 % and 70 % positions relative to the whole cell length (Figure 61). 94 of 299 cells (31 %) possessed clusters with centres in a polar region of the cell.

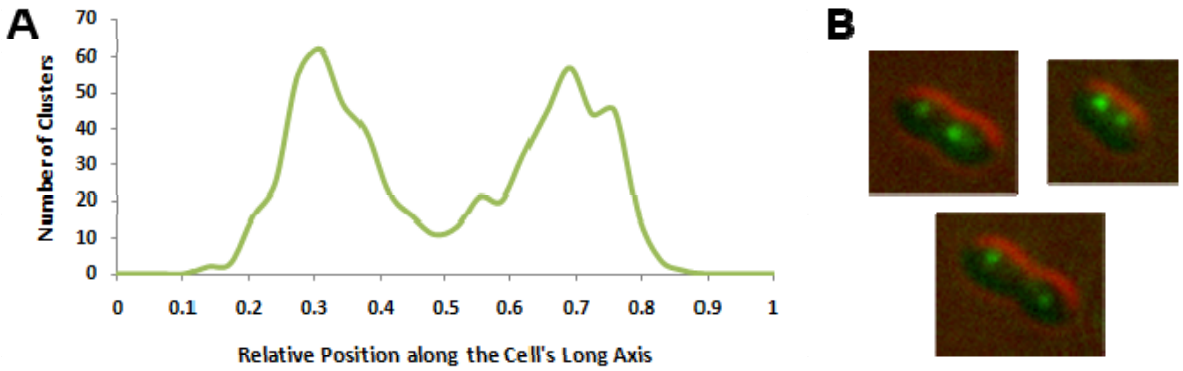


Figure 61**: A)** Two Cluster Positions: Positions of clusters in cells with two clusters, positions are calculated relative to the whole cell length **B)** Example Images**:** 3 example cells with two clusters, contrast has been adjusted for clarity.

4.3.4) Cells with more than two clusters

Cells with more than two clusters were most common in longer cells, and cells with visible septation sites. Clusters showed 3 distinct positions of localisation, at the 72.5 and 27.5 % positions, but also close to the centre of the cell (Figure 62). This is shown more clearly when plotted as distance from the nearest pole (treating each cell as if it were symmetrical about its centre), the third population of clusters was observed to be slightly off-centre, with a peak at the 45 % position (Figure 62C). 25 of 95 cells (26 %) possessed clusters with centres in a polar region of the cell.

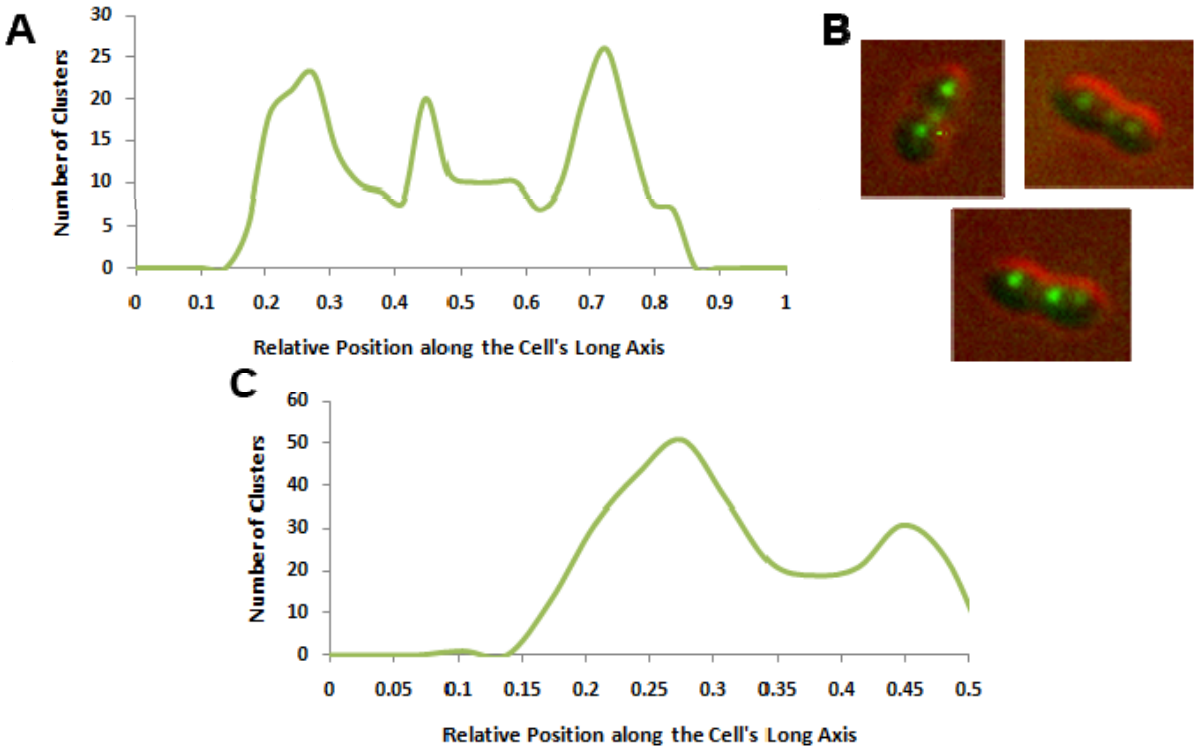


Figure 62**: A)** Greater than Two Cluster Positions**:** Positions of clusters in cells with more than two clusters, positions are calculated relative to the whole cell length **B)** Example Images: 3 example cells with more than two clusters, contrast has been adjusted for clarity. **C)** The same data as in **(A)**, but plotted as measured to the nearest pole of the cell.

4.4) Cluster separation in growing cells:

In order to measure how cluster spacing was affected by cellular grown, the positions of clusters in cells with only two clusters were measured relative to i) their nearest pole, ii) the second cluster and iii) any visible septation site (Figure 63).

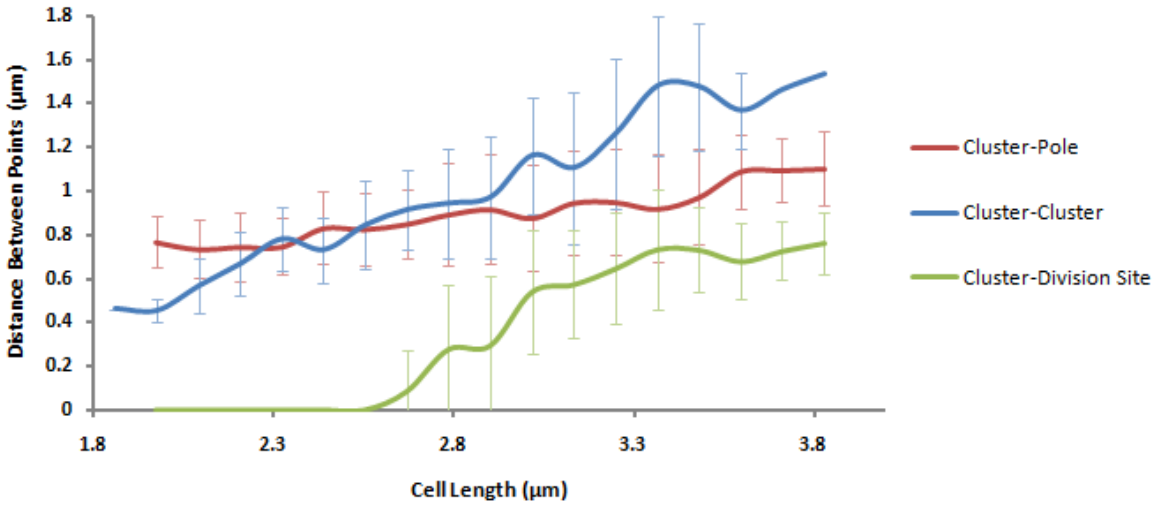


Figure 63: Cluster positioning with cell length: Distances between $CheW_4$-YFP clusters and their nearest pole (red), their neighbouring cluster (blue) and, for those cells that had them, visible septation sites (green).

The straight-line distance observed between cluster pairs increased at a greater rate than the distance between these clusters and their nearest pole, suggesting that cluster positioning is achieved relative to the cell pole, and not relative to a neighbouring cluster (Figure 63). Interestingly, cell septation was only seen in cells with clusters that were further from each other than they were from their pole. The distance between clusters and division sites increased at a similar rate to the distance between clusters alone and in long, dividing cells reached a value which was comparable to that of the cluster-pole distance in newly divided cells.

4.5) Asymmetry in polar cluster separation

Cells within a population have a tendency to acquire a second polar cluster after cell division. The result of successive rounds of cell division is that the pole inherited from the maternal bacterial cell, or “old pole”, will contain more clustered protein than the pole that was formed during the preceding round of cell division, or “new pole” (Figure 64).

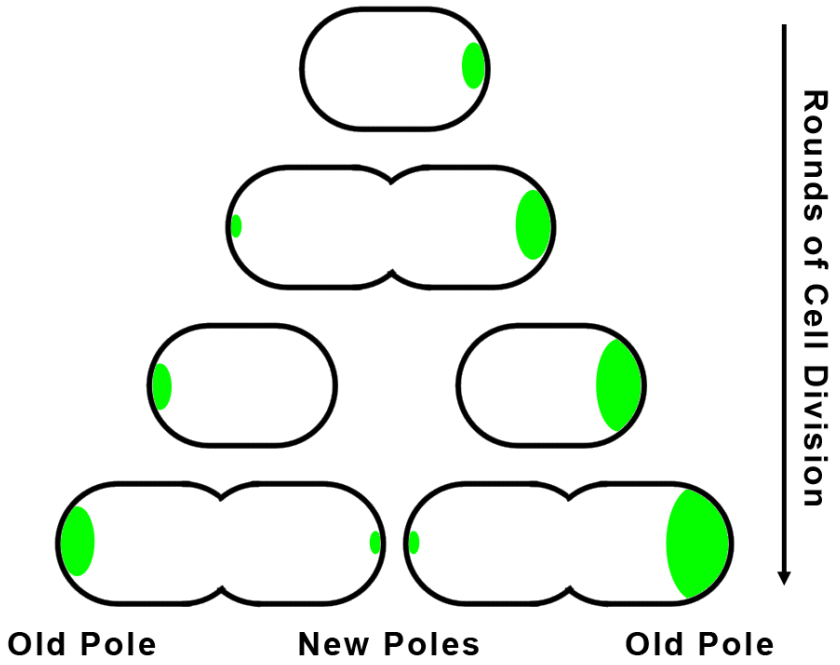


Figure 64: Old Pole/New Pole Schematic – Diagram showing how polar cluster size relates to pole age. Green dots represent YFP-CheW$_3$ protein. Old Poles will contain more protein than new poles, and protein distribution within cells is inherently asymmetric.

This mechanism of cluster separation ensures that nearly every daughter cell will have at least one polar chemotaxis cluster, and will therefore be able to undergo chemotaxis immediately after cell division. However, the secondary effect is that daughter cells will have an inherent asymmetry based on the old pole/new pole divide and that the absolute concentrations of chemotaxis proteins inherited by individual daughter cells will vary over a wide range. One factor that may limit the magnitude of this effect may be that a single pole has an upper limit to the number of proteins it can accommodate – meaning that cells with the most densely clustered poles will accumulate protein at the new pole, and not the old pole. However, when cell growth and division is at its fastest, as at mid-log phase, asymmetry will still be observed and is likely to play a crucial role in population heterogeneity (Figure 65).

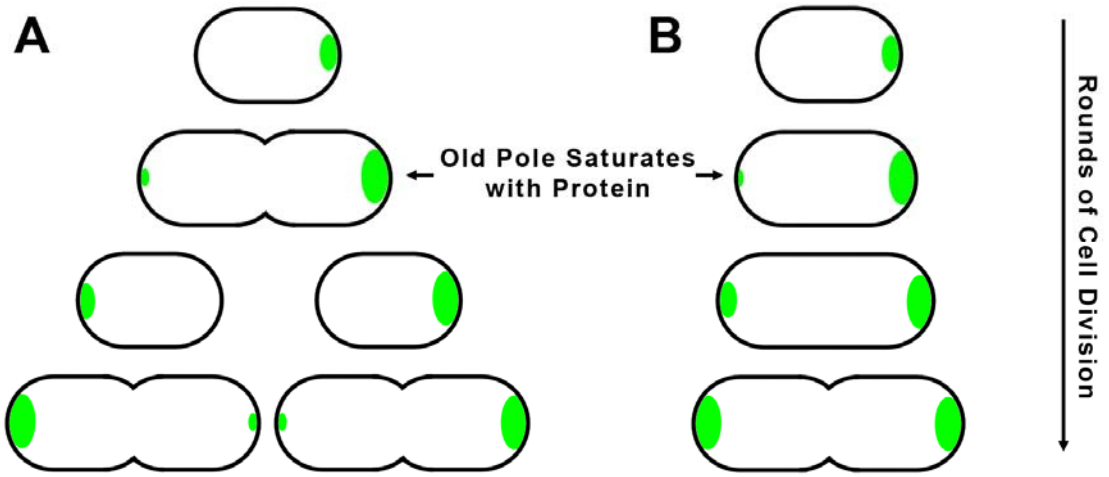


Figure 65: Affect of Growth Rate of Cluster Asymmetry – Diagram showing how saturation of the amount of protein at the pole and growth rate could affect asymmetry in a population of **A)** Fast growing cells – all cells show asymmetry **B)** Slow growing cells – not all cells show asymmetry. The size of the green dots represents the amount of YFP-$CheW_3$ protein present in each pole.

To test whether the rate of accumulation of protein at either pole was altered by the age of the pole relative to the last round of cell division, the amounts of protein found at either pole were estimated for the fluorescence image data set generated in section (4.2.1). To get a relative measure of the amounts of protein in either pole, an average value of cellular autofluorescence was calculated from pixels in WS8N, non-fluorescent cells. This value was subtracted from each YFP image as background and the sum of pixel values within a masked area covering the individual polar regions of each cell was calculated. Data was grouped into 20 equal sized bins which fell between the maximum and minimum cell lengths observed within the population, and was plotted for cells which had individual clusters at both poles (Figure 66). In this instance, the brightest pole was defined as the "old pole", and the dimmest pole was defined as the "new pole".

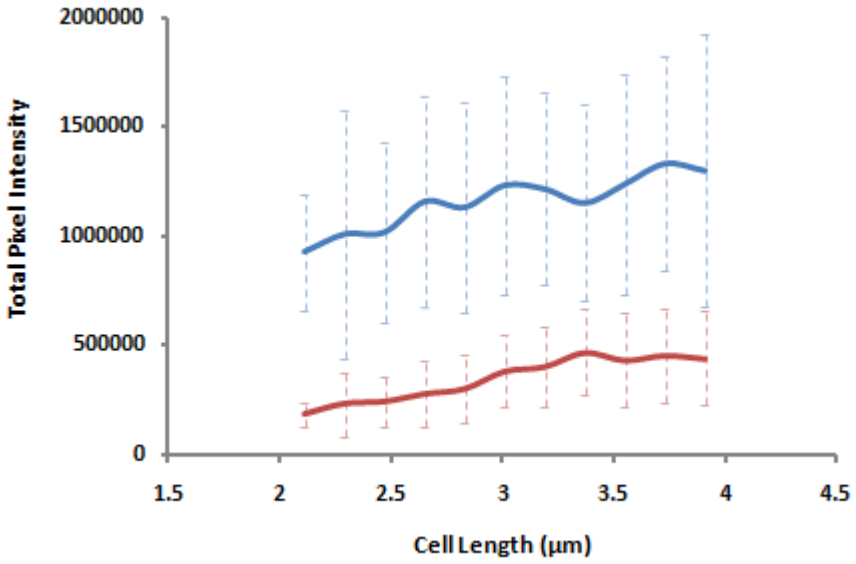


Figure 66**:** Amounts of YFP-CheW$_3$ in old and new poles with cell length. Lines show the average sum of pixel intensities for old (blue) and new (red) cell poles within each bin for cell length. Error bars represent plus or minus one standard deviation from the mean.

The absolute amounts of protein at the old pole were significantly higher than those at the new pole across all cell lengths. Protein content at both poles was seen to increase with cell length, and the rates of accumulation were similar at both new and old poles, showing that the age of the pole does not affect its ability to accumulate new protein and that there is no active mechanism for depositing protein to a newly formed pole.

After two rounds of cell division, a new pole becomes an old pole, and thus it would be predicted that the intensity of the old pole in a newly divided cell would be similar to new pole in a dividing cell. However, the observed value for an old pole in a newly divided cell was approximately double that of a dividing new pole. The explanation for this discrepancy is attributable to the fact that the old pole in a newly divided cell has been formed over several rounds of cell division.

In cells where there is no upper limit to the amount of protein that can localise at one pole, a steady increase in the amount of protein at an old pole will occur over successive rounds of cell division (Figure 67A-C). Division of cells within a population will lead to a distribution of different pole ages, which will be weighted towards cells that have divided a relatively small number of times (Figure 67D), with half being the result of one round of cell division, one quarter being a result of two rounds of division and so on.

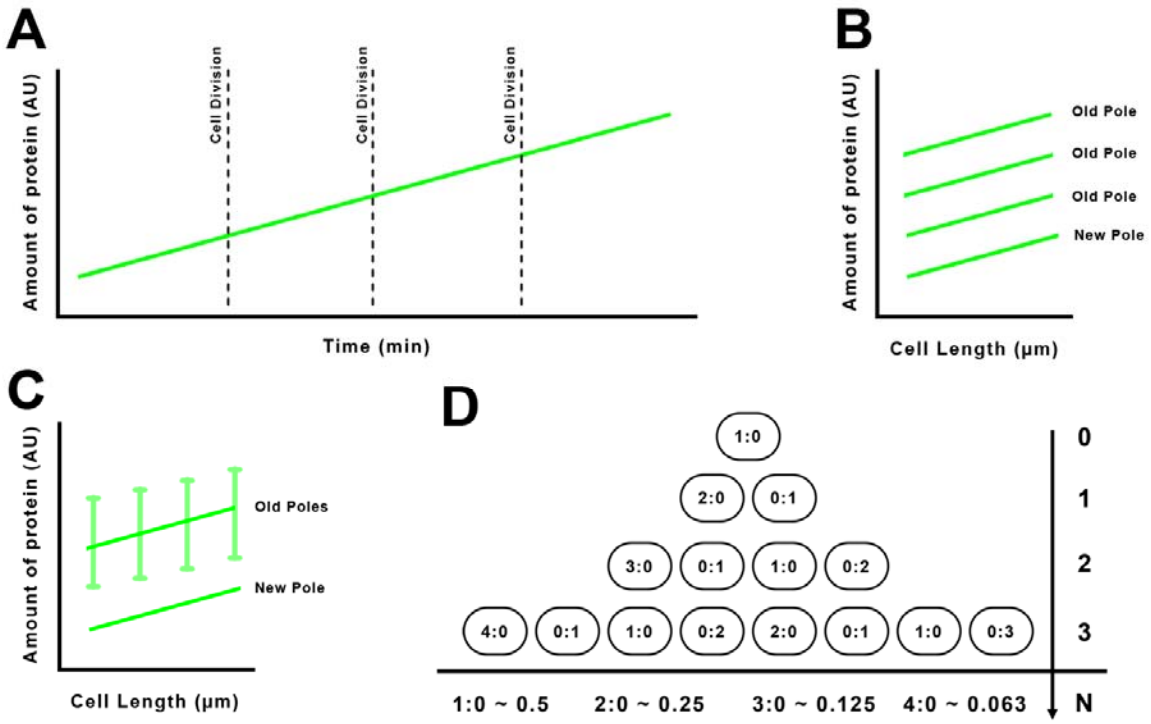


Figure 67**:** Idealised Distribution of Polar Cluster Ratios **–** Schematic representations of idealised scenarios for **A)** The intensity of a single polar cluster over its lifetime (four generations), protein accumulation at the pole is steady and linear, and polar protein concentrations do not saturate. **B)** Representation of the polar clusters in **(A)** as they would appear within a mixed population of growing cells. **C)** The same as **(B)**, but with all old poles grouped in the same way as the presented data in Figure 66. **D)** The distribution of cluster ratios within a population over 0, 1, 2, 3 and N generations of cell division, numbers represent the relative amounts of protein inherited by each daughter cell, with only one cluster being inherited, the amount of protein accumulating at each pole is 1 unit per generation, for N generations the proportion of the population that inherit the indicated ratios is estimated.

The strong similarity between the idealised distribution of cluster intensities proposed in Figure 67 and the observed data in Figure 66 indicates that polar chemotaxis protein clusters may continually grow in size over their lifetime. Within an entire population the proportion of cells that inherit a cluster which is very large will be small and therefore saturation of the amount of protein at the poles of the cell is not likely to affect protein distribution within cells. One interesting approach may be to increase the cellular *che*Op2-encoded protein concentration within the cells to see whether this results in increased polar saturation and whether this results in an increased frequency of lateral cluster formation, as is predicted by the SNM (see section (1.6.3)). This is the focus of chapters (5) and (6).

4.6) Asymmetry in cytoplasmic cluster partitioning

The cytoplasmic chemotaxis protein clusters of *R.sphaeroides* have been shown to partition between daughter cells (150). Unlike polar protein cluster formation which is not known to be an active process and may be the result of diffusion (157), the synthesis of a second cytoplasmic cluster has been shown to be an active process, which is controlled by the action of the ParA homologue, PpfA.

As seen in section (4.3), the majority of newly divided cells contain only one cytoplasmic cluster, which will (on average) become two clusters prior to cell division, giving rise to two daughter cells with one cluster each, and the process repeats. This partitioning ensures that the majority of daughter cells in a population will possess a cytoplasmic chemotaxis cluster, and will therefore be able to undergo chemotaxis immediately after cell division.

In an ideal population, separation of a single large cluster into two clusters of equal size and a steady rate of protein accumulation would allow for infinite propagation of this pattern, with no net gain of protein over successive rounds of cell division (Figure 68).

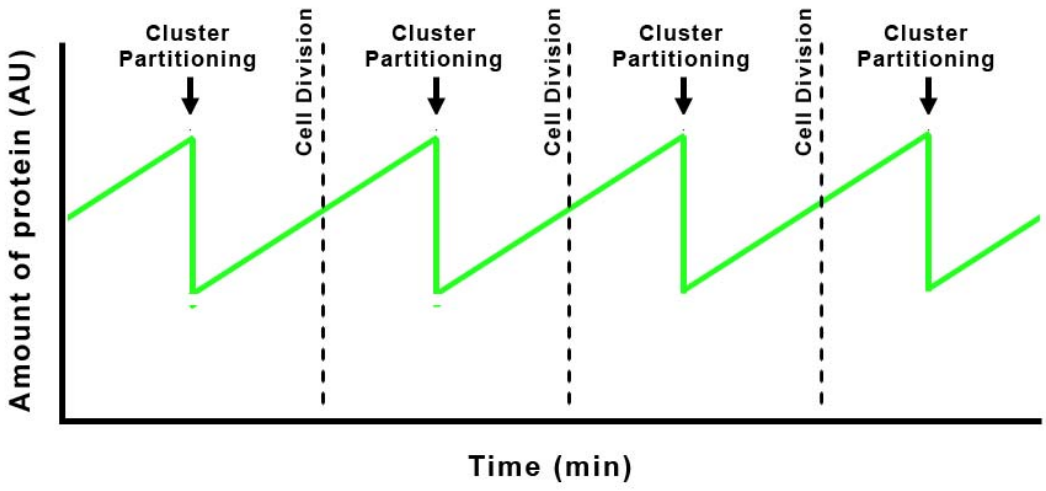


Figure 68: Ideal partitioning of cytoplasmic clusters – Schematic representation of the perfect pattern of cytoplasmic cluster partitioning relative to the cell cycle over four generations. The amount of protein represents the total amount of $CheW_4$-YFP in a particular cluster over time.

To test whether this ideal pattern was in fact observed within a wild-type population, amounts of protein in individual clusters were measured using the data set that was generated in section (4.3). Average WS8N pixel values from images taken on the same day and with the same microscope

settings were subtracted as background autofluorescence, and the sum of pixel values within a masked area around each defined cluster was calculated. The assumption made in defining how much of an area around each cluster to include was that each cluster could be approximated as a point source of light, a circle with a radius of half the wavelength of light was then drawn around the centre of the cluster, and pixels with centres that fell within this circle were included. It is possible that this method may lead to small underestimates of total cluster intensity in large clusters.

When total amounts of protein from clusters in cells containing only 2 clusters were calculated and plotted with cell length, it was observed that the average ratio between protein amounts in "daughter clusters" was approximately 2:1, showing that cluster partitioning is not perfectly even (Figure 69). As the mechanism of cluster separation is not fully understood, it is not known whether this asymmetry could be caused by a stochastic nucleation of a second cluster, or whether it is caused by an imbalance in active partitioning of one cluster into two. There was no obvious trend in the ratios between daughter clusters with cell length. As it is not known whether the intensity of each cluster increases with that cluster's age, clusters were referred to as "bright" and "dim" instead of "old" and "new".

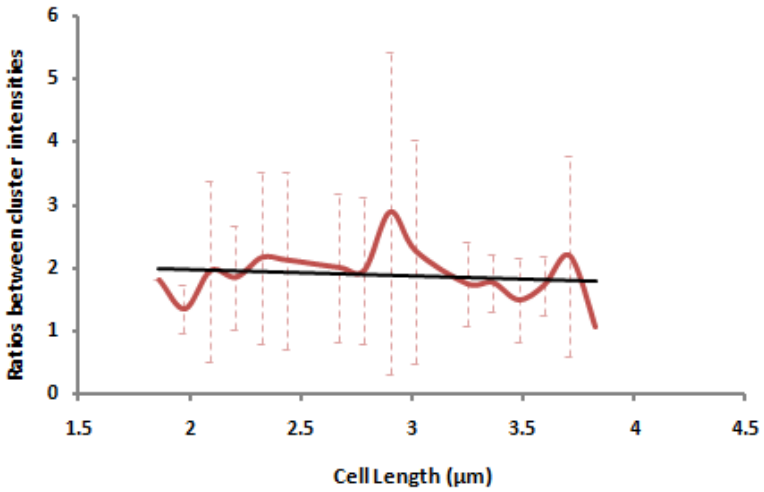


Figure 69**:** Asymmetric Cluster partitioning ratios: The average ratio between amounts of $CheW_4$-YFP protein in bright and dim cytoplasmic clusters for cells with only 2 clusters. Error bars represent plus or minus one standard deviation from the mean. The linear trend of the data is shown by a black line.

The absolute intensities of bright and dim clusters were calculated and plotted with cell length (Figure 70). Unlike polar clusters, there was a large amount of overlap in the amounts of protein found in bright and dim clusters of a population. The upwards trend of cluster intensity with cell length was similar between the two groups but was very slight, suggesting that bright and dim

clusters can accumulate new protein at similar rates, but that cluster size increases little as cells grow.

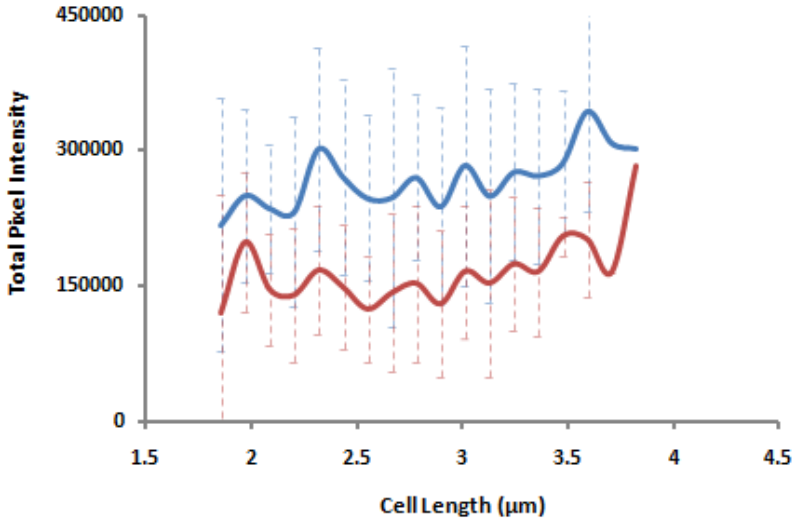


Figure 70**:** Amounts of $CheW_4$-YFP in bright and dim clusters with cell length**.** Lines show the average sum of pixel intensities for bright (blue) and dim (red) cytoplasmic clusters within each bin for cell length. Error bars represent plus or minus one standard deviation from the mean.

We see here that separated cytoplasmic daughter clusters can be similar in size, but that on average protein is asymmetrically distributed between daughter clusters by a factor of approximately 2:1. It is impossible to know when partitioning of an individual daughter cluster occurred from a single snapshot image, but analysis of a population of cells suggests that cluster partitioning happens, on average, after a cell has reached approximately 2.6 µm in length (see Figure 58), which is, on average, before cellular septation occurs. Variation of the concentration of *che*Op3 encoded proteins may offer insight into the mechanism of cytoplasmic cluster partitioning, for example, if cluster partitioning events occur at a threshold cluster size, then increasing cellular protein concentration would cause multiple-clusters with similar sizes to form within cells. This is the focus of chapters (5) and (6).

4.7) The effect of asymmetry in cluster separation

Perturbation of the chemotaxis signalling pathway of *R.sphaeroides* by mutation, yfp-fusion or deletion of individual components of either *che*Op2 or *che*Op3 has frequently resulted in a loss of chemotactic ability in mutant strains (155). This observation would suggest that the phosphorelay network controlling chemotaxis is relatively sensitive to interference. As both *che*Op2 and *che*Op3 components are essential for successful chemotaxis, it is reasonable to assume that the ratios

between the amounts of polar and cytoplasmic clustered proteins will influence the dynamics of the chemotactic response.

As both polar and cyoplasmic clusters partition unevenly between daughter cells, one mechanism for retaining similar ratios between *che*Op2 and *che*Op3 components throughout a population would be for bright cytoplasmic clusters to partition with the old cell pole, and dim cytoplasmic clusters to move to the new cell pole (Figure 71). Such an observation would also suggest that the mechanism of partitioning of the cytoplasmic cluster is actively linked to structural differences at the old and the new pole.

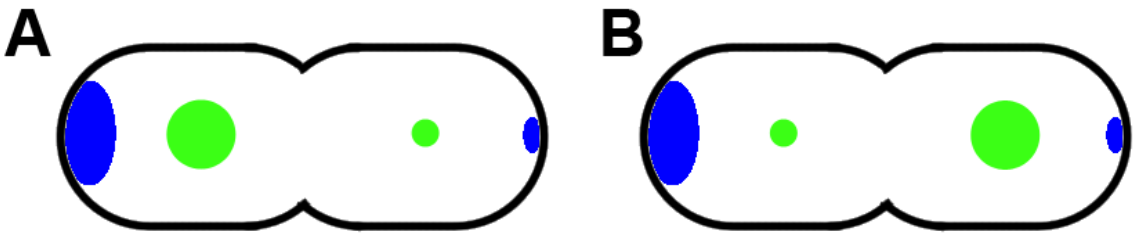


Figure 71: Inheritance of Asymmetrically partitioned clusters – Schematic diagram showing how **A)** Old-pole, Bright cytoplasmic cluster and **B)** Random partitioning of different chemotaxis protein clusters will affect the ratios of inherited components in daughter cells.

To test whether this was the case, the dual-tagged fluorescent strain JPA1447 was imaged. JPA1447 contains both $CheW_4$-YFP and CFP-$CheW_3$ genomic fusions and, like the single-fusion variants, retains full chemotactic ability on soft-agar swim plates. Subsaturating images of 398 cells from separate cultures were imaged on the same day, and background subtraction values were calculated for 80 WS8N cells captured using the same settings.

Cells were aligned by their oldest pole, and the relative positions of bright cytoplasmic clusters along the cell length were plotted for cells containing 2 cytoplasmic clusters (Figure 72).

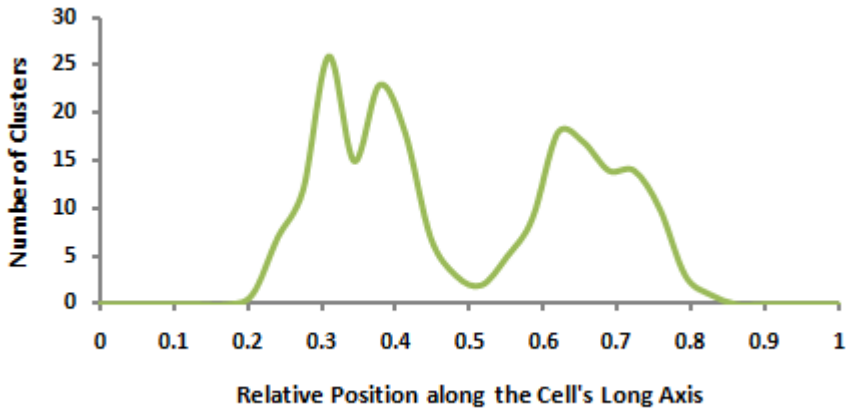


Figure 72: Alignment of bright cytoplasmic clusters with the old pole. Smoothed histogram of positions of bright cytoplasmic clusters measured from the brightest pole.

Out of 205 cells with 2 cytoplasmic clusters, 112 (55 %) bright clusters localised closer to the old pole, and 93 (45 %) localised closer to the new pole, showing that asymmetric cytoplasmic cluster separation occurs independently of the old/new pole divide. As a direct consequence of this, a population of continually dividing cells would be predicted to show a large heterogeneity in the ratios of amounts between the components of the two chemotaxis operons. Indeed, when the total concentration of protein is calculated for each cell (sum of pixel values / number of pixels in a whole cell) and $CheW_3$ concentration is plotted against $CheW_4$ concentration (Figure 73), a broad distribution of ratios is observed.

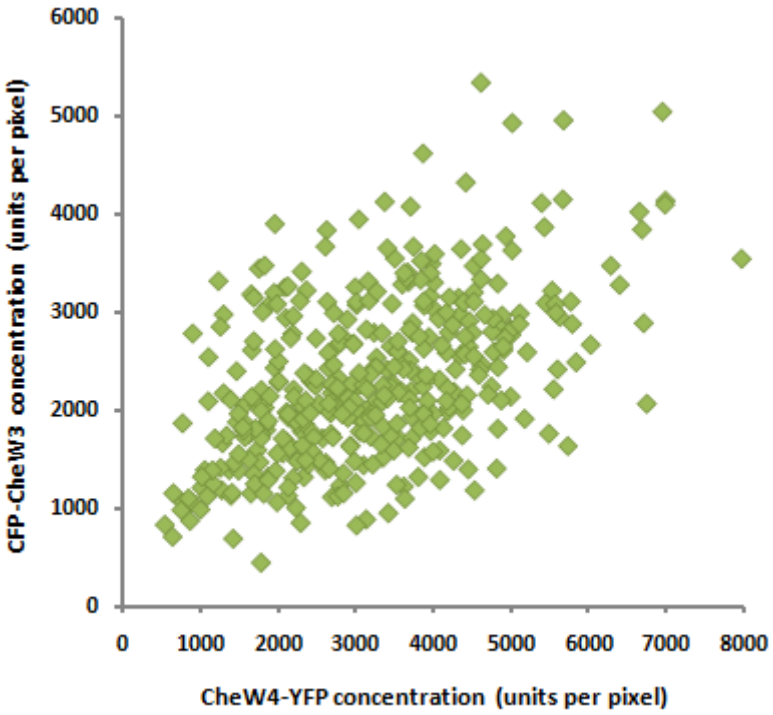


Figure 73**:** Relative concentrations of *che*Op2 and *che*Op3 components in individual cells**.** Scatter plot of the concentration of CFP-CheW$_3$ and CheW$_4$-YFP protein measured via epifluorescence from 398 JPA1447 cells.

4.8) Chapter discussion

Here, cell morphologies and protein clustering for *R.sphaeroides* populations containing functional fluorescent fusion proteins from both polar and cytoplasmic chemotaxis protein clusters were analysed using the custom software outlined in chapter (3).

Grouping cells by their length allowed assessment of how cluster numbers relate to cellular growth, and a picture of protein clustering within the average wild-type, mid-log phase cell was drawn (Figure 74).

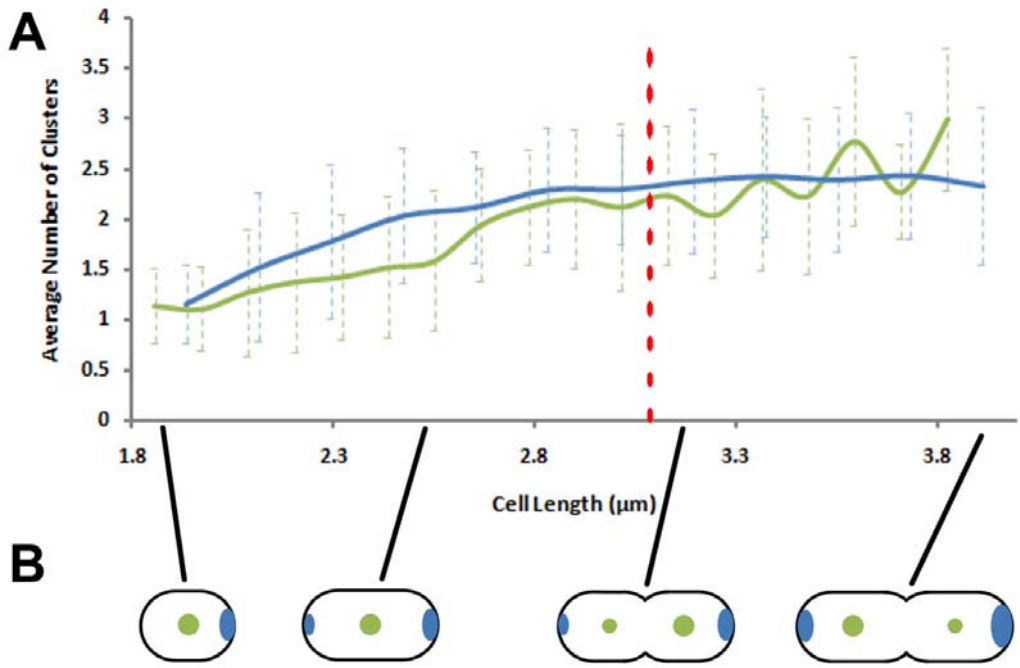


Figure 74**:** The average cell. **A)** The average number of polar/lateral (blue) and cytoplasmic (green) protein clusters are displayed. The red dashed line represents the cell length above which 80 % of cells possess visible septation sites. Error bars represent plus or minus one standard deviation from the mean. **B)** Schematic representation of the average cell at the indicated lengths, cluster sizes and positions reflect the average pattern of localisation and asymmetric separation for polar (blue) and cytoplasmic (green) clusters.

One round of cellular replication is represented by growth from approximately 1.8 µm to symmetrical division at approximately 3.6 µm. From 3.15 µm in length the majority of cells have visible septation sites. The average cell will inherit one cytoplasmic and one polar chemotaxis protein cluster after the division of its parent. This cell will develop a second polar cluster relatively early after division, and the cytoplasmic cluster will become two daughter clusters later on in the cell's growth. The two polar and two cytoplasmic clusters are separated between the two daughter cells, and the pattern repeats.

Throughout individuals of a population, discrete variation in the number of clusters is apparent, as zero, three and four-cluster cells exist for both polar and cytoplasmic clusters, however the trend in cluster inheritance ensures that the majority of newly divided cells will inherit at least one polar and at least one cytoplasmic chemotaxis protein cluster. As clustering is known to be essential to chemotaxis (155), this will ensure that the majority of cells in a mid-log phase culture are able to undergo chemotaxis immediately after cell division.

As has been observed in *E.coli*, the polar chemotaxis clusters of *R.sphaeroides* are seen to form at lateral positions within the cell, however these events are relatively rare within a mid-log phase

population and the “third” polar chemotaxis cluster is just as likely to be seen next to an existing cluster within the polar region. Interestingly, multiple clusters within a single pole were “very rare” events within *E.coli* populations (147), suggesting that the lateral diffusion dynamics of these clusters may differ between the two species, and this could be tested by time-lapse imaging of the *R.sphaeroides* lateral clusters.

The pattern of localisation observed for cytoplasmic chemotaxis protein clusters suggests that their positions are determined relative to the cell poles, and not relative to other clusters within the cell. This supports a “pull” rather than a “push” mechanism for the partitioning that is known to be mediated by the activity of the protein PpfA (150). Cells with more than two cytoplasmic clusters were seen to position the “third” cluster slightly off-centre within the cell, which may be a result of exclusion from the septation site as three-cluster cells were longer, dividing cells. The model of cluster partitioning must also explain how three-cluster patterns are achieved, and as PpfA is known to be a ParA homologue, this may be best explained by a shuttling mechanism similar to that seen in the partitioning of plasmids (see section (1.7.3)).

Measurements of the fluorescence intensities of protein clusters have shown that cluster inheritance occurs in an asymmetric manner. One daughter cell will inherit more clustered chemotaxis protein than the other, with ratios of approximately 2:1 for cytoplasmic clusters and approximately 3:1 for polar clusters, although ratios vary from cell to cell. Polar and cytoplasmic cluster inheritances are not coordinated, and therefore the ratios between cytoplasmic and polar chemotaxis protein concentrations in daughter cells vary greatly within a population as a direct result of protein clustering.

As a result of asymmetric cluster partitioning it may be predicted that the chemotaxis signalling pathway of *R.sphaeroides* would be robust to changes in the ratios between its polar and cytoplasmic components, otherwise a proportion of daughter cells would inherit an undesirable ratio of these components and be unable to undergo chemotaxis. As cluster separation cannot be controlled directly, this hypothesis would be almost impossible to test experimentally. However, computational modelling approaches may assist in our understanding of the robust nature of chemotaxis in future.

One approach for the further evaluation of both polar/lateral and cytoplasmic cluster formation and partitioning may be to vary the concentrations of the protein cluster components, as was achieved in *E.coli* by variation of the FlgM/FliA regulated expression of the chemotaxis components

(148). Manipulation of the concentrations of chemotaxis components via this method and the subsequent effects on clustering is the focus of the next two chapters.

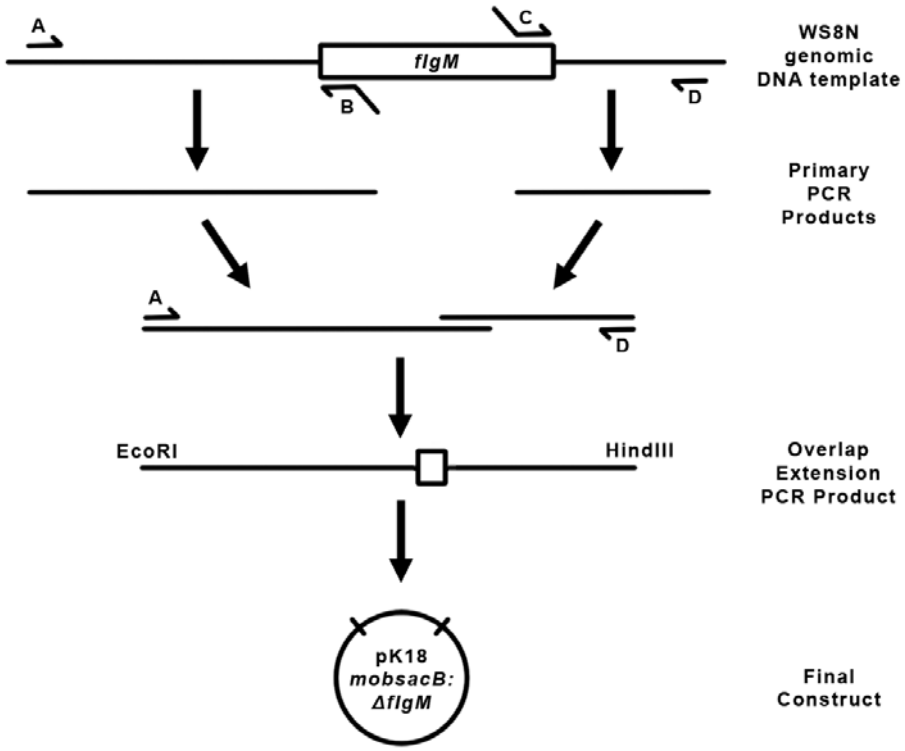
5) Regulation of expression of the chemotaxis genes by FliA and FlgM

The cytoplasmic chemotaxis protein cluster of *R.sphaeroides* is one of the only studied examples of chemotaxis proteins that localise away from the cytoplasmic membrane. Studies in other bacterial species focused on proteins that localise with transmembrane receptors to discrete polar and lateral positions within cells (discussed in section (1.6.3)). The deletion of *flgM*, the anti sigma factor, has been shown to result in an increase in the overall level of expression of the chemotaxis proteins in different bacterial species (38, 94); Studies in *E.coli* observed an 8-fold increase in expression of the chemotaxis genes, and this increase resulted in a doubling of the number of detectable protein clusters at lateral positions within the cell (147).

*che*Op2 in *R.sphaeroides*, which encodes the essential chemotaxis proteins with membrane-associated localisation, has been shown to contain a FliA promoter sequence and to have its expression regulated by FliA activity (98). Therefore, deletion of *flgM* was predicted to result in an increase in expression of *che*Op2 and hence production of components of the polar chemotaxis clusters in *R.sphaeroides*.

5.1) Deletion of *flgM*

In order not to disrupt the expression of genes either side of *flgM*, which is part of the *flgA* operon within *R.sphaeroides*, the construct for removal of *flgM* was designed to generate a chromosomal in-frame deletion which would leave behind a small gene fragment containing the first few and last few codons of the *flgM* gene. Primers were designed that incorporated 510 bp upstream of the START codon of *flgM* and 276 bp downstream of the STOP codon of *flgM*, the PCR fragments were joined via overlap extension PCR and incorporated into pK18*mobsacB* via restriction digest with *Eco*RI and *Hin*dIII followed by ligation (Figure 75). The generated plasmid, pK18*mobsacB*:*ΔflgM*, was sequenced and the sequence was checked using the Clone Manager 9 software package.



Primer	Reference Code
A	DW01
B	DW02
C	DW03
D	DW04

Figure 75 and Table 10: Deleting *flgM* - Schematic of the PCR reactions and Digest/Ligations involved in creating the pK18*mobsacB*:*ΔflgM* plasmid. The table lists the relevant primers for these reactions.

The deletion of *flgM* was incorporated into the otherwise wild-type background strain, WS8N, via the standard allelic exchange protocol (see section (2.4.13)). The successful generation of the deletion strain, JPA2001, was verified via Southern Blot analysis and DNA sequencing (data not shown).

Deletion of *flgM* did not result in an altered growth rate of the cells, suggesting that there was no detrimental effect on metabolism or cell division in the absence of this gene. Growth rate was also compared to the JPA467 (*ΔfliA*) strain, which similarly demonstrated wild-type growth rates (data not shown).

5.2) FlgM and FliA regulate *che*Op2 expression

To quantify the effect of deleting *flgM* on the expression of the components of *che*Op2, $CheW_3$ was chosen as a marker for gene expression and protein levels were calculated via western blot analysis. In order to quantify protein copy numbers known amounts of $CheW_3$ protein were also required.

To purify $CheW_3$ protein, the pQE60:*cheW₃* plasmid (Dr S.Porter) was transformed into the *E.coli*-M15 expression-optimised strain (QIAGEN). Cells were grown to an OD_{600} of 0.8, 0.1 mM IPTG was added to induce expression and cells were incubated for 6 hours at 15 °C. Cells were harvested by centrifugation and resuspended in protein wash buffer (appendix (9.2.10)). Cells were lysed by sonication and cell debris was removed by centrifugation at 15,000 x g for 30 minutes followed by filtration through a 0.4 µm filter. Protein was bound to a wash-buffer equilibrated nickel-agarose (QIAGEN) drip column, and eluted in protein elution buffer (appendix (9.2.11)). The protein concentration was calculated to be 5.80 mg ml^{-1} by Bradford Assay (see section (2.5.3)) comparison to BSA (SIGMA) protein standards, and the purity of the protein was assessed using coomassie-staining on a 15 % SDS-PAGE gel (Figure 76), and was shown to be >90 % pure and therefore pure enough for use in western-blot quantification.

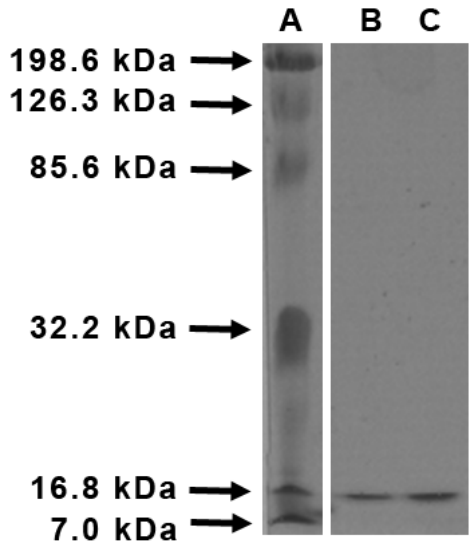


Figure 76: Purified $CheW_3$ protein**.** Coomassie-stained 12 % SDS PAGE gel (image taken with black and white camera). Lanes: A) Ladder (molecular weights are marked by arrows), B) 1 µg purified protein, C) 5 µg purified protein. The predicted molecular weight of $CheW_3$ is 18.9 kDa. Contrast has been adjusted for clarity.

Known numbers of cells were harvested from microaerobic cultures grown to OD_{700} 0.6 - 0.7, and preliminary western analysis was used to determine a suitable range of concentrations for protein standards and cell samples on the gel, as well as primary and secondary antibody concentrations. For further analysis, 1, 2, 4, 8 and 16 ng of pure protein was run in each blot. A 1/10,000 dilution of primary anti-$CheW_3$ antibody (Dr S.Porter) and a 1/20,000 dilution of secondary HRP-conjugated secondary anti-rabbit antibody (DakoCytomation) was used for blotting and Millipore Immobilon reagents were used to visualise bands in a light-emitting reaction that was observed using a Syngene G:Box gel-documentation system (Figure 77). Band intensities were quantified using the Syngene GeneTools software package.

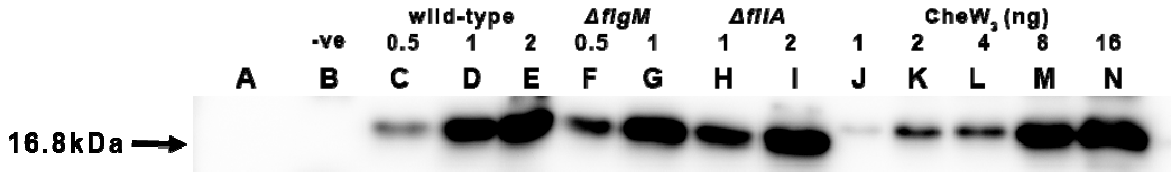


Figure 77:Example Quantitative Western Blot for $CheW_3$**:** Multiples of 10 µl of cell or protein sample that had been resuspended and boiled in 200 µl of protein loading dye were added as follows: **A)** Ladder, **B)** 2 x JPA527 (negative control), **C)** 0.5 x WS8N, **D)** 1 x WS8N, **E)** 2 x WS8N, **F)** 0.5 x JPA2001, **G)** 1 x JPA2001, **H)** 1 x JPA467, **I)** 2 x JPA467, **J)** 1 ng purified $CheW_3$, **K)** 2 ng purified $CheW_3$, **L)** 4 ng purified $CheW_3$, **M)** 8 ng purified $CheW_3$, **N)** 16 ng purified $CheW_3$. Samples were visualised using a light-emitting reaction, therefore the ladder was not visible. One ladder marker position is marked with an arrow. Pixel intensities have been inverted and contrast has been adjusted for clarity.

Quantitative western blot determination of protein copy numbers from wild-type (WS8N), *ΔflgM* (JPA2001) and *ΔfliA* (JPA467) strains revealed an increase in the expression of $CheW_3$ in the absence of *flgM* and, as expected, a decrease in expression in the absence of *fliA*, confirming that *che*Op2 is regulated by both FliA and FlgM. Copy number data are summarised in Table 11.

Strain	Description	Copy Number
WS8N	Wild-Type	10900 ± 1700
JPA2001	*ΔflgM*	15200 ± 1200
JPA467	*ΔfliA*	6000 ± 1300

Table 11: The Effect of Deleting *fliA* and *flgM* on $CheW_3$ copy number**.** Copy numbers of $CheW_3$ as calculated by quantitative western blot. The error shown is the standard error from three independently prepared samples, measured against independently prepared standard dilutions of purified $CheW_3$.

5.3) FlgM and FliA regulate *che*Op3 expression

Previous work has shown no significant change in expression of *che*Op3 upon deletion of *fliA* when studied using the activity of a promoterless *lacZ* fusion assay (98). Therefore deletion of *flgM* was predicted to result in an increase in *che*Op2 expression, and no change in *che*Op3 expression thus affecting the ratios of *che*Op2/*che*Op3 within individual cells. To test this hypothesis, $cheW_4$ was chosen as a marker of *che*Op3 expression, and copy numbers were calculated in wildtype (WS8N), *ΔflgM* (JPA2001) and *ΔfliA* (JPA467) backgrounds via western blot analysis.

As with $CheW_3$, purified $CheW_4$ was obtained via transformation of pQE60:*cheW*$_4$ (Dr S.Porter) into the expression strain of *E.coli*, M15 (QIAGEN). Expression was induced in cell cultures of OD_{600} 0.8 by the addition of 0.1 mM IPTG, and cells were harvested and lysed as before. Protein was purified on a nickel agarose drip column, as before. A final protein concentration of 40.5 mg ml^{-1} was calculated using the Bradford Assay, and the purity of the obtained protein was assessed on an SDS-PAGE gel using a coomassie stain for visualisation (Figure 78). Despite a small amount of low-molecular weight contaminant, which was visible on a coomassie stain when protein was over-loaded, the purity of the protein obtained was deemed sufficient for use in quantitative western blot analysis.

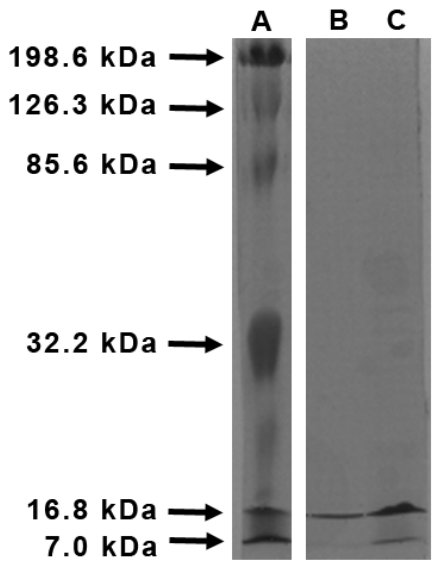


Figure 78: Purified $CheW_4$ protein. Coomassie-stained 12 % SDS PAGE gel (image taken with black and white camera). Lanes**: A)** Ladder (molecular weights are marked by arrows), **B)** 1 µg purified protein, **C)** 5 µg purified protein. The predicted molecular weight of CheW4 protein is 18.0 kDa. Contrast has been adjusted for clarity.

Preliminary western-blot analysis was used to determine suitable protein, cell sample and antibody dilutions. A 1/5,000 dilution of primary anti-$CheW_4$ (Dr S.Porter) and a 1/20,000 dilution of secondary HRP-conjugated anti-rabbit antibody were used for blotting. Blots were visualised as for $CheW_3$ (Figure 79).

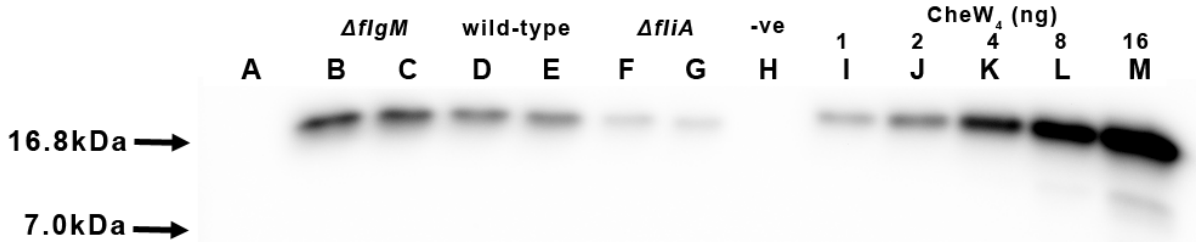


Figure 79: Example Quantitative Western Blot for $CheW_4$: Multiples of 10 µl of cell or protein sample that had been resuspended and boiled in 1000 µl of protein loading dye were added as follows: **A)** Ladder, **B)**1 x JPA2001, **C)** 1 x JPA2001, **D)** 1 x WS8N, **E)** 1 x WS8N, **F)** 1 x JPA467, **G)** 1 x JPA467, **H)** 2 x JPA1399 (negative control), **I)** 1 ng purified $CheW_4$, **J)** 2 ng purified $CheW_4$, **K)** 4 ng purified $CheW_4$, **L)** 8 ng purified $CheW_4$, **M)** 16 ng purified $CheW_4$. Samples were visualised using a light-emitting reaction, therefore the ladder was not visible. One ladder marker position is marked with an arrow. Pixel intensities have been inverted and contrast has been adjusted for clarity.

Unexpectedly, quantitative determination of protein copy numbers revealed an increase in protein expression levels of $CheW_4$ in the absence of FlgM and a decrease in expression in the absence of FliA, showing that FlgM and FliA regulate expression of *che*Op3. Copy number data is summarised in Table 12.

Strain	Description	Copy Number
WS8N	Wild-Type	8600 ± 1200
JPA2001	*ΔflgM*	13200 ± 1300
JPA467	*ΔfliA*	1500 ± 200

Table 12: The Effect of Deleting *fliA* and *flgM* on $CheW_4$ copy number. Copy numbers of $CheW_4$ as calculated by quantitative western blot. The error shown is the standard error from three independently prepared samples, measured against independently prepared standard dilutions of purified $CheW_4$

5.4) Discussion:

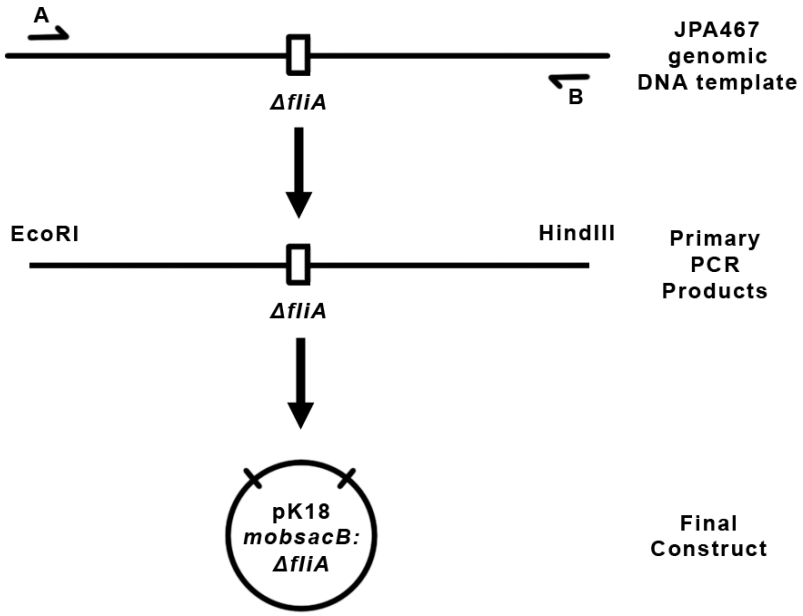
These data show that both *cheOp2* and *cheOp3* are regulated by FliA and FlgM. Although *che*Op3 was previously thought not to be regulated by FliA, $CheW_4$ expression changed more than $CheW_3$

expression in both *ΔfliA* and *ΔflgM* strains. The discrepancy with previous data may be caused by differences in growth conditions between the setups of two different laboratories. Better aeration of microaerobically grown samples in the new laboratory is achieved in an enclosed orbital shaker, and changes in gene expression relating to growth conditions may occur as a result of this difference.

5.5) Studying gene expression using fluorescence

To further study the effects of FlgM and FliA activity on chemotaxis protein concentrations, and chemotaxis protein localisation, *flgM* and *fliA* were deleted from YFP-CheW$_3$ (JPA1418) and CheW$_4$-YFP (JPA1457) fluorescent fusion background strains.

The original pK18*mobsacB*:*ΔfliA* construct had been lost, and had to be remade. In order to replicate the exact deletion in JPA467, chromosomal DNA was extracted from this strain and used as a PCR template. This allowed for a simple one-step digest and ligation (Figure 80). Successfully remade pK18*mobsacB*:*ΔfliA* was sequenced and its sequence checked using the Clone Manager 9 software package.



Primer	Reference Code
A	DW05
B	DW06

Figure 80 and Table 13: Deleting *fliA* - Schematic of the PCR reactions and Digest/Ligations involved in creating the pK18*mobsacB*:*ΔfliA* plasmid. The table lists the relevant primers for these reactions.

Deletion strains *yfp-cheW$_3$*, *ΔflgM* (JPA2018), *yfp-cheW$_3$*, *ΔfliA* (JPA2022), *cheW$_4$-YFP, ΔflgM* (JPA2019) and *cheW$_4$-YFP, ΔfliA* (JPA2023) were generated using the standard allelic exchange protocol, and successful strain isolation was confirmed via either southern blot or DNA sequencing.

The growth rates of JPA1418, JPA1457, JPA2018, JPA2019, JPA2022 and JPA2023 were not altered from wild- type (WS8N), showing that the combination of YFP-fusion and expression changes caused by the deletion of *flgM* and *fliA* had no affect on metabolism or cell division in these strains.

5.6) Development of a fluorimetry-based protocol for measuring copy number

As the further study of protein expression levels would involve multiple different growth conditions in a variety of different background strains, it was appealing to use a technique other than western blot analysis for the quantification of protein levels as western blots are relatively error prone, and thus extremely labour intensive and time consuming. Therefore a quantitative technique similar to that of western blot analysis was developed based around direct measurement of the YFP signal generated by fluorescent-fusion proteins and comparison to the signal generated by known quantities of purified YFP.

Cell preparations were grown under standard microaerobic laboratory conditions, to an OD_{700} of between 0.6 and 0.7, and fractions of liquid were collected that contained the equivalent amount of cells as 200 µl, 1 ml, 2 ml and 4 ml of cell cultures at exactly OD_{700} = 0.6. Cells were collected from these fractions via centrifugation at 13,000 xg and any excess medium was carefully removed from the cell pellet with a pipette. The harvested cells were vigorously resuspended with a pipette in 200 µl of fresh succinate medium generating 1x, 5x, 10x and 20x concentrations of cell culture, which were then placed into a 96-well plate for measurement in the plate reader. A Corning COSTAR 96-well plate was used. These have a glass bottom and black plastic sides, preventing light from scattering across samples and allowing samples to be both illuminated and scanned from underneath resulting in maximum signal detection.

For the preparation of the YFP-standards, purified YFP protein of a known concentration (a gift from George Wadhams) was diluted into 1x, 5x, 10x and 20x concentrations of wild-type WS8N cells (which contain no YFP) to generate protein concentrations representative of a wide range of cellular protein copy numbers for each preparation. Fluorescence signal was measured for these samples, and the plate reader gain setting was adjusted to a value of 2200, which gave the optimal dynamic range for these measurements. Multiple (at least 5) readings were taken of each well in order to average for machine noise. Each measurement was repeated 3 times with different cell preparations

and separately prepared dilutions of YFP protein. Values obtained for the same number of wild-type cells containing no fluorescent protein were subtracted as a zero value. A strong positive correlation was observed for copy number versus fluorescence value obtained (Figure 81), the R^2 values for a straight line fitted through the origin were 0.987, 0.999, 0.990 and 0.994 for 1x, 5x, 10x and 20x cells respectively, showing that fluorescence can be used as a good estimate of copy number. All estimations of copy number using this technique were done measuring samples of the correct cell density (1x, 5x, 10x and 20x) based on the OD_{700} reading, and by fitting the measured value to the relevant YFP standard curve (below).

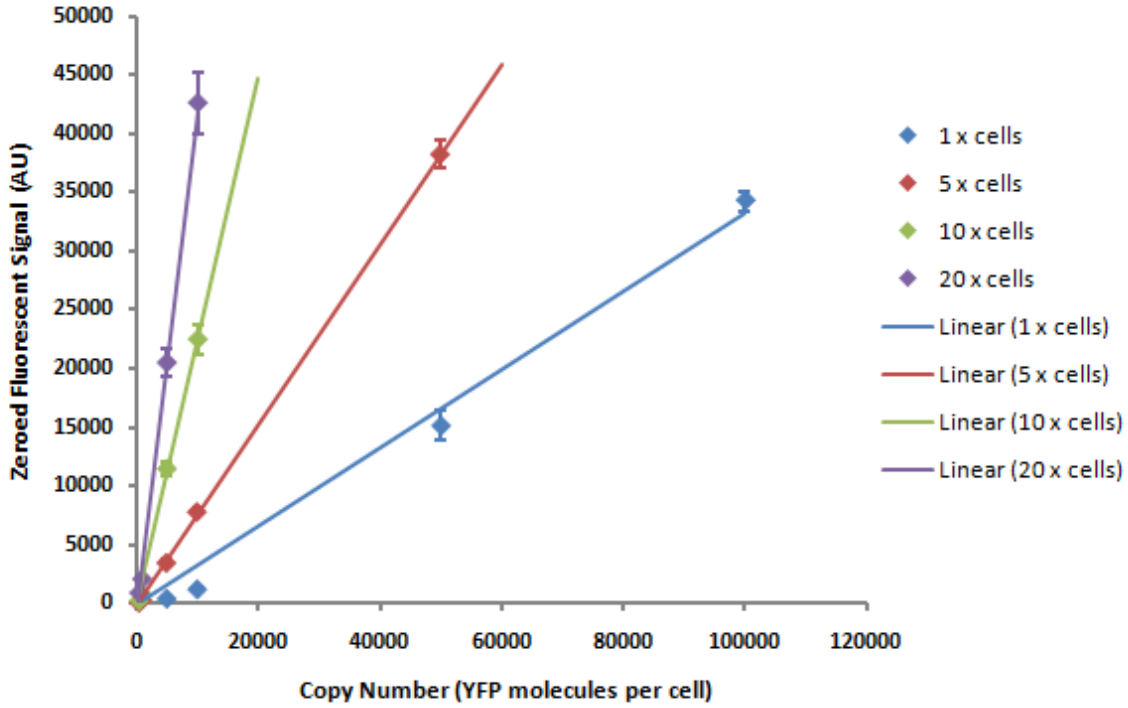


Figure 81**: Fluorescence signal with YFP copy number:** The average fluorescence signal from three independently prepared samples of 1 x (blue), 5 x (red), 10 x (green) and 20 x (purple) WS8N cells resuspended in different concentrations of YFP protein. YFP protein concentrations were plotted as the representative copy number (total number of YFP molecules/total number of WS8N cells). Error bars represent the standard deviation of mean. Best fit lines are plotted through the origin.

From these standards, an approximation of the range of copy numbers detectable via this technique was estimated based of the relative sizes of the error bars of each line and the saturation value of the detection system (Table 14). The large dynamic range of this technique (~400 – ~150,000 copies per cell) is comparable to that of western blot analysis (which also relies on the quality of the antibody used and the sensitivity of the detection system), and covers a large

proportion of the physiological protein copy numbers found in many bacteria as well as the sorts of protein copy numbers that are achieved in over-expression experiments.

	Measurable Copy Numbers	
	Lower Limit	**Upper Limit**
1 x cells	~6000	~150000
5 x cells	~3000	~75000
10 x cells	~1000	~25000
20 x cells	~400	~12000

Table 14: Dynamic Range of YFP fluorescence quantification: Estimates of the lower and upper bounds of detection of YFP copy numbers in the FLUOSTAR Optima plate reader with a gain setting of 2200 are displayed for the different measured concentrations of cell sample.

The validity of this technique as a method of quantification of cellular copy numbers is based around a few assumptions;

1) The proportion of purified protein that does not fluoresce is minimal.
2) The proportion of *in vivo* YFP protein in the un-matured non-fluorescent state is minimal.
3) Diffuse extracellular YFP gives a similar signal to intracellular YFP.

If any of these assumptions are wrong then a systematic under or over-estimate would be expected for protein copy numbers calculated via fluorimetry compared to quantitative western blot estimation. Therefore, to test the accuracy and validity of quantification via fluorimetry, numbers calculated using the two techniques were compared for all proteins where data were available (Figure 82). Numbers calculated using the two separate techniques were of comparable magnitude, and no systematic disagreement was detected, suggesting that the two techniques produced equivalent data.

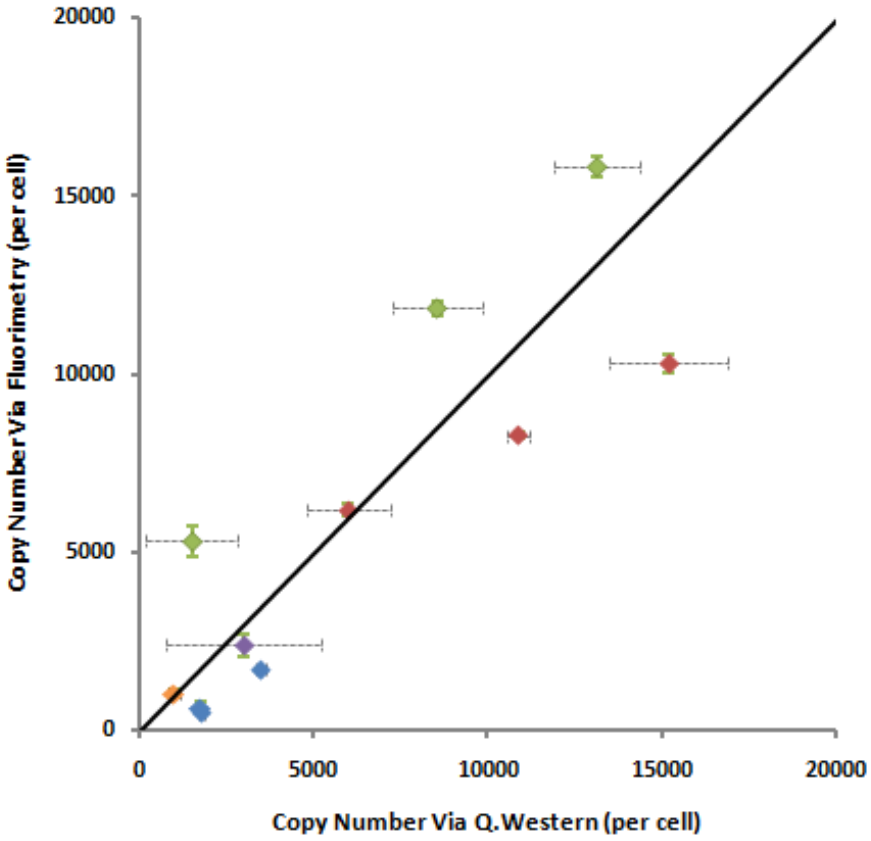


Figure 82: Fluorescence versus western copy numbers: Scatter plot of copy numbers where data were available using both western and fluorescence quantification. Data are included for MotB (JPA2006, JPA2038, and JPA2039 – blue, this study) , $CheW_3$ (JPA1418, JPA2018 and JPA2022 – red, this study), $CheW_4$ (JPA1457, JPA2019 and JPA2023 – green, this study), $CheB_2$ (JPA1444 – purple, (98) and this study) and $CheY_3$ (JPA1629 – orange, (98) and this study). Error bars represent the standard errors obtained for at least three independent repeats. A perfect correlation of 1:1 between the two techniques would give results that fall on the black line.

The genetic fusion of YFP has been seen to affect expression levels when introduced *in vivo*, and therefore the differences observed between these two techniques may be due to changes in expression caused by the introduction of YFP (Figure 82).

Interestingly, although the samples for both western blots and fluorimetry measurements were prepared in much the same way, the values obtained by fluorimetry were far more consistent across repeats, leading to very tight error bars. This suggests that the relatively large error seen for western measurements is not a result of sample variation, but is instead a representation of the experimental error that is introduced by the sample preparation, transfer and blotting stages of the protocol (Figure 83).

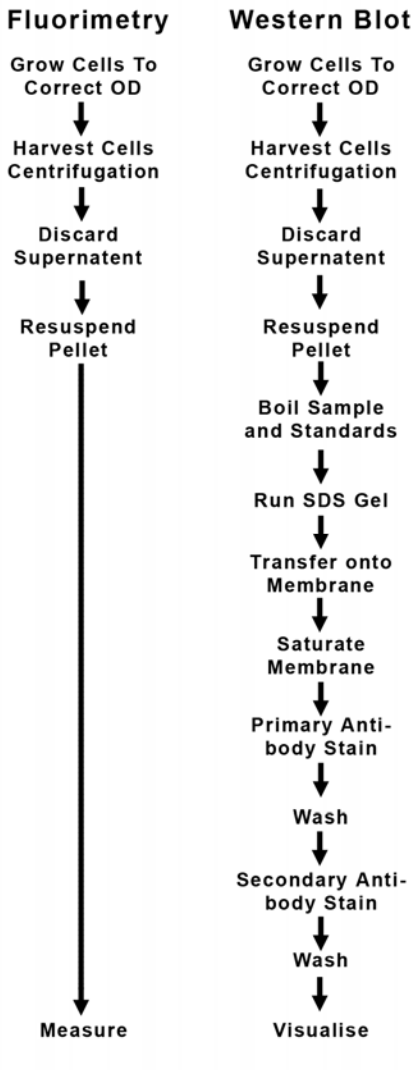


Figure 83: Comparing western and fluorimetry protocols. Schematic representation of the steps involved in the experimental procedures of Fluorimetry and Western Blot analysis. The extra steps in western blot analysis are proposed as the potential source of error that accounts for the difference in measured error between results from the two techniques.

5.7) The effect of FliA and FlgM activity on *che*Op2 and *che*Op3 expression:

The changes in expression of the chemotaxis operons that were seen in the *ΔflgM* and *ΔfliA* background strains were reanalysed for the YFP-$CheW_3$ and $CheW_4$-YFP fluorescent fusion strains using fluorimetry. Increases in copy number were seen for both YFP-$CheW_3$ and $CheW_4$-YFP in the absence of *flgM*. Decreases in copy number were seen for both YFP-$CheW_3$ and $CheW_4$-YFP in the absence of *fliA*. Copy number data are summarised in Table 15. As seen using western blots (see sections (5.2) and (5.3)), the effect on *che*Op3 is greater than on *che*Op2.

	Copy Numbers			Fold Change from WT	
Protein	**Wild-Type**	***ΔfliA***	***ΔflgM***	***ΔfliA***	***ΔflgM***
YFP-$CheW_3$	8300 ± 300	6200 ± 200	10300 ± 200	0.75	1.25
$CheW_4$-YFP	11900 ± 300	5300±100	15800 ± 400	0.45	1.33

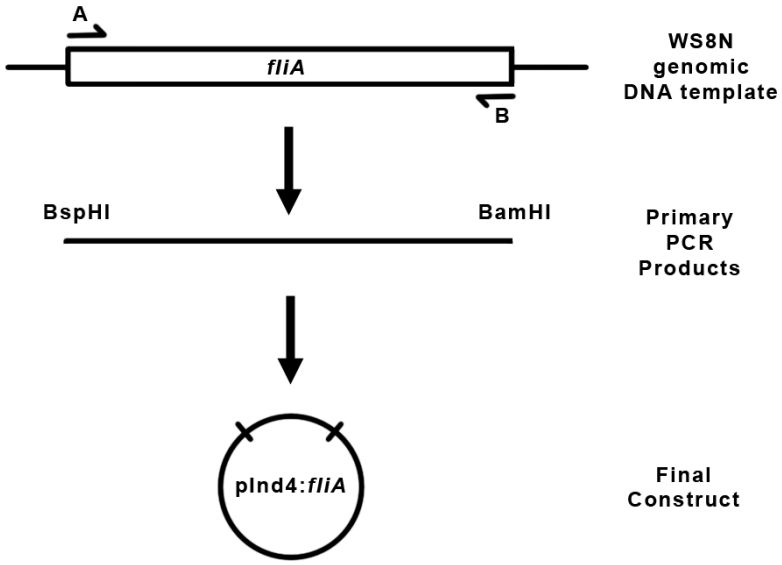
Table 15**:** YFP-$CheW_3$ and $CheW_4$-YFP copy numbers in wild-type, *ΔfliA* and *ΔflgM* backgrounds: Protein copy numbers calculated by the fluorescence-based technique discussed in section (5.6). Error values are the standard error of three independently prepared repeats. Numbers are also summarised as fold-changes from the wild-type parental strain

5.7.1) Complementing the deletion of *fliA*

One important aspect of the analysis of any chromosomal deletion mutant is determining whether or not that deletion can be complemented by the reintroduction of that gene on an inducible expression vector. Reintroduction of the deleted gene without the restoration of that gene's encoded activity may be an indication that the changes made to the chromosome have had a knock-on effect on the activity or expression of genes surrounding the deletion, or may indicate a more complex epigenetic effect where the correct function of a gene is dependent on its coordinated expression.

Until recently complementation studies have been difficult in *R.sphaeroides*, as no inducible expression vector was available. The pInd4 IPTG inducible expression vector became available in 2009, and allows the reintroduction and controlled expression of genes within *R.sphaeroides* (65).

In order to incorporate *fliA* into pInd4, primers were designed that incorporated the *Bam*HI and *Bsp*HI restriction sites at the START and STOP codons of the gene. Digestion of the PCR fragment with these enzymes and of pInd4 with *Bam*HI and *Nco*I allowed in-frame incorporation of this gene through ligation (Figure 84). Successfully generated pInd4:*fliA* was sequenced and its sequence was checked using the Clone Manager 9 software package.



Primer	Reference Code
A	DW07
B	DW08

Figure 84 and Table 16: Complementing *ΔfliA* - Schematic of the PCR reactions and Digest/Ligations involved in creating the pInd4:*fliA* plasmid. The table lists the relevant primers for these reactions.

pInd:empty and pInd:*fliA* were conjugated into the YFP-CheW$_3$ and CheW$_4$-YFP fusion strains in the presence and absence of *fliA* (JPA1418, JPA1457, JPA2022 and JPA2023) to see whether the effect of *fliA* deletion could be reversed by reintroduction of *fliA*. Cells were grown in the presence of kanamycin to ensure that the pInd vectors were retained within the cells and copy numbers were calculated by fluorimetry. The introduction of pInd:empty into the parental strains did not affect the copy number of YFP-CheW$_3$ or CheW$_4$-YFP proteins. Even in the absence of IPTG, leaky expression from the promoter on the introduction of pInd:*fliA* was sufficient to restore wild-type levels of expression to the deletion backgrounds, showing that FliA activity can be restored by complementation from pInd4. Copy number data is summarised in Table 17 and 18.

Strain	Genotype Description	$CheW_3$-YFP Copy Number ($cell^{-1}$)
JPA1418	*yfp-cheW$_3$*	8300 ± 300
JPA1418 + pInd:empty	*yfp-cheW$_3$*	8700 ± 300
JPA2022 + pInd:empty	*yfp-cheW$_3$*, *ΔfliA*	7000 ± 300
JPA2022 + pInd:*fliA*	*yfp-cheW$_3$*, *ΔfliA*	9000 ± 300

Strain	Genotype Description	$CheW_4$-YFP Copy Number ($cell^{-1}$)
JPA1457	*cheW$_4$-yfp*	11900 ± 300
JPA1457 + pInd:empty	*cheW$_4$-yfp*	12700 ± 700
JPA2023 + pInd:empty	$CheW_4$-yfp, *ΔfliA*	5400 ± 900
JPA2023 + pInd:*fliA*	$CheW_4$-yfp, *ΔfliA*	13000 ± 2000

Table 17 and Table 18**:** YFP-$CheW_3$ and $CheW_4$-YFP copy numbers with pInd:empty and pInd:*fliA***:** Copy numbers were calculated via fluorimetry, error bars represent the standard error of three independently prepared repeats. No induction of gene expression was induced, as no IPTG was added to these samples.

As both *che*Op2 and *che*Op3 components were seen to be regulated by FliA and FlgM, overexpression of *fliA* was predicted to result in an increase in expression of both YFP-$CheW_3$ and $CheW_4$-YFP. To test this hypothesis, cells were grown in the presence of kanamycin and different levels of expression of *fliA* were achieved by the addition of different amounts of IPTG (Figure 85).

Levels of YFP-$CheW_3$ in JPA2022(pInd:*fliA*) did not significantly increase above those seen in JPA1418 below IPTG concentrations of 10 μM, however statistically significant increases corresponding to 1.1, 2.6 and 3.6 – fold increases in expression were seen at IPTG concentrations of 10 μM, 50 μM and 100 μM, respectively (Figure 85).

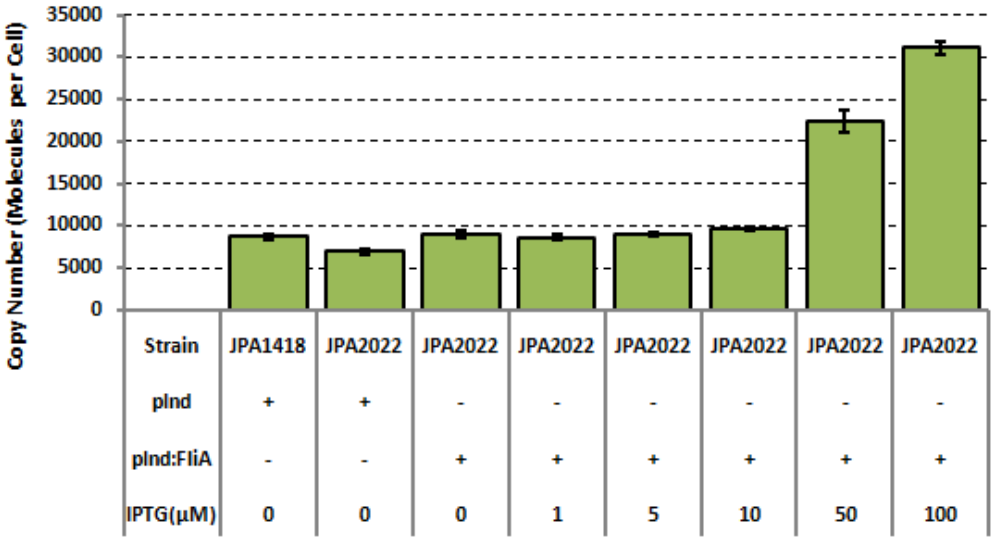


Figure 85: Effect of overexpression of FliA on YFP-CheW$_3$ copy number**:** YFP-CheW$_3$ copy number calculated via fluorimetry at different expression levels of FliA in an otherwise wild-type (JPA1418) or in a *fliA* deletion background (JPA2022). Error bars represent the standard error of three independent repeats.

Similarly, the levels of CheW$_4$-YFP in JPA2023(pInd:*fliA*) did not significantly increase at the lower induction levels of IPTG, however 2.5 and 4.3-fold increases in expression were observed at IPTG concentrations of 50 μM and 100 μM, respectively (Figure 86).

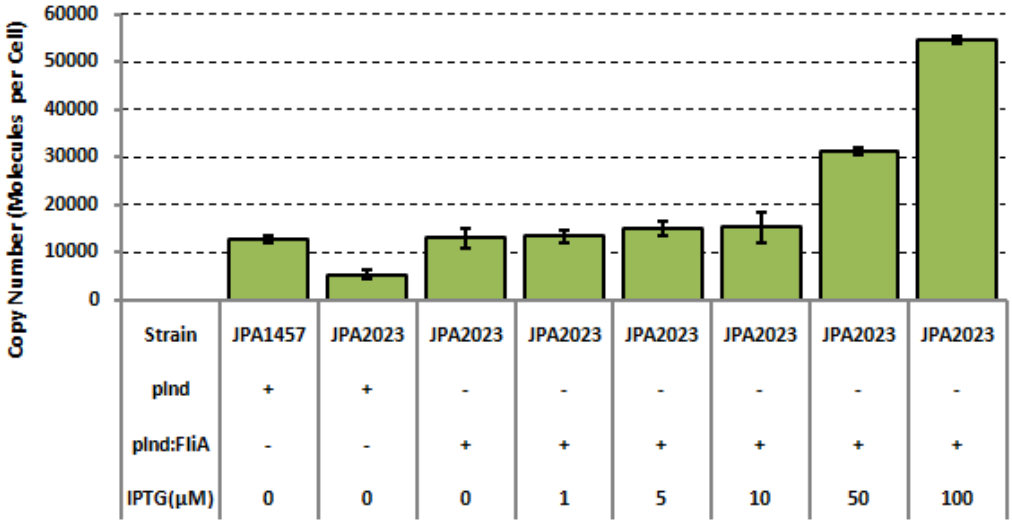


Figure 86: Effect of overexpression of FliA on CheW$_4$-YFP copy number**:** CheW$_4$-YFP copy number calculated via fluorimetry at different expression levels of FliA in an otherwise wild-type (JPA1457) or in a *fliA* deletion background (JPA2023). Error bars represent the standard error of three independent repeats.

Increases in FliA activity at 50 and 100 µM concentrations of IPTG were also accompanied by reductions in the rate of growth in all cellular backgrounds (Figure 87), demonstrating toxicity of FliA in high concentrations.

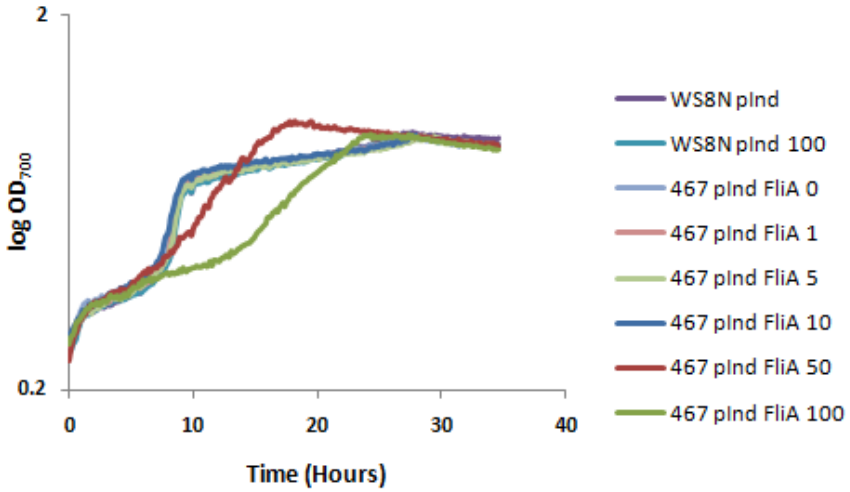
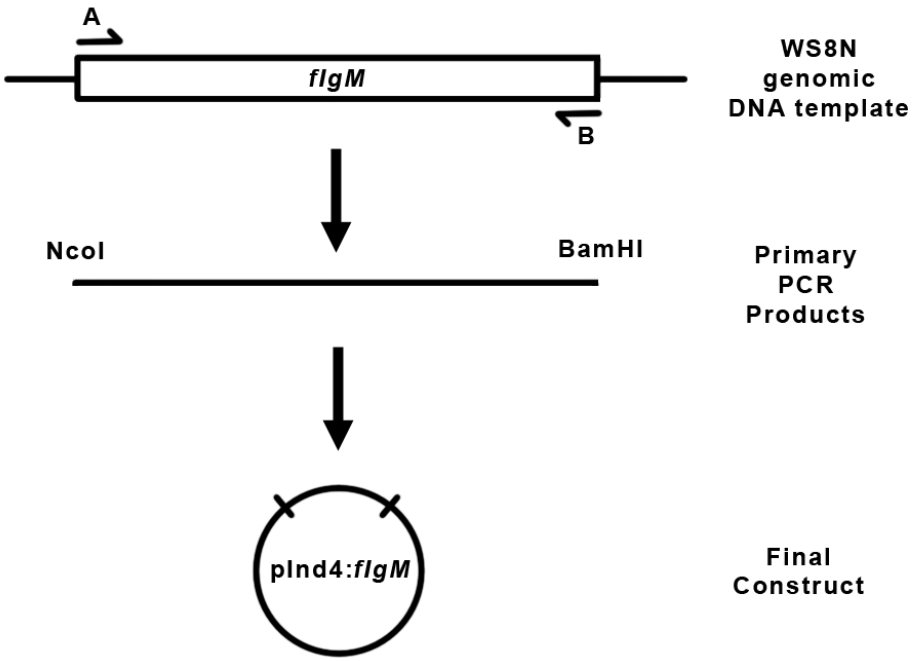


Figure 87: Effect of FliA overexpression on growth: Example growth curves of cells expressing FliA at different concentrations of IPTG. Wild-type (WS8N) and *ΔfliA* (JPA467) strains are displayed. Curves represent the average of three repeats, error bars are left off for clarity.

5.7.2) Complementing the *flgM* deletion

In order to see whether the effect of deletion of the *flgM* gene could be complemented by reintroduction of *flgM* on a plasmid, primers were designed to incorporate the *Nco*I and *Bam*HI restriction sites at the start and stop positions of the *flgM* gene. Digestion of the obtained PCR product with these enzymes, and digestion of pInd4 with *Nco*I and *Bam*HI allowed an in-frame expression construct to be synthesised for the expression of *flgM* (Figure 88). Successfully generated pInd4:*flgM* was sequenced and its sequence was verified using the Clone Manager 9 Software package.



Primer	Reference Code
A	DW09
B	DW10

Figure 88 and Table 19: Complementing *ΔflgM* - Schematic of the PCR reactions and Digest/Ligations involved in creating the pInd4:*flgM* plasmid. The table lists the relevant primers for these reactions.

pInd:empty and pInd4:*flgM* were conjugated into JPA1418, JPA1457, JPA2018 and JPA2019, cell cultures were grown containing kanamycin and *flgM* was reintroduced into the deletion background strains at varying levels by the addition of different concentrations of IPTG. Copy numbers were determined in these strains via fluorimetry.

Levels of YFP-$CheW_3$ were not restored to wild-type levels under any of the tested concentrations of IPTG, and the copy numbers under these conditions remained comparable to that of the *flgM* deletion strain, JPA2018 (Figure 89).

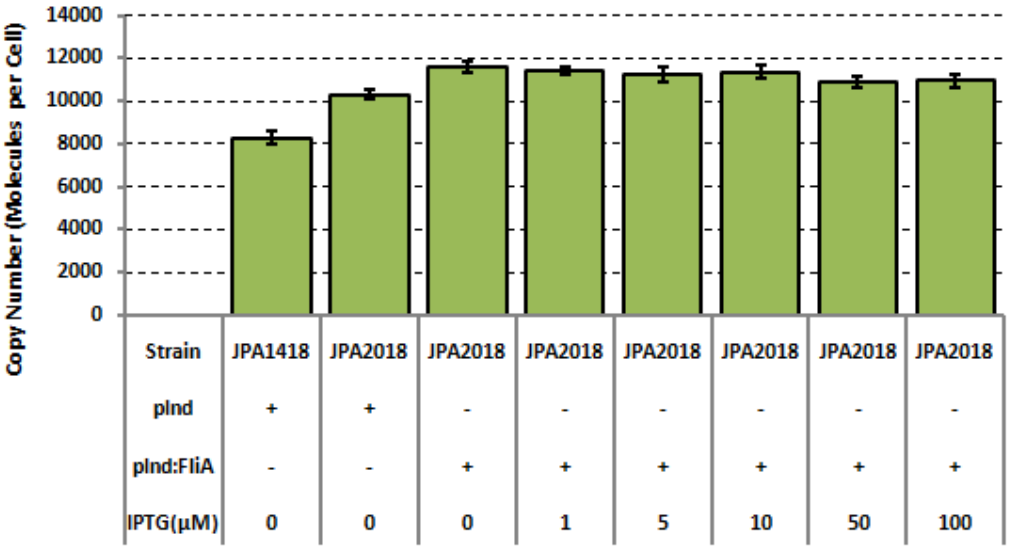


Figure 89: Effect of overexpression of FlgM on YFP-CheW$_3$ copy number: YFP-CheW$_3$ copy number calculated via fluorimetry at different expression levels of FlgM in an otherwise wild-type (JPA1418) or in a *flgM* deletion background (JPA2018). Error bars represent the standard error of three independent repeats.

Similarly, when levels of CheW$_4$-YFP were calculated at the same concentrations of IPTG, no recovery of the *flgM* deletion phenotype was observed, and expression levels of CheW$_4$-YFP remained statistically above wild-type (Figure 90).

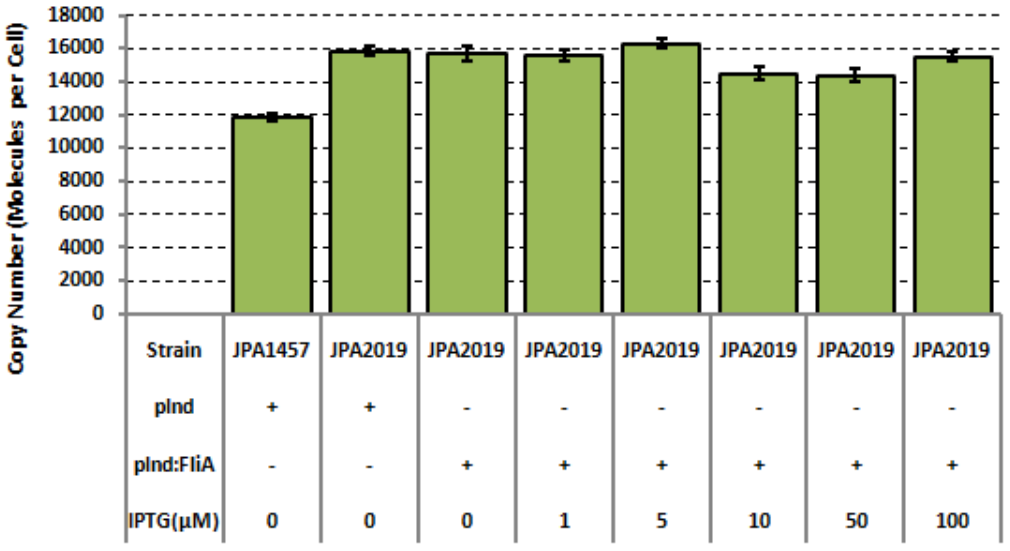


Figure 90: Effect of overexpression of FlgM on CheW$_4$-YFP copy number**:** CheW$_4$-YFP copy number calculated via fluorimetry at different expression levels of FlgM in an otherwise wild-type (JPA1457) or in a *flgM* deletion background (JPA2019). Error bars represent the standard error of three independent repeats.

Taken together these observations suggest that the *flgM* deletion is not complemented under these sets of conditions, therefore it is inconclusive as to whether the changes in gene expression observed in the *flgM* deletion mutant is directly linked to the activity of the *flgM* gene.

However, during the preliminary experiments for the development of a fluorescence-based technique for the measurement of protein levels (section 5.6), many experiments, including the *flgM*-complementation experiment were carried out where cells were grown at 30 °C within the plate-reader itself. There were several limitations to this technique due to machine sensitivity, and it was abandoned as a general method for fluorescence analysis.

Despite the limitations of this technique, data was obtained for YFP-$CheW_3$ and $CheW_4$-YFP levels where *flgM* was deleted and then reintroduced. Fluorescence values obtained from non-fluorescing WS8N cells were subtracted from those of the test strains, and the resultant fluorescent intensity was corrected for cell number by normalisation based on the optical density of the culture at that time. The final value gave a representation of the relative concentration of fluorescent protein within cells of that sample.

Under these sets of conditions, changes in expression of *che*Op2 and *che*Op3 were, again, seen to result from deletion of *flgM* and *fliA* and were in agreement with previous results (section (5.2)). Between 0 and 20 hours, the fluorescence per cell was seen to increase, then to plateau as cells reached stationary phase after 20 hours of growth. In the presence of 50 µM IPTG, both YFP-$CheW_3$ levels and $CheW_4$-YFP wild-type expression levels were seen to be restored to the *flgM* deletion mutants by the introduction of the pInd4:*flgM* vector (Figure 91).

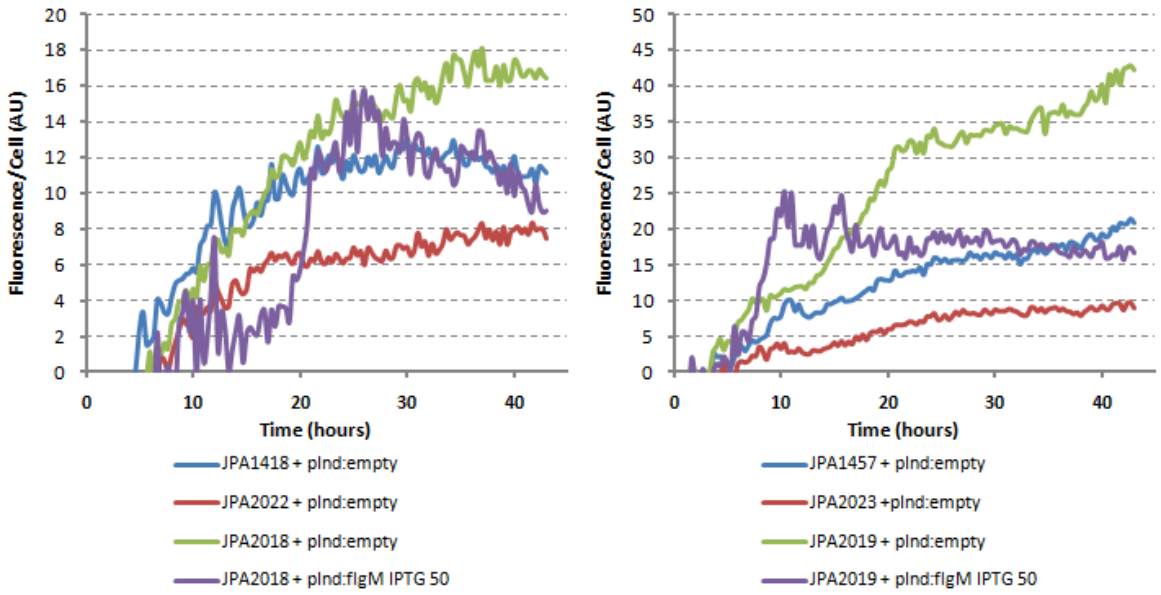


Figure 91: Complementing YFP-CheW$_3$ and CheW$_4$-YFP expression levels in the *ΔflgM* background: Fluorescence per cell was calculated as: Total fluorescent signal minus that obtained in a WS8N non-fluorescent background, then normalised for cell number. Values represent the average obtained for dual repeats.

Importantly, the ability to complement the *flgM* deletion (even only under a certain set of growth conditions) shows that the *flgM* gene itself is responsible for the increase in chemotaxis gene expression observed in the deletion mutant. The most likely explanation for the discrepancy between those data obtained via the two fluorimetry techniques is that the level of expression achieved from the pInd4:*flgM* vector is low and insufficient to match wild-type levels of *flgM* within cells grown quickly under microaerobic conditions, whereas slow growth in the plate reader allowed for sufficient expression of *flgM*. In support of this hypothesis, the growth rates of cells grown in the plate-reader whilst expressing *flgM* in the presence of 100 µM IPTG were reduced in comparison to the same cells, but grown under standard microaerobic conditions (Figure 92), suggesting that expression achieved in the plate reader was higher.

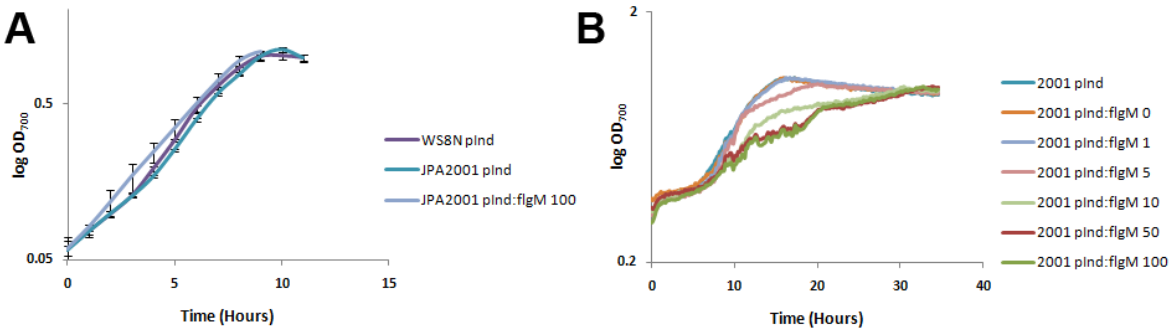
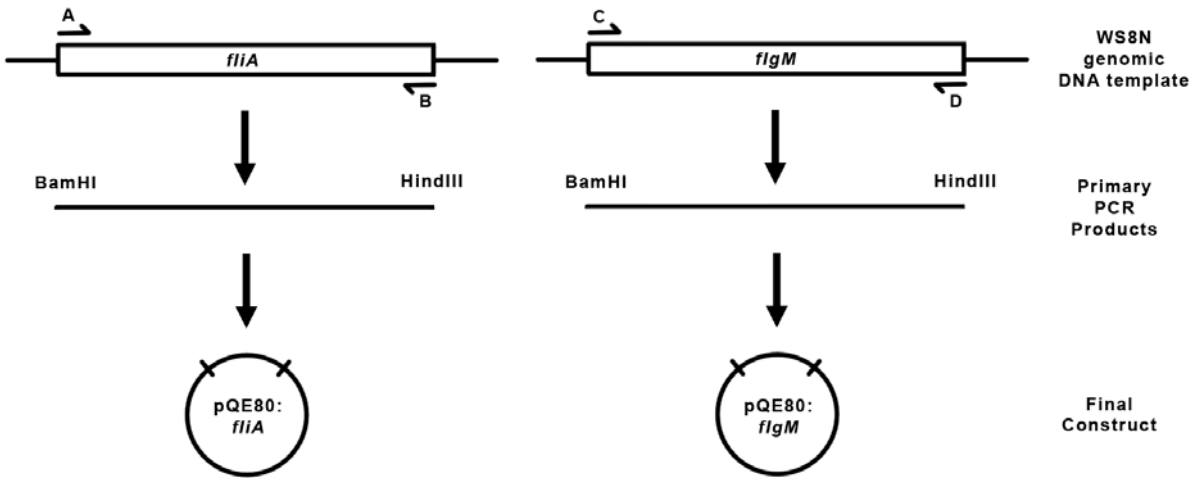


Figure 92: Expression of pInd:*flgM* under different growth conditions. The effect on growth of expressing FlgM from pInd4:*flgM* at different concentrations of IPTG under **A)** microaerobic conditions, values represent the average of three repeats, bars represent the standard deviation of three repeats. **B)** Plate reader conditions, values represent the average of three repeats, error bars are left off for clarity.

Primers were designed which allowed for the successful incorporation of both *flgM* and *fliA* into the pQE80 his-tag purification expression vector (Figure 93). Unfortunately, attempts to purify FlgM and FliA protein and produce anti-FlgM and anti-FliA antibodies were unsuccessful. Therefore, no method was available for the direct measurement of FlgM or FliA concentrations.



Primer	Reference Code
A	DW11
B	DW12
C	DW13
D	DW14

Figure 93 and Table 20: Purifying FlgM and FliA - Schematic of the PCR reactions and Digest/Ligations involved in creating the pQE80:*fliA* and pQE80:*flgM* plasmids. The table lists the relevant primers for these reactions.

5.7.3) The effect of FlgM on the activity of FliA:

FlgM has been demonstrated to competitively inhibit the binding of FliA to DNA by attachment to the DNA-binding motif of FliA in *E.coli* (67), this interaction has also been demonstrated in the FliA-FlgM co-crystal structure from *Aquifex aeolicus* (144). Sequence similarities between *E.coli* and *R.sphaeroides* suggest that the FliA and FlgM proteins of *R.sphaeroides* would behave in the same manner; however, no direct evidence is available for this behaviour in *R.sphaeroides*.

If FlgM is acting to inhibit FliA, then the increase in expression at FliA-regulated genes observed in FliA over-expression studies would be predicted to be greater in the absence of FlgM than in the presence of FlgM. Unfortunately attempts to generate *ΔfliA*, *ΔflgM* double chromosomal deletion strains were unsuccessful, so in order to test this hypothesis and to assess the magnitude of the effect of FlgM inhibition on FliA activity, pInd4:*fliA* was transformed into JPA1418 and JPA2018. The level of expression of *fliA* in these strains was varied by adding different concentrations of IPTG and the concentrations chosen were adjusted based on the results that had previously been observed in the *ΔfliA* complementation experiments (5.7.1). Copy numbers of YFP-$CheW_3$ were measured by fluorimetry in mid-log phase cells under these conditions.

When plotted as absolute values (Figure 94A), the copy numbers seen under identical growth conditions were higher in the *ΔflgM* background than in the wild-type background, supporting an increase in expression of YFP-$CheW_3$ in the absence of *flgM*. When the same numbers were displayed relative to the levels of expression obtained from the parental background strains (Figure 94B), the increase in expression seen at each concentration of IPTG was seen to be greater in the absence of *flgM*, supporting the theory that FlgM is an inhibitor of FliA.

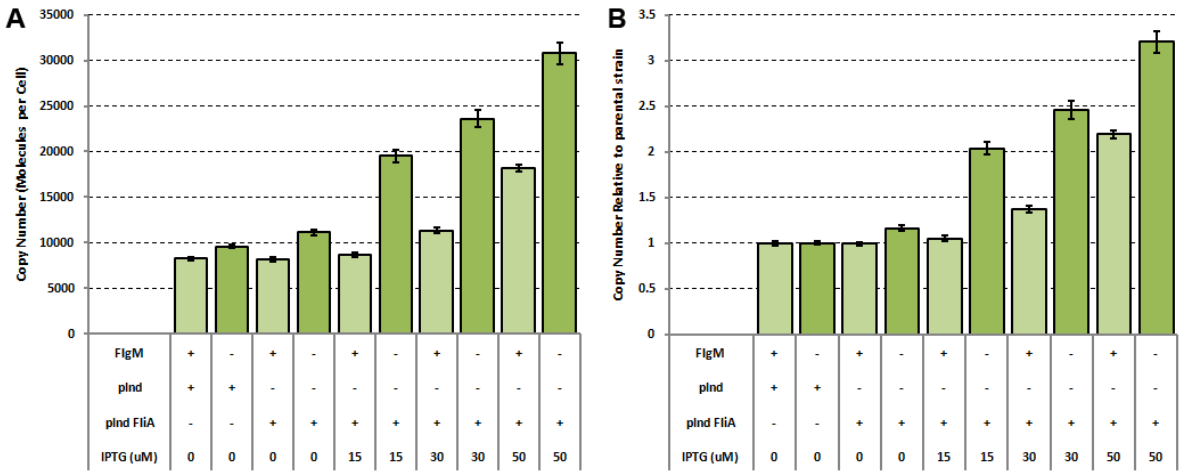


Figure 94**:** FliA activity with and without FlgM: **A)** Absolute copy numbers of YFP-CheW$_3$ calculated via fluorimetry for wild-type background JPA1418 (light green) and *ΔflgM* background JPA2018 (dark green) with increasing expression of FliA. **B)** The same data as in **(A)**, but plotted relative to copy number of the corresponding parental strain. Error bars represent the standard error of three independent repeats.

The maximum reduction in expression caused by FlgM was seen to be 2.24 fold under these conditions (Table 21).

Condition	Fold-Increase in *ΔflgM* background
pInd:empty, IPTG 0 µM	1.16
pInd:*fliA*, IPTG 0 µM	1.36
pInd:*fliA*, IPTG 15 µM	2.24
pInd:*fliA*, IPTG 30 µM	2.08
pInd:*fliA*, IPTG 50 µM	1.69

Table 21: The magnitude of FlgM inhibition of FliA**:** Fold-changes in YFP-CheW$_3$ copy number caused by the absence of FlgM at different expression levels of FliA.

The same process was repeated for the expression of CheW$_4$-YFP via conjugation of pInd4:*fliA* into JPA1457 and JPA2019. Again, absolute copy numbers were increased under all conditions in the absence of *flgM* (Figure 95A), and the increases in expression observed relative to the parental background were also observed to be higher in the *ΔflgM* background (Figure 95B).

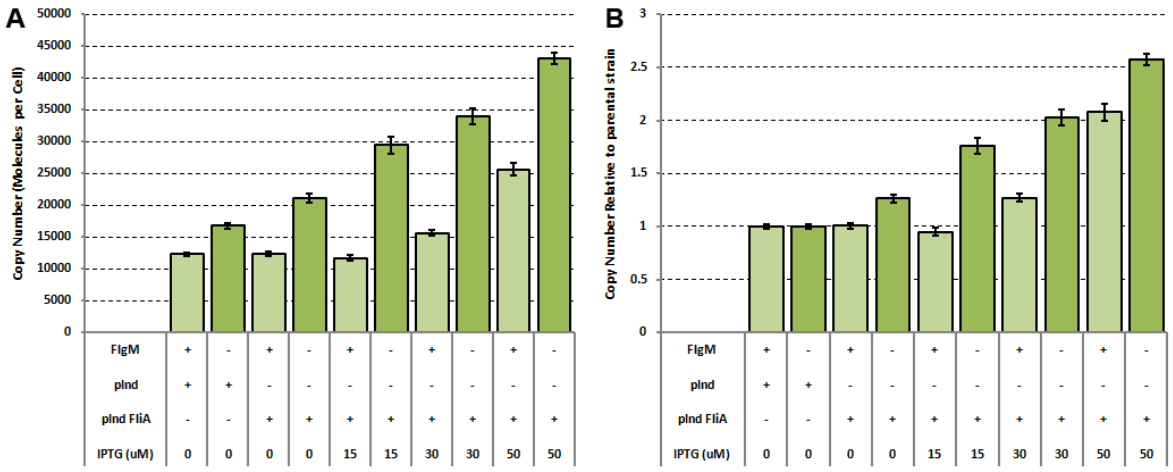


Figure 95**:** FliA activity with and without FlgM: **A)** Absolute CheW$_4$-YFP copy numbers calculated via fluorimetry for wild-type background JPA1457 (light green) and *ΔflgM* background JPA2019 (dark green) with increasing expression of FliA. **B)** The same data as in **(A)**, but plotted relative to copy number of the corresponding parental strain. Error bars represent the standard error of three independent repeats.

The maximum reduction in expression caused by FlgM was seen to be 2.51 fold under these conditions (Table 22).

Condition	Fold-Increase in *ΔflgM* background
pInd:empty, IPTG 0 µM	1.35
pInd:*fliA*, IPTG 0 µM	1.71
pInd:*fliA*, IPTG 15 µM	2.51
pInd:*fliA*, IPTG 30 µM	2.17
pInd:*fliA*, IPTG 50 µM	1.68

Table 22: The magnitude of FlgM inhibition of FliA**:** Fold-changes in CheW$_4$-YFP copy number caused by the absence of FlgM at different expression levels of FliA.

5.8) Chapter discussion

Here, it has been demonstrated that fluorimetry can be used to measure the copy number of YFP-fused proteins in mid-log phase cultures of *R.sphaeroides* cells, and that the results obtained via this technique are comparable to those obtained using western blot analysis. Having established this technique, the effects of *fliA* and *flgM* deletion and induced expression on the copy numbers of chemotaxis genes from the two essential chemotaxis operons of *R.sphaeroides* were studied.

It was demonstrated that the expression of components from both of the essential chemotaxis operons of *R.sphaeroides* is regulated by the activity of FliA, making the chemotaxis genes of *R.sphaeroides* class IV flagellar genes. Co-regulating chemotaxis gene expression with the final stage of flagellar assembly is common to many bacterial species, but these results contradict a previous analysis (98), which saw no change in *che*Op3 expression upon deletion of *fliA*. The discrepancy between the two data sets has been attributed to a change in growth conditions. Indeed, under the tested conditions the magnitude of changes in expression observed with changes in FliA activity was greater for components of *che*Op3 than for components of *che*Op2.

FlgM has been observed to inhibit FliA activity. Deletion of *fliA* resulted in small decreases in expression of the chemotaxis genes, and deletion of *flgM* resulted in a less than 1.5-fold increase chemotaxis gene expression, suggesting that the class IV regulation of chemotaxis gene expression is not the main determinant of the level of expression of these genes. It has previously been suggested that *che*Op2 expression is co-regulated by FliA and the house-keeping gene, σ^{70}, whereas *che*Op3 expression is thought to be regulated by the activity of the σ^{54} homologue, RpoN3 (98). It is likely that FliA activity fine-tunes the basal level of expression of the chemotaxis genes in response to the completion of HBBs and the secretion of FlgM. The magnitude of the chemotaxis gene regulation by FliA is much less than in *E.coli*, where a 3 - 8 fold increase in chemotaxis gene expression is observed upon deletion of *flgM* (147, 148). The response to *flgM* deletion is, however, similar in magnitude to that of the monoflagellate *Y.pseudotuberculosis*, which demonstrates a 2-fold increase in chemotaxis gene expression over wild-type (38).

Importantly, it has been demonstrated that the copy numbers of the both essential sets of chemotaxis proteins can be increased by over-expression of FliA from pInd4. This is used in the next chapter to study chemotaxis protein clustering at increasing cellular copy numbers, as proposed in section (4.8). The effects of varying FliA activity on chemotaxis itself have also been addressed, and are discussed in section (7.1).

6) Protein clustering in non-wild-type strains

Having shown that the concentration of proteins from both essential chemotaxis operons can be altered by the activity of FlgM and FliA in *R.sphaeroides*, the effect of concentration on protein clustering was investigated for both polar/lateral and cytoplasmic proteins.

6.1) The effect of concentration on polar protein clusters

From the copy numbers calculated under different growth conditions and different levels of FliA expression in section (5.7.1), four conditions were chosen that gave a broad range of cellular concentrations of *che*Op2 components (Table 23). Cells contained either pInd4:empty, or pInd4:*fliA*, and were grown in the presence of kanamycin. Cells were grown microaerobically to an OD_{700} of 0.55-0.65, reproducing the conditions under which copy numbers were calculated in section (5.7.1). Cells were immobilised on agarose pads, and visualised using epifluorescence microscopy. In order for intensity values to be directly comparable between image datasets, images were acquired on the same day for three separate cultures for each condition, using the same microscope settings. Images of WS8N containing pInd4:empty were also acquired, and used for background subtraction.

Strain	Genotype Description	Vector	IPTG (μM)	YFP-CheW$_3$ Copy Number
JPA2022	*yfp-cheW$_3$, ΔfliA*	pInd4:empty	0	6200 ± 200
JPA1418	*yfp-cheW$_3$*	pInd4:empty	0	8300 ± 200
JPA2018	*yfp-cheW$_3$, ΔflgM*	pInd4:empty	0	9600 ± 200
JPA2018	*yfp-cheW$_3$, ΔflgM*	pInd4:*fliA*	30	23600 ± 1000

Table 23: Chosen growth conditions and copy numbers: Strains, growth conditions and YFP-CheW$_3$ copy numbers chosen for the study of the affect of concentration on polar-clustering.

The protein concentrations in cells from each population were calculated as the sum of pixel intensities within each cell divided by the area of that cell in pixels. The average protein concentration of each population showed a positive correlation with protein copy number from each strain although there was a large amount of overlap at lower protein copy numbers (Figure 96).

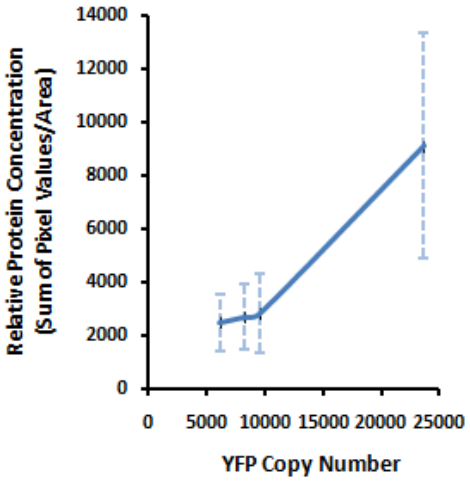


Figure 96: Relative fluorescent protein levels with YFP-CheW$_3$ copy number: Relative protein concentration is calculated as the sum of WS8N background-corrected pixel values within each cell, divided by the pixel-area of that cell, plotted against the average YFP protein copy number calculated via fluorimetry. Error bars represent plus and minus one standard deviation from the mean value of all cells.

To determine how protein concentration affects cluster number, the proportions of each population with 0, 1, 2, 3 and >3 distinguishable clusters were calculated (Table 24).

Strain	**YFP-CheW$_3$ Copy Number**	**Frequency of Number of Clusters (% of population)**					**Total Number of Cells**	**Average No. of clusters**
		0	1	2	3	>3		
JPA2022	6200 ± 200	0	19.7	44.1	37.4	8.8	497	2.27 ±0.93
JPA1418	8300 ± 200	0	17.3	60.0	18.8	4.5	624	2.09 ± 0.76
JPA2018	9600 ± 200	0	9.2	66.9	19.8	4.1	293	2.19 ± 0.66
JPA2018 pInd4:*fliA*	23600 ± 1000	0	5.3	52.2	27.5	14.9	469	2.33 ± 0.95

Table 24: Numbers of YFP-CheW$_3$ clusters per cell observed in populations with different average copy numbers. Average numbers of clusters are calculated as the mean value in all cells, error represents plus or minus one standard deviation from the mean of all cells.

The average number of clusters within a cell did not change greatly with protein copy number. However, the proportion of each population with only one polar cluster was seen to decrease with increasing copy number. Interestingly the two cell populations with the lowest and highest copy numbers had a higher proportion of cells with three or more clusters than the two populations with

mid-range copy numbers - which had similar distributions of cluster numbers, with the majority of cells containing either one or two clusters.

To further investigate the patterns in number of cluster, the proportions of cells with different numbers of clusters were plotted with cell length (Figure 97).

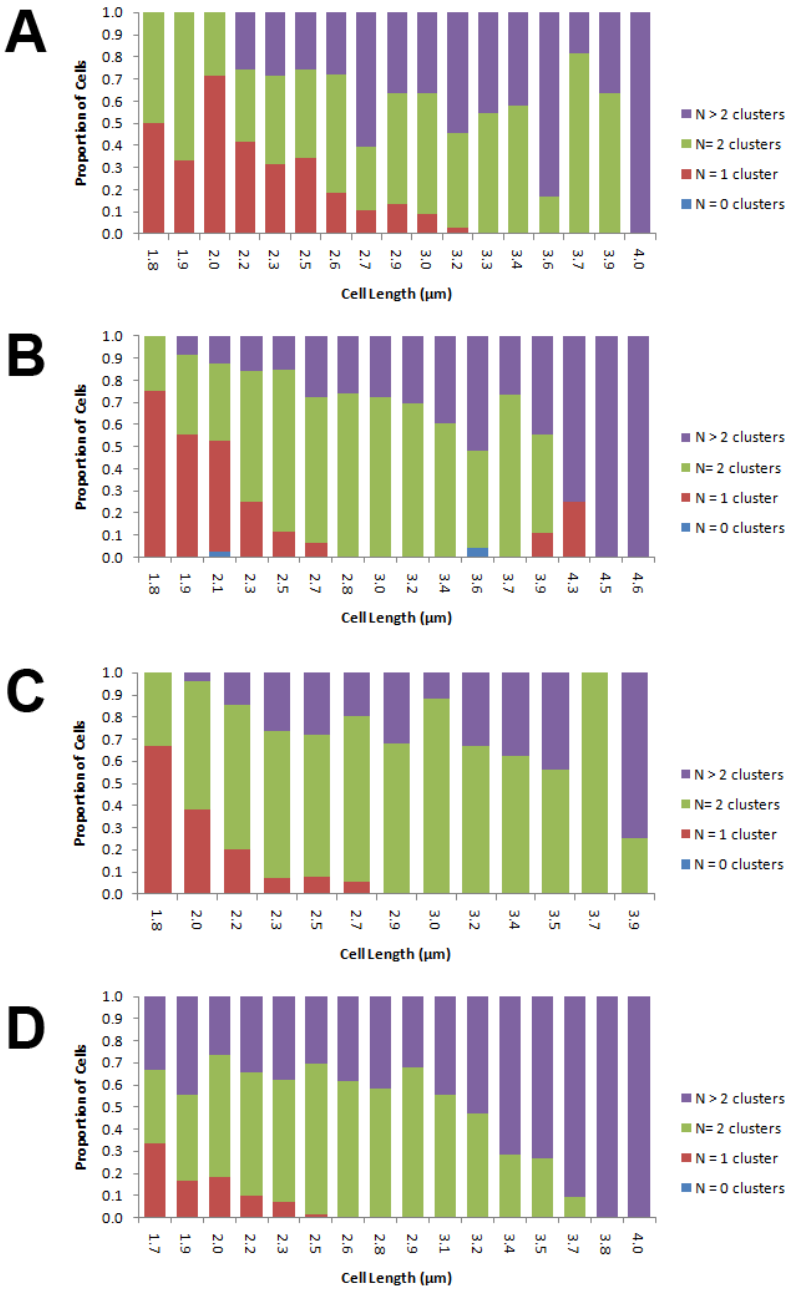


Figure 97**:** Cluster Numbers With Cell Length: The proportion of cells found in each bin with zero (blue), one (red), two (green) and more than two (purple) YFP-CheW$_3$ clusters for the four different conditions analysed, average YFP-CheW$_3$ protein copy numbers increase from **A-D**. **A)** *ΔfliA* (JPA2022), **B)** wild-type (JPA1418), **C)** *ΔflgM* (JPA2018), **D)** high *fliA* activity (JPA2018,pInd4:*fliA*). Results are plotted for cell lengths between 2 and 4 µm.

As the copy number of each population increased, the range of lengths where cells with one cluster were observed decreased. This suggests that newly divided cells which inherit a single cluster will acquire a second cluster more quickly, or where a second cluster is present, high protein concentration will mean the cluster becomes visible earlier. This in turn suggests that rate of polar cluster nucleation events is concentration dependant.

Interestingly, the proportion of each population with more than two clusters was seen to increase with cell length, however at mid-range copy numbers (Figure 97B and Figure 97C) the majority of cells contained two clusters over all cell lengths, whereas at the lowest measured copy number (Figure 97A) two or three clusters were equally likely between cell lengths of 2.5 and 4 µm in length. This pattern differed again at high protein copy number (Figure 97D), where cells with more than two clusters replaced cells with two clusters gradually with cell growth.

To investigate the possibility of poles becoming saturated with protein at higher levels of expression, the average amounts of protein in the old and new poles of the cell as well as the total amount of cellular protein (calculated as the sum of pixel intensities within specified masked areas) were plotted with cell length for each population (Figure 98).

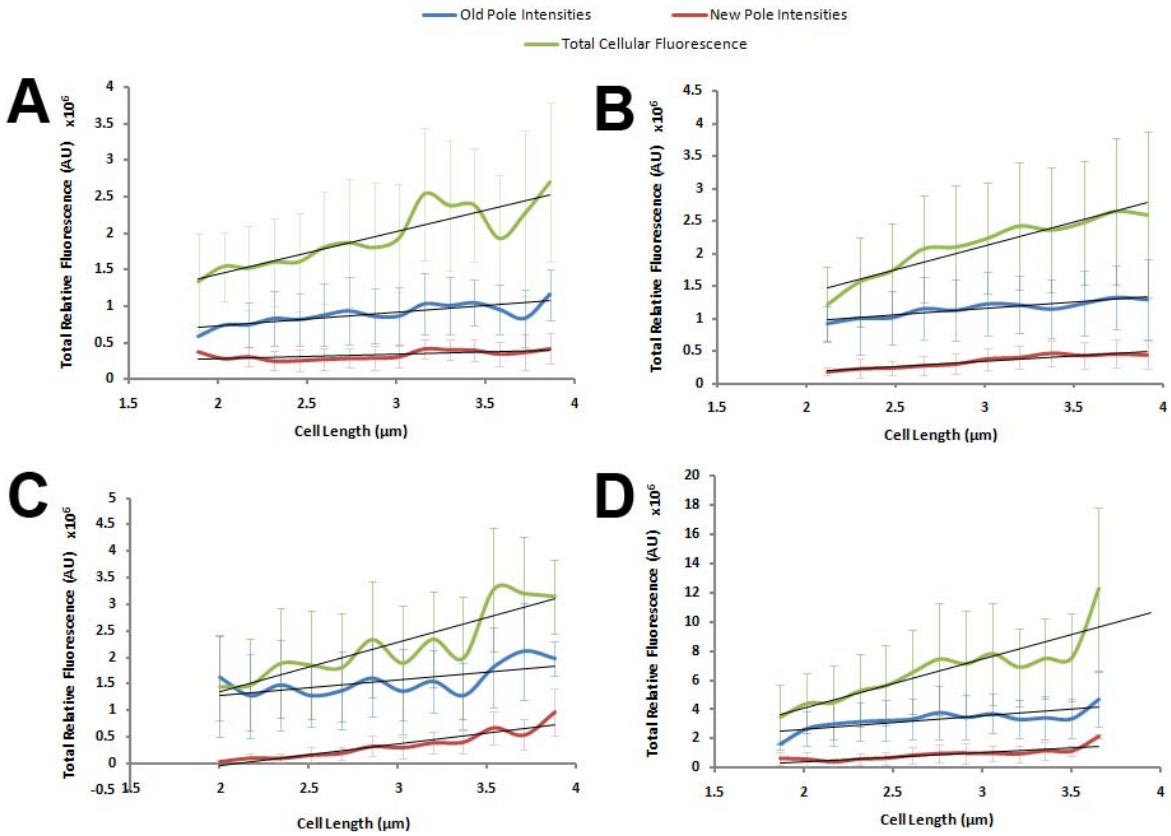


Figure 98**:** Amounts of YFP-CheW$_3$ with cell length. Lines show the average sum of pixel intensities in old cell poles (blue), new cell poles (red) and the total cell area (green) within each bin for cell length. Error bars represent plus or minus one standard deviation from the mean of all points. Best-fit trend lines are displayed in black. Average YFP-CheW$_3$ protein copy numbers increase from **A-D**. **A)** *ΔfliA* (JPA2022), **B)** wild-type (JPA1418), **C)** *ΔflgM* (JPA2018), **D)** high *fliA* activity (JPA2018,pInd4:*fliA*).

The trend in amount of protein found in old and new poles was similar for all populations, but total pixel sums were higher at higher copy numbers of protein, suggesting that the amount of protein that can be incorporated into the pole of the cell does not saturate across the tested range of protein copy numbers.

Another predicted effect of the polar regions of the cell becoming saturated with protein is that an increased amount of protein would be seen in the lateral section of the cell. To investigate whether this was the case, the total amount of protein found in the cell poles was calculated as a percentage of the total amount of protein in the whole cell. In order to exclude cells that contained lateral protein in clusters analysis was limited to the subsection of the populations that had only two clusters, one at each pole (Figure 99).

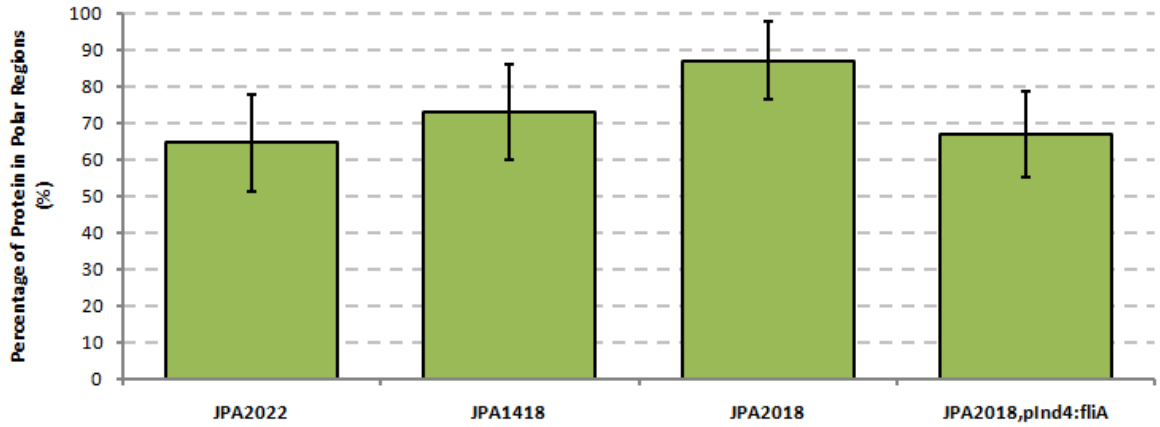


Figure 99**:** Percentage of total YFP-CheW$_3$ protein found in the cell poles **–** Amounts of protein were calculated as the pixel sums of defined masked areas within fluorescence images. Results display the mean of all cells within each population, error bars represent plus or minus one standard deviation of the mean of all points. Average YFP-CheW$_3$ copy numbers increase from left to right.

The majority of protein was seen to localise to the poles of the cell within each population, and the percentage of protein in the poles was similar at different copy numbers. Again, this suggests that even at increased copy number the cell poles do not become saturated with protein.

As both a small reduction in *che*Op2 encoded protein levels caused by the deletion of *fliA* and a relatively large increase in *che*Op2 encoded protein levels caused by over-expression of *fliA* resulted in an increase in the proportion of cells with more than 2 distinguishable clusters, the difference between the two populations was further investigated by looking at the localisation patterns of these clusters.

Of 399 cells with two or more visible clusters in the *ΔfliA* strain JPA2022, 32.3 % were seen to have multiple clusters at a single pole; this figure was only 18.9 % of 444 cells for cells over-expressing *fliA* (JPA2018,pInd4:*fliA*), suggesting that closely positioned clusters are less likely to be observed at higher concentrations of protein, this is not unexpected as the more protein there is in one pole, the greater the total area of that pole that will contain protein, leaving less physical space for the formation of a second cluster.

Interestingly, when the positions of lateral clusters from these two populations were measured relative to the septation sites of dividing cells from each population, the positions of these clusters showed two distinct patterns of localisation. At higher protein copy number, lateral clusters were found near to the sites of cell division, whereas at low protein copy number, they were positioned

closer to the polar region of the cell (Figure 100). Distances were measured relative to the whole length of the cell and the cell lengths of dividing cells were similar for each population.

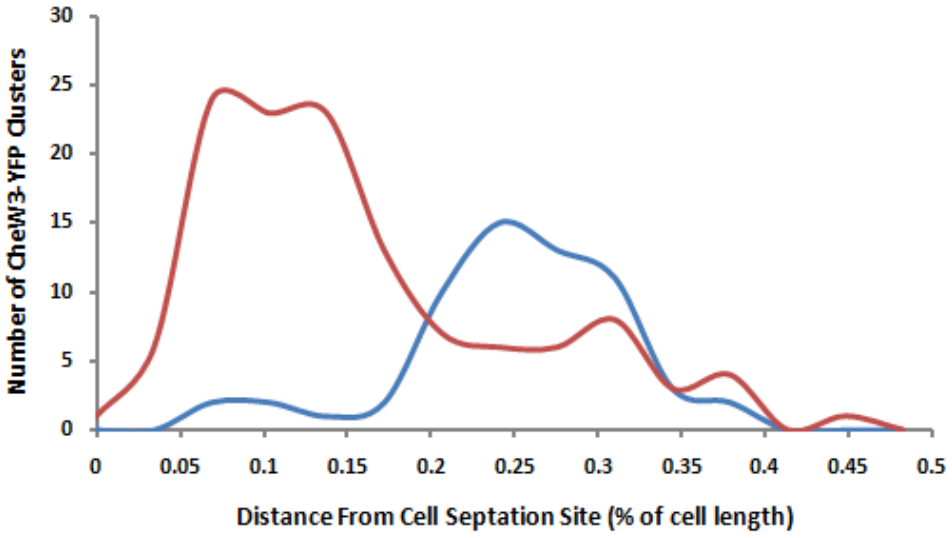


Figure 100**:** Positions of lateral clusters relative to cell septation loci**:** Distances of laterally positioned YFP-$CheW_3$ clusters measured from cell septation sites, and relative to the whole cell length. Data is presented for the *ΔfliA* strain (JPA2022- blue) and cells over-expressing *fliA* (JPA2018, pInd4:*fliA* - red).

We see here that the likelihood of formation of a new protein cluster near a septation site increases with protein copy number, explaining the reduction in the proportion of newly divided cells that have only one polar cluster at high copy number (see Figure 97). At low copy numbers, the total amount of fluorescent protein in the poles of the cell decreased. This, in turn, may have facilitated the visualisation of clusters that appear near to the poles, explaining the observed pattern in localisation (above).

6.2) The effect of concentration on cytoplasmic protein clusters

As with the *che*Op2 components, four conditions were chosen that gave a broad range of cellular concentrations of *che*Op3 components from the copy numbers calculated in section (5.7.1) (Table 25). Cells contained either pInd4:empty or pInd4:*fliA* and were grown in the presence of kanamycin. Cells were grown reproducing the conditions under which copy numbers were calculated in section (5.7.1), immobilised on agarose pads, and visualised via epifluorescence microscopy. In order for intensity values to be directly comparable between image datasets, images were acquired with the same microscope settings and on the same day for three separate cultures for each condition.

Background subtraction values were calculated from images of WS8N containing pInd4:empty, which were grown in the same way.

Strain	Genotype Description	Vector	IPTG (μM)	$CheW_4$-yfp Copy Number
JPA2023	*$cheW_4$-yfp, ΔfliA*	pInd4:empty	0	5300 ± 100
JPA1457	*$cheW_4$-yfp*	pInd4:empty	0	12300 ± 300
JPA2019	*$cheW_4$-yfp, ΔflgM*	pInd4:empty	0	16800 ± 400
JPA2019	*$cheW_4$-yfp, ΔflgM*	pInd4:*fliA*	30	34000 ± 1300

Table 25: Chosen growth conditions and copy numbers**:** Strains, growth conditions and $CheW_4$-YFP copy numbers chosen for the study of the affect of concentration on polar-clustering.

Cellular protein concentrations were calculated as in section (6.1). The average protein concentration of a population showed a positive correlation with protein copy number from each strain (Figure 101).

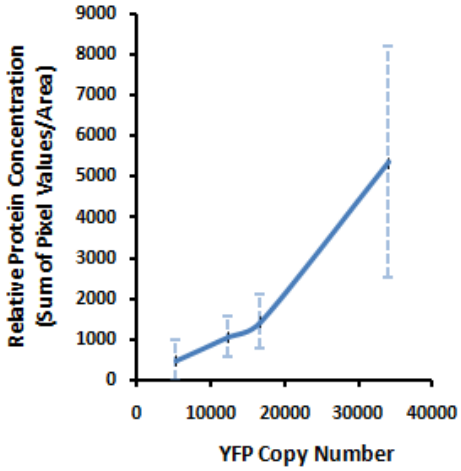


Figure 101: Relative fluorescent protein levels with $CheW_4$-YFP copy number**:** Relative protein concentration is calculated as the sum of WS8N background-corrected pixel values within each cell, divided by the pixel-area of that cell, plotted against the average YFP protein copy number calculated via fluorimetry. Error bars represent plus and minus one standard deviation from the mean value of all cells.

To determine how protein concentration affects cluster number, the proportions of each population with 0, 1, 2, 3 and >3 distinguishable clusters were calculated (Table 26).

Strain	$CheW_4$-yfp Copy Number	Frequency of Number of Clusters (% of population)					Max. Cluster Number	Total Number of Cells	Average No. of clusters
		0	1	2	3	>3			
JPA2023	5300 ± 100	18.7	67.1	13.4	0.2	0	3	402	0.95 ± 0.58
JPA1457	12300 ± 300	4.6	28.5	50.8	14.9	1.2	4	589	1.79 ± 0.79
JPA2019	16800 ± 400	0.6	10.7	53.3	27.8	7.4	5	474	2.31 ± 0.82
JPA2019 pInd4:*fliA*	34000 ± 1300	0.4	17.7	42.6	28.2	11.0	6	446	2.33 ± 0.95

Table 26**:** Numbers of $CheW_4$-YFP clusters observed in populations with different average copy numbers. Average numbers of clusters are calculated as the mean value in all cells, error represents plus or minus one standard deviation from the mean.

Unlike the observations of polar/lateral clusters, the trend in average number of clusters within a cell was an obvious increase with cellular copy number. The proportion of each population with no clusters decreased to near zero at three times wild-type copy numbers. The maximum number of clusters seen within an individual cell also increased with copy number to a maximum of 6. Interestingly, the change in average cluster number was most pronounced close to the wild-type copy number of protein, and the ~2-fold increase in protein copy number between the *ΔflgM* (JPA2019) and *fliA* over-expressing (JPA2019,pInd4:*fliA*) populations had little effect. The resolution limit of epifluorescence microscopy limits the observable number of clusters within cells, as two objects can only be distinguished if they are separated by a distance of approximately half the wavelength of fluorescent light (~250 nm separation). In cells which range in length between 2 and 4 µm, it is perhaps unsurprising that the average number of detectable clusters did not increase further. However, it is interesting that the formation of cytoplasmic clusters showed a more obvious concentration dependence than polar clusters, especially given that cytoplasmic cluster formation is thought to occur by active partitioning and not nucleation.

To further investigate the patterns in numbers of cluster, the proportions of cells with different numbers of clusters were plotted with cell length (Figure 102).

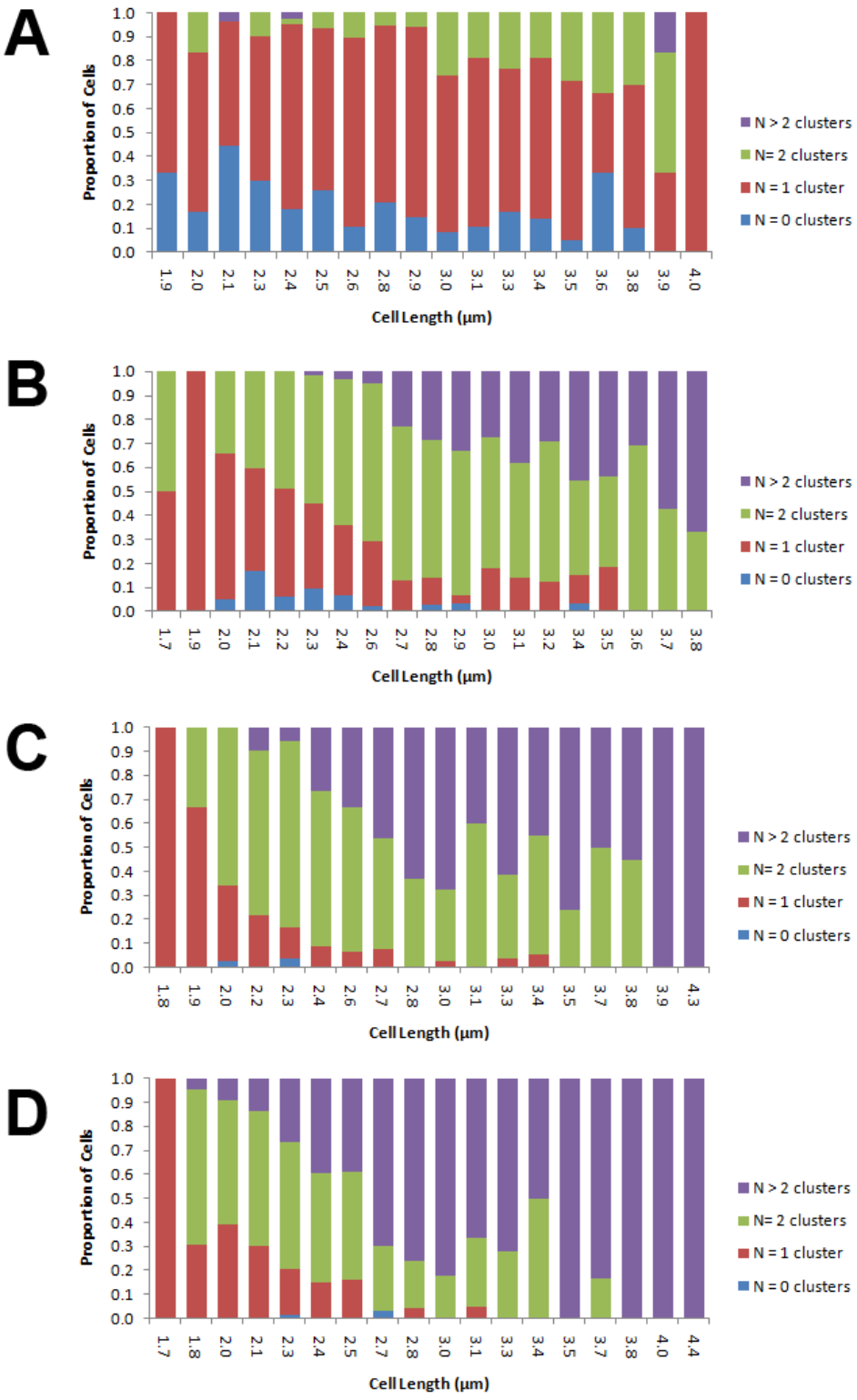


Figure 102**:** Cluster Numbers With Cell Length: The proportion of cells found in each bin with zero (blue), one (red), two (green) and more than two (purple) CheW$_4$-YFP clusters for the four different conditions analysed, average CheW$_4$-YFP protein copy numbers increase from **A-D**. **A)** *ΔfliA* (JPA2023), **B)** wild-type (JPA1457), **C)** *ΔflgM* (JPA2019), **D)** high *fliA* activity (JPA2019,pInd4:*fliA)*. Results are plotted for cell lengths between 2 and 4 µm.

The number of cells with zero or one clusters was high over all cell lengths in the strain with the lowest copy number of $CheW_4$-YFP protein (Figure 102A), whereas in the other populations, cells with two clusters were common over all cell lengths. The range of lengths over which cells with only one cluster was observed decreased with increasing copy number, suggesting that newly divided cells which inherit only one cluster will form a second cluster more rapidly when they have a higher protein concentration.

To compare cytoplasmic and polar protein localisation and to investigate whether cytoplasmic clusters became saturated with protein at higher protein concentrations, the average amounts of protein in the bright and dim clusters, as well as the total amount of cellular protein were plotted with cell length, as was done in section (6.1). Analysis was restricted to cells within the population that contained only two clusters, and as the proportion of cells with 2 clusters in the *ΔfliA* strain (JPA2023) was so low, this population was excluded from the analysis (Figure 103).

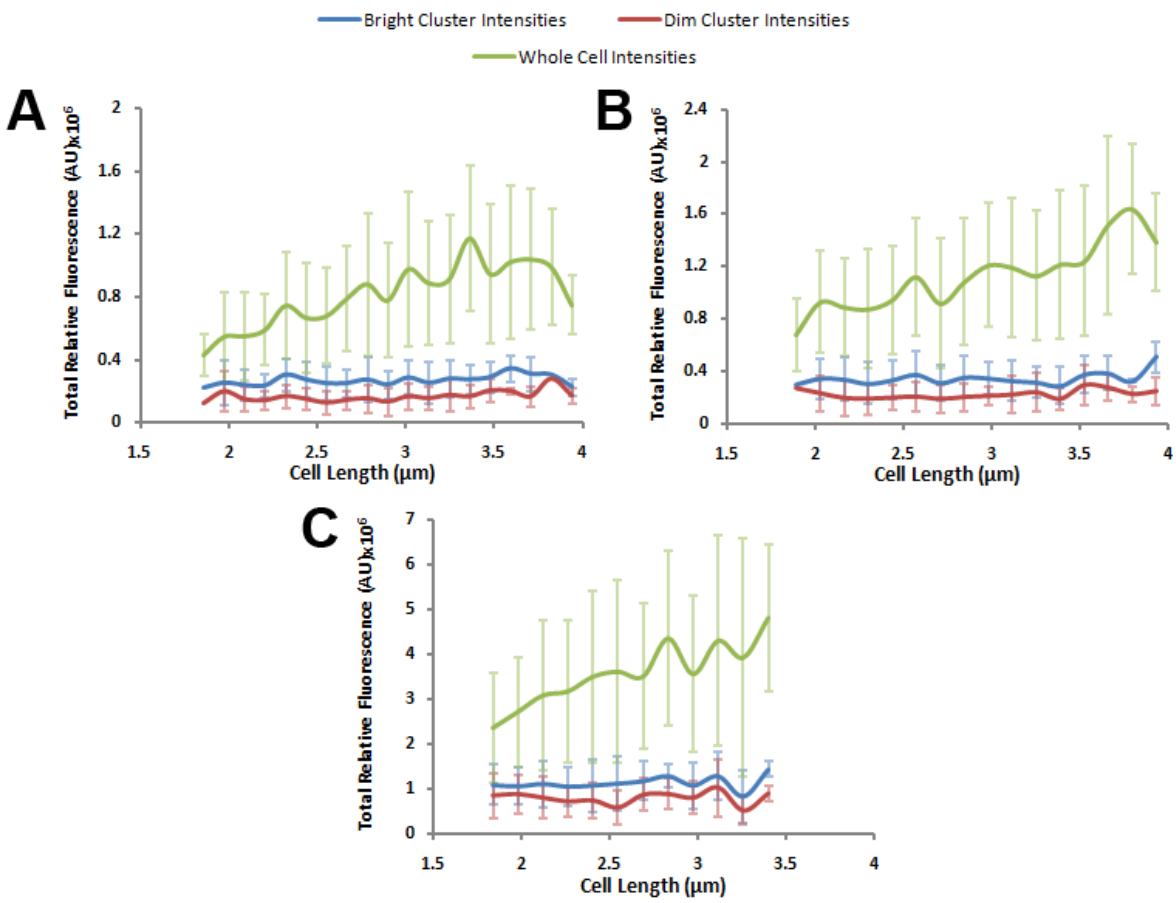


Figure 103**:** Amounts of $CheW_4$-YFP with cell length**.** Lines show the average sum of pixel intensities in bright clusters (blue), dim clusters (red) and the total cell area (green) for all cells within each bin for cell length. Error bars represent plus or minus one standard deviation from the mean of all points. Average $CheW_4$-YFP protein copy numbers increase from **A-C**. **A)** wild-type (JPA1418), **B)** *ΔflgM* (JPA2019), **C)** high *fliA* activity (JPA2019,pInd4:*fliA)*.

The average cluster intensity of bright and dim cytoplasmic clusters increased with total protein copy number, suggesting that the size of a cytoplasmic cluster is dependent on the protein concentration in the cell and that a maximum cluster size is not reached within the tested concentration range. Unlike polar clusters (see section (6.1)), the amount of protein that accumulates in cytoplasmic clusters does not necessarily increase as the cells grow, suggesting that cluster size is also relatively stable.

When the concentration of protein in the cell (as opposed to total amount of protein) was calculated with cell length (Figure 104), this too was seen to remain relatively stable across cell lengths.

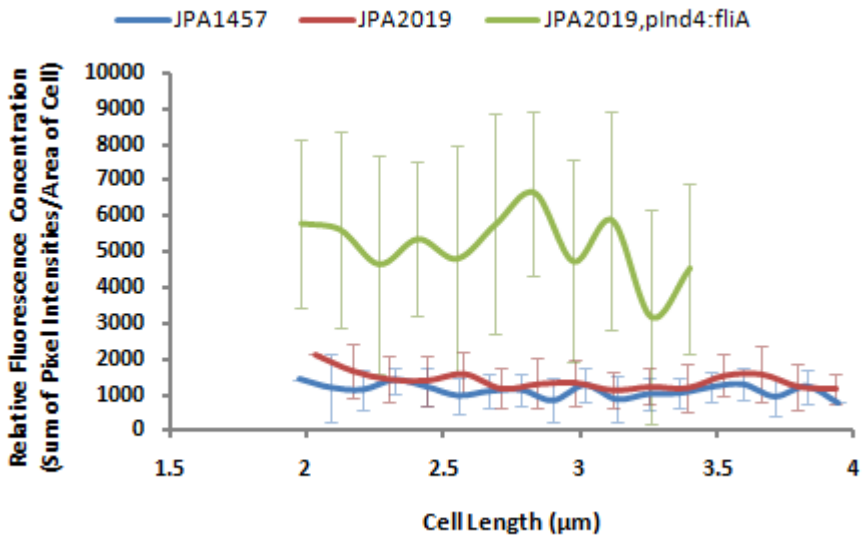


Figure 104: $CheW_4$-YFP cellular protein concentrations with cell length: Fluorescent protein concentration, calculated as the average pixel intensity within each cell plotted with cell length for cells containing two cytoplasmic clusters in wild-type (JPA1457 - blue), *ΔflgM* (JPA2019 - red) and *fliA* over-expressing (JPA2019,pInd4:*fliA* - green) backgrounds. Error bars represent plus or minus one standard deviation from the mean of all points within each bin.

Taken together, these observations suggest a different form of behaviour to that suggested in section (4.6), where the size of a cytoplasmic cluster is determined by the concentration of the *che*Op3 components within the cell. As the cell grows the total concentration of cellular protein remains stable, and thus the cluster does not grow in size. The rate of exchange between proteins in the cluster and proteins in the cytoplasm, along with any energetic barriers to the addition of new protein to a cytoplasmic cluster will determine the rate at which the size of each cluster changes after a cluster partitioning or cluster nucleation event.

6.3) Patterns of localisation in filamentous cells

6.3.1) Dual-tagged strains for simultaneous study of polar and cytoplasmic clusters

The lateral chemotaxis protein clusters of *E.coli* have been suggested to form at predefined future division sites, in association with an unknown cytoskeletal element that is not MreB or FtsZ and that is also involved in the positioning of the DNA replication machinery (147, 148). Protein clusters are suggested to form at the 1/8, 1/4, 1/2, 3/4 and 7/8 positions along the cell length.

The treatment of bacterial cells with the antibiotic cephalexin, inhibiting FtsZ and cell division, generates long, filamentous cells. Lateral chemotaxis protein cluster localisation in *E.coli* filamentous cells has also been suggested to show a regular pattern of localisation, with lateral clusters forming approximately 1 µm apart from each other (147). Protein cluster localisation was therefore investigated in filamentous *R.sphaeroides* cells.

The growth of cephalexin-treated, filamentous *R.sphaeroides* is, at best, unpredictable. The optimal growth conditions for *R.sphaeroides* filamentous cell preparations were determined in 2006 (150). However the quality of the obtained preparations grown in this manner varies from day to day and it is therefore difficult to obtain comparable data sets across multiple days. Cephalexin-treated cultures must also be grown at low-density, and cells have a tendency to stick together meaning that obtaining usable data from non-touching cells within fields of view is a challenge. Because of these complications, bacterial strains were developed to obtain a maximum amount of comparable data from images taken on the same day.

The dual-tagged strain JPA1447 (discussed in section (4.7)) was chosen as the reporter strain for cephalexin experiments, as it allows the simultaneous acquisition of data for both polar/lateral and cytoplasmic protein clusters.

The protein partitioning factor, PpfA, has been shown to actively partition cytoplasmic chemotaxis protein clusters in *R.sphaeroides*, and its deletion results in a reduction in the number of cytoplasmic clusters in filamentous cells (150). In order to study the effects of cellular protein concentration and the role of PpfA in cephalexin-treated cells, three further derivatives of this strain were generated by allelic exchange, and the successful generation of these mutants was verified by southern blot or DNA sequencing.

Strain	Genotype Description
JPA1447	*cfp-cheW$_3$, cheW$_4$-yfp*
JPA2002	*cfp-cheW$_3$, cheW$_4$-yfp, ΔflgM*
JPA2003	*cfp-cheW$_3$, cheW$_4$-yfp, ΔppfA*
JPA2004	*cfp-cheW$_3$, cheW$_4$-yfp, ΔppfA, ΔflgM*

Table 27**:** Chosen strains for the study of cephalexin-treated cells: A list of strain numbers and genotype descriptions for strains generated for the study of cephalexin-treated cells.

Gene expression changes in these strains were measured via the YFP fluorimetry protocol detailed in section (5.6), unfortunately attempts to obtain similar measurements for CFP were unsuccessful due to machine sensitivity, and expression changes were only measured for the *cheOp3* marker protein, CheW$_4$-YFP (Table 28).

Strain	**CheW4-YFP Copy Number**
JPA1447	10500 ± 400
JPA2002	15600 ± 400
JPA2003	7600 ± 200
JPA2004	14500 ± 2000

Table 28: CheW$_4$-YFP copy numbers in dual-tagged strains. CheW$_4$-YFP copy numbers calculated using fluorimetry for wild-type strain (JPA1447), *ΔflgM* strain (JPA2002), *ΔppfA* strain (JPA2003) and *ΔppfA, ΔflgM* strain (JPA2004). Error represents the standard error of three independently acquired repeats. All strains also contained CFP-CheW$_3$.

As expected, deletion of *flgM* resulted in an increase in CheW$_4$-YFP protein copy number. Unexpectedly, deletion of *ppfA* was seen to decrease the expression of cheW$_4$-YFP by 27 %.

6.3.2) Cell morphology analysis of filamentous cells

After several attempts, good preparations of filamentous cells were achieved simultaneously for all strains by the addition of cephalexin to cells at an OD_{700} ~ 0.2 and subsequent incubation at 30 °C for a further 4 hours. Images were acquired on the same day with the same microscope settings in order for fluorescence intensities to be comparable across images, and images of WS8N cells were also acquired for background-subtraction. Exposure times for YFP and CFP were adjusted in order to obtain maximal signal with minimal camera-saturation, however the range of cytoplasmic cluster sizes (and therefore brightness) was such that it was not possible to eliminate saturation from all YFP

images. In images of more than 850 cephalexin-treated cells, cell lengths ranged between 1.8 and 10.9 µm, and very few cell septation sites were visible (Figure 105).

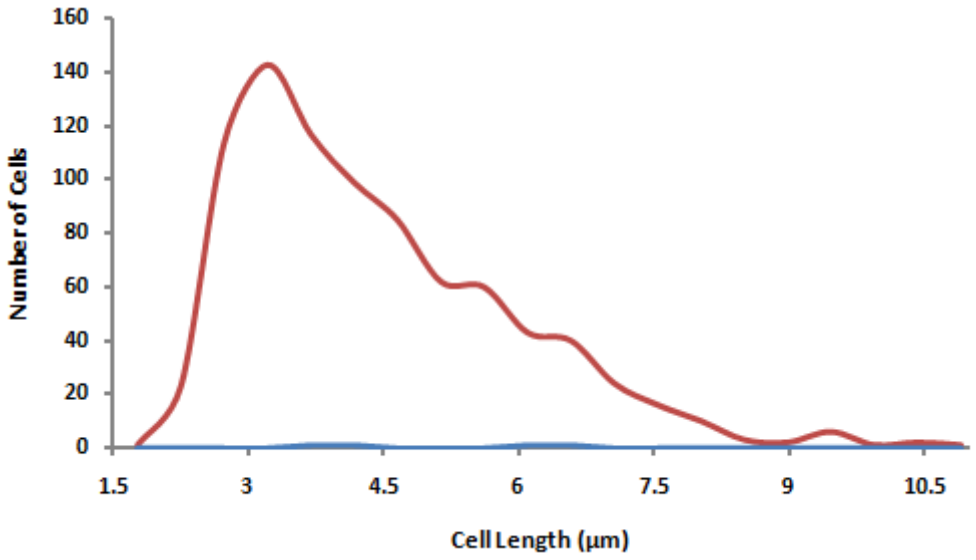


Figure 105**:** Cell Lengths in cephalexin-treated cultures: Smoothed histogram of non-dividing (red) and dividing (blue) cell lengths for 856 cells from separate cultures of JPA1447, JPA2002, JPA2003 and JPA2004.

The analysis of the fluorescent clusters of filamentous cells proved difficult using the MATLAB software from chapter (3), which was very successful for the analysis of wild-type cell populations (see section (4.2)). The crowded nature of cytoplasmic clusters, the dim appearance of many of the lateral clusters, along with the high degree of variability in the amounts of apparently diffuse protein for both YFP and CFP channels meant that the accuracy of cluster definition using the method laid out in section (3.8.3) was limited. For this reason there was a relatively high “discard rate” for cells that could not be assigned satisfactory cluster positions, even though extra care was taken at this stage.

6.3.3) Polar/Lateral clusters in filamentous cells

Polar and lateral CFP-CheW$_3$ clusters were clearly visible in cephalexin-treated strains (Figure 106) for all visualised strains, although lateral clusters were dimmer than polar clusters.

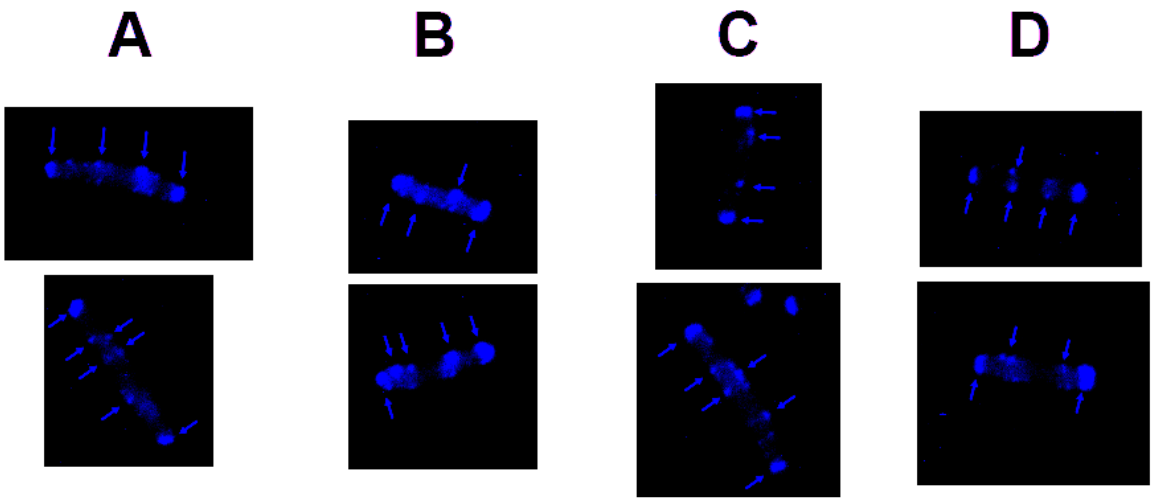
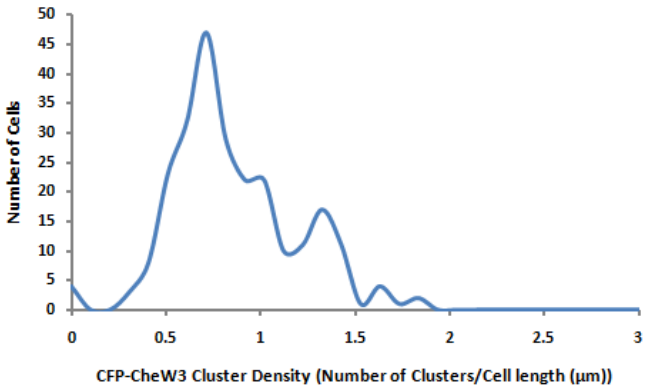
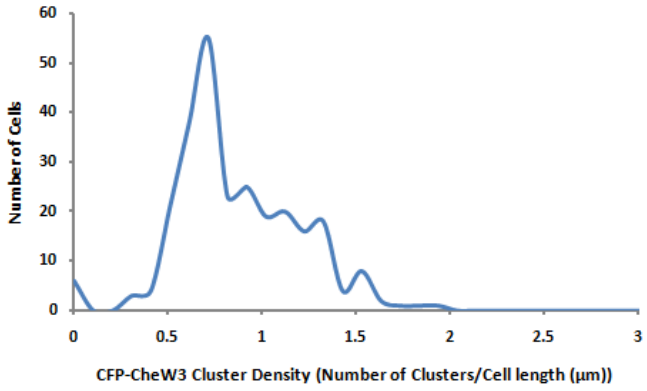
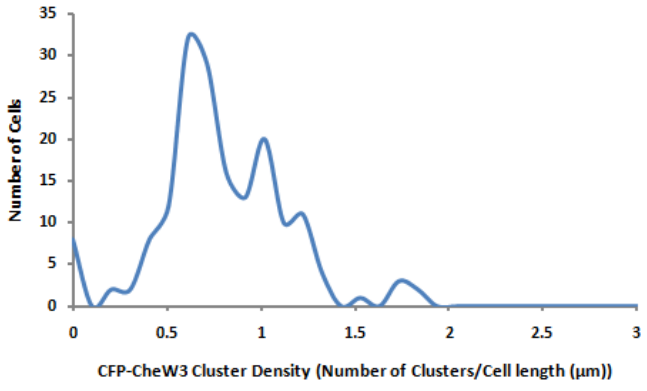
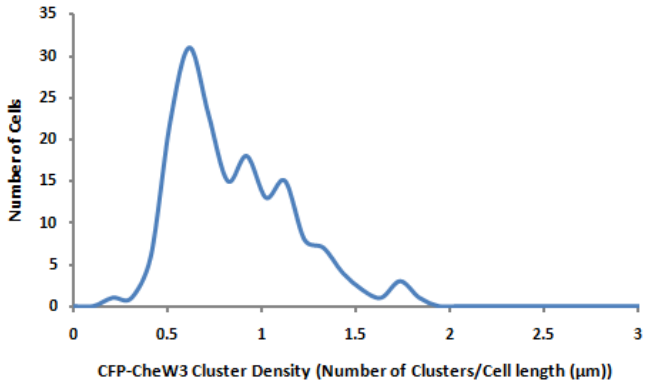


Figure 106: Example Cephalexin-Treated cells with CheW$_3$-CFP polar and lateral clusters. Two example long cells from populations of **A)** wild-type (JPA1447), **B)** *ΔflgM* (JPA2002), **C)** *ΔppfA* (JPA2003), and **D)** *ΔflgM, ΔppfA* (JPA2004). Contrast has been enhanced equally for all images, DIC image information is deliberately dark for printing purposes. Arrows highlight the positions of clusters within cells.

To test whether the level of expression or total numbers of clusters had been altered in these populations of cephalexin-treated cells, CFP fluorescence signal-based protein concentration and CFP-CheW$_3$ cluster densities were calculated for all four strains (Table 29).

Table 29 (next page): CFP-CheW$_3$ Cluster densities and protein concentrations in populations of cephalexin-treated cells: Cluster densities within each population, calculated as the number of clusters divided by the cell length, are displayed as a histogram. Protein concentrations, calculated as the sum of pixel intensities within each cell divided by the number of pixels that fall within each cell, are calculated as the mean value of all cells, error represents the standard deviation of this mean.

Background Strain	CFP-CheW$_3$ Cluster Densities, µm^{-1} (Number of Clusters/Cell Length)	Cellular Protein Concentrations (Sum of Pixel Values/Pixel area)	Total Number of Cells
wild-type (JPA1447)		2800 ± 1100	247
ΔflgM (JPA2002)		3100 ± 1200	265
ΔppfA (JPA2003)		2100 ± 900	173
ΔflgM, ΔppfA (JPA2004)		2800 ± 1000	171

Very little change was observed in cluster densities or cellular protein concentrations of CFP-$CheW_3$ across all four cephalexin-treated strains, suggesting that the gene expression change induced by the deletion of *flgM* or the deletion of *ppfA* did not greatly interfere with the amount of *che*Op2 encoded proteins in the cells under these conditions. *flgM* deletion was seen to increase expression of *che*Op2 components under microaerobic conditions (see section (5.7)), and the difference is likely due to growth-related regulation of *che*Op2.

Preliminary analysis of the positions of polar and lateral clusters in the four studied strains suggested that each strain showed the same, or very similar, pattern of cluster localisation. For this reason, data from each strain were grouped to provide more data points across the whole range of cell lengths.

Interestingly, plotting the positions of all CFP-$CheW_3$ protein clusters showed a clear pattern to the distribution of positions of lateral clusters, which localised to the 1/3 and 2/3 positions along the cells' long axes (Figure 107).

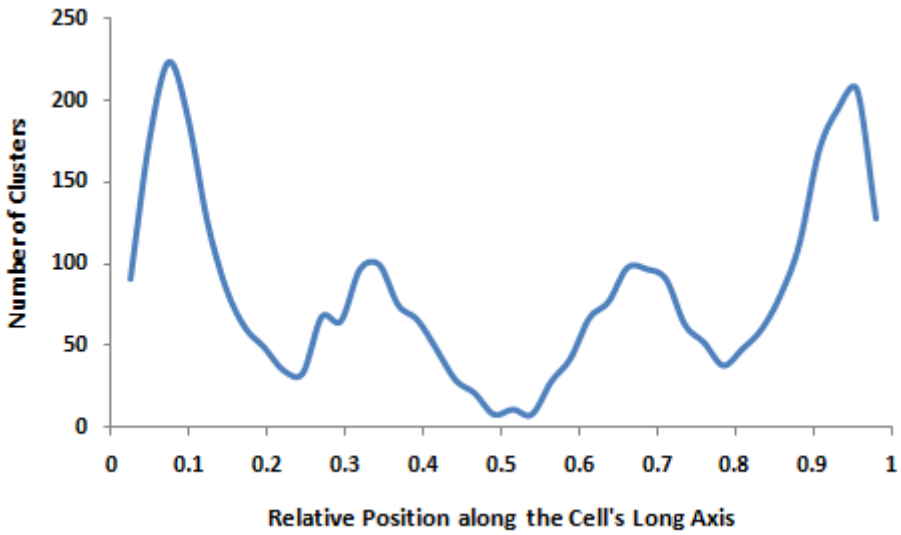


Figure 107: Positions of Polar and Lateral CFP-$CheW_3$ protein clusters - Positions of clusters in cells are calculated relative to the whole cell length. Data are grouped in 40 equal sized bins between 0 and 100 %.

The same pattern was also clearly visible when plotted with cell length (Figure 108), and was observed over the majority of cell lengths apart from in shorter cells which had lengths which were similar to those of wild-type cell populations (2 – 4 µm).

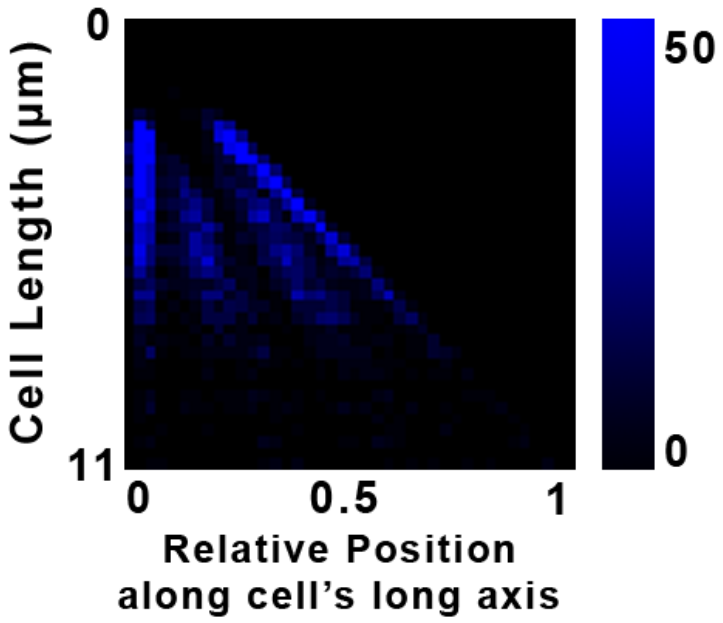


Figure 108: Positions of polar and lateral CFP-CheW$_3$ protein clusters with different length cells: 2D Histogram of relative cluster positions with cell length. Data are grouped in 40 equal sized bins for cell length and 40 equal sized bins for relative cluster position. The occupancy of each bin is indicated by an intensity of blue which corresponds to a number of clusters, indicated by the colour-bar to the right of the figure.

6.3.4) Cytoplasmic clusters in filamentous cells:

Cytoplasmic clusters were clearly seen at multiple loci within filamentous cells with *ppfA* and, as expected, at fewer loci in cells where *ppfA* had been deleted (Figure 109).

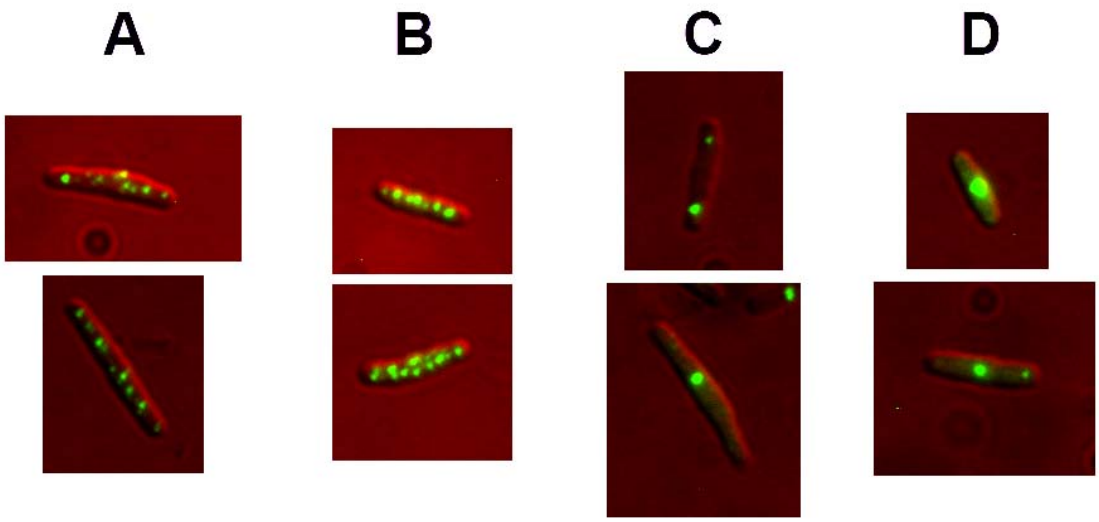
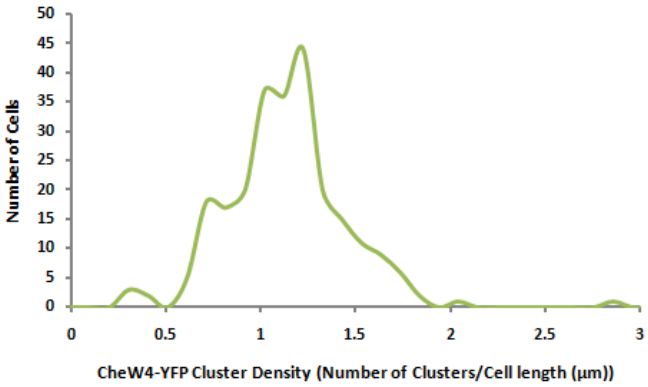
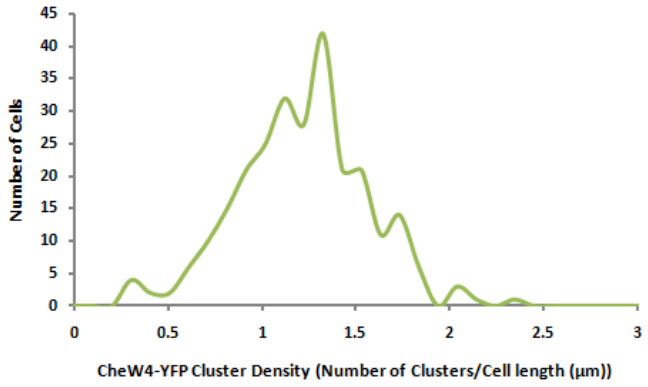
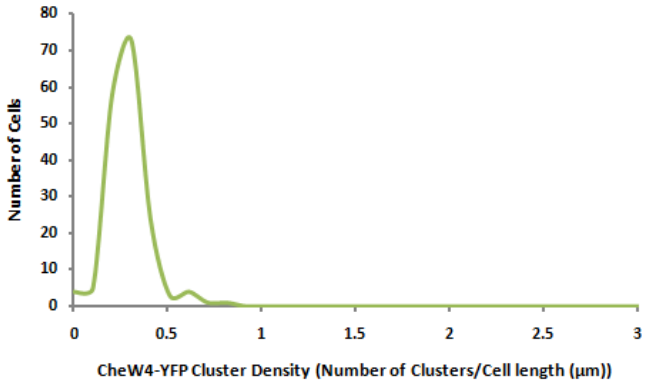
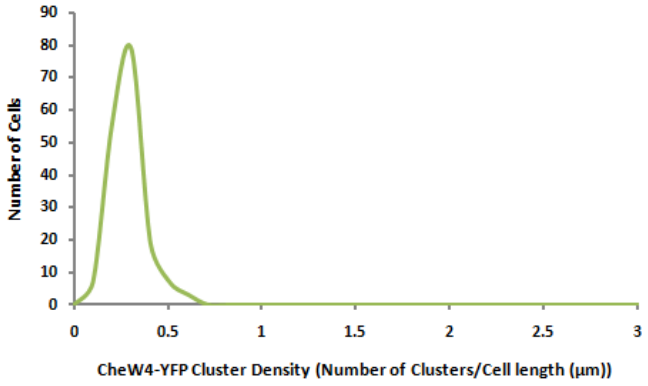


Figure 109**:** Example cephalexin-treated cells with $CheW_4$-YFP cytoplasmic clusters. Two randomly chosen example long cells from populations of **A)** wild-type (JPA1447), **B)** *ΔflgM* (JPA2002), **C)** *ΔppfA* (JPA2003), and **D)** *ΔflgM, ΔppfA* (JPA2004). Contrast has been enhanced equally for all images. NB. The cells displayed are the same as those in Figure 106.

To investigate the differences in levels of expression and total numbers of clusters between the different populations of cephalexin-treated cells, YFP fluorescence signal-based protein concentration and $CheW_4$-YFP cluster densities were calculated for all four strains (Table 30).

Table 30 (next page): $CheW_4$-YFP Cluster densities and protein concentrations in populations of cephalexin-treated cells: Cluster densities within each population, calculated as the number of clusters divided by the cell length, are displayed as a histogram. Protein concentrations, calculated as the sum of pixel intensities within each cell divided by the number of pixels that fall within each cell, are calculated as the mean value of all cells, error represents the standard deviation of this mean.

Background Strain	$CheW_4$-YFP Cluster Densities, μm^{-1} (Number of Clusters/Cell Length)	Cellular Protein Concentrations (Sum of Pixel Values/Pixel area)	Total Number of Cells
wild-type (JPA1447)		5100 ± 1800	247
ΔflgM (JPA2002)		8700 ± 2700	265
ΔppfA (JPA2003)		2400 ± 1700	173
ΔflgM, ΔppfA (JPA2004)		6000 ± 2400	171

As expected, the effect of deleting *ppfA* from these cells (JPA2003 and JPA2004) was to drastically reduce the total number of clusters within each cell. As had been observed with the protein copy numbers from mid-log phase microaerobic cultures (see section (6.3.1)), cellular $CheW_4$-YFP protein concentrations were lower in the strains lacking *ppfA*.

The deletion of *flgM* (JPA2002 and JPA2004) had a much greater effect on the concentrations of protein within each population for $CheW_4$-YFP than had been observed for CFP-$CheW_3$ (see section (6.3.3)), suggesting that gene expression changes induced by FliA and FlgM still play a role in cells treated with cephalexin, however this had little effect on cluster density between strains.

In the absence of *ppfA* and cluster partitioning, cluster nucleation is predicted to be a stochastic and concentration dependent process, similar to that of polar cluster formation. The change in concentration of protein caused by deletion of *flgM* in the *ppfA* deletion strains JPA2003 and JPA2004 allowed the effect of concentration on cluster nucleation to be investigated. 22 % of cephalexin-treated cells that contained FlgM (JPA2003) had more than one visible cluster with a maximum cluster number of 3. 29 % of cephalexin-treated cells that lacked FlgM (JPA2004) had more than one visible cluster with a maximum cluster number of 4. If higher concentrations of protein do facilitate the formation of clusters in the absence of PpfA, then this effect was subtle in the two strains tested. Unfortunately, attempts to over-express FliA in cephalexin cells were unsuccessful as, for an unknown reason, cells did not elongate in the presence of both IPTG and cephalexin.

For the analysis of cluster positioning, data were grouped for cells that contained PpfA (JPA1447 and JPA2002), and cells that lacked PpfA (2003 and 2004). Grouping these data was deemed reasonable because of the similar cluster density measurements between these strains (Table 30). The localisation of $CheW_4$-YFP cytoplasmic clusters was not observed to form a regular pattern in cephalexin-treated cells in the presence of PpfA (Figure 110A). Interestingly, in the absence of PpfA a noticeable proportion of clusters localised around the mid-point of the cell, but clusters could be found across the whole length of cephalexin-treated cells (Figure 110B).

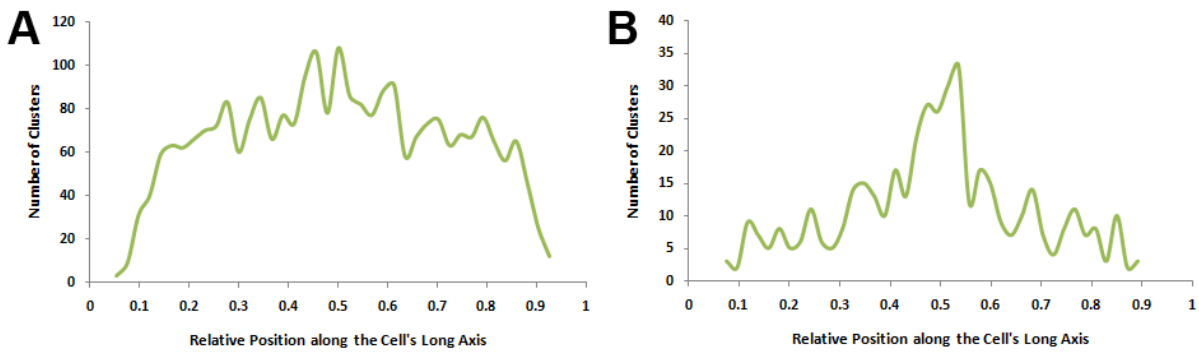


Figure 110: Positions of CheW$_4$-YFP cytoplasmic clusters in cephalexin-treated cells. - Positions of clusters in cells are calculated relative to the whole cell length for **A)** cells with PpfA (JPA1447 and JPA2002), and **B)** cells without PpfA (JPA2003 and JPA2004) Data are grouped in 40 equal sized bins between 0 and 100 %.

Similarly, no obvious pattern was visible when cluster localisation was visualised with cell length (Figure 111).

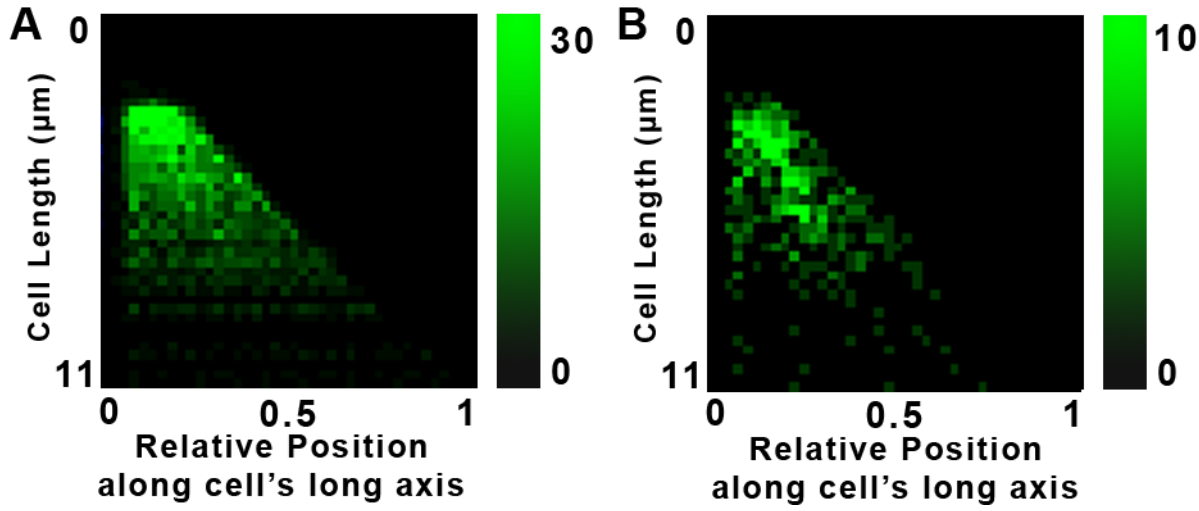


Figure 111: Positions of cytoplasmic CheW$_4$-YFP protein clusters with different length cells: 2D Histogram of relative cluster positions with cell length. Data are grouped in 40 equal sized bins for cell length and 40 equal sized bins for relative cluster position. The occupancy of each bin is indicated by an intensity of green which corresponds to a number of clusters, indicated by the colour-bar to the right of the figure. Positions are plotted for **A)** cells containing PpfA (JPA1447 and JPA2002), and **B)** cells with PpfA deleted (JPA2003 and JPA2004).

6.4) Observation and isolation of long-cell spontaneous mutants:

During the course of this study, an intriguing observation was made. A 2 % agar plate inoculated with 5 µl of cell culture from the FliA deletion strain, JPA467, produced a flare after more than a week of incubation at 30 °C. The flare had moved further across the agar surface than the rest of the colony (Figure 112), this was unexpected as *R.sphaeroides* is not known to exhibit the surface-motility of other bacterial species and thus this phenomenon was further investigated.

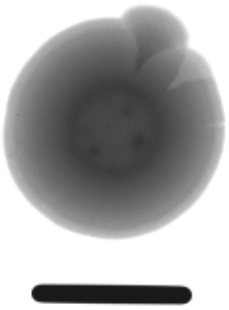


Figure 112: Example flaring colony from strain JPA467. Scale Bar represents 10 mm.

Cells from this flare were picked from the plate, resuspended in 50 µl of succinate medium, and directly visualised under a phase-contrast light microscope. Cells from the flare were generally longer than cells picked from the edge of the same colony that had not flared (Figure 113).

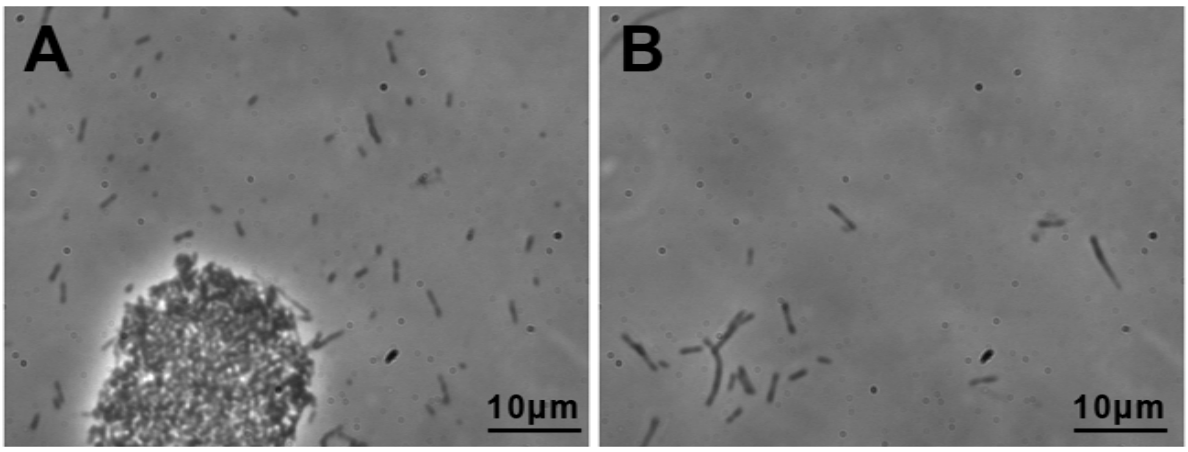


Figure 113: The appearance of flared cells – Example fields of view from phase contrast images taken at 40-fold magnification. The cells imaged are picked from the edge of the flaring colony shown in Figure 112. Cells were picked from **A)** the edge that did not flare and from **B)** the flare.

To see whether this phenomenon was the product of an isolatable spontaneous mutation, the flaring cells were inoculated back onto a second 2 % agar plate, and left to grow for 4-5 days. Cells from the edge of this colony were picked again and observed via phase contrast microscopy (Figure 114). The proportion of cells that appeared long from this second round was greater than those from the first round, suggesting that the long ph/genotype had been enriched by this second round of growth.

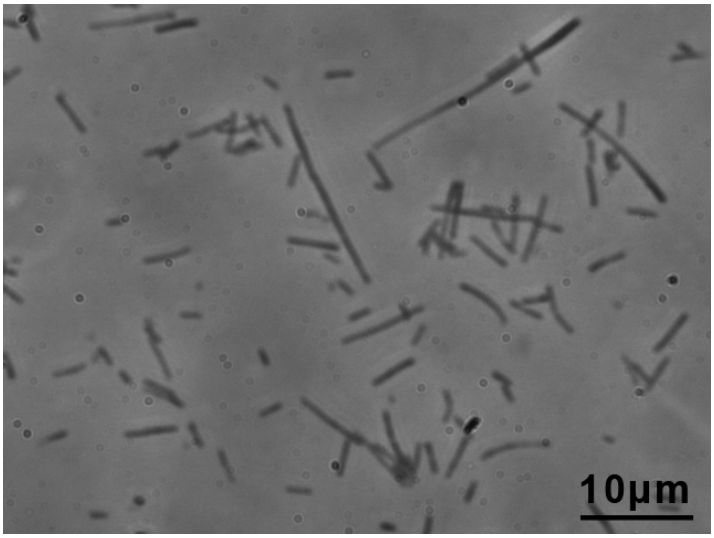


Figure 114: The appearance of flared cells after enrichment – example field of view from phase contrast images taken at 40-fold magnification.

Cells from this population were streaked onto a third 2 % agar plate, and from the resulting single colonies of enriched long cells, cultures were grown photoheterotrophically at 30 °C in succinate medium overnight, and to mid-log phase microaerobically the next day. The long phenotype was seen to persist in liquid culture and it was therefore concluded to be the product of a spontaneous chromosomal mutation.

To further investigate the relevance of this observed mutation, the same long-cell isolation protocol as above (Figure 115) was repeated for several different *R.sphaeroides* mutant strains. Colony flares were regularly seen on plates, and long cells could frequently be observed under phase contrast, however the isolation of strains where the long-phenotype persisted in liquid culture was only successful for seven different strains (Table 31).

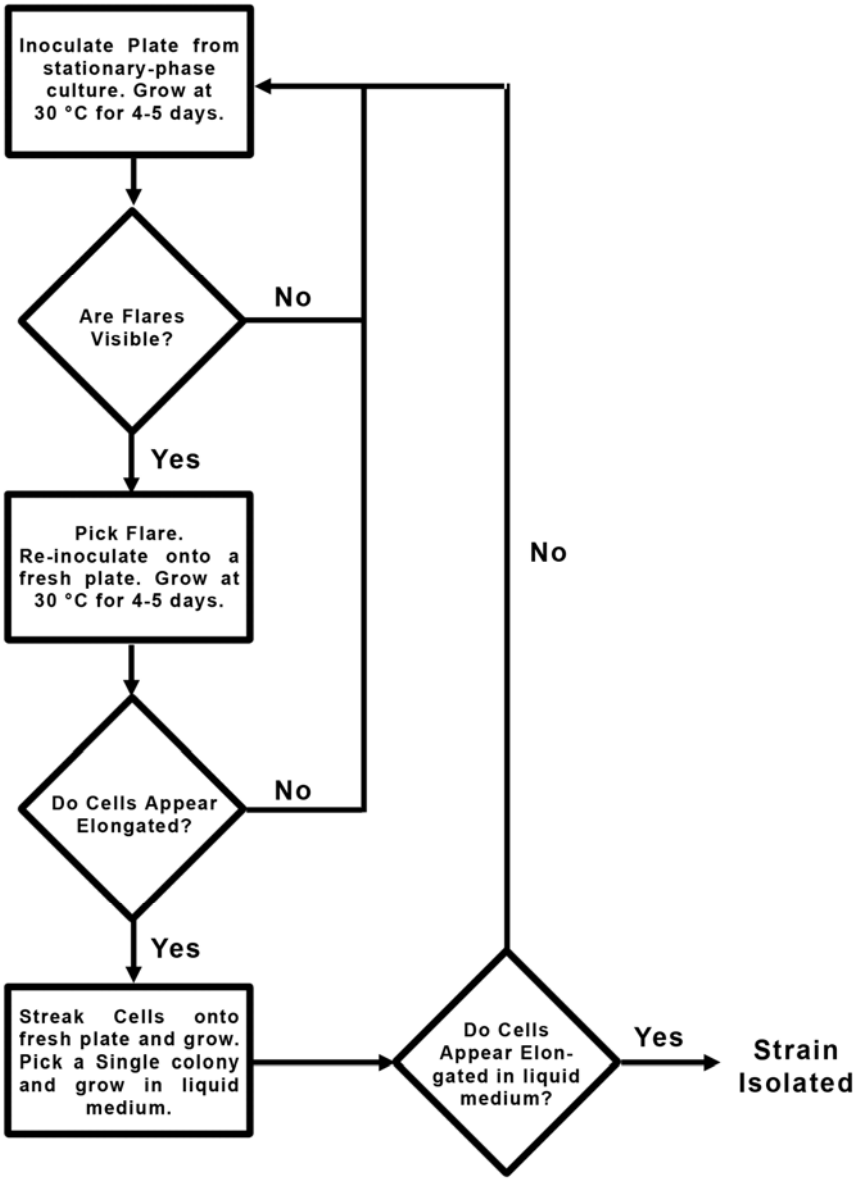


Figure 115: Isolating long-cell mutant strains. Schematic flow chart of the method employed to isolate long-cell mutant strains

Strain	Genotype Description	Flares Obtained	Strain Isolated	Designated Strain Name
WS8N	Wild-type	✓	×	-
JPA2001	*ΔflgM*	✓	×	-
JPA467	*ΔfliA*	✓	✓	JPA2050
JPA1638	*yfp-fliM*	✓	✓	JPA2051
JPA2032	*yfp-fliM, ΔfliA*	✓	✓	JPA2052
JPA1447	*cheW$_4$-yfp, cfp-cheW$_3$*	✓	×	-
JPA2002	*cheW$_4$-yfp, cfp-cheW$_3$, ΔflgM*	✓	×	-
JPA2003	*cheW$_4$-yfp, cfp-cheW$_3$, ΔppfA*	✓	×	-
JPA2004	*chew$_4$-yfp, cfp-cheW$_3$, ΔppfA, ΔflgM*	✓	✓	JPA2053
JPA2006	*yfp-motB*	✓	×	-
JPA1418	*cheW$_4$-yfp*	✓	✓	JPA2054
JPA2018	*cheW$_4$-yfp, ΔflgM*	✓	×	-
JPA1457	*yfp-cheW$_3$*	✓	×	-
JPA2019	*yfp-cheW$_3$, ΔflgM*	✓	✓	JPA2055
JPA500	*mcpG-gfp*	✓	✓	JPA2056

Table 31: Isolated Spontaneous mutants – List of strains for which spontaneous long cell mutants were seen and of those which were isolated.

6.5) Phenotypic characterisation of long-cell Mutants:

Colony morphologies were analysed in successfully isolated strains from 2 % agar plates that had been inoculated with 5 µl of stationary-phase cell culture and left to grow for 4 days at 30 °C (Table 32).

		Colony Diameter (mm)		
Parental Strain Name	**Genotype Description**	**Parental Strain**	**Mutant Derivative**	**50:50 mix**
JPA467	*ΔfliA*	12.7 ± 0.6	15.3 ± 0.6	16.0 ± 0
JPA1638	*yfp-fliM*	12.0 ± 0	13.7 ± 0.6	15.0 ± 0
JPA2032	*yfp-fliM, ΔfliA*	12.0 ± 1.0	14.7 ± 0.6	15.0 ± 0
JPA2004	*chew$_4$-yfp, cfp-cheW$_3$, ΔppfA, ΔflgM*	12.3 ± 0.6	15.3 ± 0.6	16.0 ± 0
JPA1418	*cheW$_4$-yfp*	12.0 ± 0	15.0 ± 0	15.3 ± 0.6
JPA2019	*yfp-cheW$_3$, ΔflgM*	12.0 ± 0	13.3 ± 0.6	14.3 ± 0.6
JPA500	*mcpG-gfp*	12.0 ± 0	14.7 ± 0.6	15.7 ± 0.6

Table 32: Colony diameters: Diameters of colonies on 2 % agar plates. Diameters were measured to the nearest millimetre and values show the average and standard deviation of the diameters measured from three repeats.

For each isolated mutant, colonies spread slightly further on plates than cells from the parental background strain. Colonies grown from a 50:50 mixture of the parental strain and the long-cell isolate showed the same colony phenotype as the long-cell isolate.

Having determined the effect of the isolated mutations on the colony formation at hard-agar surfaces, the swimming efficiency of the mutants was also analysed and compared to each parental strain on soft-agar swim plates (Figure 116).

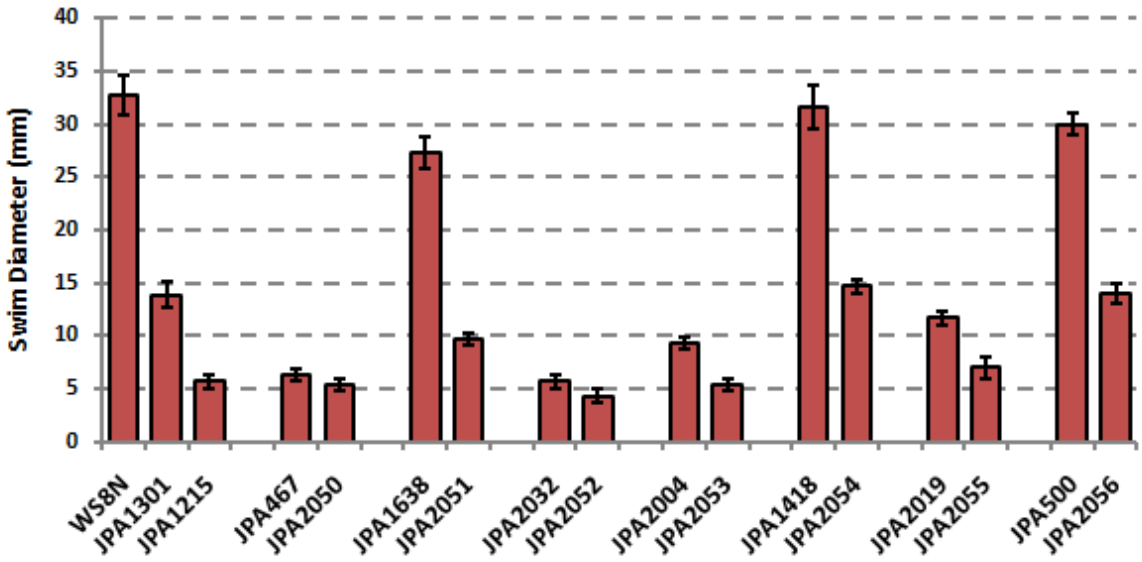


Figure 116: Chemotactic efficiency of flaring mutants: Chemotactic efficiency was measured on soft-agar swim plates containing 100 µM propionate. Parental strains and their long-cell derivatives are displayed in pairs. Average diameters from three repeats on separate days were calculated for parental strains and their flared-mutant derivatives. Values were compared to wild-type (WS8N), non-chemotactic (JPA1301) and non-motile (JPA1215) controls. Error bars represent the standard deviation of the three repeats.

All long-cell isolates showed a reduced swimming efficiency on soft-agar swim plates compared to the parental strain, except where the parent was non-motile and did not swim.

In order to characterise cellular morphologies of each long-cell isolate cell lengths and septation sites were measured for approximately 100 cells from each population using the MATLAB program discussed in chapter (3) (Figure 117). Three different subtypes of long-cell morphology were observed via this analysis:

1) Populations that appeared long, but had one or several clearly visible septation sites, these included *yfp-fliM* (JPA2051), *yfp-fliM,ΔfliA* (JPA2052) and *mcpG-gfp* (JPA2056).

2) Populations that had wild-type-like cells with a proportion of the population displaying an increased length, cells could be septate or not. These included *cheW$_4$-yfp* (JPA2054), *yfp-cheW$_3$ ΔflgM* (JPA2055) and *ΔfliA* (JPA2050).
3) Populations that had predominantly long non-septate populations. *chew$_4$-yfp,cfp-cheW$_3$, ΔppfA,ΔflgM* (JPA2053) was the only example of this.

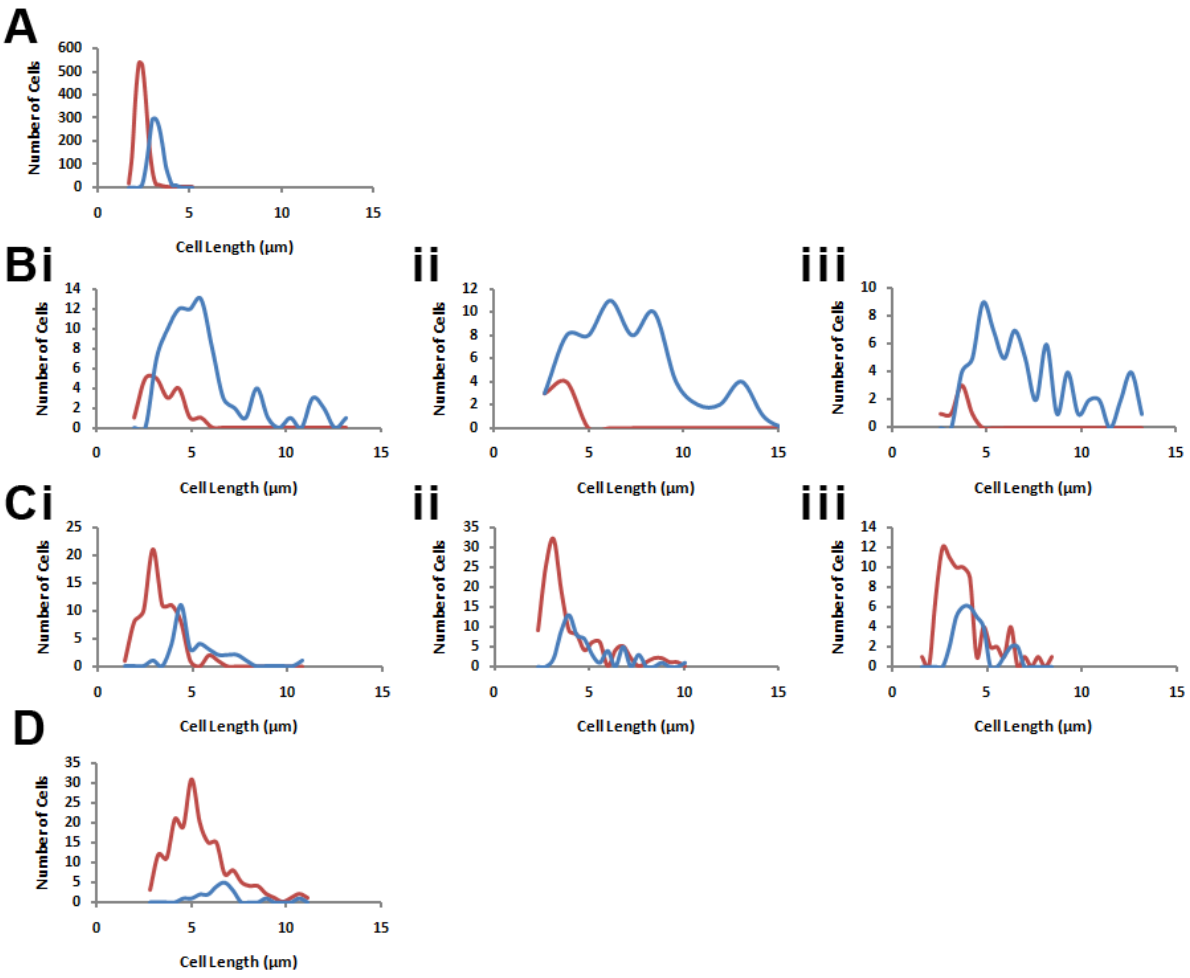


Figure 117: Flare-mutant cell lengths: Smooth histograms of cell lengths of non-dividing (red) and dividing (blue) cells within each population. **A)** Wild-type cell data from section (4.1). **Bi)** *yfp-fliM* (JPA2051, n=100) **ii)** *yfp-fliM,ΔfliA* (JPA2052, n=72) **iii)** *mcpG-gfp* (JPA2056, n=73). Cells had increased lengths, but had many visible septation sites. **Ci)** *cheW$_4$-yfp* (JPA2054, n=108) **ii)** yfp-*cheW$_3$ ΔflgM* (JPA2055, n=193) **iii)** *ΔfliA* (JPA2050, n=110). Cells had both wild-type subpopulations and long-cell subpopulations. **D)** *chew$_4$-yfp,cfp-cheW$_3$,ΔppfA,ΔflgM* (JPA2053, n=202) Cells had a reduced proportion of dividing cells, with the majority of the population being long and non-septate. Histograms are plotted between cell lengths of 0 and 15 µm. Data are grouped into 20 equal sized bins which span the total range of lengths obtained from each population.

Further observation of the long-cell isolated mutants via electron microscopy (Figure 118) did not provide any evidence for active surface motility in these mutants. Populations generally possessed few or no flagella and although cells appeared distinguishably longer than a wild-type

population, there were no defining features visible further to those observed via epifluorescence imaging.

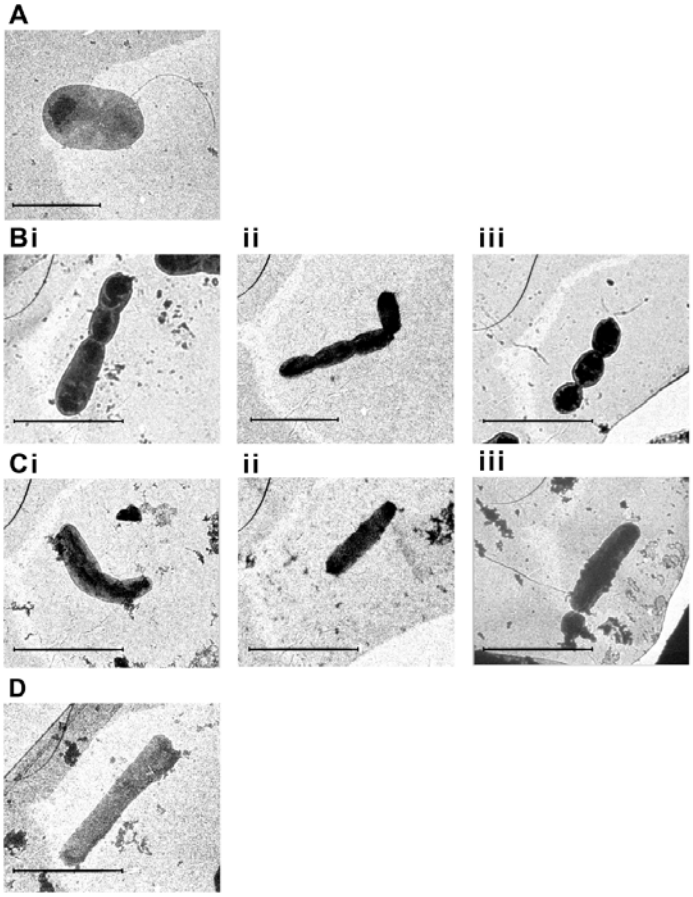


Figure 118: Flare-mutant electron microscopy: Example electron micrographs of cells at various magnifications. **A)** wild-type (WS8N), scale bar represents 2 µm. **Bi)** *yfp-fliM* (JPA2051) **ii)** *yfp-fliM,ΔfliA* (JPA2052) **iii)** *mcpG-gfp* (JPA2056). Cells had increased lengths, but had many visible septation sites. **Ci)** *cheW$_4$-yfp* (JPA2054) **ii)** *yfp-cheW$_3$,ΔflgM* (JPA2055) **iii)** *ΔfliA* (JPA2050). Cells had both wild-type subpopulations and long-cell subpopulations. **D)** *chew$_4$-yfp,cfp-cheW$_3$,ΔppfA,ΔflgM* (JPA2053) Cells had a reduced proportion of dividing cells, with the majority of the population being long and non-septate. Scale bars represent 5 µm.

6.6) Cytoplasmic cluster positioning in long-cell mutants

The localisation of protein clusters was revisited for the obtained long-cell mutants that contained cytoplasmic cheW$_4$-YFP clusters (JPA2055), and polar CFP-cheW$_3$ clusters (JPA2053) was addressed to see if localisation patterns differed between cephalexin-treated long cells and those generated via spontaneous mutation.

The lengths of the cells in the 193 measured cells of strain JPA2054 were comparable to those generated in the cephalexin-treated cells (section (6.3.2)). As had been observed for cytoplasmic

CheW$_4$-YFP clusters in sections (6.3.4) and (4.3), no obvious pattern in localisation relative to the cell's length was observed. A plot of all points showed a relatively uniform distribution across all non-polar subcellular locations (Figure 119).

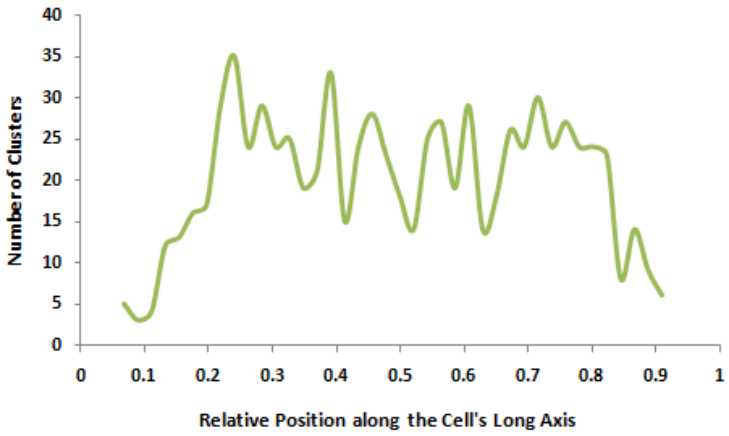


Figure 119: Positions of CheW$_4$-YFP cytoplasmic clusters in long-cell mutant strain JPA2054. Positions of clusters in cells are calculated relative to the whole cell length.

Similarly, no pattern was observed when these positions were plotted with cell length (Figure 120).

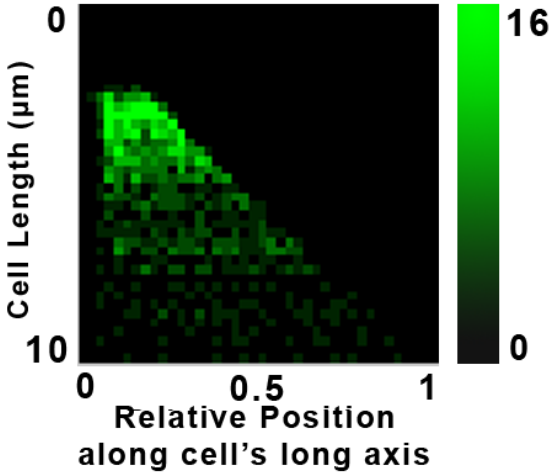


Figure 120: Positions of cytoplasmic CheW$_4$-YFP protein clusters with different length cells: 2D Histogram of relative cluster positions with cell length. Data are grouped in 40 equal sized bins for cell length and 40 equal sized bins for relative cluster position. The occupancy of each bin is indicated by an intensity of green which corresponds to a number of clusters, indicated by the colour-bar to the right of the figure.

6.7) Polar/Lateral cluster positioning in long-cell mutants

Positions of polar/lateral CFP-CheW$_3$ clusters from 202 cells of the long-cell mutant strain JPA2053 were analysed in the same manner (Figure 121 and Figure 122). This strain was chosen as data were obtained for positioning of CFP-CheW$_3$ cluster in cephalexin-treated cells of its parental strain, JPA2004, in section (6.3). The regular pattern of localisation that had been observed in cephalexin cells (Figure 107 and Figure 108) was no longer observed in the spontaneous long-cell derivative strain JPA2053.

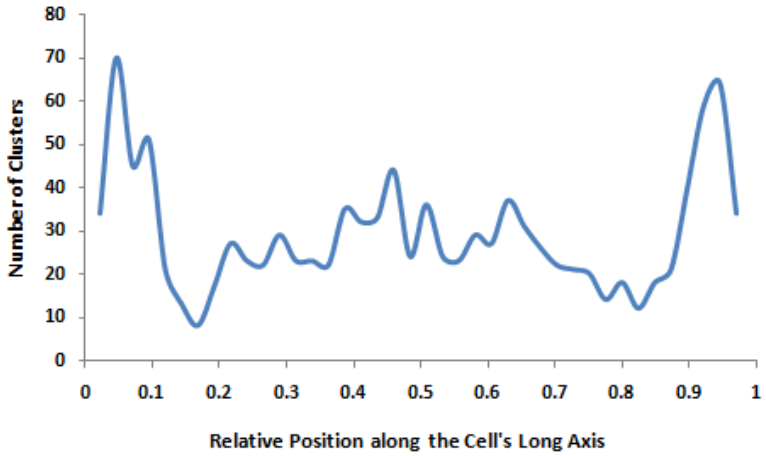


Figure 121: Positions of CFP-CheW$_3$ polar/lateral clusters in long-cell mutant strain JPA2053. Positions of clusters in cells are calculated relative to the whole cell length.

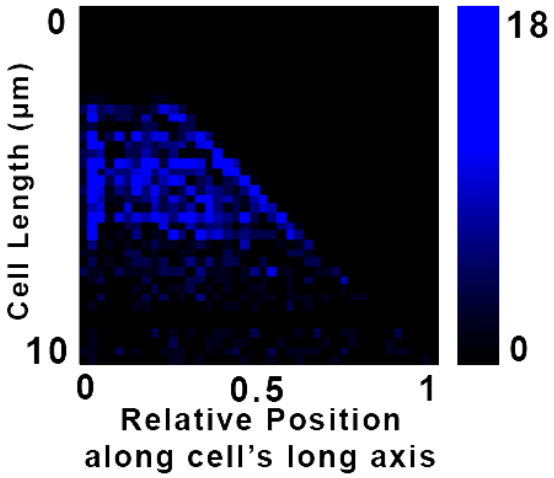


Figure 122: Positions of polar/lateral CFP-CheW$_3$ protein clusters with different length cells: 2D Histogram of relative cluster positions with cell length. Data is grouped in 40 equal sized bins for cell length and 40 equal sized bins for relative cluster position. The occupancy of each bin is indicated by an intensity of blue which corresponds to a number of clusters, indicated by the colour-bar to the right of the figure.

6.8) Chapter Discussion

Here, deletion of *fliA* and *flgM* combined with plasmid-induced expression of *fliA* enabled the study of chemotaxis protein clustering at different cellular copy numbers of the chemotaxis proteins. Interestingly, where polar/lateral chemotaxis protein clusters were observed to form more frequently at higher protein concentrations in *E.coli* (148), *R.sphaeroides* did not show an obvious increase in cluster number over the ~4-fold variation of protein copy number that was tested. At lower, near-wild-type levels of chemotaxis proteins, lateral cluster formation was rare and clusters were not necessarily observed to form near to septation sites or with any pattern in positioning that may indicate the presence of the "pre-division site" as in *E.coli* (147). At nearly three times wild-type concentrations of chemotaxis protein, however, lateral clusters were observed to form close to septation sites. One factor that may influence chemotaxis protein cluster positioning is that of membrane curvature. Transmembrane MCPs are the structural components that anchor polar/lateral chemotaxis clusters to the membrane, and there is evidence that MCPs may localise preferentially to curved membranes, such as those at the poles of rod-shaped bacterial cells (55). Membrane curvature is not factored into the stochastic nucleation model, but a preferential localisation to curved membrane regions may explain the observed patterns in cluster localisation made in this study, as dividing cells will have curved membrane regions at newly forming cell poles. Increased protein concentrations may simply facilitate the visualisation of what are otherwise small protein clusters.

The polar/lateral chemotaxis protein clusters of *R.sphaeroides* also show other interesting and subtle differences to those reported in *E.coli*. Lateral clusters were commonly observed in filamentous cephalexin-treated *R.sphaeroides* cells, and showed regular localisation patterns at the 1/3 and 2/3 positions within the cells. When similar shaped cells were generated by an undefined spontaneous mutation, this pattern was abolished suggesting that lateral chemotaxis protein positioning maybe controlled by some form of cytoskeletal element. Additionally, multiple lateral clusters were frequently observed to form at similar loci on either side of the cell's membrane. When available images of *E.coli* cells were inspected, this was not seen to be the case – and lateral clusters seemed solitary, forming on only one side of the cell (Figure 123).

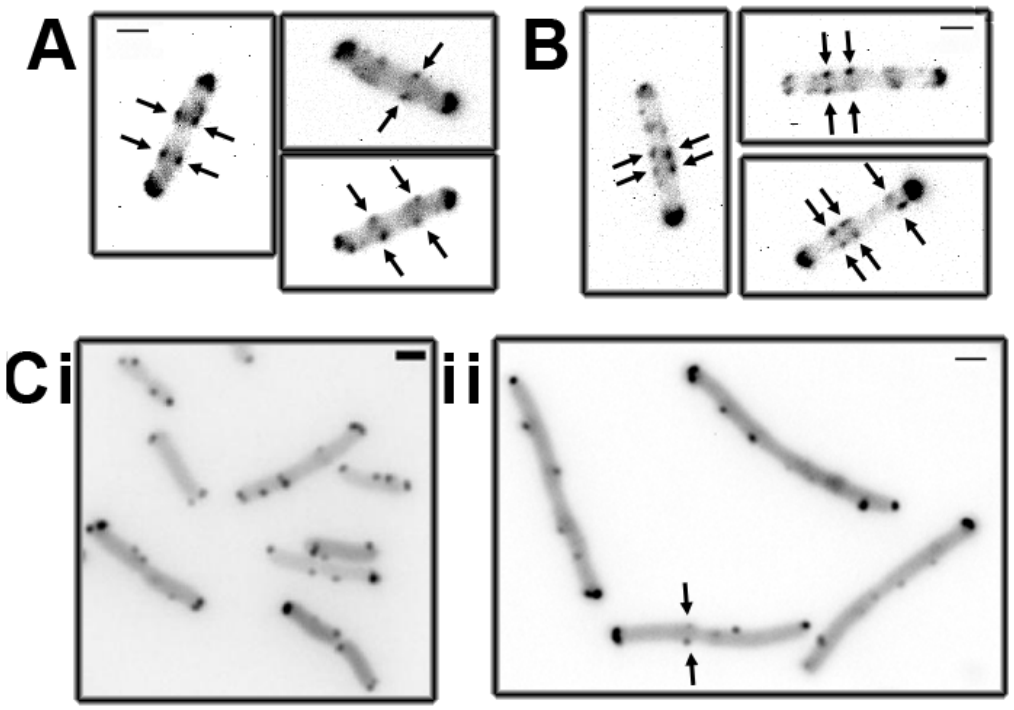


Figure 123: Lateral cluster positioning in *R.sphaeroides* and *E.coli*. Images of polar and laterally positioned protein clusters in **A)** cephalexin-treated *R.sphaeroides* JPA2004, **B)** spontaneous long-cell mutant JPA2004-derivative (JPA2053), **Ci)** *E.coli* wild-type cells **Cii)** *E.coli* cephalexin-treated cells. Multiple lateral clusters that occupy the same locus are highlighted with arrows, they appear commonly in *R.sphaeroides* cells, and less common in *E.coli* cells (from available published pictures). Scale bars represent 1 µm.

This observation and the observation that multiple clusters form frequently at individual poles in *R.sphaeroides* and rarely in *E.coli*, suggest that the diffusion dynamics and mechanism control of cluster localisation differs between the two species.

When the cytoplasmic chemotaxis protein clusters of *R.sphaeroides* were studied, cluster number showed a clearer positive correlation with protein copy number than polar/lateral protein clusters. This is intriguing as cytoplasmic cluster partitioning is an active process controlled by the protein PpfA and therefore cytoplasmic clusters are not predicted to undergo stochastic nucleation in the same way as polar chemotaxis protein clusters. Instead, clusters are divided into separate "daughter clusters" by an unknown mechanism. The data presented here suggest that the rate of cluster partitioning is dependent on chemotaxis protein concentration. In addition, average cluster sizes (measured as fluorescence intensities) were seen to increase with protein copy number. This suggests that the partitioning of an individual cluster is not dependent on the size of that cluster. If cluster partitioning happened when a cluster reached a certain "threshold size", then clusters across all populations would be predicted to have similar sizes, and only the number of clusters would vary

with protein copy number. A better explanation may be that increased expression of PpfA, which is encoded with the chemotaxis genes of *che*Op3, results in an increased rate of cluster partitioning.

Multiple cytoplasmic clusters were observed to form along the length of elongated cells for both cephalexin-treated and spontaneous mutant populations. By-eye, the spacing of these clusters appeared regular, however, analysis of the positions of these cytoplasmic protein clusters showed no regular subcellular pattern to their positioning. In elongated cells that lacked PpfA small numbers of clusters were seen, showing that the separation and distribution of these clusters is controlled by the same active partitioning mechanism responsible for the inheritance of daughter clusters in wild-type cells. The mechanism by which this is achieved is not clear and it is difficult to rationalise how cluster positions can be determined relative to the cell pole (see section (4.4)) and simultaneously be affected along the length of long cells that lack adjacent poles and new-pole septation sites.

The cause of the long-cell spontaneous mutations remains unknown, but resulted in flaring on solid agar surfaces. Cells presumably encountered a form of selection pressure on solid agar plates that allowed elongated cells to have better access to nutrients. This was not seen to be achieved via active motility, but directly as a consequence of cell length. This is highlighted by the variety of different long-cell phenotypes that were observed microscopically in the isolated mutants; cells could be long and septate – presumably with a mutation in the genes controlling cytokinesis. Isolated mutants could also be long without being septate, which may be the result of a genetic defect in the apparatus of cell division or a cytoskeletal element. Due to the large numbers of possible mutations that may have led to the observed long-cell phenotypes, gene sequencing to identify the responsible point mutations was not attempted. The *R.sphaeroides* WS8N genome is, however, currently being sequenced and its availability may allow for a Solexa™-style method of sequencing that could identify the mutations involved in this long-cell phenotype by whole-genome analysis.

7) FliA and FlgM regulate motility and flagellar genes

7.1) Motility and Chemotaxis

7.1.1) Chemotaxis with varying FliA activity

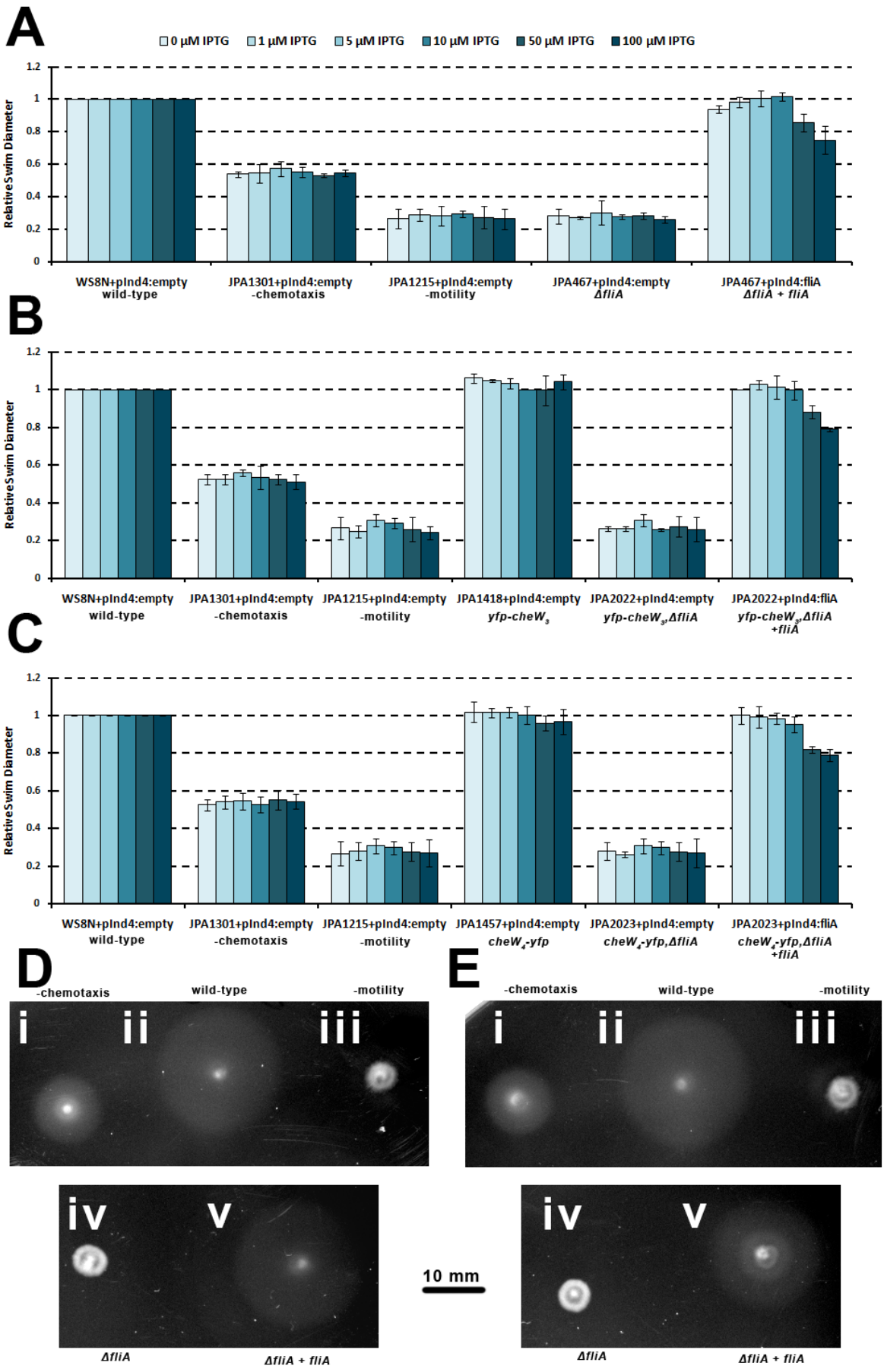
Having determined the effect of varying FliA and FlgM activity on the level of expression of components from the two essential chemotaxis operons of *R.sphaeroides*, it was decided to examine whether these changes affect chemotaxis itself.

Deletion of *fliA* has previously been shown to result in a loss of motility, which is caused by the lack of ability to express and synthesise the flagellar filament, composed of thousands of copies of the FliC protein (98).

To investigate how FliA activity affects chemotaxis, pInd4:*fliA* was conjugated into the *fliA* deletion strain (JPA467) and deletion strains expressing YFP-$CheW_3$ (JPA2022) and $CheW_4$-YFP (JPA2023). Recovery and over-expression of *fliA* was achieved by the addition of different concentrations of IPTG to soft agar swim plates which contained kanamycin and 100 µM propionate. Inoculated plates were left to grow for two days and the swim-diameter of each colony was measured in order to assess the chemotactic efficiency of these strains (Figure 124).

In each of these strains, and similarly to the complementation experiment performed in section (5.7), the activity restored to the *fliA* deletion backgrounds was sufficient to restore near-wild-type efficiency of chemotaxis to each strain, even in the absence of IPTG. At higher induction levels (50 and 100 µM IPTG) a slight decrease in chemotactic efficiency was observed, however, this decrease was most likely caused by a reduction in growth rate, which had previously been observed to be reduced under these conditions (section (5.7.1)).

Figure 124 (next page): Chemotaxis with complement and overexpression of FliA: Swim diameters show chemotactic efficiency of cells on soft-agar swim plates after deletion, and reintroduction of FliA at different concentrations of IPTG (indicated by the shade of blue) for **A)** wild-type background strains, **B)** YFP-$CheW_3$ background strains and **C)** $CheW_4$-YFP background strains. All numbers are calculated relative to a WS8N wild-type control which was included on each plate. Chemotactic efficiency is also compared to non-chemotactic (JPA1301) and non-motile (JPA1215) control strains. Error bars represent the standard deviation of three repeats taken on three separate days. Example images of swim colonies are shown at **D)** 0 µM IPTG, for **i)** non-chemotactic , **ii)** wild-type, **iii)** non-motile, **iv)** *fliA* deletion and **v)** *fliA* recovered strains. Example images of swim colonies are shown at **E)** 100 µM IPTG for **i)** non-chemotactic, **ii)** wild-type, **iii)** non-motile, **iv)** *fliA* deletion and **v)** *fliA* recovered strains. Image contrast has been adjusted for clarity.


A
0 µM IPTG
1 µM IPTG
5 µM IPTG
10 µM IPTG
50 µM IPTG
100 µM IPTG
RelativeSwim Diameter
1.2
1
0.8
0.6
0.4
0.2
0
WS8N+pInd4:empty
wild-type
JPA1301+pInd4:empty
-chemotaxis
JPA1215+pInd4:empty
-motility
JPA467+pInd4:empty
ΔfliA
JPA467+pInd4:fliA
ΔfliA + fliA
B
RelativeSwim Diameter
WS8N+pInd4:empty
wild-type
JPA1301+pInd4:empty
-chemotaxis
JPA1215+pInd4:empty
-motility
JPA1418+pInd4:empty
yfp-cheW3
JPA2022+pInd4:empty
yfp-cheW3,ΔfliA
JPA2022+pInd4:fliA
yfp-cheW3,ΔfliA +fliA
C
RelativeSwim Diameter
WS8N+pInd4:empty
wild-type
JPA1301+pInd4:empty
-chemotaxis
JPA1215+pInd4:empty
-motility
JPA1457+pInd4:empty
cheW4-yfp
JPA2023+pInd4:empty
cheW4-yfp,ΔfliA
JPA2023+pInd4:fliA
cheW4-yfp,ΔfliA +fliA
D
-chemotaxis
wild-type
-motility
i
ii
iii
iv
v
ΔfliA
ΔfliA + fliA
10 mm
E
-chemotaxis
wild-type
-motility
i
ii
iii
iv
v
ΔfliA
ΔfliA + fliA

As soft-agar swim plates are sensitive to affects on rates of metabolism and growth. Growth rate, growth conditions and IPTG uptake efficiency are likely to differ between soft-agar swim plate and liquid growth conditions. Therefore, in order to test the chemotactic response of FliA-expressing cells grown in liquid culture, the phototaxis assay and tethered-cell assays were used. *ΔfliA* cells (JPA467) containing the pInd4:*fliA* plasmid that were grown in the absence of IPTG showed a wild-type response to the removal of propionate in tethered cell experiments (Figure 125) and reacted normally to the phototaxis assay, showing that the motility and chemotaxis are restored under these conditions.

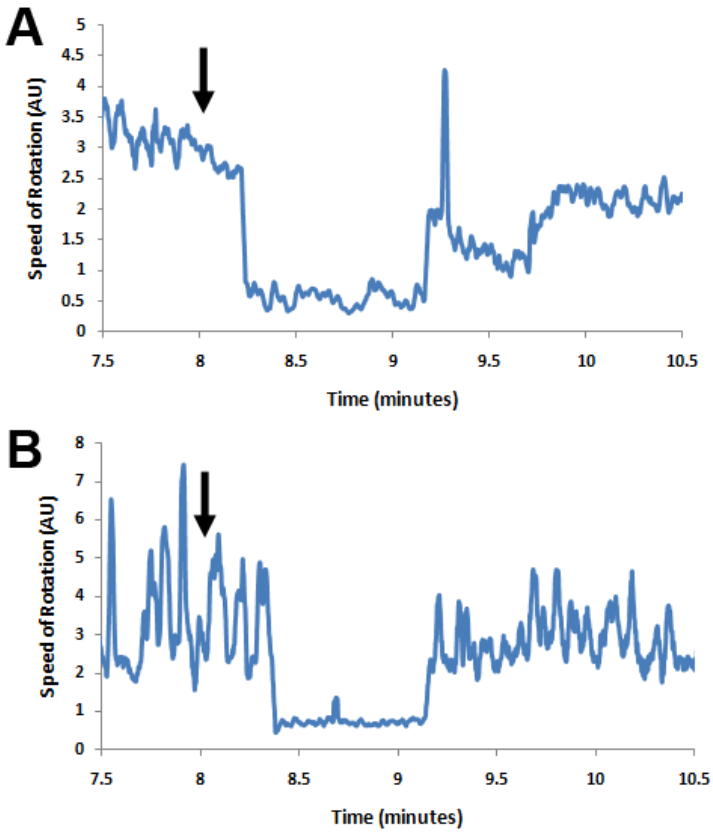


Figure 125: Tethered-cell analysis with FliA complementation. Representative example single cell speed/time traces for **A)** a wild-type (WS8N+pInd4:empty) cell, and **B)** a *ΔfliA* cell with FliA activity restored from a plasmidin the absence of IPTG (JPA467,pInd4:*fliA*). Traces are displayed for the section of each video where propionate is removed (marked by an arrow) inducing the chemotactic response. Cellular rotation stops shortly after the removal of propionate and rotation restarts approximately one minute after this stimulus.

Interestingly, overexpression of *fliA* in microaerobic liquid media at 50 and 100 µM concentrations of IPTG resulted in a non-motile phenotype when observed microscopically.

7.1.2) Chemotaxis with varying FlgM activity

As the only physiological role of FlgM is believed to be as the FliA anti-sigma factor, and increased activity of FliA did not greatly affect chemotactic efficiency on swarm plates, the *flgM* deletion was predicted not to have an effect on the chemotactic response of *R.sphaeroides*. However, soft-agar swim plate analysis of strain JPA2001 showed a reduced chemotactic efficiency, comparable to that of the non-chemotactic control strain JPA1301 (Table 33).

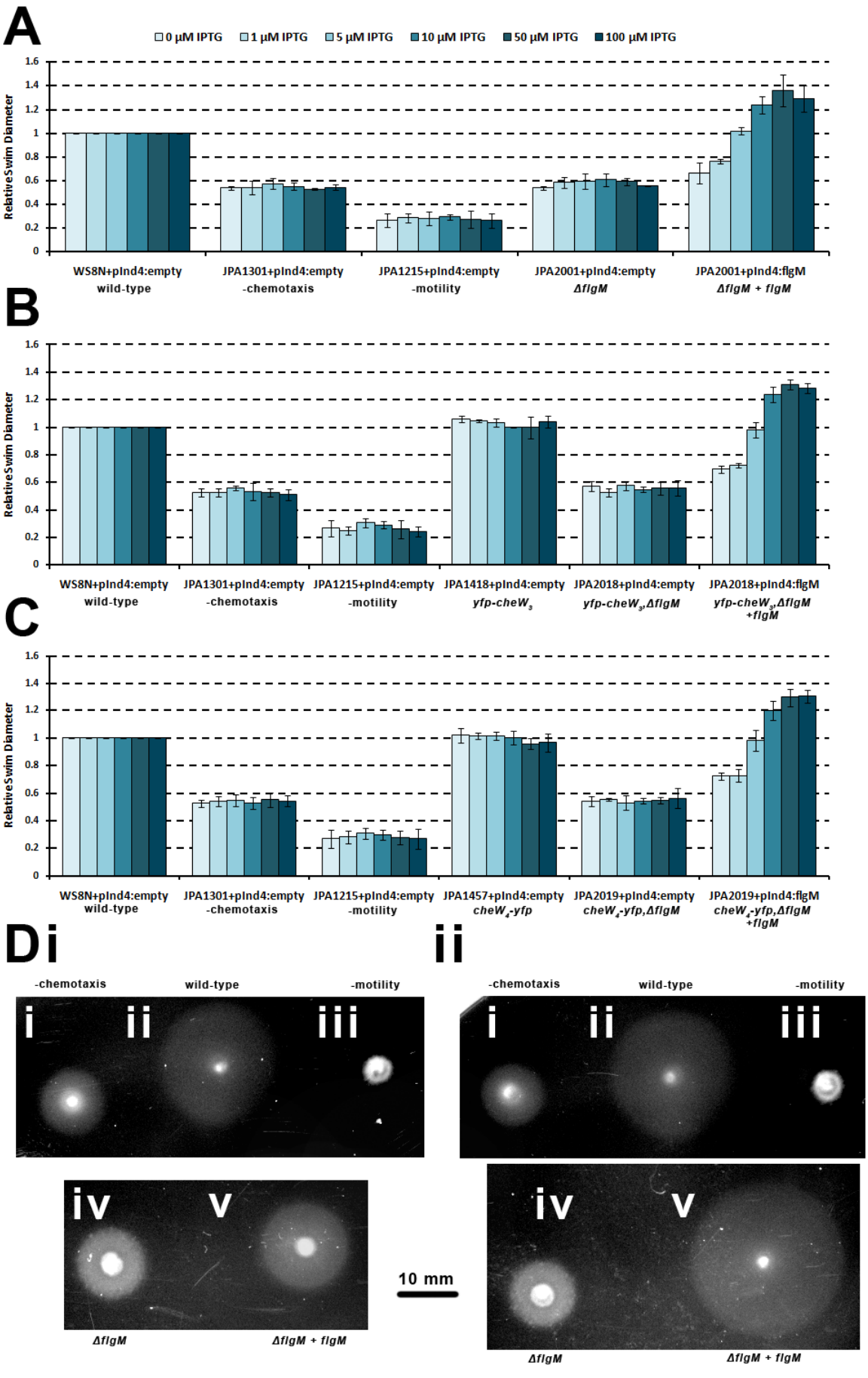
Strain	Description	Swim Diameter (mm)
WS8N	Wild-type	19.0 ± 1.0
JPA1301	Non-chemotactic control	9.7 ± 1.5
JPA1215	Non-motile control	5.3 ± 0.6
JPA467	*ΔfliA*	5.7 ± 0.6
JPA2001	*ΔflgM*	12.7 ± 0.6

Table 33: Chemotactic efficiency of *ΔflgM* and *ΔfliA* mutants. Swim diameters on soft-agar swim plates were calculated and compared to wild-type (WS8N), non-chemotactic (JPA1301) and non-motile (JPA1215) strains. Error represents the standard deviation of three repeats from three separate days.

pInd4:*flgM* was then conjugated into the *flgM* deletion strain (JPA2001) and deletion strains expressing YFP-$CheW_3$ (JPA2018) and $CheW_4$-YFP (JPA2019). The ability to complement the deletion was tested by reintroduction and over-expression of *flgM* at different concentrations of IPTG.

In the absence of IPTG, pInd4:*flgM* expression was not sufficient to restore wild-type chemotaxis to these mutants, however, at higher concentrations of IPTG (5 μM) the chemotactic efficiency of the *flgM* deletion strains was restored to wild-type levels. Interestingly, at further increased levels of expression of *flgM* (10, 50 100 μM), a super-swarming phenotype was observed which was consistent across different background strains (Figure 126).

Figure 126 (next page): Chemotaxis with complement and overexpression of FlgM: Swim diameters show chemotactic efficiency of cells on soft-agar swim plates after deletion, and reintroduction of FlgM at different concentrations of IPTG (indicated by the shade of blue) for **A)** wild-type background strains, **B)** YFP-$CheW_3$ background strains and **C)** $CheW_4$-YFP background strains. All numbers are calculated relative to a WS8N wild-type control which was included on each plate. Chemotactic efficiency is also compared to non-chemotactic (JPA1301) and non-motile (JPA1215) control strains. Error bars represent the standard deviation of three repeats taken on three separate days. Example images of swimming colonies are shown at **D)** 0 μM IPTG for **i)** non-chemotactic, **ii)** wild-type, **iii)** non-motile, **iv)** *flgM* deletion and **v)** *flgM* recovered strains. Example images of swimming colonies are shown at **E)** 100 μM IPTG for **i)** non-chemotactic, **ii)** wild-type, **iii)** non-motile, **iv)** *flgM* deletion and **v)** *flgM* recovered strains. Image contrast has been adjusted for clarity.


A
0 µM IPTG
1 µM IPTG
5 µM IPTG
10 µM IPTG
50 µM IPTG
100 µM IPTG
RelativeSwim Diameter
WS8N+pInd4:empty
wild-type
JPA1301+pInd4:empty
-chemotaxis
JPA1215+pInd4:empty
-motility
JPA2001+pInd4:empty
ΔflgM
JPA2001+pInd4:flgM
ΔflgM + flgM
B
RelativeSwim Diameter
WS8N+pInd4:empty
wild-type
JPA1301+pInd4:empty
-chemotaxis
JPA1215+pInd4:empty
-motility
JPA1418+pInd4:empty
yfp-cheW3
JPA2018+pInd4:empty
yfp-cheW3,ΔflgM
JPA2018+pInd4:flgM
yfp-cheW3,ΔflgM +flgM
C
RelativeSwim Diameter
WS8N+pInd4:empty
wild-type
JPA1301+pInd4:empty
-chemotaxis
JPA1215+pInd4:empty
-motility
JPA1457+pInd4:empty
cheW4-yfp
JPA2019+pInd4:empty
cheW4-yfp,ΔflgM
JPA2019+pInd4:flgM
cheW4-yfp,ΔflgM +flgM
Di
ii
-chemotaxis
wild-type
-motility
i
ii
iii
iv
v
ΔflgM
ΔflgM + flgM
10 mm

The reduction in chemotactic efficiency observed in the absence of *flgM* may be linked to the increased activity of FliA, as overexpression of *fliA* in microaerobic liquid conditions results in the loss of motility. However, increased expression of *fliA* on soft-agar swim plates does not result in the same phenotype, even though the activity of FliA at these concentrations is predicted to exceed that of the *flgM* deletion strain under microaerobic conditions (see section (5.7.1)). This discrepancy may result from a difference in expression levels achieved in microaerobic conditions to those of cells on soft-agar swim plates, as has been observed previously between different growth conditions and the over-expression from pInd4:*flgM* (see section (5.7.2)).

To further investigate the reduction in chemotactic efficiency in the *flgM* deletion strains, cells were grown to mid-log phase in a microaerobic liquid culture, observed under a phase contrast light microscope wild-type (WS8N). When compared to wild-type, a large proportion of JPA2001 cells appeared non-motile. Cells that were seen to swim did not appear to behave differently to wild-type, and responded to the phototaxis assay in the wild-type manner. However, the proportion of cells that remained motile was so low that tethered cell analysis of the mutant's response to the removal of propionate was impossible.

Upon reintroduction of *flgM* via conjugation of pInd4:*flgM* into the *flgM* deletion strain JPA2001, and its over-expression by the addition of 50 µM IPTG, motility was restored to this strain in mid-log phase microaerobic liquid culture, and cells were seen to respond normally to the removal of propionate in tethered cell experiments (Figure 127). This is unexpected, as YFP-$CheW_3$ and $CheW_4$-YFP protein expression levels were not restored to wild-type under the same growth conditions (see section (5.7.2)), suggesting that there is a distinction between the quantity of FlgM required to restore wild-type swimming and the quantity of FlgM required to restore FliA-regulated gene expression.

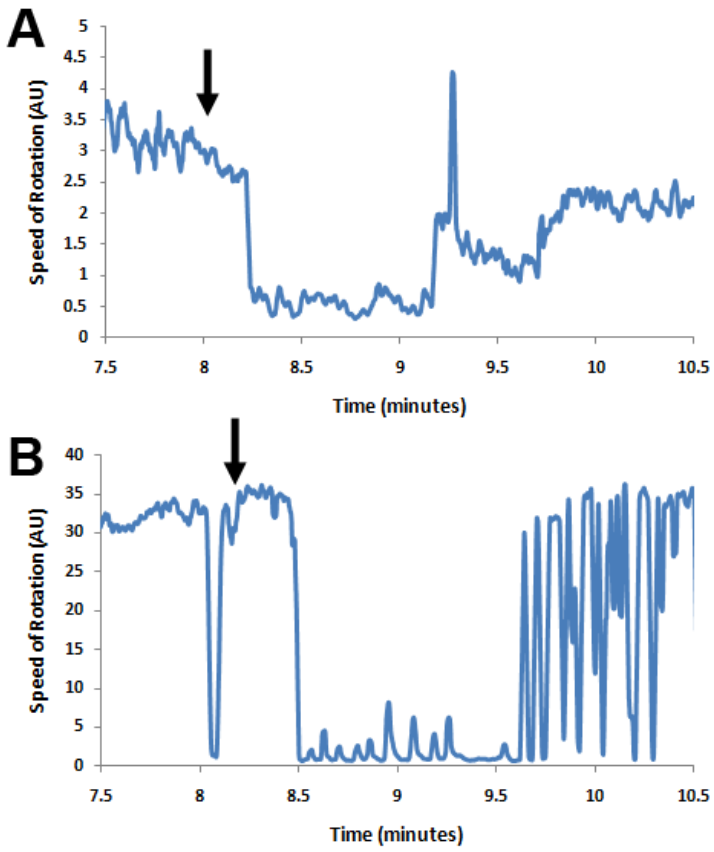


Figure 127: Tethered-cell analysis with FlgM complementation. Representative example single cell speed/time traces for **A)** a wild-type (WS8N,pInd4:empty) cell, and **B)** a *ΔflgM* cell with FlgM activity restored from a plasmid (JPA467,pInd4:*fliA,*50 µM IPTG). Traces are displayed for the section of each video where propionate is removed (marked by an arrow) inducing the chemotactic response. Cellular rotation stops shortly after the removal of propionate and rotation restarts approximately one minute after this stimulus.

7.1.3) FliA and FlgM regulate flagellar number in *R.sphaeroides*

Deletion of *flgM* has been seen to affect flagella number and morphology in other bacterial species, so to test for the presence of flagella in JPA2001, qualitative western blot analysis was performed using the anti-FliC antibody (142) and by comparison to wild-type (WS8N) and *ΔfliA* (JPA467) strains. As expected, no FliC was detected in JPA467, whereas JPA2001 was seen to produce near wild-type levels of FliC, suggesting that filaments can be formed in the absence of *flgM* (Figure 128 and Figure 129), despite an increased non-motile population.

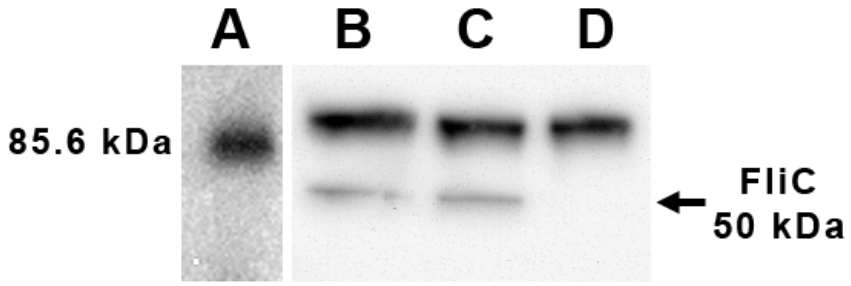


Figure 128: Example Anti-FliC Western Blot – 10 µl of the same quantities of cellular sample resuspended in 200 µl of loading dye was added to each lane. **A)** Ladder, **B)** Wild-type (WS8N), **C)** *ΔflgM* (JPA2001) and **D)** *ΔfliA* (JPA467). Samples were visualised via a light-emitting reaction, the black and white light image of the ladder is overlaid to the left, and the visible molecular weight marker is labelled. Molecular weights decrease from top to bottom on this figure. The anti-FliC antibody has previously been shown to bind to a second, high molecular weight protein (142), which was visible in each lane. The lower band corresponds to FliC, which has a molecular weight of approximately 50 kDa. Pixels values have been inverted where appropriate and contrast has been enhanced for clarity.

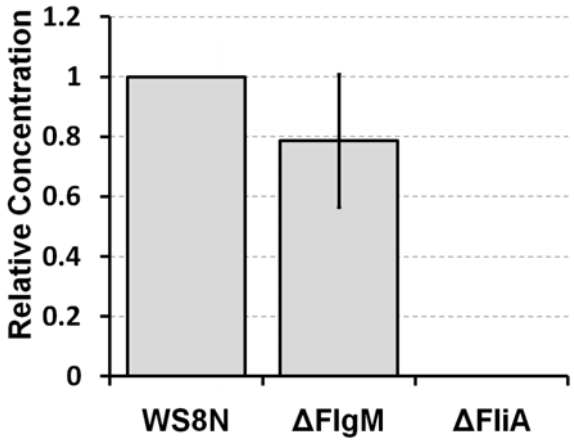


Figure 129: Qualitative measurement of the amount of FliC in JPA2001 and JPA467. Band intensities, such as those in Figure 128, were quantified for three repeats using the Syngene Genetools software package. Expression of FliC was calculated relative to wild-type for each blot. Error represents the standard error of three repeats from independently prepared samples.

In order to confirm that the FliC protein present in the *ΔflgM* strain, JPA2001, was incorporated into functional flagellar filaments and to compare flagella phenotypes to wild-type and *ΔfliA* strains, wild-type (WS8N), *ΔflgM* (JPA2001) and *ΔfliA* (JPA467) cells from mid-log phase cultures were visualised via electron microscopy after being negatively stained using uranyl acetate (Figure 130). At 2,300-fold magnification wild-type cells could clearly be seen to possess long flagellar filaments that protruded from the cell surface. As expected, *fliA* deletion cells had almost no filaments. From a tally of over 100 cells, the number of *flgM* deletion cells which had visible filaments was reduced from wild-type, but remained above that of the *fliA* deletion (Table 34).

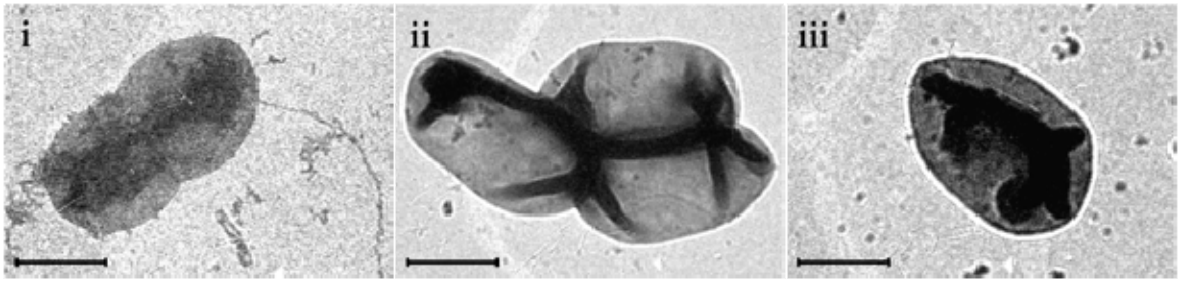


Figure 130: Flagella phenotypes. Electron micrographs of **(i)** Wild-type (WS8N) **(ii)** *ΔflgM* (JPA2001) and **(iii)** *ΔfliA* (JPA467) at 2300-fold magnification. Scale bars represent 1 µm

No. of Visible flagella	**Wild-Type**	***ΔflgM***	***ΔfliA***
0	4 %	76 %	98 %
1	84 %	23 %	2 %
2	12 %	1 %	0 %
Total no of cells	**100 (27)**	**107 (30)**	**102 (20)**

Table 34: Flagella phenotypes. Tally count of flagella visible within populations of bacteria viewed at 2300-fold magnification on electron micrographs. Cells with clearly visible septation sites are shown in brackets.

When cells were observed at 11,500-fold magnification it became clear that filaments were present in JPA2001 cells. In several cases where no flagella filament was seen to protrude from the cell at low magnification, higher magnification revealed that filaments were adhered to the surface of the cell body (Figure 131). However, there was no way to distinguish whether the filaments seen belonged to filament-possessing cells, or were sheared filaments that had attached to the cell bodies (although the same was not observed in the wild-type strain), but this observation suggested that cell surface properties may be altered in the *ΔflgM* mutant.

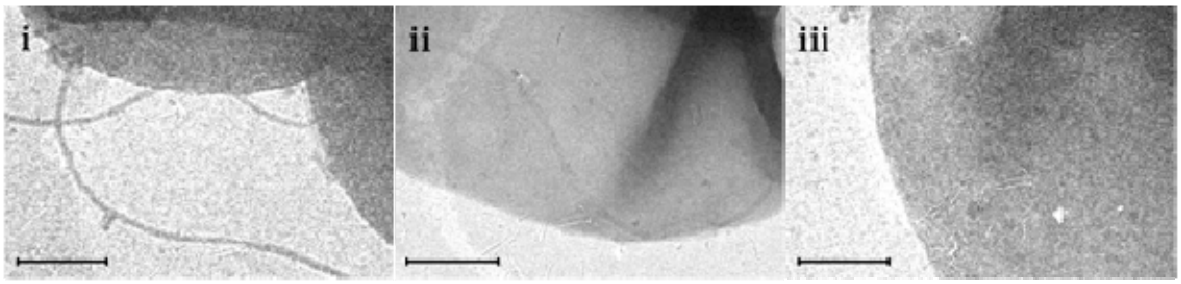


Figure 131: Flagella phenotypes. Electron micrographs of **(i)** Wild-type (WS8N), **(ii)** *ΔflgM* (JPA2001) and **(iii)** *ΔfliA* (JPA467) cells at 11500-fold magnification. Scale bars represent 200 nm.

To further investigate the non-motile phenotype observed at high expression levels of *fliA*, a tally of flagella seen in 100 JPA467+pInd4:*fliA* cells from microaerobic cultures containing 50 µM IPTG was taken using the same electron microscope setup as above (Figure 132 and Table 35).

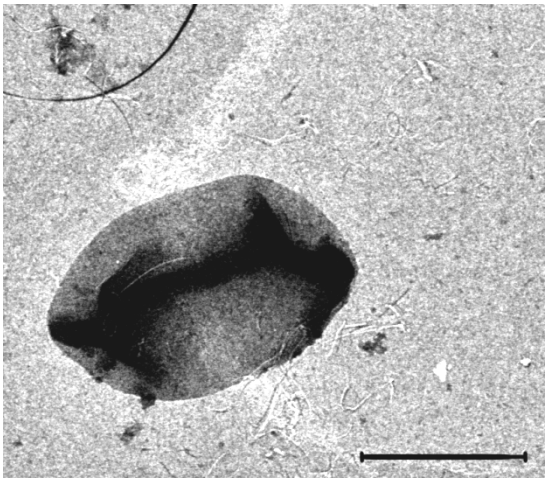


Figure 132: Flagella phenotypes. Example electron micrograph of a *ΔfliA* (JPA467) cell over expressing *fliA* from pInd4:*fliA* at 3000-fold magnification. Scale bar represents 1 µm.

No. of Visible flagella	***ΔfliA* + pInd4:*fliA***
0	91 %
1	8 %
2	1 %
Total no of cells	**101 (40)**

Table 35: Flagella phenotypes. Tally count of flagella visible within populations of bacteria viewed at 3000-fold magnification on electron micrographs. Cells with clearly visible septation sites are shown in brackets.

At high levels of *fliA* activity, few flagella were visible within the population, suggesting that the non-motile behaviour of these cells was due to a lack of flagella filaments.

To further investigate the super-swarming phenotype observed on soft-agar swim plates for cells over-expressing *flgM*, a tally of flagella numbers in 100 cells seen under an electron microscope was taken for JPA2001+pInd4:flgM cells grown microaerobically in the presence of 100 µM IPTG (Figure 133 and Table 36).

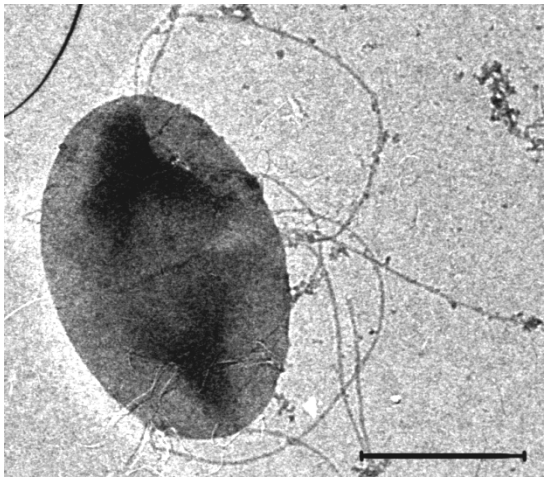


Figure 133: Flagella phenotypes. Example electron micrograph of a *ΔflgM* (JPA2001) cell over expressing *flgM* from pInd4:*flgM* at 3000-fold magnification. Scale bar represents 1µm.

No. of Visible flagella	***ΔflgM* + pInd4:*flgM***
0	15 %
1	62 %
>1	23 %
Total no of cells	**104 (34)**

Table 36: Flagella phenotypes. Tally count of flagella visible within populations of bacteria viewed at 3000-fold magnification on electron micrographs. Cells with clearly visible septation sites are shown in brackets.

At high levels of *flgM* activity flagella were clearly visible throughout the whole population, there was also an increased proportion of cells that had more than one clearly protruding filament, however, it was difficult to distinguish between single filaments that curled round, filaments that overlapped from adjacent cells on the grid and single cells with multiple filaments using this technique.

7.1.4) FliA and FlgM are required for biofilm formation

There is extensive evidence across different species that flagellar expression and biofilm formation are often regulated by the same system, flagella being repressed when biofilm genes are activated (74, 83, 118). The results above suggested that *fliA* and *flgM* activity could be controlling both motility and cell-surface properties, and thus biofilm formation was investigated in the *ΔfliA* and *ΔflgM* mutants. This work was done in collaboration with S.Chacko (Wadhams Laboratory, Oxford) who is currently investigating *R.sphaeroides* biofilm formation.

To do this, either CFP or YFP was expressed in the different strains and their ability to form biofilms was measured. pInd4:*ecfp* and pInd4:*eyfp* (gifts from Dr.M.Roberts) were conjugated into wild-type (WS8N), *ΔflgM* (JPA2001) and *ΔfliA* (JPA467) strains. Biofilms were grown on glass microscope-slides in a static-culture of succinate medium for 48 – 72 hours from mixed populations of mutant and wild-type cells expressing either CFP or YFP protein from pInd4. Biofilms were visualised on a Nikon A1R confocal microscope. The proportion of YFP expressing and CFP expressing cells in each mixed population gave a measure of the competency of that strain to either form biofilms, or to grow within biofilms. In wild-type/wild-type mixed cultures approximately 50% of cells were CFP-expressing, and the remaining 50% were YFP-expressing. However, both Δ*fliA* (Figure 134A) and Δ*flgM* (Figure 134B) mutants showed a drastically reduced ability to form biofilms or to grow within wild-type biofilms and represented only a small proportion of cells in wild-type/mutant mixed populations. Interestingly, the Δ*flgM* mutant could form micro-colonies individually, whereas the Δ*fliA* mutant could not, suggesting that this is a motility-linked phenotype.

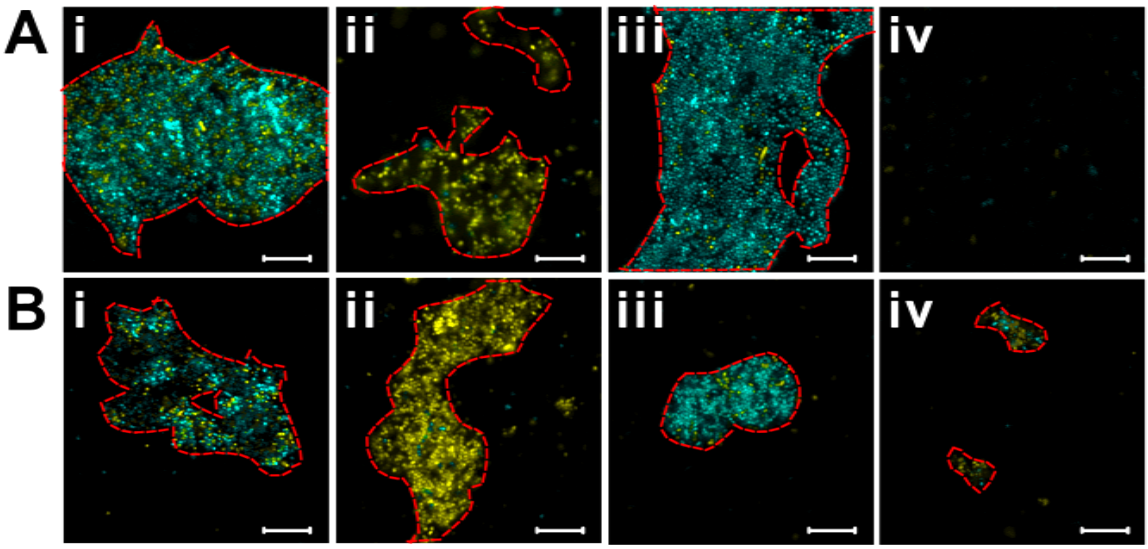


Figure 134: Two-colour confocal microscopy data for *ΔflgM* and *ΔfliA* mutants. **A)** Biofilms of mixed cultures; **(i)** WT:yfp/WT:cfp, **(ii)** WT:yfp/*ΔfliA*:cfp, **(iii)** *ΔfliA*:yfp/WT:cfp and **(iv)** *ΔfliA:yfp/ΔfliA:cfp*. **B)** Biofilms of mixed cultures; **(i)** WT:yfp/WT:cfp, **(ii)** WT:yfp/*ΔflgM*:cfp, **(iii)** *ΔflgM*:yfp/WT:cfp and **(iv)** *ΔflgM:yfp/ΔflgM:cfp.* Microcolony/biofilm outlines obtained from a transmitted light channel are marked in red. Scale bars represent 10 µm. Images were acquired by S.Chacko. The contrast of individual images has been adjusted for clarity. NB/ Work by S.Chacko is currently ongoing to obtain quantitative measurements of the data from these images.

7.1.5) Discussion:

Here we have seen that both FliA and FlgM are essential to wild-type motility in *R.sphaeroides* and that motility is a requirement for the formation of biofilms.

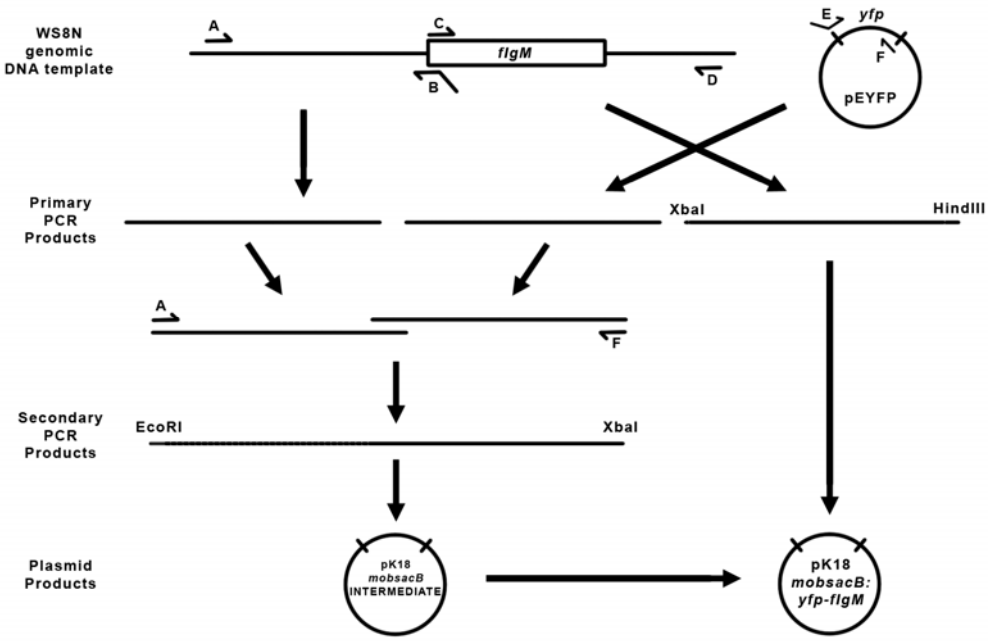
Intriguingly, FliA and FlgM activity affects motility and chemotaxis in *R.sphaeroides* by regulating flagellar number. Unsurprisingly, deletion of *fliA* results in the inability to express class IV FliC and no filaments are made in these strains. However, the upregulation of FliA activity either by deletion of *flgM* or by the over-expression of FliA from a plasmid also results in reduced motility and a reduction in flagellar number when observed by electron microscopy. This is evidence that FliA, and the FlgM secretion checkpoint, is involved in the regulation of the flagellar assembly hierarchy. A reduction in flagellar number is evidence for repression of the synthesis of flagellar components, as FliA promotes class IV flagellar gene expression, this must be achieved via negative feedback repression of class I, II or III genes. To investigate this further, the FliA/FlgM mediated regulation of known class III genes *flgM*, *motB* and *fliM* was investigated (see below).

7.2) FlgM

7.2.1) Making the YFP-tagged FlgM reporters

To detect FlgM protein, and as attempts to generate antibody were unsuccessful (see section (5.7.3)), chromosomal *flgM-yfp* fusion mutants were generated. Primers were designed to incorporate the necessary restriction sites for independent fusion of YFP to both the N and C termini of FlgM.

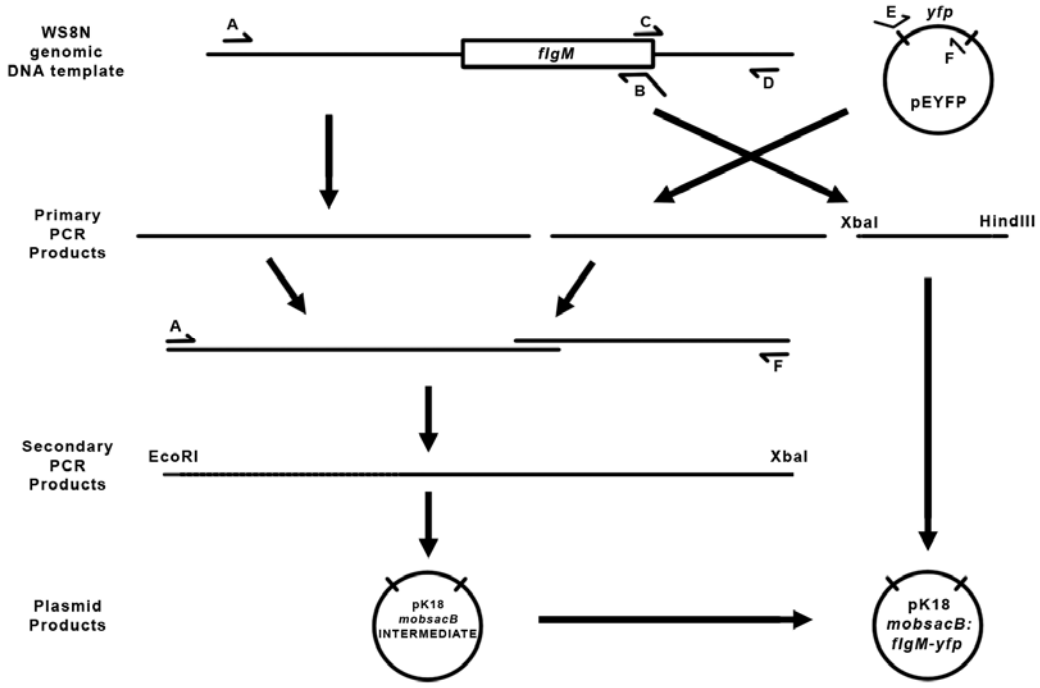
For the N-terminal fusion, the 510 bp upstream of *flgM* including the *flgM* start codon were joined to *yfp* lacking both the start and stop codons via overlap extension PCR. This sequence was then digested with *Eco*RI and *Xba*I and inserted into pK18*mobsacB*. The remaining 312 bp of *flgM* were amplified along with 275 bp of downstream residues, and this sequence was digested with *Xba*I and *Hind*III and inserted into the pK18*mobsacB* that contained the first fragment (Figure 135).



Primer	Reference Code
A	DW01
B	DW15
C	DW18
D	DW04
E	DW16
F	DW17

Figure 135 and Table 37: Generating the N-terminal *yfp-flgM* fusion - Schematic of the PCR reactions and digest/ligations involved in creating the pK18*mobsacB*:*yfp-flgM* plasmid. The table lists the relevant primers for these reactions.

For the C-terminal fusion, the *flgM* sequence lacking the stop codon along with 510 bp of upstream residues was joined to the *yfp* sequence, lacking both start and stop codons, via overlap extension PCR. This sequence was digested using *Eco*RI and *Xba*I, and inserted into pK18*mobsacB*. The *flgM* stop codon, along with 275 bp of downstream residues was amplified and digested with *Xba*I and *Hin*dIII and inserted into the pK18*mobsacB* that contained the first fragment (Figure 136).



Primer	Reference Code
A	DW01
B	DW19
C	DW20
D	DW04
E	DW21
F	DW17

Figure 136 and Table 38: Generating the C-terminal *flgM-yfp* fusion - Schematic of the PCR reactions and digest/ligations involved in creating the pK18*mobsacB*: *flgM-yfp* plasmid. The table lists the relevant primers for these reactions.

Successfully generated pK18*mobsacB*:*yfp-flgM* and pK18*mobsacB*:*flgM-yfp* plasmids were sequenced and their sequences were checked using the Clone Manager 9 software package.

yfp-flgM and *flgM-yfp* were incorporated into the *R.sphaeroides* chromosome via allelic replacement, and the successful generation of *yfp-flgM* (JPA2034) and *flgM-yfp* (JPA2035) mutants were confirmed via southern blot and DNA sequencing.

7.2.2) Phenotypic analysis of YFP-tagged FlgM mutants:

As motility and chemotaxis had been observed to be affected in the *ΔflgM* mutant (JPA2001, see section (7.1.2)), the functionality of *yfp-flgM* and *flgM-yfp* mutants was assessed on soft agar swim plates (Figure 137).

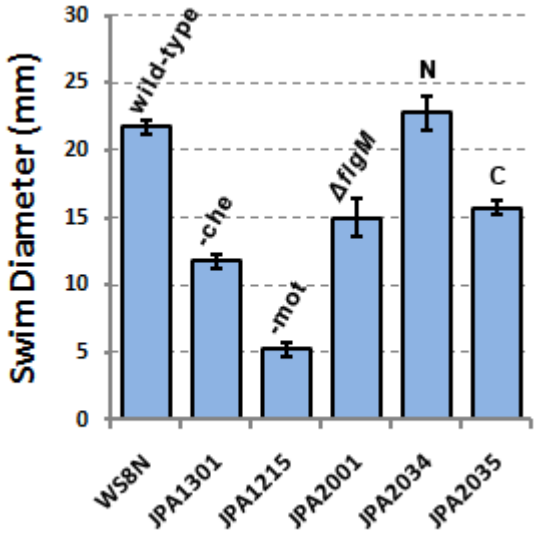


Figure 137: YFP-tagged FlgM mutant swimming phenotypes: Swim diameters were measured on soft-agar swim plates for *yfp-flgM* (JPA2034) and *flgM-yfp* (JPA2035) strains and compared to wild-type (WS8N), non-chemotactic (JPA1301), non-motile (JPA1215) and *ΔflgM* (JPA2001) controls. Error bars represent the standard deviation of three repeats from three separate days.

The C-terminally tagged FlgM-YFP mutant showed reduced chemotaxis which was comparable to that of the *ΔflgM* mutant suggesting that this protein-fusion was non-functional and that the C-terminal tag may interfere with FliA-binding. The N-terminally tagged YFP-FlgM mutant retained wild-type levels of chemotaxis, suggesting that this protein fusion was functional. In addition to swim-plate analysis, the free-swimming behaviour of the each mutant was observed using phase-contrast microscopy, where again the N-terminal mutant behaved as wild-type and the C-terminal mutant mimicked the deletion. Growth rates in these mutants were unaffected.

7.2.3) FliA regulates FlgM expression

The protein copy numbers of YFP-FlgM and FlgM-YFP were calculated using the fluorimetry. N-terminally tagged YFP-FlgM had a copy number of 767 ± 20 copies per cell. C-terminally tagged FlgM-YFP had a much higher copy number of 7580 ± 375 copies per cell. The difference in expression level may be explained by the activity of FliA, or by the relative stabilities of the YFP-fusion proteins in the different strains (see below).

Previous work in other bacterial species has shown that FlgM inhibits FliA activity, but also that FliA activity up-regulates expression of FlgM (47, 63). *flgM* is part of a divergent operon which is regulated by FleQ/FleT mediated activity but is suggested to contain a separate FliA promoter sequence, although this sequence has not been identified. To test whether *R.sphaeroides flgM* expression was affected by the activity of FliA, *fliA* was deleted from the FlgM fluorescent fusion mutants to generate strains JPA2036 and JPA2037. Successful deletion of *fliA* in these strains was verified via southern blot and DNA sequencing.

FlgM-YFP and YFP-FlgM copy numbers were calculated via fluorimetry. Both FlgM fluorescent fusion proteins showed decreases in their level of expression in the absence of FliA, but expression levels remained above zero, showing that *R.sphaeroides flgM* is regulated by FliA (Table 39). In the absence of FliA, expression levels of YFP-FlgM and FlgM-YFP would be predicted to be the same, as the loss of function in FlgM-YFP should have no effect (there is no longer a σ-factor for the anti σ-factor). The observed difference in copy numbers in these strains is an indication that there is a change in either protein-stability or gene expression which differs between the two fusion strains.

	YFP Protein Copy Numbers		
Protein	**Wild-Type Background**	***ΔfliA* Background**	**Fold Change from Wild-Type**
YFP-FlgM	770 ± 20	620 ± 40	0.80
FlgM-YFP	7580 ± 375	1340 ± 135	0.18

Table 39: FlgM Fluorescent Fusion Copy Numbers: Copy Numbers calculated via fluorimetry for *yfp-flgM* (JPA2034), *flgM-yfp* (JPA2035), *yfp-flgM,ΔfliA* (JPA2036) and *flgM-yfp,ΔfliA* (JPA2037). Error represents the standard error of three independent repeats.

In order to determine how varying FliA activity affected expression levels of the two FlgM fluorescent-fusions, pInd:*fliA* was conjugated into YFP-FlgM (JPA2034) and FlgM-YFP (JPA2035) strains, and the level of expression of *fliA* was varied by the addition of different concentrations of

IPTG (Figure 138), concentrations were chosen which were analogous to those tested in section (5.7.3). The expression level of both proteins was seen to increase with *fliA* expression, providing further evidence for FliA dependent regulation of *flgM*. Expression of the non-functional C-terminal YFP fusion increased by a greater amount than that of the functional N-terminal YFP fusion at higher concentrations of IPTG, providing further evidence for the inhibition of FliA by FlgM.

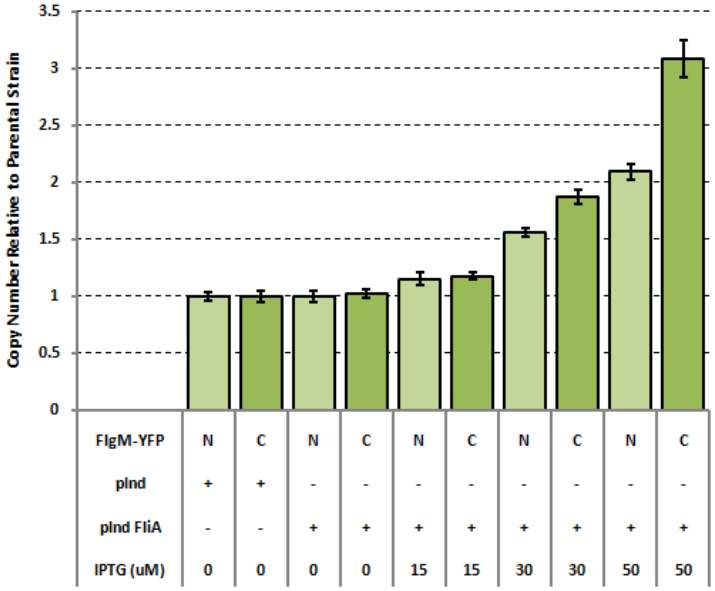


Figure 138: FliA activity in YFP FlgM-fusion backgrounds**:** Copy numbers of YFP-tagged protein in FlgM-YFP (JPA2035, indicated by 'C') and YFP-FlgM (JPA2034, indicated by 'N') strains at different induction levels of *fliA*, plotted relative to copy number of the corresponding parental strain. Error bars represent the standard error of three independent repeats.

7.2.4) FlgM is secreted

Most species of bacteria achieve FlgM-mediated gene regulation through the secretion of FlgM on completion of the HBB complex, however *Helicobacter pylori* has been shown not to secrete, but to sequester FlgM to the C-ring of the flagellar motor as a means of regulation.

To test whether FlgM is secreted from *R.sphaeroides* cells, intracellular and extracellular fractions were analysed by western blot. As no anti-FlgM antibody was available (see section (5.7.3)), this analysis was carried out using the YFP FlgM fusion strains JPA2034 and JPA2035, and monoclonal mouse anti-GFP antibody (Invitrogen) was used to visualise protein bands. The YFP-labelled small molecule, YFP-$CheW_3$, which is not secreted from the cell, was used as a control. Cells were grown in 200 ml of succinate medium, and harvested at mid-log phase. The whole-cell fraction was harvested by centrifugation, and the remaining liquid (representing the extracellular fraction)

was filtered through a 0.2 µm filter to remove any remaining cells. As the concentration of protein in the extracellular fraction of growth media was very low in comparison to the whole-cell fraction, the liquid was protein-concentrated ~1000-fold using a Biolabs VIVASPIN 20 with a molecular weight cut-off of 10 kDa.

Western Blot analysis of these preparations (Figure 139) showed clearly detectable bands in both whole-cell and extracellular fractions for both fusion strains, whereas $CheW_3$-YFP was only detectable in the whole-cell fraction, suggesting that cell lysis was minimal and supporting the theory of active secretion of FlgM by *R.sphaeroides*.

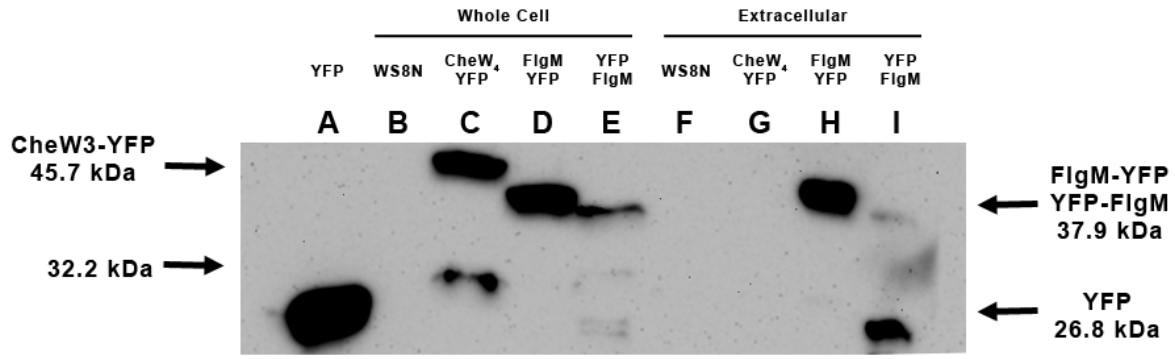


Figure 139: Anti-GFP western Blot of intracellular and extracellular environments – Monoclonal mouse anti-GFP antibody was used to visualise YFP protein in protein and cellular samples of **A)** ~25 ng purified YFP, **B)** 1x WS8N whole-cell fraction, **C)** 1x JPA1418 (YFP-$CheW_3$) whole-cell fraction, **D)** 1x JPA2035 (FlgM-YFP) whole-cell fraction, **E)** 3x JPA2034 (YFP-FlgM) whole-cell fraction, **F)** 1000x WS8N extracellular fraction, **G)** 1000x JPA1418 extracellular fraction, **H)** 1000x JPA2035 extracellular fraction, **I)** 1000x JPA2034 extracellular fraction. The ladder is not shown. Molecular weights decrease from top to bottom on this figure. The position of one molecular weight marker is shown to the left of the figure, and estimated molecular weights of YFP-$CheW_3$, FlgM-YFP, YFP-FlgM and purified YFP protein are also shown. Pixel Intensities have been inverted and contrast adjusted for clarity.

Low molecular weight protein cleavage products of approximately the same size as purified YFP were also detectable for YFP-FlgM and YFP-$CheW_3$ fractions; the ratio of YFP-FlgM cleavage product to full length protein was higher in the extracellular fraction, suggesting that the cleavage event is likely to be occurring extracellularly or after cell lysis. Although only detected at low levels, it is possible that the cleavage product of YFP-FlgM is present within cells and that the functionality of this mutant could be conferred by a YFP-cleaved FlgM protein *in vivo*.

7.2.5) Observation of FlgM via epifluorescence

Both fluorescent FlgM fusion strains were immobilised on agarose pads and visualised via epifluorescence microscopy (Figure 140). Both YFP-FlgM (JPA2034) and FlgM-YFP (JPA2035)

appeared diffuse within cells, and displayed no form of localisation. As expected, JPA2035 appeared much brighter at a given exposure setting than JPA2034 due to its higher copy number of protein.

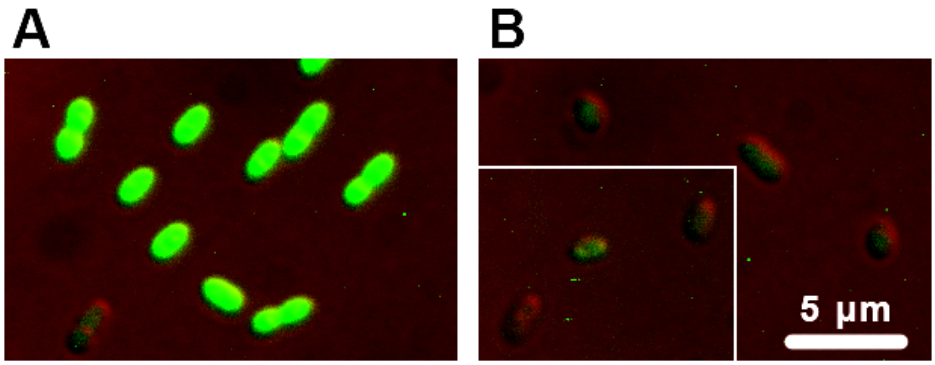


Figure 140: Example images of the diffuse nature of **A)** FlgM-YFP (JPA2035) and **B)** YFP-FlgM (JPA2034). Images were taken with the same YFP-exposure time of 2 seconds. Contrast has been adjusted equally in each image.

7.2.6) Discussion

Development of two YFP-tagged variants of FlgM allowed for the observation of FlgM protein expression, secretion and localisation. We have seen that *R.sphaeroides* FlgM is a diffuse protein that is secreted from the cell and that FlgM inhibits FliA activity, indicating that the role of FlgM in *R.sphaeroides* is similar to that in other bacterial species – an anti-σ factor that is involved in the regulation of FliA activity in response to flagellar synthesis.

The *flgA* operon that encodes *flgM* has been shown to possess a class III FleQ/FleT binding sequence (116), but the *flgM* gene is thought to possess its own class IV promoter sequence (122) (although FliA consensus sequence is not known). Similar to the situation in other bacteria (94), this means that a basal level of FlgM is present in cells and expressed with FliA from its class III promoter. Secretion of FlgM upon flagellar synthesis increases FliA-regulated expression of FlgM, and cellular FlgM is replaced – FliA essentially inhibits its own activity. The data obtained here supports a similar model of *R.sphaeroides* FliA/FlgM regulation, as low cellular levels of FlgM were detected in functional YFP-fusion strains and increased FliA activity in the non-functional FlgM-YFP fusion and in FliA induced-expression experiments led to large increases in the protein copy number of FlgM. The magnitude of the observed changes in FlgM copy number suggests that FlgM is predominantly class IV regulated, despite being part of a class III operon.

7.3) MotB

7.3.1) Introduction

MotB is an essential element of the stator complex of the flagellar motor. The production of torque which drives flagellar rotation is achieved by proton flow through the $MotA_4B_2$ complex, which transfers this energy to the flagella through interaction with the FliG protein (14). MotB and MotA function together as the stator and their genes are commonly encoded as part of the same operon in many different bacterial species (11).

There are two *motB*-like genes in the genome of *R.sphaeroides*. Alignment of the MotB protein products of *R.sphaeroides* with those from *E.coli* and *Salmonella* using ClustalW and UniProt (Figure 141) shows that MotB2 has a more classical sequence than that of MotB1, which has an additional 33 amino acid sequence in its periplasmic domain. However, the *motB1* gene is found in the *motAB* operon, and, importantly, has been shown to have its expression regulated by FleQ/FleT dependant σ_{54} activity (116). This first *motB* gene is therefore a flagellar system 1 linked class III gene. The second *motB* gene is orphaned, that is to say – it is not found on the genome with its *motA* partner (although there is a second *motA* gene, which is also orphaned) and there is nothing known about its expression or function.

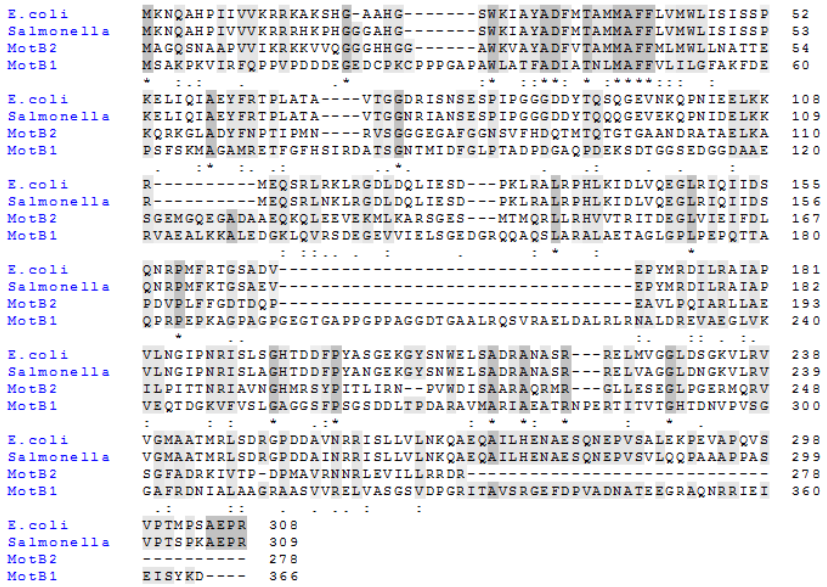


Figure 141: MotB protein sequence alignment – UniProt sequence alignment of the MotB proteins from *E.coli*, *Salmonella*, *R.sphaeroides* MotB1 and *R.sphaeroides* MotB2.

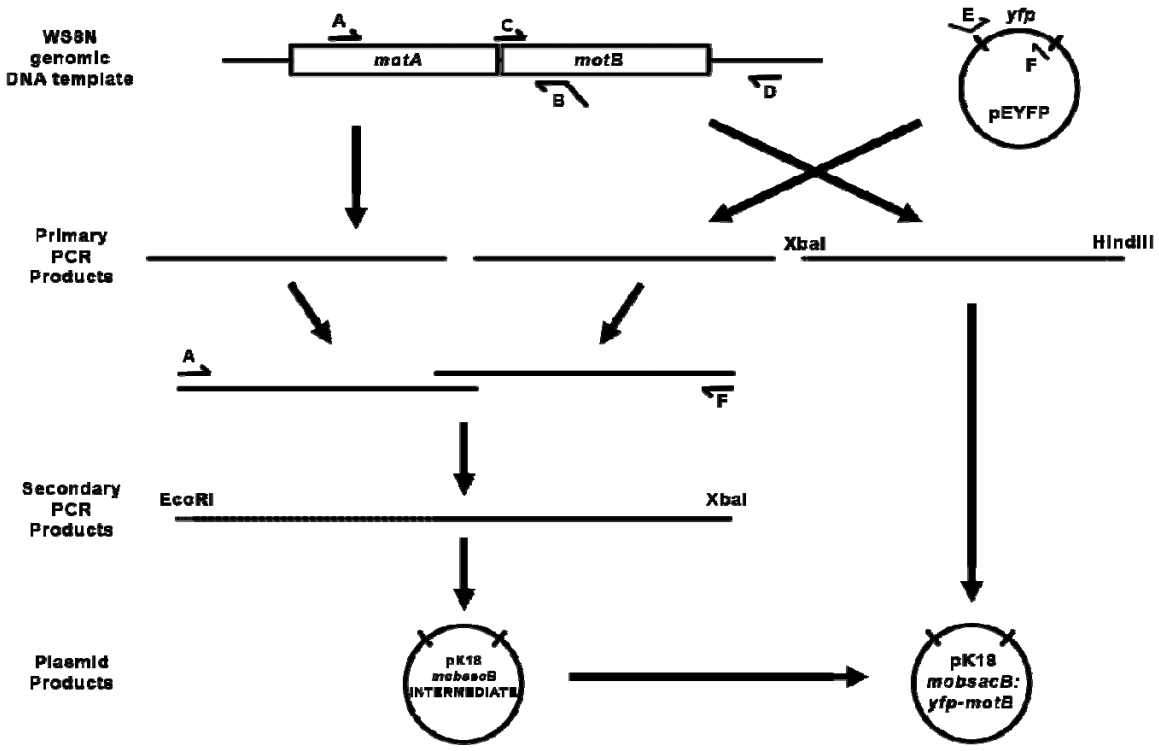
Based on the understanding of the expression of the *motAB* operon and the fact that *R.sphaeroides* has two flagellar systems of which only one is expressed under laboratory conditions,

it is assumed that the *motB2* gene is associated with the function of flagellar system 2 and not of flagellar system 1. Therefore, this study focused on the *motB1* gene and not the second.

7.3.2) Making the YFP-tagged MotB reporter

In order to characterise the role of FliA and FlgM in regulating the expression of MotB in *R.sphaeroides*, a chromosomal *yfp-motB* fusion was made.

The design of the *yfp-motB* fusion construct was based around that used for *E.coli motB* by Chandler *et al.* (81). The MotB protein is found as part of the $MotA_4B_2$ complex in the inner membrane of bacteria. Many proteins that are found within the inner membrane have their positions determined by an N-terminal amino acid sequence which targets that protein for membrane insertion, which for motor components is achieved exclusively via the Sec pathway (42). In order to maintain the correct cellular positioning of the N-terminally fused YFP-MotB protein, the first 84 base pairs (28 codons) of the MotB sequence are repeated before insertion of the *yfp* sequence. The rest of the *motB* gene, including a repeat of the first 27 codons (but not the START codon), is joined to the *yfp* sequence (Figure 142).



Primer	Reference Code
A	DW22
B	DW23
C	DW17
D	DW26
E	DW24
F	DW25

Figure 142 and Table 40: Generating the *yfp-motB* fusion - Schematic of the PCR reactions and digest/ligations involved in creating the pK18*mobsacB*:*yfp-motB* plasmid. The table lists the relevant primers for these reactions.

Primers were designed that incorporated sufficient lengths of DNA both upstream and downstream of the fusion sequence to allow chromosomal replacement via sequential rounds of recombination. The fusion of the *yfp* sequence to the *motB1* sequence was achieved by overlap extension PCR of products. The resulting product was digested with restriction enzymes *Eco*RI and *Xba*I, and ligated into pK18*mobsacB*, generating an intermediate plasmid construct. The sequence downstream of the *yfp* fusion and the intermediate plasmid construct were digested with *Xba*I and *Hin*dIII, and joined via ligation. The pk18*mobsacB*:*yfp-motB1* construct was sequenced and its sequence was checked for point mutations using the Clone Manager 9 software package.

Successfully generated pK18*mobsacB*:*yfp-motB* was transformed into the *E.coli* S-17 strain, and conjugated into WS8N *R.sphaeroides* cells. Allelic exchange was performed to obtain a chromosomal *yfp-motB* fusion. The successful incorporation of the *yfp* fusion was checked by southern blot and DNA sequencing.

7.3.3) Phenotypic Analysis of the YFP-MotB1 Strain:

In order to assess whether the YFP-MotB1 (JPA2006) fusion protein was functional, soft-agar swim plates were used to assess chemotactic efficiency towards propionate, and swim diameters were compared to wild-type (WS8N), non-chemotactic (JPA1301) and non-motile (JPA452) control strains (Figure 143). The YFP-MotB1 fusion demonstrated a reduced swimming efficiency under these conditions suggesting that the mutant protein retains partial activity. Growth rates were as wild-type in this strain.

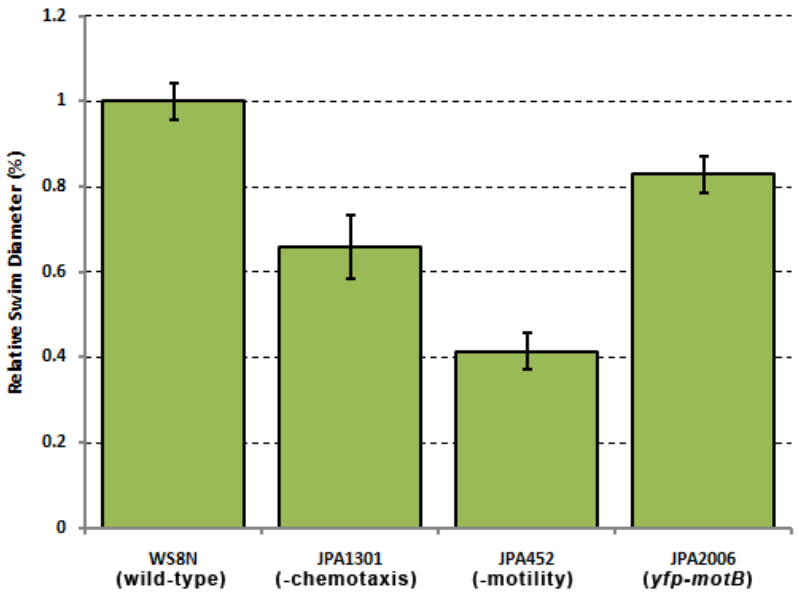


Figure 143: YFP-MotB1 Swimming Efficiency: Chemotactic efficiency measured on soft-agar swim plates. Measurements represent the diameters of colonies after 48 hours of growth, measured relative to the average wild-type value. Error bars represent the standard deviation of three repeats taken on three separate days.

7.3.4) YFP-MotB1 cannot be seen in the motor

To verify whether or not the YFP-MotB1 localised to the flagellar motor, motile cells from mid-log phase microaerobic cultures were tethered to tunnel slides using a 1/1000 dilution of the anti-FliC polyclonal antibody and observed under Total Internal Reflection Fluorescence (TIRF) illumination on a custom-built in-house microscope. Tethered cells were observed to rotate about a centre, confirming that the flagellar motor is functional under these conditions (Figure 144A). However, in >30 cells observed on two separate days, no distinct fluorescence signal above that of the background cellular fluorescence was observed at the centre of rotation for these cells (Figure 144B), suggesting that the YFP-MotB1 protein is not present at the motor or that, if YFP-MotB1 does localise to the motor, the number of proteins in that region is very small.

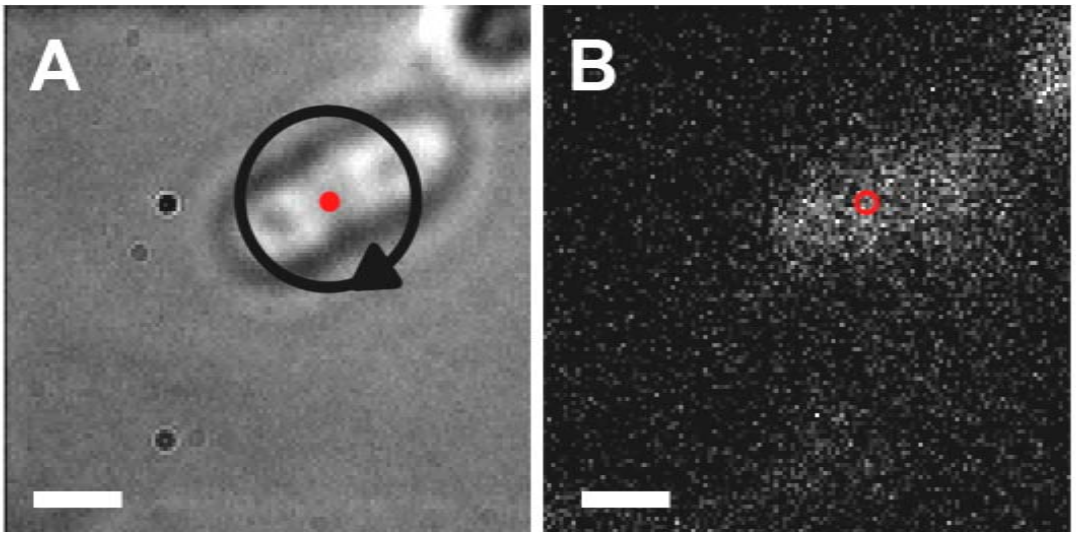


Figure 144: Example TIRF Images of JPA2006: A) Bright-field image of a single frame from a video of a tethered cell rotating about its centre. The centre of rotation is marked by a red spot, and the direction of rotation is indicated by an arrow. B) TIRF-illuminated fluorescence image of a single frame from a video of the same cell in (A), the centre of rotation estimated from (A) is marked by a red ring. Contrast has been enhanced for clarity. Bars represent 1 μm.

7.3.5) FliA/FlgM regulation of MotB1 expression

In other bacteria, FliA activity up-regulates the expression of Class I and Class II genes via direct and indirect positive feedback loops (27, 134). To investigate whether or not the expression of FleQ/FleT dependent σ_{54}-regulated genes was affected by the FliA/FlgM activity in *R.sphaeroides*, YFP-MotB1 expression was studied in the different deletion background strains. First, in order to

generate these mutant strains the pK18:*ΔflgM* and pK18:*ΔfliA* were conjugated into strain JPA2006, and the genes were deleted via successive rounds of recombination. Successful deletions of *flgM* (JPA2038) and *fliA* (JPA2039) were confirmed by southern blot analysis and DNA sequencing.

After successful isolation of the deletion mutants, YFP-MotB 1copy numbers were determined in each strain via fluorimetry. The overall copy number of YFP-MotB1 in an otherwise wild-type background was approximately 600 copies per cell, this value is low in comparison to the copy numbers of the chemotaxis proteins (section 5.7), but it is 3-fold above that of the larger, peritrichous bacterium *E.coli* (81).YFP-MotB1 expression was slightly decreased in the absence of FlgM, and a 2.8-fold increase in YFP-MotB1 expression was observed in the absence of FliA (Table 41). From these data, FliA is seen to be a strong inhibitor of *motAB* (class III) gene expression.

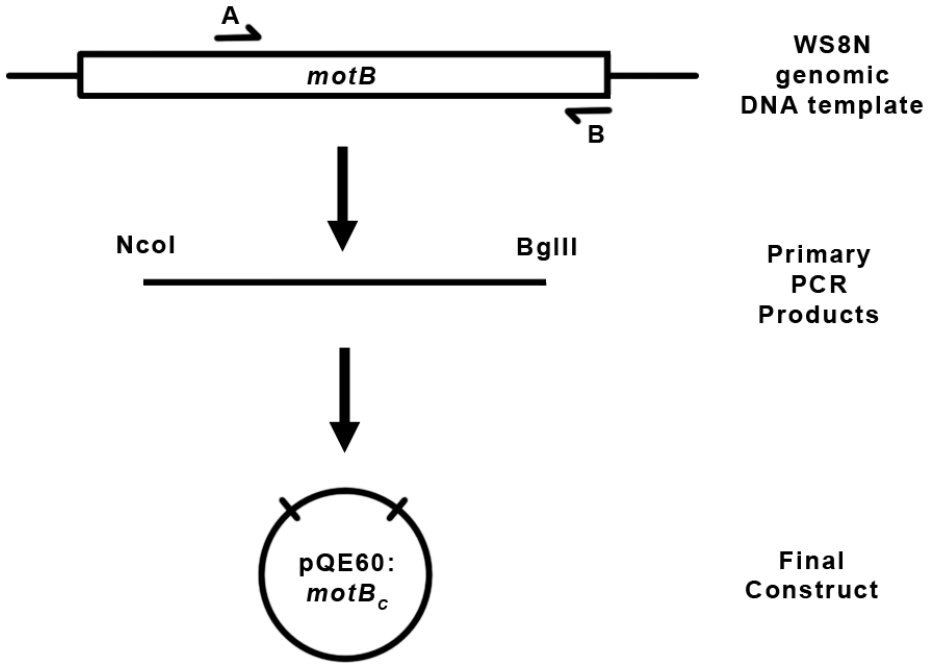
Strain	Genotype Description	YFP Protein Copy Number
JPA2006	*yfp-motB1*	600 ± 20
JPA2038	*yfp-motB1, ΔflgM*	500 ± 25
JPA2039	*yfp-motB1, ΔfliA*	1700 ± 30

Table 41: The Effect of Deleting *fliA* and *flgM* on YFP-MotB1 copy number. Copy numbers of YFP-MotB1 as calculated by the fluorimetry. The error shown is the standard error from three independently prepared samples.

7.3.6) Quantitative Western Blot Analysis MotB expression levels:

Expression studies of MotAB in *E.coli* have revealed that the number of diffuse proteins available in the inner membrane (or "membrane pool") of the cell has an effect on the speed of rotation of the flagellar motor (130). It is therefore important to have an accurate measure of the copy number of MotB in order to estimate the amount of protein available in the membrane pool. Although the copy number estimate for YFP-MotB1 has been shown to be an accurate measure of the number of mutant proteins in the cell, YFP fusion mutants have also been shown to have altered expression (see section (5.6)), and therefore a method of detection for wild-type protein was desirable. Antibody was therefore raised against purified MotB protein. Protein purification of membrane proteins is notoriously difficult; therefore in order to purify MotB1, primers were designed to amplify the cytoplasmic region of the MotB ($MotB_C$) only (Figure 145). Incorporation of the *Nco*I and *Bgl*II restriction sites allowed the amplified gene fragment to be incorporated into the pQE60 his-tag purification vector. Upon successful isolation of the pQE60:$motB_C$ plasmid, it was sequenced at

Geneservice, and the obtained sequence was checked for point mutations in the Clone Manager 9 software package.



Primer	Reference Code
A	DW27
B	DW28

Figure 145 and Table 42: Purifying $MotB_C$ - Schematic of the PCR reactions and Digest/Ligations involved in creating the pQE60:*$motB_C$* plasmid. The table lists the relevant primers for these reactions.

Purification of $MotB_C$ was achieved via the standard protein purification protocol. The obtained protein concentration was determined to be 9.1 mg ml^{-1} via the Bradford Assay, and protein purity was assessed on an SDS-PAGE gel via staining with coomassie blue (Figure 146). The protein was judged to be sufficiently pure for antibody preparation, and was therefore diluted and sent the Eurogentec for rabbit injection as per the company's instructions.

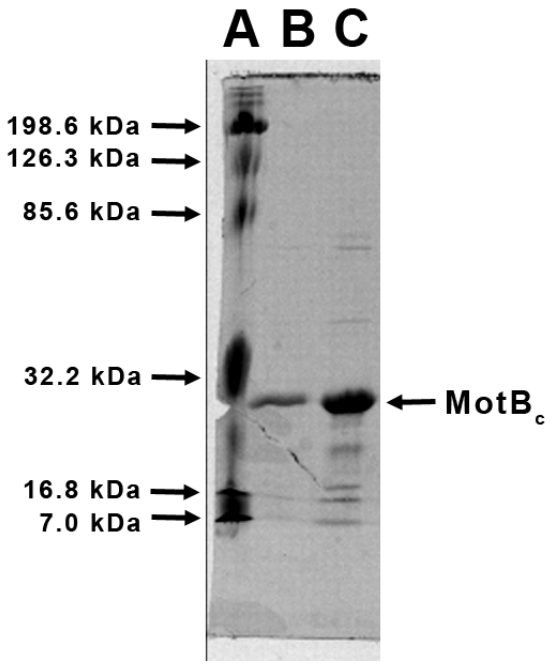


Figure 146: Purified MotBc protein. Coomassie-stained 12 % SDS PAGE gel (image taken with black and white camera). Lanes: **A)** Ladder, **B)** 1 µg purified protein, **C)** 5 µg purified protein. Ladder molecular weight markers and MotBc (predicted to be a 30 kDa protein) are marked with arrows. Contrast has been adjusted for clarity.

Quantitative western blot analysis of MotB protein levels was performed for wild-type (WS8N), *ΔflgM* (JPA2001) and *ΔfliA* (JPA467) cells, protein concentrations were calculated in the Syngene GeneTools software package by comparing band intensities to those obtained for known amounts of $MotB_C$ protein. In this instance, use of the YFP-MotB strain (JPA2006) as a negative control on the gel allowed for successful identification of the band that corresponded to MotB (Figure 147).

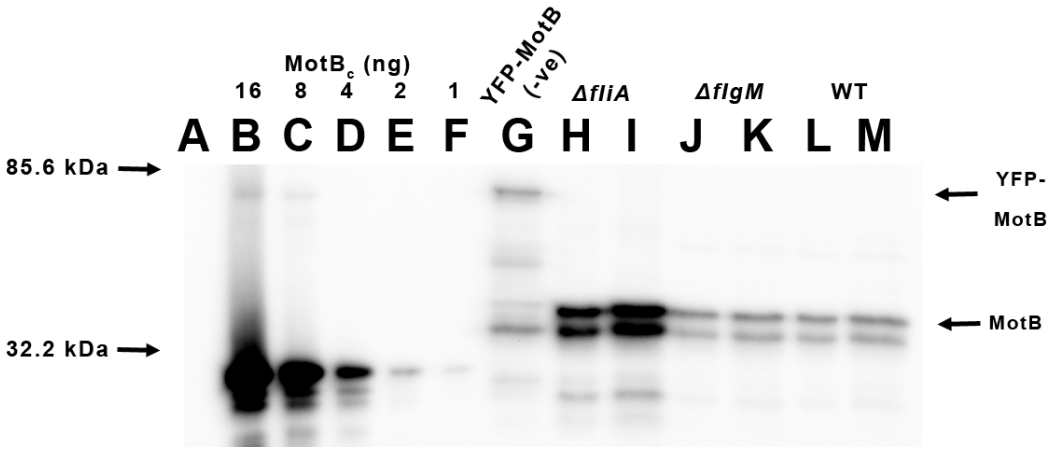


Figure 147: Example Quantitative Western Blot for MotB**:** Multiples of 10 µl of cell or protein sample that had been resuspended and boiled in 200 µl of protein loading dye were added as follows: **A)** Ladder, **B)** 16 ng purified $MotB_C$, **C)** 8 ng purified $MotB_C$, **D)** 4 ng purified $MotB_C$, **E)** 2 ng purified $MotB_C$,, **F)** 1 ng purified $MotB_C$, **G)** 1 x JPA2006 (negative control), **H)** 1 x JPA467, **I)** 1 x JPA467, **J)** 1 x JPA2001, **K)** 1 x JPA2001, **L)** 1 x WS8N, **M)** 1 x WS8N. Samples were visualised via a light-emitting reaction, therefore the ladder was not visible. Two weight-marker positions and the predicted positions of YFP-MotB and MotB proteins are marked with arrows. Pixel values have been inverted and contrast has been adjusted for clarity.

The western blot results showed that levels of wild-type, unfused MotB protein are higher than those of YFP-MotB protein (Table 43), but that the trends in changes of expression between wild-type (WS8N), *ΔflgM* (JPA2001) and *ΔfliA* (JPA467) mutants remained similar, however the magnitude increase in expression for the *ΔfliA* mutant was 2-fold, lower than that observed in the YFP-mutant. No statistical difference could be seen between wild-type and the Δ*flgM* strain. The copy number of MotB in wild-type cells was calculated to be approximately 1700, which is nearly 9 times that of similar data for *E.coli*. Anti-MotBc antibody was seen to bind to a second protein of a similar size to MotB within all cellular samples, however the YFP-MotB negative control strain allowed for identification of the relevant band corresponding to MotB.

Strain	Genotype Description	YFP Protein Copy Number
WS8N	*Wild-type*	1800 ± 100
JPA2001	*ΔflgM*	1800 ± 100
JPA467	*ΔfliA*	3500 ± 300

Table 43: The Effect of Deleting *fliA* and *flgM* on MotB copy number. Copy numbers of MotB as calculated by quantitative western blot. The error shown is the standard error from three independently prepared samples, measured against independently prepared standard dilutions of purified $MotB_C$.

7.3.7) Discussion

Here, antibody and fluorescent-fusion reporters were generated for the stator protein, MotB. Analysis of copy numbers in the presence and absence of FliA and FlgM showed that MotB, which has previously been identified as a class III flagellar gene (116), has its expression repressed by the activity of FliA. This repression may be linked to the reduction in flagellar number observed with increased FliA activity (see section (7.1.3)), however, the stator itself is not an essential element of the HBB complex and a reduction in stator number would not be predicted to result in reduced flagellar formation. To test whether FliA-mediated repression was common to different class III genes, another class III marker, FliM, was studied (see section 7.4).

Intriguingly, but not central to this study, YFP-MotB fluorescent fusion resulted in reduced motility but YFP-MotB proteins could not be visualised in a functional flagellar motor. This shows that the tagged MotB is involved in wild-type flagellar rotation but leaves open the possibility that the second *motB* gene of *R.sphaeroides*, which was previously thought to be part of a separate flagellar system, may also function within wild-type rotating flagella.

7.4) FliM

7.4.1) Introduction

The FliM protein is an integral part of the flagellar motor C-Ring, in the paradigm of chemotaxis established in *E.coli* and *Salmonella*, binding of the phosphorylated form of response regulator protein CheY to FliM is responsible for the switch between swimming and tumbling (164). Binding of the CheYs to FliM has previously been demonstrated in *R.sphaeroides* (45), and the role of FliM is believed to be the same as that in *E.coli*.

As with the *motB* genes, *R.sphaeroides* has two gene variants of *fliM*, one of which has more classical sequence than the other. Alignment of the FliM protein products of *R.sphaeroides* with the known FliM proteins of *E.coli* and *Salmonella* using ClustalW and UniProt shows that the FliM1 of *R.sphaeroides* better resembles the classical FliM proteins, whereas FliM2 is missing a distinct 26 amino acid section from its N-terminal region (Figure 148).

```
E.coli        --MGDSILSQAEIDALLNGDSEVKDEPTASVSGESDIRPYDPNTQRRVVRERLQALEIIN  58
Salmonella    --MGDSILSQAEIDALLNGDSDTKDEPTPGIASDSDIRPYDPNTQRRVVRERLQALEIIN  58
FliM1         MAATPRKLSSKEVAALVG--NLMEASESTSLENGLEVRPYAFGENELNQLGDYHALRIIN  58
FliM2         ------MWEATVTESVLRRKIEAKARPAVPQVTGAERG--------------------LR  34
                           .      :::      :    :        :                :.
E.coli        ERFARHFRMGLFNLLRRSPDITVGAIRIQPYHEFARNLPVPTNLNLIHLKPLRGTGLVVF  118
Salmonella    ERFARQFRMGLFNLLRRSPDITVGAIRIQPYHEFARNLPVPTNLNLIHLKPLRGTGLVVF  118
FliM1         ERFCRTARDVFLPMLRLQPRISSFPPEVRSFDDYRSSQDNFVSITASRIEELRGNQMIVI  118
FliM2         LALARAGRDAFALALD-VPRLDCLRRSLTELLEMPEDLS----FHAVLEGPGEGLGLIAL  89
                :.*  *  :   *   * :       :    :   .      :         .*  ::.:
E.coli        SPSLVFIAVDNLFGGDGRFPTKVEGREFTHTEQRVINRMLKLALEGYSDAWKAINPLEVE  178
Salmonella    SPSLVFITVDNLFGGDGRFPTKVEGREFTHTEQRVINRMLKLALEGYSDAWKAINPLEVE  178
FliM1         PPPFISLLTDSYYGGQIRHVPTTR-TEFTATEERVIELVTDRLNVALQVAWRDLMALTFT  177
FliM2         APPLLAGLVEVQTVGRVRPQPAAT-RKPSRTDAAMVQDFIDRALAEL-----DLLLLEDE  143
              .*.::   .:    *  *  . .   : : *:  ::: . .              :  *
E.coli        YVRSEMQVKFTNITTSP---NDIVVNTPFHVEIGNLT-GEFNICLPFSMIEPLR---ELL  231
Salmonella    YVRSEMQVKFTNITTSP---NDIVVNTPFHVEIGNLT-GEFNICLPFSMIEPLR---ELL  231
FliM1         VVSREESMQFASFVDG----EDMVVNCSFMVQLPNTEPASFDILYPLQTLKPIS---SQL  230
FliM2         DRVWASGFRYASFLDDPRPLGLLLEDCSYRVLVAEVELALGAKSGQLLLAVPADGRGERP  203
                    .:::.:  .      :: : .: * : :   .      :    *     .
E.coli        VNPPLENSRNEDQNWRDNLVRQVQHSQLELVANFADISLRLSQILKLNPGDVLPIEKP--  289
Salmonella    VNPPLENSRHEDQNWRDNLVRQVQHSELELVANFADIPLRLSQILKLKPGDVLPIEKP--  289
FliM1         RSRMQSDFVDDDRSWREKLERAILSIPLTLSARLCEPEVPLRQLMQMQPGDVLPVHLT--  288
FliM2         HDGSQDEAPPPLTDFGAALAEQVMAVDCRLEAVLGRITLPLAQVLDWQPGDTVPLPQATL  263
               .   .:      .:   * . :      * * :    : * *::. :***.:*:  .
E.coli        -DRIIAHVDGVPVLTSQYGTLNGQYALRIEHLINPILN-SLNEEQPK-------------  334
Salmonella    -DRIIAHVDGVPVLTSQYGTVNGQYALRVEHLINPILN-SLNEEQPK-------------  334
FliM1         -EAVSLLVEGQPIFEAAPGERGGQAALNLTRRHVRS-----------------------  323
FliM2         DHILFVGLDGHPVAEGRLGQHRGMRAVRLAPMAAADPPPALPEFDPEDEDLSFAMTGTDE  323
               . :   ::* *:  .  *   *  *:.:
```

Figure 148: FliM protein sequence alignment – UniProt sequence alignment of the FliM proteins from *E.coli*, *Salmonella*, *R.sphaeroides* FliM1 and *R.sphaeroides* FliM2.

FliM1 has also been shown to be essential for *R.sphaeroides* motor formation (121) and to be part of the flagellar system one-associated FliK operon (50) which is regulated by FleQ/FleT dependant σ_{54} activity (116), thus making it a flagellar system one class III gene. FliM2, however is part another predicted operon, which encodes a 10-transmembrane metabolite symporter protein with a kinase domain, a single response regulator and FliM2 (Figure 149). No FleQ/FleT binding sequence is present in the promoter region of this operon. Current ongoing work is investigating the role of this potential two-component signalling pathway. Further study focused on FliM1 and its regulation as a class III gene.

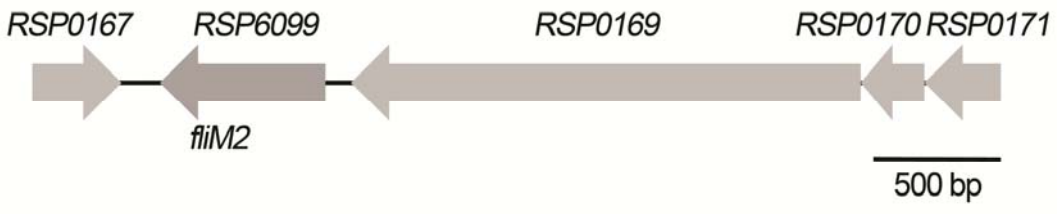


Figure 149: The *fliM2* operon. Schematic diagram of the genes encoded within the fliM2 operon. RSP171 is a predicted single-domain response regulator. RSP170 is an unknown hypothetical. RSP169 is a predicted 10-transmembrane symporter protein. RSP6099 is *fliM2*.

7.4.2) YFP-FliM is detectable within the motor

The YFP-FliM fluorescent protein fusion strain, JPA1638, consists of an N-terminal *yfp* gene fusion, a short (6 base pair) linker region and the full length *fliM* gene including the STOP codon. Due to the small (3 base pair) separation between *fliL* and *fliM* in the *fliK* operon, the START codon of this genetic fusion was carefully positioned at the same location as the native *fliM* gene, so as not to disrupt the expression of the rest of the components in this operon.

JPA1638 has previously been shown to have wild-type chemotactic efficiency on soft-agar swim plates, showing that the flagellum remains functional in the presence of the YFP-fusion. Despite this observation, the presence of FliM2 in the genome of *R.sphaeroides* left open the possibility that FliM1 may not participate in the motor itself. As with YFP-MotB, to verify whether or not the YFP-FliM localised to the flagellar motor, motile cells were tethered to tunnel slides using the anti-FliC polyclonal antibody and observed under TIRF illumination on a custom-built in-house microscope. Tethered cells were observed to rotate about a centre, confirming that the flagellar motor is functional under these conditions (Figure 150A). In more than 30 cells observed on 3 separate days, all contained distinct fluorescent spots which localised to the centre of rotation of the tethered cells (Figure 150B), showing that the YFP-FliM protein is present in functional motors.

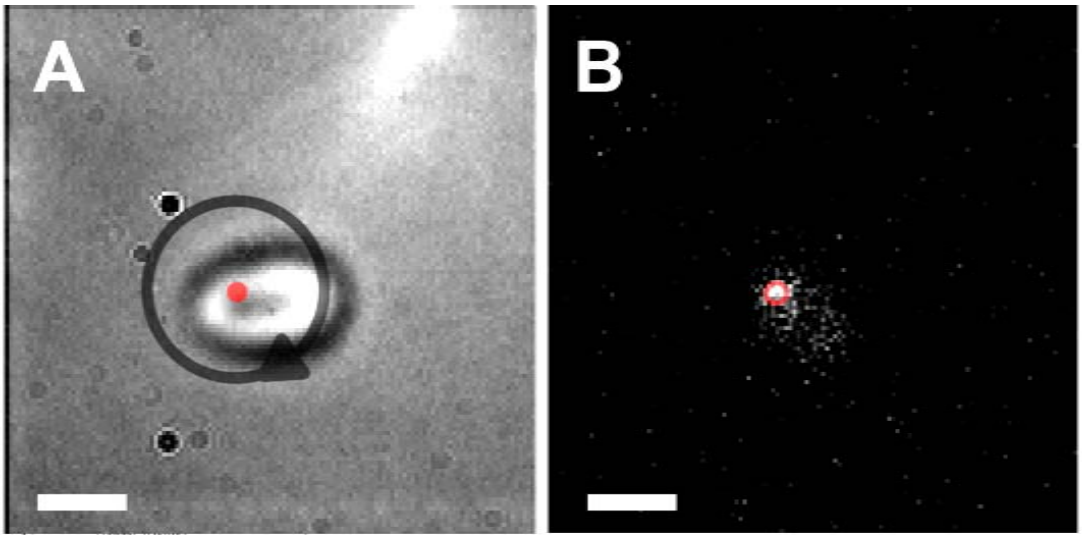


Figure 150: Example TIRF Images of JPA1638: **A)** Bright-field image of a single frame from a video of a tethered cell rotating about its centre. The centre of rotation is marked by a red spot, and the direction of rotation is indicated by an arrow. **B)** TIRF-illuminated fluorescence image of a single frame from a video of the same cell in **(A)**, the centre of rotation estimated from **(A)** is marked by a red ring. Contrast has been enhanced for clarity; the transparency of the arrow and red markers has been increased so as not to obscure the fluorescence data. Bars represent 1 μm.

7.4.3) FliA/FlgM regulation of FliM expression

The *motAB* operon and *fliK* operon have both been shown to be regulated by FleQ/FleT dependant σ_{54} activity (116), and the two operons are found in distinct locations of chromosome 1 of *R.sphaeroides*. To test whether the effect of *fliA* and *flgM* deletion observed for MotB was FleQ/FleT mediated, these genes were also deleted from the YFP-FliM background strain, JPA1638. To achieve this, pK18:*ΔflgM* and pK18:*ΔfliA* plasmids were conjugated into strain JPA1638 and the genes were deleted via successive rounds of recombination. Deletion of *flgM* (JPA2033) and *fliA* (JPA2032) was confirmed by southern blot analysis and DNA sequencing.

Having isolated the deletion mutants, YFP-FliM copy numbers were determined in each strain using fluorimetry (Table 44). The copy number of YFP-FliM in an otherwise wild-type background was approximately 1000 copies per cell. As was observed for YFP-MotB, YFP-FliM expression was decreased in the *ΔflgM* mutant (JPA2033) and a 3.6-fold increase in YFP-FliM expression was observed in the *ΔfliA* mutant (JPA2032), showing that FliA activity inhibits the expression of the *fliK* operon and that this inhibition is most likely achieved via a down-regulation of FleQ/FleT mediated σ_{54} activity.

Strain	Genotype Description	YFP Protein Copy Number
JPA1638	*yfp-fliM*	1000 ± 200
JPA2033	*yfp-fliM, ΔflgM*	600 ± 20
JPA2032	*yfp-fliM, ΔfliA*	3800 ± 60

Table 44**:** The Effect of Deleting *fliA* and *flgM* on YFP-FliM copy number**.** Copy numbers of YFP-FliM as calculated by the fluorimetry. The error shown is the standard error from three independently prepared samples.

7.4.4) Discussion

Here, the study for YFP-FliM expression levels in the presence and absence of FlgM and FliA has shown that FliA activity represses FliM expression. FliM is an essential component of the *R.sphaeroides* flagellar motor (121), and has been shown to be a FleQ/FleT-regulated class III flagellar gene (116). This provides further evidence that the FliA-mediated repression of class III genes may be responsible for the reduction in motility and flagellar number observed in *ΔflgM*

mutants and strains overexpressing FliA, and therefore that the FliA/FlgM checkpoint may regulate the monotrichous nature of *R.sphaeroides*.

7.5) Chapter Discussion

Previously, we had observed that FliA and FlgM regulate the expression of the chemotaxis genes of *R.sphaeroides*. Investigating chemotaxis in the *fliA* and *flgM* deletion mutants revealed that both strains have reduced motility, which is most likely the result of a reduction in flagellar number which was observed by electron microscopy. This phenotype was also observed with the over-expression of FliA, suggesting that FliA activity represses flagellar assembly.

Measuring changes in expression showed that deletion of *fliA* results in an increase of expression of class III flagellar genes, showing that FliA regulates flagellar number via a negative feedback loop within the flagellar assembly pathway. As FliA itself is unlikely to form an inhibitory protein:protein interaction, the mechanism of inhibition is likely to involve the FliA-regulated expression of an inhibitory protein. However, BLAST search analyses showed that *R.sphaeroides* lacks the majority of known flagellar-number repressors, most notably – FleN, which is known to repress FleQ master regulator activity in *P.aeruginosa*. Thus, FliA feedback repression of flagellar synthesis is achieved via a new, unknown mechanism, and could repress any combination of class I, class II and class III genes. Unfortunately, no reporters were available for the direct measurement of class II or class I gene expression so further analysis was not possible within the given time frame.

Interestingly FlgM, which is known to be part of a class III flagellar gene operon, appears to be highly regulated by FliA, and is therefore predominantly class IV.

Additionally, motility defects in both *fliA* and *flgM* deletion strains resulted in a reduced ability to form biofilms, adding *R.sphaeroides* to the long list of bacterial species where motility is essential to complex processes such as biofilm formation, virulence and symbiosis.

It appears that one of FliA's main functions is to inhibit its own activity, both transcriptionally by repression of its own expression (and the expression of other class III flagellar genes), and by promoting the expression of its inhibitor – FlgM. This auto-repression allows FliA activity to be regulated exclusively by the temporal secretion of FlgM, and a wave of FliA activity would be predicted to follow the completion of the HBB. One interesting question is what happens after cell division? As *R.sphaeroides* cells are monotrichous, after division many daughter cells would be predicted to have no flagella – It is possible that the change from having a single flagellum to having no flagella is sufficient to prevent FlgM secretion in a way that allows for repression of FliA activity,

and expression of the class III flagellar genes and completion of a HBB, thus linking the cell cycle to the regulation of flagellar gene expression. One method of addressing this question was by the generation of the FlgM-YFP fluorescent fusions in this chapter - as a short build up of FlgM would be expected in the cytoplasm directly after cell septation (and this could be observed by microscopy), however this was not observed (data not shown). Without directly linking flagellar assembly to the cell cycle – the monotrichous nature of *R.sphaeroides* (and other bacteria) must be explained by a careful balance of gene expression levels that result in one flagellum being made (on average) per cell division – indeed, the perturbation of gene expression in the mutants made in this study was sufficient to affect flagellar number, but no obviously peritrichous phenotype was observed. The only way to determine whether this is indeed the case is via the use of improved time-lapse microscopic observation, which would require the synthesis of many different fluorescently labelled mutant strains, and could well be addressed in the future.

Without being directly linked to the cell cycle, the complex regulatory system identified here is sufficient to regulate flagella number in *R.sphaeroides*, and the FlgM secretion checkpoint would be a logical control mechanism for regulating flagella number in all monotrichous bacterial species.

8) General Discussion

Chemotaxis and motility in *R.sphaeroides* are achieved in ways that diverge from the well established paradigm of *E.coli*. The multi-component phosphorelay network that controls chemotactic signalling in *R.sphaeroides* is a step up in complexity from that of *E.coli*, and is still not fully understood. Additionally, the mechanism of motor braking (the output signal of chemotaxis) and the rationale behind the run-stop swimming behaviour of *R.sphaeroides* when compared to the run-tumble swimming behaviour of *E.coli* are yet to be fully described.

Key to bacterial chemotaxis and motility is the assembly of macromolecular protein complexes that possess the ability to relay signals (chemotaxis) and adaptively harness energy (motility). While assembly of the complex macromolecular structure of the flagellar motor is known to be regulated in a logical and step-wise manner, vast polar arrays of chemoreceptor proteins are thought to form spontaneously in a stochastic process that is dependent on diffusion, protein:protein interaction and, potentially, a cytoskeletal "capture" mechanism. Work in *E.coli* and other species has suggested that transcription-level regulation of protein expression is the mechanism that controls the rate of assembly of these protein complexes and, therefore, the number of flagella and chemoreceptor arrays that are present in a cell at any given moment. The regulation of protein expression is therefore achieved in many and varied ways, which reflect the need for different amounts of the different components of each system.

The complexity of *R.sphaeroides* chemotaxis is enhanced by the presence of an additional structural component which is not observed in other bacterial systems; the cytoplasmic chemoreceptor cluster. Although little is known about the structural organisation or dynamics of cytoplasmic chemotaxis cluster proteins, cluster assembly is believed to occur in a similar manner to that of the polar chemotaxis proteins but where new clusters are not spontaneously nucleated, but formed from the active partitioning of old clusters.

This study has focused on the regulation and positioning of the macromolecular protein components involved in chemotaxis and motility in *R.sphaeroides*. In order to achieve this, two new, fluorescence-based, analytical techniques were developed which facilitated the measurement of protein expression levels in fluorescent-fusion mutants and the analysis of protein cluster intensities and positions from epifluorescence images. These techniques were used to demonstrate how the

activity of the σ-factor, FliA, and its anti-σ-factor, FlgM, regulate flagellar and chemotaxis protein expression and cluster number.

It was demonstrated that expression of flagellin (FliC), FlgM and the components of both cytoplasmic and polar chemoreceptor clusters in *R.sphaeroides* are upregulated by the activity of FliA. The co-regulation of chemotaxis and flagellar genes at the final stage of flagellar assembly is common in many bacterial species, and reflects the closely-linked functions of the two sets of genes. Indeed, the chemotaxis proteins are useless in the absence of flagella, and the functionality of the flagellar motor is reduced in the absence of the chemotaxis proteins, thus their co-regulation is an efficient use of a single σ-factor.

Using this knowledge, it was possible to vary the concentration of chemotaxis proteins within cells by inducing expression of FliA from a plasmid and to quantify the amount of protein and cluster positions for polar and cytoplasmic fluorescent reporter strains.

Observation of polar-localising chemoreceptor arrays revealed some intriguing differences between these protein clusters and the analogous clusters of *E.coli*. In *E.coli*, these clusters are known to form at lateral positions along the cell corresponding to the 1/8, 1/4, 1/2, 3/4 and 7/8 positions of the cell, successive rounds of cell division result in laterally-positioned clusters moving to newly formed poles of the cell. Additionally, the number of clusters seen in cells increases with protein concentration – an observation that led to the development of the stochastic nucleation model of cluster formation. In *R.sphaeroides*, laterally-positioned clusters were rare and increasing protein concentration did not cause a significant increase in the number of clusters seen, although high protein concentrations did result in the detection of clusters at newly-forming poles.

Furthermore, *E.coli* laterally-positioned clusters were seen to form in elongated cells, demonstrating that the polar curvature of the membrane, which has previously been suggested to influence the positioning of MCPs, is not a requirement for cluster formation. In *R.sphaeroides*, multiple clusters commonly formed at similar loci on the long axis of the cell, with clusters being visible on either side of the cell. A distinct pattern of lateral protein localisation was also observed in *R.sphaeroides* cephalexin-treated cells, with clusters occupying the 1/3 and 2/3 positions along the cell axis. This pattern was not seen in long-cells that were the product of spontaneous mutation. Additionally, multiple clusters were observed to form within individual cell poles across all studied conditions, with as many as three distinct clusters being visible in one polar region. Taken together, these observations may be an indication that polar/lateral protein clusters associate with a cytoskeletal element, similar to the “pre-division site” proposed in *E.coli*, and that this element may

form a band around the cell diameter, allowing positioning on either side of the cell. Association with an immobile cytoskeletal element would also explain why multiple clusters could be seen in individual poles, as reduced mobility of these clusters would allow for close association without clusters combining by random diffusion.

Excitingly, the isolation of filamentous cells with altered lateral chemotaxis cluster localisation may well provide a method for the study of the hypothetical "pre-division site". Identification of genes responsible for the observed filamentous growth, although impossible in the time-frame of this project, would provide a starting-block for this study. From the observations made within this project, *R.sphaeroides* requirement for lateral chemotaxis clusters is predicted to differ from that of *E.coli*. The methods used for the arrangement of proteins along the lateral edge of the cell are also predicted to differ between the two species, and these may well be linked to the formation of cellular division sites.

Unlike the polar/lateral clusters, the number of cytoplasmic clusters in an *R.sphaeroides* cell was seen here to be affected by the concentration of the cluster components. This was most obvious with reduced expression of the *che*Op3 components, where cells with two clusters were much less common within a population. This observation suggests that the rate of cluster partitioning is concentration dependant, and therefore the cellular level of *che*Op3 proteins must be kept high enough in order for cluster partitioning to occur before cell division so that daughter cells inherit at least one cytoplasmic chemotaxis cluster. This raises the interesting question as to whether chemotaxis gene expression may be linked to the cell cycle, and whether bacterial cells regulate their cell cycle in the same way as eukaryotic organisms.

In wild-type populations, cytoplasmic cluster positions were seen to be determined relative to the poles of the cell; however no obvious pattern was seen in the cluster positions of elongated cells. It is not clear how PpfA-mediated partitioning and positioning of cytoplasmic clusters generates the observed cluster distributions in *R.sphaeroides*. Although positions are linked in some way to the cell poles, partitioning can occur along the length of elongated cells, far away from existing poles, and in the absence of newly-forming cell division sites. Further investigation into the function of PpfA, which is essential to cluster partitioning, will no doubt help to determine how cluster partitioning is achieved.

One benefit of the image analysis software that was custom-designed as part of this study, aside from an increase in the number of analysed cells, was the potential to simultaneously extract image

data relating to cluster positions and intensity data relating to cluster sizes. This allowed the nature of asymmetrical distribution of protein clusters to be addressed.

Bacterial cells that localise proteins to their poles will have an inherent asymmetry in protein distribution due to successive rounds of cell division. The data in this study suggest that polar chemotaxis proteins accumulate at poles over time, and thus, the old pole contains more protein than the new pole of the cell. The ratio of chemotaxis proteins at old pole and new pole positions was seen to be, on average, approximately 3:1. In *R.sphaeroides* this asymmetry may cause an additional problem for the successful signalling of the chemotaxis pathway, as daughter cells inherit both polar and cytoplasmic chemotaxis protein clusters. If cytoplasmic were split exactly 50:50, then each set of daughter cells would have 3-fold differing ratios of polar/cytoplasmic chemotaxis proteins. In fact, it was seen here that cytoplasmic cluster partitioning does not result in a 50:50 split, but on average clusters are split 2:1, and that larger cytoplasmic clusters can end up in either daughter cell, irrespective of the age of that daughter cell's old pole. Interestingly, this means that as a direct consequence of chemotaxis protein clustering, the ratio of cytoplasmic/polar chemotaxis components within a population of cells will vary over a vast range. Thus, in order for the majority of *R.sphaeroides* cells to achieve chemotaxis, this signalling network must be robust to variation in the concentration of the components of each cluster.

It was demonstrated that both of the essential chemotaxis operons of *R.sphaeroides* are regulated by the activity of FliA and FlgM, but that the magnitude of this regulation suggests that FliA activity fine tunes the expression of these genes under relevant physiological conditions. *che*Op2 is also known to possess a σ^{70} promoter, whereas *che*Op3 possesses a σ^{54} promoter. The exact reason for differential regulation of these operons is unknown, but may be related to the observed asymmetry of cell division. As the cell pole ages and gains more chemotaxis protein, the cytoplasm does not and partitions to daughter cells randomly. As individual cells inherit unfavourable ratios of chemotaxis proteins, differential regulation of the chemotaxis operons may mean that cells could adjust the expression of these components in order to recover chemotaxis. In order to achieve this, components of chemotaxis operons 2 and 3 would have to directly regulate transcription, and this has not been observed to date – although, interestingly, unclassified open reading frames exist close to either end of chemotaxis operon 3.

FliA itself has been demonstrated to be part of a complex regulatory network that is involved in flagellar assembly and the regulation of the chemotaxis genes. Deletion of *fliA* combined with fluorescence-based studies of protein copy number revealed that a negative feedback loop exists within the flagellar assembly hierarchy of *R.sphaeroides*, where class IV FliA activity represses the

expression of class III genes. FliA is itself known to be a class III gene. Therefore, a novel mechanism exists with *R.sphaeroides* where the secretion of FlgM upon completion of flagellar synthesis results in increased FliA activity. This increase in activity promotes the expression of FliC, allowing completion of the flagellar filament, and expression of FlgM replacing the secreted cellular FlgM and inhibiting FliA activity. FliA activity is also repressed at the level of transcription by the repression of the class III genes, these include components of the HBB and therefore the rate of production of a second flagellum is reduced in response to the completion of the first. We observed that increased FliA activity was sufficient to repress flagellar synthesis and reduce flagellar number within a population.

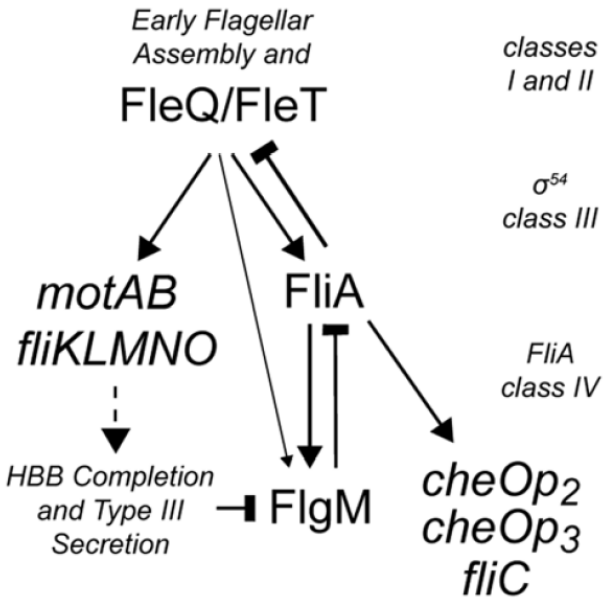


Figure 151 **Schematic representation of FliA and FlgM within the flagellar assembly pathway of *R.sphaeroides*.** Expression from FleQ/FleT dependant class III promoters encoding *motAB*, *fliKLMNO*, *flgM* and *fliA* is activated by expression of the early flagellar genes and *fleQ*. FliA activity represses early flagellar gene expression (by an unknown mechanism) and promotes expression of the class IV genes, including *fliC*, two chemotaxis operons and *flgM*. This allows completion of the first flagellum and inhibits production of a second. FlgM inhibits FliA activity, but is itself secreted from the cell temporally upon completion of the HBB complex, releasing FliA and promoting class IV gene expression.

FlgM-secretion is a logical checkpoint for the regulation of flagellar number in monotrichous bacteria, as completion of the first flagellum provides the signal to repress synthesis of a second. It would not be surprising, therefore, if other monotrichous species possessed similar mechanisms of feedback repression. This is not seen to be the case when examined in the literature, but few examples have been studied. For example, *P.aeruginosa* possesses FleN, which allows regulation of flagellar number independently of FliA but is itself part of a negative feedback loop within the flagellar assembly hierarchy (30). Recent observations in *Y.pseudotuberculosis* did show that

deletion of *flgM* resulted in suppression of flagellar synthesis without altering the activity of the class I master regulator (FlhDC), which may well be evidence for a similar repression system to that observed in this study (38).

Although the well established behavioural models of motility and chemotaxis from *E.coli* and *Salmonella* are useful, it has long been known that they do not necessarily provide an accurate representation of what is happening in other bacterial species. The chemotaxis system of *R.sphaeroides* is known to be far more complex than that of *E.coli* and scaling up the model of signalling is insufficient to provide an accurate picture of how *R.sphaeroides* achieves chemotaxis. We have seen here that chemotaxis gene regulation; chemotaxis protein cluster positioning and flagellar number regulation are all achieved in ways that differ from the situation in *E.coli* and other bacteria. Although general principles can be developed from the study of these systems, it is likely that individual species will display unique characteristics, and it is these features that distinguish one species from another.

9) Appendices

9.1) List of Primers

Primer Reference Code	Label Name	Primer Sequence 5' – 3'
DW01	DELFLGM1	CCG**GAATTC**AGCTTGAACGGCAGCTGGAG
DW02	DELFLGM2	GTGTTCGAGGTCTCAGACCATGTGACCCTC
DW03	DELFLGM3	GAGGGTCACATGGTCTGAGACCTCGAACAC
DW04	DELFLGM4	CCC**AAGCTT**CCGGTGAGCCAGTCGATGTC
DW05	dFliA1	CCG**GAATTC**TCGGCGAACTTCTCGAGGAC
DW06	dFliA2	CCC**AAGCTT**ACGCGCTGGAGGATTATGTG
DW07	FliA.AlEx1	CGC**TCATGA**TCCATGCACGCGGATATGCC
DW08	FliA.AlEx2	CGC**GGATCC**CGCGGACCTTCAGACCTCGCCCTGC
DW09	FlgM.AlEx1	CATG**CCATGG**GGAAGGAGGGTCACATGGTC
DW10	FlgM.AlEx2	CGC**GGATCC**TCAGAGCTTCAGATCGAGATAG
DW11	FliA.Pur1	CGC**GGATCC**ATCCATGCACGCGGATATG
DW12	FliA.Pur2	CCC**AAGCTT**TCAGACCTCGCCCTGCAG
DW13	FlgM.Pur1	CGC**GGATCC**AAGGAGGGTCACATGG
DW14	FlgM.Pur2	CCC**AAGCTT**TCAGAGCTTCAGATCGAG
DW15	FlgM N-YFP 2 NEW	CTCCTCGCCCTTGCTCACCATGCCTCTCGTAACGGCCC
DW16	FlgM N-YFP 3 NEW	GGGCCGTTACGAGAGGCATGGTGAGCAAGGGCGAGGAG
DW17	MotB1.FP4	CACG**TCTAGA**CTTGTACAGCTCGTCCATG
DW18	FlgM N-YFP 5	TGC**TCTAGA**AAGGAGGGTCACATGGTCG
DW19	FlgM C-YFP 2 NEW	CTCCTCGCCCTTGCTCACGAGCTTCAGATCGAGATAGTC
DW20	FlgM C-YFP 3 NEW	GACTATCTCGATCTGAAGCTCGTGAGCAAGGGCGAGGAG
DW21	FlgM C-YFP 5	TGC**TCTAGA**TGAGACCTCGAACACCGG
DW22	MotB1.FP1	CCG**GAATTC**TTCCCTCCCTGGTGATTGTC
DW23	MotB1.FP2	CTCCTCGCCCTTGCTCACGGGCGGACAT
DW24	MotB1.FP3	ATGTCCGCCCGTGAGCAAGGGCGAGGAG
DW25	MotB1.FP5	TGC**TCTAGA**TCCGCAAAGCCCAAGGTC
DW26	MotB1.FP6	CCC**AAGCTT**CATTGGTGACAGCGGGAATC
DW27	MOTB1.PUR1	CATG**CCATGG**GGGCGATGCGGGAG
DW28	MOTB1.PUR2	GGA**AGATCT**GTCCTTGTAGGAAATC

Table 45: Primer List. Reference codes, label names and primer sequences of all the primers referred to in this work. The restriction enzyme cut-sites used for cloning are marked in bold.

9.2) Solutions and Chemicals

9.2.1) Luria broth

Ingredient	Quantity
Yeast Extract	5 g
Tryptone	10 g
NaCl	5 g

Made up to 1 litre with MilliQ water. Adjusted to pH 7.0. Autoclave-sterilised.

9.2.2) Succinate medium

Ingredient	Quantity
Concentrated Base (appendix 9.2.2.1)	20 ml
Growth Factors (appendix 9.2.2.2)	2 ml
1 M Phosphate buffer pH 7.0	20 ml
$(NH_4)SO_4$	0.5 g
Sodium Succinate	2.0 g
NaCl	0.5 g
Casamino acids	1 g

Made up to 1 litre with MilliQ water. Adjusted to pH 7.2 with KOH. Autoclave-sterilised.

9.2.2.1) Concentrated base

Ingredient	Quantity
Metals 44 solution (appendix 9.2.2.3)	25 ml
$MgSO_4.7H_20$	14.5 g
Nitrilotriacetic acid (Na salt)	5.94 g
$CaCl_2.6H_2O$	2.5 g
$FeSO_4.7H_2O$	0.05 g
Ammonium molybdate tetrahydrate	0.0046 g

Made up to 1 litre with MilliQ water. Adjusted to pH 6.8.

9.2.2.2) Growth factors

Ingredient	Quantity
Nicotinic Acid	100 mg
Thiamine HCL	50 mg
$NaHCO_3$	50 mg
Biotin	2 mg

Dissolve in 100 ml of MilliQ water. Filter sterilised.

9.2.2.3) Metals 44 solution

Ingredient	Quantity
EDTA	2.5 g
$ZnSO_4.7H_2O$	11 g
$FeSO_4.7H_2O$	5 g
$MnSO_4.7H_2O$	2 g
$CuSO_4.5H_2O$	0.39 g
H_3BO_3	12 g
$CoCl_2.6H_2O$	0.2 g
3 M sulphuric acid	1.5 ml

Made up to 1 litre in MilliQ water. Filter sterilised.

9.2.3) 5 x DNA loading dye

Ingredient	Quantity
Glycerol	70 % (w/v)
Bromophenol blue	0.25 % (w/v)
Xylene cyanol	0.25 % (w/v)
RNAse	Trace

Dissolved in MilliQ water.

9.2.4) ½ x TBE

Ingredient	Quantity
Tris base	108 g
Boric acid	55 g
EDTA	7.4 g

Dissolved in 1 litre of MilliQ water. Made up to 20 litres in MilliQ water.

9.2.5) Chromoprep lysis buffer

Ingredient	Quantity
Tris-HCL pH 8.0	10 mM
EDTA	1 mM
SDS	1 %

Dissolved in MilliQ water.

9.2.6) TFB1

Ingredient	Quantity
30 mM Potassium Acetate	2.945 g
100 mM Rubidium chloride	12.09 g
10 mM $CaCl_2.2H_2O$	1.47 g
50 mM $MnCl_2.4H_2O$	9.895 g
Glycerol	150 ml

Made up to 1 litre in MilliQ water. Adjusted to pH 5.8 with 0.2M acetic acid. Filter sterilised.

9.2.7) TFB2

Ingredient	Quantity
10 mM PIPES	0.3024 g
75 mM $CaCl_2.2H_20$	1.103 g
10 mM $RbCl_2$	0.121 g
Glycerol	15 ml

Made up to 100 ml in MilliQ water. Autoclave sterilised. Stored at 4 °C.

9.2.8) M22 medium

Ingredient	Quantity
Concentrated Base (appendix 9.2.2.1)	20 ml
Growth Factors (appendix 9.2.2.2)	2 ml
1 M Phospate buffer pH 7.0	20 ml
$(NH_4)SO_4$	0.5 g
NaCl	0.5 g

Made up to 1 litre with MilliQ water. Adjusted to pH 7.2 with KOH. Autoclave-sterilised.

9.2.9) 2YT medium

Ingredient	Quantity
Yeast Extract	10 g
Tryptone	16 g
NaCl	5 g

Made up to 1 litre in MilliQ water. Adjusted to pH 7.0. Autoclave sterilised.

9.2.10) Protein wash buffer

Ingredient	Quantity
Glycerol	10 %
Tris-HCL pH 8.0	50 mM
NaCl	150 mM
DTT	1 mM
Imidazole	150 µM

Dissolved in MilliQ water.

9.2.11) Protein elution buffer

Ingredient	Quantity
Glycerol	10 %
Tris-HCL pH 8.0	50 mM
NaCl	150 mM
DTT	1 mM
Imidazole	10 mM

Dissolved in MilliQ water.

9.2.12) 5 x protein loading dye

Ingredient	Quantity
B-mercaptotethanol	750 µl
SDS	0.75 g
Bromophenol blue	75 mg
Glycerol	2.5 ml

Made up to 10 ml with MilliQ water.

9.2.13) SDS running buffer

Ingredient	Quantity
Tris	30.2 g
Glycine	188 g
SDS	10 g

Made up to 10 litres with MilliQ water.

9.3) MATLAB code

The MATLAB software developed for the analysis of epifluorescence images of *R.sphaeroides* is available on the attached CD-ROM.

10) References

1. **Abeles, A. L., S. A. Friedman, and S. J. Austin.** 1985. Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. J Mol Biol **185:**261-72.
2. **Alon, U.** 2003. Biological networks: the tinkerer as an engineer. Science **301:**1866-7.
3. **Amati, G., P. Bisicchia, and A. Galizzi.** 2004. DegU-P represses expression of the motility *fla-che* operon in *Bacillus subtilis*. J Bacteriol **186:**6003-14.
4. **Anand, G. S., P. N. Goudreau, and A. M. Stock.** 1998. Activation of methylesterase CheB: evidence of a dual role for the regulatory domain. Biochemistry **37:**14038-47.
5. **Anderson, J. K., T. G. Smith, and T. R. Hoover.** 2010. Sense and sensibility: flagellum-mediated gene regulation. Trends Microbiol **18:**30-7.
6. **Armitage, J. P.** 1992. Bacterial motility and chemotaxis. Sci Prog **76:**451-77.
7. **Armitage, J. P., and R. Schmitt.** 1997. Bacterial chemotaxis: *Rhodobacter sphaeroides* and *Sinorhizobium meliloti*--variations on a theme? Microbiology **143 (Pt 12):**3671-82.
8. **Ballado, T., L. Camarena, B. Gonzalez-Pedrajo, E. Silva-Herzog, and G. Dreyfus.** 2001. The hook gene (*flgE*) is expressed from the *flgBCDEF* operon in *Rhodobacter sphaeroides*: study of an *flgE* mutant. J Bacteriol **183:**1680-7.
9. **Bange, G., N. Kummerer, C. Engel, G. Bozkurt, K. Wild, and I. Sinning.** 2010. FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. Proc Natl Acad Sci U S A **107:**11295-300.
10. **Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, and D. A. Hopwood.** 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature **417:**141-7.
11. **Berg, H. C.** 2003. The rotary motor of bacterial flagella. Annu Rev Biochem **72:**19-54.
12. **Berg, H. C., and D. A. Brown.** 1972. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. Nature **239:**500-4.
13. **Berry, R. M., and J. P. Armitage.** 1999. The bacterial flagella motor. Adv Microb Physiol **41:**291-337.
14. **Blair, D. F.** 2003. Flagellar movement driven by proton translocation. FEBS Lett **545:**86-95.
15. **Blobel, G.** 2000. Protein targeting (Nobel lecture). Chembiochem **1:**86-102.
16. **Bourret, R. B., and R. E. Silversmith.** 2010. Two-component signal transduction. Curr Opin Microbiol **13:**113-5.
17. **Briegel, A., D. R. Ortega, E. I. Tocheva, K. Wuichet, Z. Li, S. Chen, A. Muller, C. V. Iancu, G. E. Murphy, M. J. Dobro, I. B. Zhulin, and G. J. Jensen.** 2009. Universal architecture of bacterial chemoreceptor arrays. Proc Natl Acad Sci U S A **106:**17181-6.
18. **Brown, J. D., S. Saini, C. Aldridge, J. Herbert, C. V. Rao, and P. D. Aldridge.** 2008. The rate of protein secretion dictates the temporal dynamics of flagellar gene expression. Mol Microbiol **70:**924-37.
19. **Bulyha, I., C. Schmidt, P. Lenz, V. Jakovljevic, A. Hone, B. Maier, M. Hoppert, and L. Sogaard-Andersen.** 2009. Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. Mol Microbiol **74:**691-706.
20. **Butler, S. M., and A. Camilli.** 2005. Going against the grain: chemotaxis and infection in *Vibrio cholerae*. Nat Rev Microbiol **3:**611-20.

21. **Chadsey, M. S., and K. T. Hughes.** 2001. A multipartite interaction between *Salmonella* transcription factor sigma28 and its anti-sigma factor FlgM: implications for sigma28 holoenzyme destabilization through stepwise binding. J Mol Biol **306:**915-29.
22. **Chadsey, M. S., J. E. Karlinsey, and K. T. Hughes.** 1998. The flagellar anti-sigma factor FlgM actively dissociates *Salmonella typhimurium* sigma28 RNA polymerase holoenzyme. Genes Dev **12:**3123-36.
23. **Charon, N. W., and S. F. Goldstein.** 2002. Genetics of motility and chemotaxis of a fascinating group of bacteria: the spirochetes. Annu Rev Genet **36:**47-73.
24. **Charon, N. W., E. P. Greenberg, M. B. Koopman, and R. J. Limberger.** 1992. Spirochete chemotaxis, motility, and the structure of the spirochetal periplasmic flagella. Res Microbiol **143:**597-603.
25. **Chevance, F. F., and K. T. Hughes.** 2008. Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol **6:**455-65.
26. **Chun, S. Y., and J. S. Parkinson.** 1988. Bacterial motility: membrane topology of the *Escherichia coli* MotB protein. Science **239:**276-8.
27. **Clarke, M. B., and V. Sperandio.** 2005. Transcriptional regulation of flhDC by QseBC and sigma (FliA) in enterohaemorrhagic *Escherichia coli*. Mol Microbiol **57:**1734-49.
28. **Coulter, A. H.** 2006. Godless : the church of liberalism, 1st ed. Crown Forum, New York.
29. **Crosson, S., H. McAdams, and L. Shapiro.** 2004. A genetic oscillator and the regulation of cell cycle progression in *Caulobacter crescentus*. Cell Cycle **3:**1252-4.
30. **Dasgupta, N., S. K. Arora, and R. Ramphal.** 2000. *fleN*, a gene that regulates flagellar number in *Pseudomonas aeruginosa*. J Bacteriol **182:**357-64.
31. **Dasgupta, N., E. P. Ferrell, K. J. Kanack, S. E. West, and R. Ramphal.** 2002. *fleQ*, the gene encoding the major flagellar regulator of *Pseudomonas aeruginosa*, is sigma70 dependent and is downregulated by Vfr, a homolog of *Escherichia coli* cyclic AMP receptor protein. J Bacteriol **184:**5240-50.
32. **Dasgupta, N., and R. Ramphal.** 2001. Interaction of the antiactivator FleN with the transcriptional activator FleQ regulates flagellar number in *Pseudomonas aeruginosa*. J Bacteriol **183:**6636-44.
33. **Dasgupta, N., M. C. Wolfgang, A. L. Goodman, S. K. Arora, J. Jyot, S. Lory, and R. Ramphal.** 2003. A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*. Mol Microbiol **50:**809-24.
34. **De Mot, R., and J. Vanderleyden.** 1994. The C-terminal sequence 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 Microbiol **12:**333-4.
35. **Dean, G. E., R. M. Macnab, J. Stader, P. Matsumura, and C. Burks.** 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:**991-9.
36. **Delisi, C.** 1984. Random-Walks in Biology - Berg,Hc. Nature **310:**258-258.
37. **Diederich, B., J. F. Wilkinson, T. Magnin, M. Najafi, J. Erringston, and M. D. Yudkin.** 1994. Role of interactions between SpoIIAA and SpoIIAB in regulating cell-specific transcription factor sigma F of *Bacillus subtilis*. Genes Dev **8:**2653-63.
38. **Ding, L., Y. Wang, Y. Hu, S. Atkinson, P. Williams, and S. Chen.** 2009. Functional characterization of FlgM in the regulation of flagellar synthesis and motility in *Yersinia pseudotuberculosis*. Microbiology **155:**1890-900.
39. **Driessen, A. J., and N. Nouwen.** 2008. Protein translocation across the bacterial cytoplasmic membrane. Annu Rev Biochem **77:**643-67.
40. **Duke, T. A., N. Le Novere, and D. Bray.** 2001. Conformational spread in a ring of proteins: a stochastic approach to allostery. J Mol Biol **308:**541-53.
41. **Egelman, E. H.** 2009. Divergence of the flagellar hook and filament. Structure **17:**1425-6.

42. **Egelman, E. H.** 2010. Reducing irreducible complexity: divergence of quaternary structure and function in macromolecular assemblies. Curr Opin Cell Biol **22:**68-74.
43. **Engstrom, P., and G. L. Hazelbauer.** 1980. Multiple methylation of methyl-accepting chemotaxis proteins during adaptation of *E. coli* to chemical stimuli. Cell **20:**165-71.
44. **Fan, F., and R. M. Macnab.** 1996. Enzymatic characterization of FliI. An ATPase involved in flagellar assembly in *Salmonella typhimurium*. J Biol Chem **271:**31981-8.
45. **Ferre, A., J. De La Mora, T. Ballado, L. Camarena, and G. Dreyfus.** 2004. Biochemical study of multiple CheY response regulators of the chemotactic pathway of *Rhodobacter sphaeroides*. J Bacteriol **186:**5172-7.
46. **Francis, N. R., V. M. Irikura, S. Yamaguchi, D. J. DeRosier, and R. M. Macnab.** 1992. Localization of the *Salmonella typhimurium* flagellar switch protein FliG to the cytoplasmic M-ring face of the basal body. Proc Natl Acad Sci U S A **89:**6304-8.
47. **Frisk, A., J. Jyot, S. K. Arora, and R. Ramphal.** 2002. Identification and functional characterization of *flgM,* a gene encoding the anti-sigma 28 factor in *Pseudomonas aeruginosa*. J Bacteriol **184:**1514-21.
48. **Frye, J., J. E. Karlinsey, H. R. Felise, B. Marzolf, N. Dowidar, M. McClelland, and K. T. Hughes.** 2006. Identification of new flagellar genes of *Salmonella enterica serovar Typhimurium*. J Bacteriol **188:**2233-43.
49. **Galkin, V. E., X. Yu, J. Bielnicki, J. Heuser, C. P. Ewing, P. Guerry, and E. H. Egelman.** 2008. Divergence of quaternary structures among bacterial flagellar filaments. Science **320:**382-5.
50. **Garcia, N., A. Campos, A. Osorio, S. Poggio, B. Gonzalez-Pedrajo, L. Camarena, and G. Dreyfus.** 1998. The flagellar switch genes fliM and fliN of *Rhodobacter sphaeroides* are contained in a large flagellar gene cluster. J Bacteriol **180:**3978-82.
51. **Gerdes, K., M. Howard, and F. Szardenings.** 2010. Pushing and pulling in prokaryotic DNA segregation. Cell **141:**927-42.
52. **Gillen, K. L., and K. T. Hughes.** 1991. Negative regulatory loci coupling flagellin synthesis to flagellar assembly in *Salmonella typhimurium*. J Bacteriol **173:**2301-10.
53. **Gillen, K. L., and K. T. Hughes.** 1993. Transcription from two promoters and autoregulation contribute to the control of expression of the *Salmonella typhimurium* flagellar regulatory gene *flgM*. J Bacteriol **175:**7006-15.
54. **Giuliano, G., D. Pollock, H. Stapp, and P. A. Scolnik.** 1988. A Genetic-Physical Map of the *Rhodobacter-Capsulatus* Carotenoid Biosynthesis Gene-Cluster. Molecular & General Genetics **213:**78-83.
55. **Greenfield, D., A. L. McEvoy, H. Shroff, G. E. Crooks, N. S. Wingreen, E. Betzig, and J. Liphardt.** 2009. Self-organization of the *Escherichia coli* chemotaxis network imaged with super-resolution light microscopy. PLoS Biol **7:**e1000137.
56. **Hamblin, P. A., B. A. Maguire, R. N. Grishanin, and J. P. Armitage.** 1997. Evidence for two chemosensory pathways in *Rhodobacter sphaeroides*. Mol Microbiol **26:**1083-96.
57. **Hao, S., D. Hamel, H. Zhou, and F. W. Dahlquist.** 2009. Structural basis for the localization of the chemotaxis phosphatase CheZ by CheAS. J Bacteriol **191:**5842-4.
58. **Hay, N. A., D. J. Tipper, D. Gygi, and C. Hughes.** 1997. A nonswarming mutant of *Proteus mirabilis* lacks the Lrp global transcriptional regulator. J Bacteriol **179:**4741-6.
59. **Hazelbauer, G. L., J. J. Falke, and J. S. Parkinson.** 2008. Bacterial chemoreceptors: high-performance signaling in networked arrays. Trends Biochem Sci **33:**9-19.
60. **Hazelbauer, G. L., and W. C. Lai.** 2010. Bacterial chemoreceptors: providing enhanced features to two-component signaling. Curr Opin Microbiol **13:**124-32.
61. **Hu, Z., and J. Lutkenhaus.** 1999. Topological regulation of cell division in *Escherichia coli* involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE. Mol Microbiol **34:**82-90.

62. **Hughes, K. T., K. L. Gillen, M. J. Semon, and J. E. Karlinsey.** 1993. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. Science **262:**1277-80.
63. **Ikebe, T., S. Iyoda, and K. Kutsukake.** 1999. Promoter analysis of the class 2 flagellar operons of *Salmonella*. Genes Genet Syst **74:**179-83.
64. **Imada, K., T. Minamino, M. Kinoshita, Y. Furukawa, and K. Namba.** 2010. Structural insight into the regulatory mechanisms of interactions of the flagellar type III chaperone FliT with its binding partners. Proc Natl Acad Sci U S A **107:**8812-7.
65. **Ind, A. C., S. L. Porter, M. T. Brown, E. D. Byles, J. A. de Beyer, S. A. Godfrey, and J. P. Armitage.** 2009. Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*. Appl Environ Microbiol **75:**6613-5.
66. **Irie, Y., and M. R. Parsek.** 2008. Quorum sensing and microbial biofilms. Curr Top Microbiol Immunol **322:**67-84.
67. **Iyoda, S., and K. Kutsukake.** 1995. Molecular dissection of the flagellum-specific anti-sigma factor, FlgM, of *Salmonella typhimurium*. Mol Gen Genet **249:**417-24.
68. **Jarrell, K. F., and M. J. McBride.** 2008. The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol **6:**466-76.
69. **Jensen, R. B., and K. Gerdes.** 1999. Mechanism of DNA segregation in prokaryotes: ParM partitioning protein of plasmid R1 co-localizes with its replicon during the cell cycle. EMBO J **18:**4076-84.
70. **Jiang, R. M., D. Crookes, N. Luo, and M. W. Davidson.** 2010. Live-cell tracking using SIFT features in DIC microscopic videos. IEEE Trans Biomed Eng **57:**2219-28.
71. **Jorgensen, B. B.** 2010. Big sulfur bacteria. ISME J **4:**1083-4.
72. **Jyot, J., N. Dasgupta, and R. Ramphal.** 2002. FleQ, the major flagellar gene regulator in *Pseudomonas aeruginosa*, binds to enhancer sites located either upstream or atypically downstream of the RpoN binding site. J Bacteriol **184:**5251-60.
73. **Kagawa, H., S. I. Aizawa, S. Yamaguchi, and J. I. Ishizu.** 1979. Isolation and characterization of bacterial flagellar hook proteins from *salmonellae* and *Escherichia coli*. J Bacteriol **138:**235-40.
74. **Karatan, E., and P. Watnick.** 2009. Signals, regulatory networks, and materials that build and break bacterial biofilms. Microbiol Mol Biol Rev **73:**310-47.
75. **Keenan, R. J., D. M. Freymann, R. M. Stroud, and P. Walter.** 2001. The signal recognition particle. Annu Rev Biochem **70:**755-75.
76. **Kentner, D., and V. Sourjik.** 2009. Dynamic map of protein interactions in the *Escherichia coli* chemotaxis pathway. Mol Syst Biol **5:**238.
77. **Khan, S., M. Dapice, and T. S. Reese.** 1988. Effects of mot gene expression on the structure of the flagellar motor. J Mol Biol **202:**575-84.
78. **Kiley, P. J., and S. Kaplan.** 1988. Molecular genetics of photosynthetic membrane biosynthesis in *Rhodobacter sphaeroides*. Microbiol Rev **52:**50-69.
79. **Kutsukake, K., T. Ikebe, and S. Yamamoto.** 1999. Two novel regulatory genes, *fliT* and *fliZ*, in the flagellar regulon of *Salmonella*. Genes Genet Syst **74:**287-92.
80. **Larsen, R. A., C. Cusumano, A. Fujioka, G. Lim-Fong, P. Patterson, and J. Pogliano.** 2007. Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. Genes Dev **21:**1340-52.
81. **Leake, M. C., J. H. Chandler, G. H. Wadhams, F. Bai, R. M. Berry, and J. P. Armitage.** 2006. Stoichiometry and turnover in single, functioning membrane protein complexes. Nature **443:**355-8.
82. **Lee, L. K., M. A. Ginsburg, C. Crovace, M. Donohoe, and D. Stock.** 2010. Structure of the torque ring of the flagellar motor and the molecular basis for rotational switching. Nature **466:**996-1000.

83. **Lemon, K. P., A. M. Earl, H. C. Vlamakis, C. Aguilar, and R. Kolter.** 2008. Biofilm development with an emphasis on *Bacillus subtilis*. Curr Top Microbiol Immunol **322:**1-16.
84. **Lim, S. K., S. J. Kim, S. H. Cha, Y. K. Oh, H. J. Rhee, M. S. Kim, and J. K. Lee.** 2009. Complete genome sequence of *Rhodobacter sphaeroides* KD131. J Bacteriol **191:**1118-9.
85. **Lindner, A. B., R. Madden, A. Demarez, E. J. Stewart, and F. Taddei.** 2008. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. Proc Natl Acad Sci U S A **105:**3076-81.
86. **Lintner, R. E., P. K. Mishra, P. Srivastava, B. M. Martinez-Vaz, A. B. Khodursky, and R. M. Blumenthal.** 2008. Limited functional conservation of a global regulator among related bacterial genera: Lrp in *Escherichia, Proteus and Vibrio*. BMC Microbiol **8:**60.
87. **Liu, J., T. Lin, D. J. Botkin, E. McCrum, H. Winkler, and S. J. Norris.** 2009. Intact flagellar motor of *Borrelia burgdorferi* revealed by cryo-electron tomography: evidence for stator ring curvature and rotor/C-ring assembly flexion. J Bacteriol **191:**5026-36.
88. **Liu, X., and P. Matsumura.** 1994. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. J Bacteriol **176:**7345-51.
89. **Liu, Z., T. Miyashiro, A. Tsou, A. Hsiao, M. Goulian, and J. Zhu.** 2008. Mucosal penetration primes *Vibrio cholerae* for host colonization by repressing quorum sensing. Proc Natl Acad Sci U S A **105:**9769-74.
90. **Locsei, J. T., and T. J. Pedley.** 2009. Run and tumble chemotaxis in a shear flow: the effect of temporal comparisons, persistence, rotational diffusion, and cell shape. Bull Math Biol **71:**1089-116.
91. **Lutkenhaus, J.** 2007. Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring. Annu Rev Biochem **76:**539-62.
92. **Mackenzie, C., M. Choudhary, F. W. Larimer, P. F. Predki, S. Stilwagen, J. P. Armitage, R. D. Barber, T. J. Donohue, J. P. Hosler, J. E. Newman, J. P. Shapleigh, R. E. Sockett, J. Zeilstra-Ryalls, and S. Kaplan.** 2001. The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. Photosynth Res **70:**19-41.
93. **Macnab, R. M.** 1992. Genetics and biogenesis of bacterial flagella. Annu Rev Genet **26:**131-58.
94. **Macnab, R. M.** 2003. How bacteria assemble flagella. Annu Rev Microbiol **57:**77-100.
95. **Maki-Yonekura, S., K. Yonekura, and K. Namba.** 2010. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. Nat Struct Mol Biol **17:**417-22.
96. **Maki, N., J. E. Gestwicki, E. M. Lake, L. L. Kiessling, and J. Adler.** 2000. Motility and chemotaxis of filamentous cells of *Escherichia coli*. J Bacteriol **182:**4337-42.
97. **Maquat, L. E., and W. S. Reznikoff.** 1978. In vitro analysis of the *Escherichia coli* RNA polymerase interaction with wild-type and mutant lactose promoters. J Mol Biol **125:**467-90.
98. **Martin, A. C., M. Gould, E. Byles, M. A. Roberts, and J. P. Armitage.** 2006. Two chemosensory operons of *Rhodobacter sphaeroides* are regulated independently by sigma 28 and sigma 54. J Bacteriol **188:**7932-40.
99. **Martin, A. C., G. H. Wadhams, and J. P. Armitage.** 2001. The roles of the multiple CheW and CheA homologues in chemotaxis and in chemoreceptor localization in *Rhodobacter sphaeroides*. Mol Microbiol **40:**1261-72.
100. **McCarter, L. L.** 2006. Regulation of flagella. Curr Opin Microbiol **9:**180-6.
101. **McClure, W. R.** 1985. Mechanism and control of transcription initiation in prokaryotes. Annu Rev Biochem **54:**171-204.
102. **Merritt, P. M., T. Danhorn, and C. Fuqua.** 2007. Motility and chemotaxis in *Agrobacterium tumefaciens* surface attachment and biofilm formation. J Bacteriol **189:**8005-14.
103. **Min, K. T., C. M. Hilditch, B. Diederich, J. Errington, and M. D. Yudkin.** 1993. Sigma F, the first compartment-specific transcription factor of B. subtilis, is regulated by an anti-sigma factor that is also a protein kinase. Cell **74:**735-42.

104. **Minamino, T., Y. Saijo-Hamano, Y. Furukawa, B. Gonzalez-Pedrajo, R. M. Macnab, and K. Namba.** 2004. Domain organization and function of *Salmonella* FliK, a flagellar hook-length control protein. J Mol Biol **341:**491-502.
105. **Mitchell, J. G., M. Martinez-Alonso, J. Lalucat, I. Esteve, and S. Brown.** 1991. Velocity changes, long runs, and reversals in the *Chromatium minus* swimming response. J Bacteriol **173:**997-1003.
106. **Moorthy, S., and P. I. Watnick.** 2005. Identification of novel stage-specific genetic requirements through whole genome transcription profiling of *Vibrio cholerae* biofilm development. Mol Microbiol **57:**1623-35.
107. **Motaleb, M. A., M. S. Sal, and N. W. Charon.** 2004. The decrease in FlaA observed in a flaB mutant of *Borrelia burgdorferi* occurs posttranscriptionally. J Bacteriol **186:**3703-11.
108. **Mueller, C. A., P. Broz, and G. R. Cornelis.** 2008. The type III secretion system tip complex and translocon. Mol Microbiol **68:**1085-95.
109. **Murray, H., H. Ferreira, and J. Errington.** 2006. The bacterial chromosome segregation protein Spo0J spreads along DNA from parS nucleation sites. Mol Microbiol **61:**1352-61.
110. **Nelson, S. S., P. P. Glocka, S. Agarwal, D. P. Grimm, and M. J. McBride.** 2007. *Flavobacterium johnsoniae* SprA is a cell surface protein involved in gliding motility. J Bacteriol **189:**7145-50.
111. **Ninfa, A. J.** 2010. Use of two-component signal transduction systems in the construction of synthetic genetic networks. Curr Opin Microbiol **13:**240-5.
112. **Packer, H. L., and J. P. Armitage.** 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:**6041-5.
113. **Paul, K., and D. F. Blair.** 2006. Organization of FliN subunits in the flagellar motor of *Escherichia coli*. J Bacteriol **188:**2502-11.
114. **Paulick, A., A. Koerdt, J. Lassak, S. Huntley, I. Wilms, F. Narberhaus, and K. M. Thormann.** 2009. Two different stator systems drive a single polar flagellum in *Shewanella oneidensis* MR-1. Molecular Microbiology **71:**836-850.
115. **Pearson, M. M., D. A. Rasko, S. N. Smith, and H. L. Mobley.** 2010. Transcriptome of swarming *Proteus mirabilis*. Infect Immun **78:**2834-45.
116. **Pena-Sanchez, J., S. Poggio, U. Flores-Perez, A. Osorio, C. Domenzain, G. Dreyfus, and L. Camarena.** 2009. Identification of the binding site of the {sigma}54 hetero-oligomeric FleQ/FleT activator in the flagellar promoters of *Rhodobacter sphaeroides*. Microbiology **155:**1669-79.
117. **Penfold, R. J., and J. M. Pemberton.** 1992. An improved suicide vector for construction of chromosomal insertion mutations in bacteria. Gene **118:**145-6.
118. **Pesavento, C., G. Becker, N. Sommerfeldt, A. Possling, N. Tschowri, A. Mehlis, and R. Hengge.** 2008. Inverse regulatory coordination of motility and curli-mediated adhesion in *Escherichia coli*. Genes Dev **22:**2434-46.
119. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu Rev Biochem **55:**339-72.
120. **Poggio, S., C. Abreu-Goodger, S. Fabela, A. Osorio, G. Dreyfus, P. Vinuesa, and L. Camarena.** 2007. A complete set of flagellar genes acquired by horizontal transfer coexists with the endogenous flagellar system in *Rhodobacter sphaeroides*. J Bacteriol **189:**3208-16.
121. **Poggio, S., A. Osorio, G. Corkidi, G. Dreyfus, and L. Camarena.** 2001. The N terminus of FliM is essential to promote flagellar rotation in *Rhodobacter sphaeroides*. J Bacteriol **183:**3142-8.
122. **Poggio, S., A. Osorio, G. Dreyfus, and L. Camarena.** 2005. The flagellar hierarchy of *Rhodobacter sphaeroides* is controlled by the concerted action of two enhancer-binding proteins. Mol Microbiol **58:**969-83.
123. **Porter, S. L., M. A. Roberts, C. S. Manning, and J. P. Armitage.** 2008. A bifunctional kinase-phosphatase in bacterial chemotaxis. Proc Natl Acad Sci U S A **105:**18531-6.

124. **Porter, S. L., G. H. Wadhams, and J. P. Armitage.** 2008. *Rhodobacter sphaeroides*: complexity in chemotactic signalling. Trends Microbiol **16:**251-60.
125. **Porter, S. L., A. V. Warren, A. C. Martin, and J. P. Armitage.** 2002. The third chemotaxis locus of *Rhodobacter sphaeroides* is essential for chemotaxis. Mol Microbiol **46:**1081-94.
126. **Prouty, M. G., N. E. Correa, and K. E. Klose.** 2001. The novel sigma54- and sigma28-dependent flagellar gene transcription hierarchy of *Vibrio cholerae*. Mol Microbiol **39:**1595-609.
127. **Rao, C. V., J. R. Kirby, and A. P. Arkin.** 2004. Design and diversity in bacterial chemotaxis: a comparative study in *Escherichia coli* and *Bacillus subtilis*. PLoS Biol **2:**E49.
128. **Rappas, M., D. Bose, and X. Zhang.** 2007. Bacterial enhancer-binding proteins: unlocking sigma54-dependent gene transcription. Curr Opin Struct Biol **17:**110-6.
129. **Raskin, D. M., and P. A. de Boer.** 1999. MinDE-dependent pole-to-pole oscillation of division inhibitor MinC in *Escherichia coli*. J Bacteriol **181:**6419-24.
130. **Reid, S. W., M. C. Leake, J. H. Chandler, C. J. Lo, J. P. Armitage, and R. M. Berry.** 2006. The maximum number of torque-generating units in the flagellar motor of *Escherichia coli* is at least 11. Proc Natl Acad Sci U S A **103:**8066-71.
131. **Ringgaard, S., J. van Zon, M. Howard, and K. Gerdes.** 2009. Movement and equipositioning of plasmids by ParA filament disassembly. Proc Natl Acad Sci U S A **106:**19369-74.
132. **Ritchings, B. W., E. C. Almira, S. Lory, and R. Ramphal.** 1995. Cloning and phenotypic characterization of fleS and fleR, new response regulators of *Pseudomonas aeruginosa* which regulate motility and adhesion to mucin. Infect Immun **63:**4868-76.
133. **Rosu, V., and K. T. Hughes.** 2006. sigma28-dependent transcription in *Salmonella enterica* is independent of flagellar shearing. J Bacteriol **188:**5196-203.
134. **Saini, S., J. D. Brown, P. D. Aldridge, and C. V. Rao.** 2008. FliZ Is a posttranslational activator of FlhD4C2-dependent flagellar gene expression. J Bacteriol **190:**4979-88.
135. **Samatey, F. A., H. Matsunami, K. Imada, S. Nagashima, T. R. Shaikh, D. R. Thomas, J. Z. Chen, D. J. Derosier, A. Kitao, and K. Namba.** 2004. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. Nature **431:**1062-8.
136. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
137. **Schulmeister, S., M. Ruttorf, S. Thiem, D. Kentner, D. Lebiedz, and V. Sourjik.** 2008. Protein exchange dynamics at chemoreceptor clusters in *Escherichia coli*. Proc Natl Acad Sci U S A **105:**6403-8.
138. **Shapiro, L., H. H. McAdams, and R. Losick.** 2009. Why and how bacteria localize proteins. Science **326:**1225-8.
139. **Shimizu, T. S., N. Le Novere, M. D. Levin, A. J. Beavil, B. J. Sutton, and D. Bray.** 2000. Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis. Nat Cell Biol **2:**792-6.
140. **Shiomi, D., M. Yoshimoto, M. Homma, and I. Kawagishi.** 2006. Helical distribution of the bacterial chemoreceptor via colocalization with the Sec protein translocation machinery. Mol Microbiol **60:**894-906.
141. **Slovak, P. M., G. H. Wadhams, and J. P. Armitage.** 2005. Localization of MreB in *Rhodobacter sphaeroides* under conditions causing changes in cell shape and membrane structure. J Bacteriol **187:**54-64.
142. **Sockett, R. E., and J. P. Armitage.** 1991. Isolation, characterization, and complementation of a paralyzed flagellar mutant of *Rhodobacter sphaeroides* WS8. J Bacteriol **173:**2786-90.
143. **Sogaard-Andersen, L.** 2004. Cell polarity, intercellular signalling and morphogenetic cell movements in *Myxococcus xanthus*. Curr Opin Microbiol **7:**587-93.
144. **Sorenson, M. K., S. S. Ray, and S. A. Darst.** 2004. Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. Mol Cell **14:**127-38.

145. **Sourjik, V., and J. P. Armitage.** 2010. Spatial organization in bacterial chemotaxis. EMBO J **29:**2724-33.
146. **Taylor, B. L.** 2007. Aer on the inside looking out: paradigm for a PAS-HAMP role in sensing oxygen, redox and energy. Mol Microbiol **65:**1415-24.
147. **Thiem, S., D. Kentner, and V. Sourjik.** 2007. Positioning of chemosensory clusters in *E. coli* and its relation to cell division. EMBO J **26:**1615-23.
148. **Thiem, S., and V. Sourjik.** 2008. Stochastic assembly of chemoreceptor clusters in *Escherichia coli*. Mol Microbiol **68:**1228-36.
149. **Thomas, D. R., N. R. Francis, C. Xu, and D. J. DeRosier.** 2006. The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of *Salmonella enterica serovar Typhimurium*. J Bacteriol **188:**7039-48.
150. **Thompson, S. R., G. H. Wadhams, and J. P. Armitage.** 2006. The positioning of cytoplasmic protein clusters in bacteria. Proc Natl Acad Sci U S A **103:**8209-14.
151. **Trachtenberg, S., L. M. Dorward, V. V. Speransky, H. Jaffe, S. B. Andrews, and R. D. Leapman.** 2008. Structure of the cytoskeleton of Spiroplasma melliferum BC3 and its interactions with the cell membrane. J Mol Biol **378:**778-89.
152. **Vande Broek, A., M. Lambrecht, and J. Vanderleyden.** 1998. Bacterial chemotactic motility is important for the initiation of wheat root colonization by *Azospirillum brasilense*. Microbiology **144 (Pt 9):**2599-606.
153. **Wadhams, G. H., and J. P. Armitage.** 2004. Making sense of it all: bacterial chemotaxis. Nat Rev Mol Cell Biol **5:**1024-37.
154. **Wadhams, G. H., A. C. Martin, and J. P. Armitage.** 2000. Identification and localization of a methyl-accepting chemotaxis protein in *Rhodobacter sphaeroides*. Mol Microbiol **36:**1222-33.
155. **Wadhams, G. H., A. C. Martin, A. V. Warren, and J. P. Armitage.** 2005. Requirements for chemotaxis protein localization in *Rhodobacter sphaeroides*. Mol Microbiol **58:**895-902.
156. **Wadhams, G. H., A. V. Warren, A. C. Martin, and J. P. Armitage.** 2003. Targeting of two signal transduction pathways to different regions of the bacterial cell. Mol Microbiol **50:**763-70.
157. **Wang, H., N. S. Wingreen, and R. Mukhopadhyay.** 2008. Self-organized periodicity of protein clusters in growing bacteria. Phys Rev Lett **101:**218101.
158. **Wang, S., R. T. Fleming, E. M. Westbrook, P. Matsumura, and D. B. McKay.** 2006. Structure of the *Escherichia coli* FlhDC complex, a prokaryotic heteromeric regulator of transcription. J Mol Biol **355:**798-808.
159. **Werner, J. N., E. Y. Chen, J. M. Guberman, A. R. Zippilli, J. J. Irgon, and Z. Gitai.** 2009. Quantitative genome-scale analysis of protein localization in an asymmetric bacterium. Proc Natl Acad Sci U S A **106:**7858-63.
160. **West, M. A., and G. Dreyfus.** 1997. Isolation and ultrastructural study of the flagellar basal body complex from *Rhodobacter sphaeroides* WS8 (wild type) and a polyhook mutant PG. Biochem Biophys Res Commun **238:**733-7.
161. **Williams, S. M., Y. T. Chen, T. M. Andermann, J. E. Carter, D. J. McGee, and K. M. Ottemann.** 2007. *Helicobacter pylori* chemotaxis modulates inflammation and bacterium-gastric epithelium interactions in infected mice. Infect Immun **75:**3747-57.
162. **Winkler, J., A. Seybert, L. Konig, S. Pruggnaller, U. Haselmann, V. Sourjik, M. Weiss, A. S. Frangakis, A. Mogk, and B. Bukau.** 2010. Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. EMBO J **29:**910-23.
163. **Wolfe, A. J., D. S. Millikan, J. M. Campbell, and K. L. Visick.** 2004. *Vibrio fischeri* sigma54 controls motility, biofilm formation, luminescence, and colonization. Appl Environ Microbiol **70:**2520-4.

164. **Yamaguchi, S., H. Fujita, A. Ishihara, S. Aizawa, and R. M. Macnab.** 1986. Subdivision of flagellar genes of *Salmonella typhimurium* into regions responsible for assembly, rotation, and switching. J Bacteriol **166:**187-93.
165. **Yonekura, K., S. Maki-Yonekura, and K. Namba.** 2001. Structure analysis of the flagellar cap-filament complex by electron cryomicroscopy and single-particle image analysis. J Struct Biol **133:**246-53.
166. **Young, K. D.** 2010. Bacterial shape: two-dimensional questions and possibilities. Annu Rev Microbiol **64:**223-40.

11) Publications

Scott, K.A., Porter, S.L., Bagg, E.A., Hamer, R, Hill, J.L., **Wilkinson, D.A.,** Armitage, J.P. **2010**. Specificity of localization and phosphotransfer in the CheA proteins of *Rhodobacter sphaeroides*. Mol Microbiol. Apr;**76**(2):318-30.

Wilkinson, D.A., Chacko, S.J., Venian-Bryan, C., Wadhams, G.A., Armitage, J.P. **2010**. FlgM and FliA: roles in motility, chemotaxis and biofilm formation in *Rhodobacter sphaeroides*. **In preparation**.

Roberts, M.A., **Wilkinson, D.A.**, Armitage, J.P., **2011**. The localisation of chemotaxis proteins in *Rhodobacter sphaeroides*. **In preparation**.

Wood, T., Wilkinson, D.A., Yates, C., Wadhams, G.A., Armitage, J.P., **2011**. Tracking free-swimming *Rhodobacter sphaeroides* using a deterministic algorithm. **In preparation**.