University of Oxford and St Peter's College
Department of Biochemistry
Systems Biology Doctoral Training Centre

Complexity in *Rhodobacter sphaeroides* chemotaxis

Andrea Szöllössi

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Trinity Term 2015

Pentru Bunica mea dragă
To my beloved Grandmother

Acknowledgements

I am grateful to my supervisor, Professor Judith P. Armitage, for the opportunity to work in her laboratory. The trust, guidance, and support received from her and the lab as a whole have made my time as a DPhil student one I will always fondly remember.

I would like to thank the EPSRC-funded Systems Biology Doctoral Training Centre for offering me a place on a programme that allowed me to develop into a well-rounded scientist.

I am especially grateful to Dr. Kathryn Scott, Professor Elspeth Garman, Dr. Jennifer A. De Beyer, Elaine Byles, Dr. Christopher W. Jones, and Dr. George H. Wadhams, for mentoring, proofreading, and being generally helpful in matters of the lab and beyond.

I will always be grateful to St Peter's College, my home away from home. Thank you for providing the warmth one needs after having left home.

And last, however certainly not least, I would like to thank my family and Richard for their continued support, particularly during the last part of my DPhil. Thank you for the unshakable faith you have always had in me. Thank you for the words of encouragement, the muffins, and for making sure I made it in time to submit my thesis.

Declaration

The work described in this thesis was undertaken in the Department of Biochemistry at the University of Oxford. Work was performed between October 2012 and October 2015 under the supervision of Professor Judith P. Armitage. All the work in this thesis is my own unless otherwise stated and has not been submitted for a degree at this or any other university.

Abstract

Complexity in *Rhodobacter sphaeroides* chemotaxis

ANDREA SZÖLLÖSSI

St Peter's College, University of Oxford
Submitted for the degree of Doctor of Philosophy, Trinity Term 2015

Perceiving and responding to the environment is key to survival. Using the prokaryotic equivalent of a nervous system – the chemotaxis system – bacteria sense chemical stimuli and respond by adjusting their movement accordingly.

In chemotactic bacteria, such as the well-studied *E. coli*, environmental nutrient sensing is achieved through a membrane embedded protein array that specifically clusters at the cell poles. Signalling to the motor is performed by activation of the CheA kinase, which phosphorylates CheY and CheB. CheY–P tunes the activity of the flagellar motor while CheB–P, together with CheR is involved in adaptation to the stimulus. In *E. coli*, a dedicated phosphatase terminates the signal.

Most bacterial species however, have a much more complex chemotaxis network. *Rhodobacter sphaeroides*, a model organism for complex chemotaxis systems, has one membrane-embedded chemosensory array and one cytoplasmic chemosensory array, plus several homologs of the *E. coli* chemotaxis proteins. Signals from both arrays are integrated to control the rotation of a single start-stop flagellar motor.

The phosphorelay network has been studied extensively through *in vitro* phosphotransfer while *in vivo* studies have established the components of each array and the requirements for formation. Mathematical modelling has also contributed towards inferring connectivities within the signalling network.

Starting by constructing a two-hybrid-based interaction network focused on the components of the cytoplasmic chemosensory array, this thesis further addresses its associated adaptation network through a series of *in vivo* techniques.

The swimming behaviour of series of deletion mutants involving the adaptation network of *R. sphaeroides* is characterised under steady state conditions as well as upon chemotactic stimulation. New connectivities within the *R. sphaeroides* chemotaxis network are inferred from analysing these data together with results from *in vivo* photoactivation localisation microscopy of $CheB_2$.

The experimental results are used to propose a new model for chemotaxis in *R. sphaeroides*.

Abbreviations

–P	Phosphorylated
ATP	Adenosine triphosphate
bp	Base pair(s)
BACTH	Bacterial adenylate cyclase two-hybrids
CCD	Charge-coupled device
CCW	Counter-clockwise
Che	Chemotaxis
*che*Op	Chemotaxis operon
CW	Clockwise
D^*	Apparent diffusion coefficient
DIC	Differential interference contrast
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside-5'-triphosphate
EDTA	Ethylene diamine tetra acetic acid
EM	Electron microscopy
HAMP	Histidine kinase, adenyl cyclase, methyl binding protein and phosphatase domain
HCD	Highly conserved domain
HEPES	Sodium N-2-hydroxyethylpiperazine-N-2-ethanesulphuric acid
HMM	Hidden Markov model
IPTG	Isopropyl β-D-thiogalactoside
Kan	Kanamycin
kb	Kilo base
kDa	Kilo Dalton
lac	Lactose
LB	Luria-Bertani
LBA	Luria-Bertani agar
MAC	Mean angle change
MBR	Minimum bounding radius
MCP	Methyl-accepting chemotaxis protein
MilliQ	Ultrapure water

MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSD	Mean squared displacement
Nal	Nalidixic acid
NEMS	Normalised effective mean speed
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
OD_{600}	Optical density at 600nm
OD_{700}	Optical density at 700nm
PALM	Photoactivatable localisation microscopy
PAmCherry	Photo activatable mCherry
PBS	phosphate buffered saline
PCR	Polymerase chain reaction
PMF	Proton motive force
PSF	Point spread function
PVDF	Polyvinylidene fluoride
RNA	Ribonucleic acid
SAM	S-adenosyl-L-methionine
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
Sux	Succinate medium
T_m	Melting temperature
TBE	Tris-borate-EDTA buffer
TFB	Transformation buffer
Tris	2-amino-2(hydroxymethyl)-1,3-propandiol
UV	Ultraviolet
WT	Wild-type
YFP	Yellow fluorescent protein

Standard one letter abbreviations used for amino acids.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Professor Howard C. Berg famously refers to himself as a neurobiologist, one who has "had the courage to deal only with the simplest single-celled nervous systems". He is of course referring to his work on bacterial chemotaxis and he is also not the first scientist to make such a claim. Regarding bacterial chemotaxis systems as the prokaryotic equivalent of a nervous system is a parallel first drawn by Professor Julius Adler more than 30 years ago [2]. This may seem a bold claim to many – particularly neuroscientists. But perhaps the claims of Adler and Berg are not as far fetched considering that the emergence of chemotaxis systems in bacteria marks the fundamental transition from passive sensing to active exploration of the environment, and chemotaxis systems are also some of the most complex signal transduction networks in prokaryotes with respect to component design.

Thus without attempting to claim a doctorate in neuroscicence, the present thesis takes an interdisciplinary approach to offer answers to outstanding questions regarding *Rhodobacter sphaeroides* chemotaxis signal transduction.

1.1 Motility and taxis

Bacteria have developed various strategies for swimming through liquid or moving on solid surfaces *e.g. Flavobacterium johnsoniaeglides* moves rapidly over surfaces while *Myxococcus xanthus* has two differnt mechanisms for surface mobility. Other organisms like *Pseudomonas aeruginosa*, can perform both surface as well as aqueous movement [46]. Movement on surfaces is often mediated by type IV pili, while movement through liquid environments is predominantly mediated by the rotation of helical filaments extruded from the cell called flagella [46].

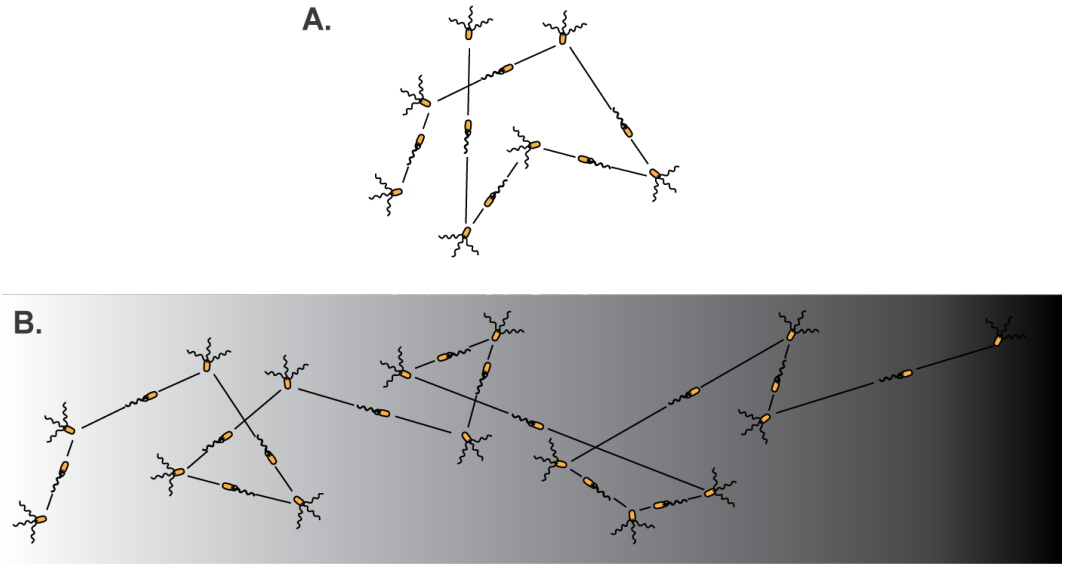

Figure 1.1: Bacterial biased random walk. A. In the absence of a gradient, the bacterium performs a random walk. B. When a gradient of attractant is present, there is a directed movement up the gradient.

A great proportion of bacteria have the evolutionary advantage of directed movement towards conditions that are more favourable for growth, a process known as taxis. Such parameters include nutrient concentration in their environment (*chemo*taxis), light (*photo*taxis), and optimum temperature (*thermo*taxis). However bacteria living in diverse environments, that is outside the conical culture flask carefully placed in the incubator, will move towards environments that are optimal for metabolic ac-

tivity. This is a process called energy taxis which uses the chemotaxis pathway to integrate signals from a series of chemosensors and metabolic sensors in order to find the optimum conditions.

Tracking bacterial movement reveals a characteristically erratic trajectory consisting of a series of runs and tumbles, showing that bacteria are unable to move directly towards a nutrient source, for example. Because of their small size and the time scales involved, it is generally accepted that bacteria cannot sense gradients along their cell body. The strategy they use instead is comparing current environmental conditions with the ones experienced just before. If environmental conditions become less favourable, the bacterium will randomly change direction. However, if conditions are improving, it will continue swimming in the same direction for longer. This movement pattern is knows as a *biased random walk* and causes the bacterium to move, on average, towards favourable environmental conditions.

1.2 Bacterial chemotaxis

The emergence of the chemotaxis system marks a fundamental transition in signal transduction: from passive sensing to active exploration of the environment.

The bacterial chemotaxis system does not have equivalent systems present in eukaryotes, hinting that eukaryotes and prokaryotes have developed navigation systems that are similar in basic principles (use of dedicated receptors and protein kinases), but very different in therms of component design *e.g.* MCPs vs. G-protein coupled receptors.

1.2.1 An evolutionary perspective

The simplest transduction system consists of a single protein which does both the sensing and directly effects a cellular response *e.g.* ligand binding transcription regulator. The sensory and regulatory domains are in the same protein. This is the one-component system. A two-component system has the sensor (histidine kinase) separate from the response regulator. One-component systems mostly detect signals in the cytoplasm while most two-component systems detect signals in the extracellular environment. The chemotaxis system is a special case of two-component signal transduction.

Some of the components of the chemotaxis system present similarities to regulatory proteins, but most elements have not been found in any other type of signal transduction system. The overall design of the chemotaxis system is substantially more complex than that of any other prokaryotic signal transduction system.

Two-component systems probably originated in Bacteria and were laterally transferred to Archaea [117]. The same pattern of evolution was proposed for chemotaxis systems and confirmed by Briegel *et al.* in 2015 [20]. It is also proposed that the chemotaxis system originated gradually from two-component systems through simple, incremental innovations such as domain acquisition and protein recruitment.

The basic chemotaxis system includes seven different types of proteins, making it the most complex signal transduction network in prokaryotes with respect to component design.

As pointed out by Wuichet and Zhulin, the choice of *E. coli* as a model for chemotaxis [1] was unknowingly excellent: the presence of all central components and the lack of most auxiliary components subsequently enabled the detailed molecular

mechanisms of chemotaxis to be determined.

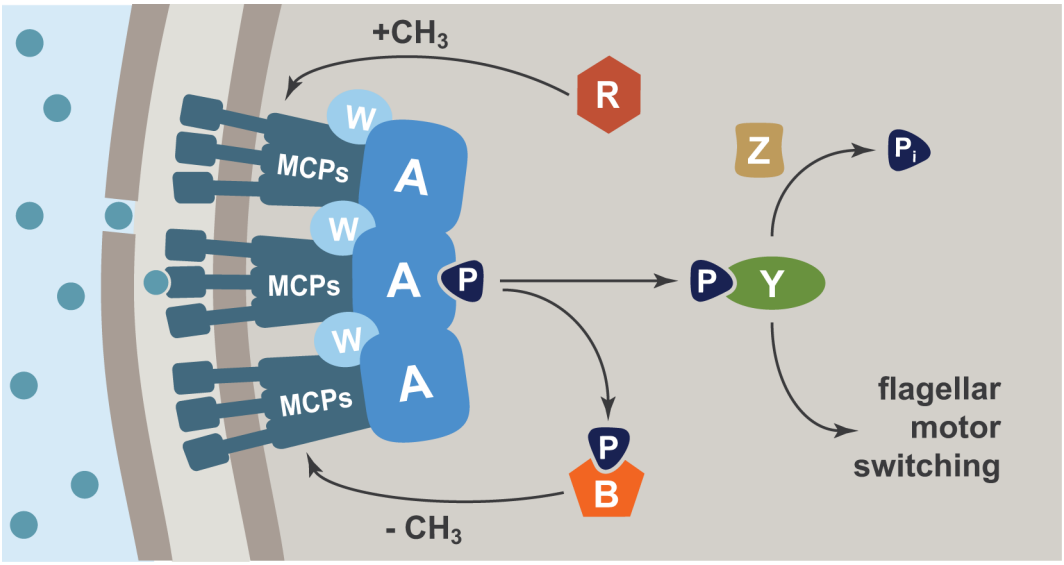

Figure 1.2: The model for chemotaxis in *E. coli*. Decreased attractant concentrations are sensed by the methyl-accepting chemotaxis proteins (MCPs) which initiate signalling by activating the CheA histidine kinases through conformational changes. CheA–P phosphorylates its two cognate response regulators: CheY–P interacts with the flagellar motor increasing the tumble bias, while CheB–P desensitizes the MCPs by demethylation, suppressing the signal that led to the activation of CheA. Signalling to the motor is terminated by CheY–P dephosphorylation catalysed by CheZ. CheR methylates the MCPs, increasing their sensitivity, and together with the antagonistic demethylation by CheB–P enables the MCPs to adapt to environmental conditions.

Ligand binding causes a chemoreceptor conformational change which tunes the CheA kinase activity. In response to decreasing attractant, CheA phosphorylates two response regulators, CheY, which controls flagellar motor switching, and CheB, which is involved in resetting chemoreceptor sensitivity. The signal is terminated through CheY-P dephosphorylation, which needs to happen rapidly to allow continous sensing, and is catalysed by phosphatases.

1.2.2 Sensing

The external environment is sampled by thousands of membrane-embedded chemoreceptors. They are found as trimers of receptor dimers (Figure 1.3A), which form signalling complexes with CheA, the chemotaxis histidine kinase (Figure 1.3B), and CheW, the scaffold protein. These complexes are organised into large arrays usually located at the cell poles. Allosteric interactions between array components allow for signal amplification, providing bacteria with the ability to respond to small relative changes in the environment over a broad range of background concentrations.

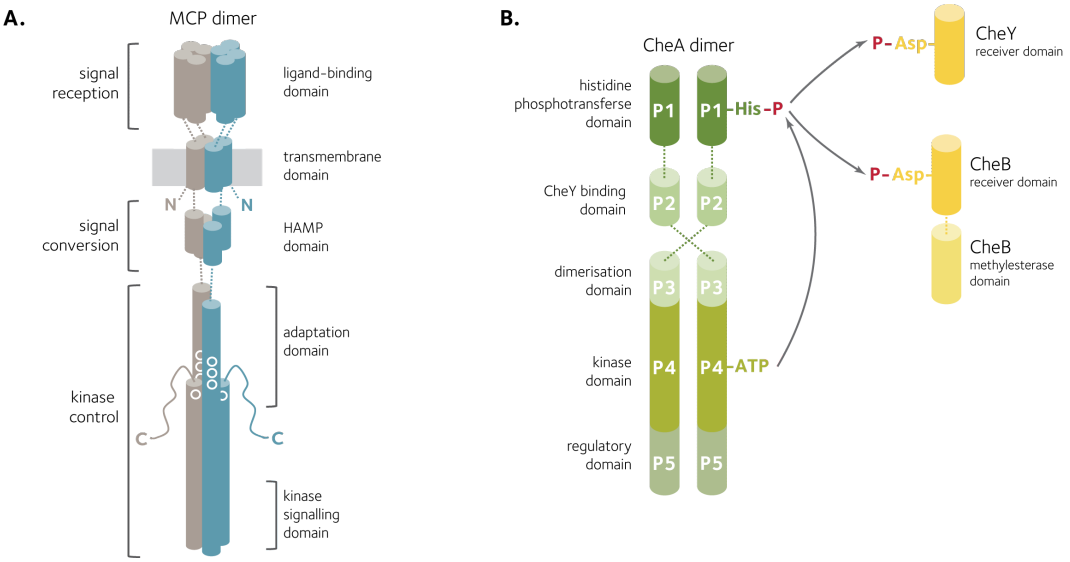

Figure 1.3: Domain structure for **A.** chemoreceptor dimers and **B.** CheA dimers.

In a cryo-EM study comprising different chemoreceptor classes and many different species representing major bacterial phyla, a highly conserved 12 nm hexagonal array was observed [21]. Chemoreceptor clustering appears universal among bacterial species and was probably conserved during evolution. Combining cryo-EM data of the array with X-ray structures of its components showed that the chemoreceptor trimers of dimers lie at the vertices of a hexagonal lattice surrounding a ring of alternating CheA P5 domains and CheW scaffold proteins [19]. These hexago-

nal rings are linked through CheA P3 (dimerisation) domains to form an extended, stable array Figure 1.4.

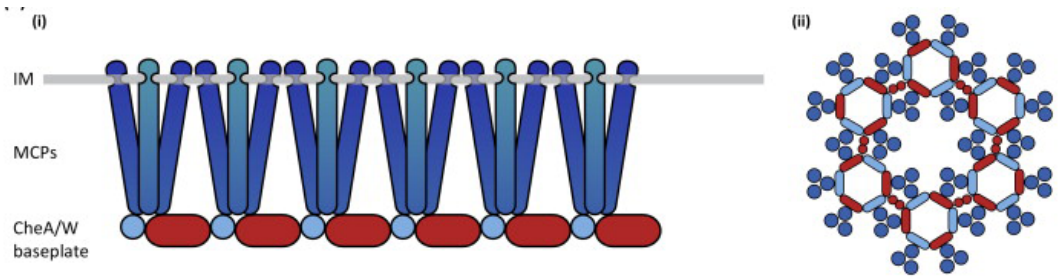

Figure 1.4: Model of the architecture of membrane chemosensory arrays. i) Side view showing trimers of MCP dimers embedded in the inner membrane (IM) at one end and connected through the CheA/W baseplate. ii) Top view showing the hexagonal rings formed by CheA/W, connected by the P3 domains. The MCP trimers of dimers surround these hexagonal rings. Colours represent MCPs (dark blue), CheW (light blue), and CheA (red). The figure was taken from [49].

Chemosensory arrays reconstructed *in vitro* retained CheA kinase activity for up to 20 days, displaying kinetic and thermodynamic ultrastability [32, 33]. This property is proposed to arise from the tight packing of the array which establishes multiple linkages between its components, also facilitating network cooperativity and increasing the sensitivity of the array [69, 97]. Ultrastability is believed to contribute to cell survival during starvation as it can remain stable for days while the cell finds a more favourable environment. Through its high sensitivity and stability, the array also provides an attractive platform for a new generation of ultrastable, ultrasensitive biosensors.

The number of types of chemoreceptors varies greatly between species: some species have a single type of chemoreceptor *e.g. Methanococcoides burtonii* and *Mesorhizobium loti* while others have as many as 65 putative chemoreceptors, as is the case for *Magnetospirillum magnetotacticum* [4]. *E. coli* has four chemoreceptors, each specialised in recognising aspartate (Tar), serine (Tsr), ribose and galactose (Trg), or dipeptides (Tap). The expression level for the various chemoreceptors can vary,

as is the case for *E. coli* Tar and Tsr with a copy number of several thousand [25], in contrast to the ten-fold lower copy number of Trp and Tap [41].

Chemoreceptors are transmembrane alpha-helical proteins. Most chemoreceptors are methyl-accepting chemotaxis proteins (MCPs) and share the overall structure described in Figure 1.3A. The periplasmic ligand-binding domain is flanked by transmembrane (TM) regions. Chemoreceptors that lack the transmembrane domain are termed transducer-like proteins (TLPs). A HAMP (histidine kinase, adenylate cyclase,methyl-accepting chemotaxis protein and phosphatase) domain connects TM2 with the cytoplasmic region, known as the highly conserved domain (HCD). The HCD is highly conserved across bacterial species and can be used to classify the MCPs using phylogeny [4]. The HCD region is further divided into two adaptation and two signalling domains. The C-terminus of the MCP often contains a conserved recognition site for CheR binding, known as the pentapeptide tail.

1.2.3 Signalling

Bacterial chemotaxis relies on on a two-component phosphotransfer pathway to integrate information from multiple chemosensors to produce an appropriate change in swimming behaviour. Two-component phosphotransfer pathways contain histidine kinase homodimers and response regulators [102]. Within each homodimer, the kinase domain transphosphorylates the receiver domain at a conserved histidine residue using ATP as the phosphodonor. The rate of autophosphorylation is controlled by sensory stimuli and only phosphorylated histidine kinases are able to transfer the phosphoryl group from their conserved histidine residue onto a conserved aspartate residue of their cognate response regulator.

Two-component protein domains are modular, meaning that there is a variety of

ways in which they can be integrated into proteins and pathways while maintaining core structures and activities.

CheA, the bacterial chemotaxis histidine kinase, is a sensorless histidine kinase and its response regulator, CheY, lacks an output domain. Despite the similarity between the CheA-CheY pair and classical two-component systems, the structural features of CheA place it in a separate class of histidine kinases, classII [15] (Figure 1.5). A typical CheA kinase, like the one in *E. coli*, has five domains:

- **P1** contains the conserved histidine residue involved in transautophosphorylation; this is a histidine phosphotransferase domain.
- **P2** interacts with the response regulator and accelerates phosphotransfer to the conserved aspartate residue; it is not essential for phosphotransfer [44, 101].
- **P3** is the dimerisation domain.
- **P4** binds ATP and phosphorylates the conserved histidine residue in the P1 domain; this is the kinase domain [105].
- **P5** interacts with the chemoreceptors and adaptor protein CheW; this is the regulatory domain [17, 53].

The rate of CheA autophosphorylation is increased by conformational changes in associated chemoreceptors as a response to a less favourable chemical environment. Its cognate response regulators are CheY proteins which tune the activity of the flagellar motor, and typically CheB proteins involved in resetting the sensitivity of the chemoreceptors (Section 1.2.5). CheY has a greater affinity for the P2 domain than CheB, ensuring that the signal is transmitted to the flagellar motor before adaptation interrupts the signalling. The time it takes for the signal to travel from the chemoreceptor cluster to the motor is approximatively 100 ms.

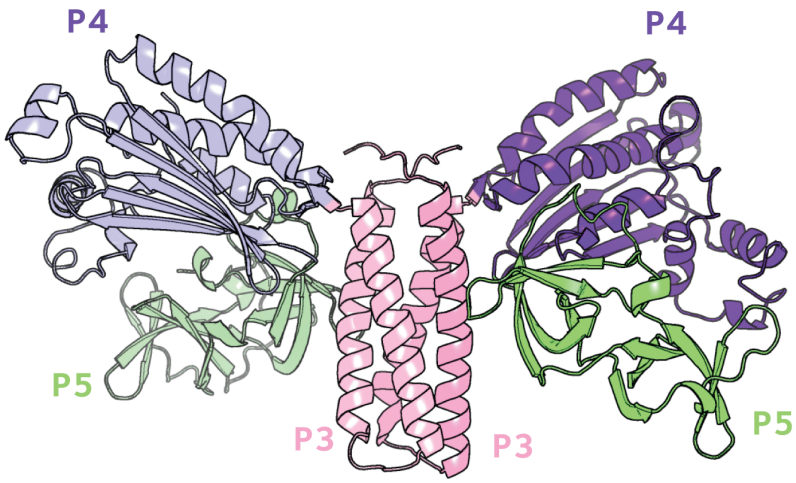

Figure 1.5: Structure of CheA dimer showing the simerisation (P3), kinase (P4), and regulatory (P5) domains. [15]

CheY is one of the CheA response regulators, getting phosphorylated at a conserved aspartate residue (D57) once it has bound to the CheA P2 domain. With a lower affinity for CheA, CheY-P dissociates and diffuses to the flagellar motor. CheY-P binding to FliM in the flagellar motor increases the probability of motor switching to CW rotation, and thus for the cell to start tumbling.

1.2.4 Signal termination

The signal initiated by the chemosensory array is terminated through the removal of the phosphoryl group from CheY-P. This is achieved either through autodephosphorylation or through the catalytic action of the CheZ phosphatase. According to the study by Wuichet and Zhulin, >20% of the genomes with chemotaxis genes do not have an identifiable CheY phosphatase [117].

E. coli responds to a nutrient concentration drop by starting to tumble within 0.2 s from the stimulus [91]. The half life of CheY–P is 14 s for autodephosphorylation, but the action of the CheZ phosphatase reduces the half life of this species 100-

fold, down to 0.14 s [96]. This faster rate of depleting CheY–P from the cytoplasm allows the motor to return to CCW rotation once the chemoreceptors switch back to the inactive state as a result of methylation-based adaptation.

1.2.5 Adaptation

To allow for chemoreceptor sensitivity over a broad range of chemoeffector concentrations, the signalling state of the MCPs needs to be reset periodically. Strategies bacteria use to achieve adaptation include methylation-based adaptation, the CheC-CheD system, and the CheY-CheV system.

Methylation-based adaptation refers to the tuning of chemoreceptor sensitivity by addition or removal of methyl groups from key glutamate resides in the HCD. Increased methylation promotes closer packing through the removal of negative charges, leading to more sensitive chemoreceptors that are more likely to start signalling through CheA. Conversely, demethylation reintroduces negative charges, disrupting the packing of the array and making signalling through CheA less probable.

All *E. coli* chemoreceptors have four methylation sites. In the case of Tsr, for example, two of the sites are expressed as glutamine residues and are post-translationally modified to glutamate, presumably during the first adaptation response. The receptors are thus expressed in what can be considered a half-methylated state (two negative and two neutral charges), which presumably allows incorporation into the signalling array in a neutral signalling conformation. Mutation of these sites to alanine leads to a subtractive loss in chemotaxis ability. It has been predicted that methylation could occur in the majority of chemoreceptors [4].

CheR is a methyltransferase which uses S-adenosyl-L-methionine (SAM) to methy-

late glutamate residues. Work on enterobacteria has shown that in addition to binding at the methylation site, some chemoreceptors bind CheR through additional high-affinity sites. A highly conserved pentapeptide tail at the chemoreceptor C-terminus is important for CheR binding in *E. coli* and *S. typhimurium*, as indicated by its deletion leading to a great reduction in chemoreceptor methylation. In other organisms, such as *T. maritima*, the removal of the pentapeptide tail did now show a reduction in chemoreceptor methylation *in vitro*. In a more recent study on *P. aeruginosa*, it was revealed that there are two families of CheR proteins – those that bind pentapeptide-containing chemoreceptors, and those that do not. The difference between these two families is a three amino acid insertion in the β-subdomain of CheR [34].

As mentioned before, CheB is a CheA response regulator. It contains an N-terminal regulatory domain that is homologous to CheY, and a C-terminal methylesterase domain that can also perform deamidation. The methylesterase activity is increased by phosphorylation of the regulatory domain. CheB is thought to also bind to the chemoreceptor C-terminal pentapeptide tether in *E. coli* [10], but it remains unclear whether this applies to other species. Methanol release is an indicator for methylation-based adaptation, as it is a side product from the CheB-catalysed methyl de-esterification [120].

The CheC-CheD system has been studied in *Bacillus subtilis* [77]. CheC is a CheY-P phosphatase similar to CheZ, while CheD is an MCP deamidase. It is proposed that CheD interacting with the MCPs enhances CheA activation upon attractant binding. CheY-P is produced, which diffuses to the flagellar motor and initiates a run, but also interacts with CheC forming a complex that recruits CheD away from the MCPs. The absence of CheD results in a decrease in CheA activity and hence a decrease in CheY-P, which in *B. subtilis* causes cell tumbling.

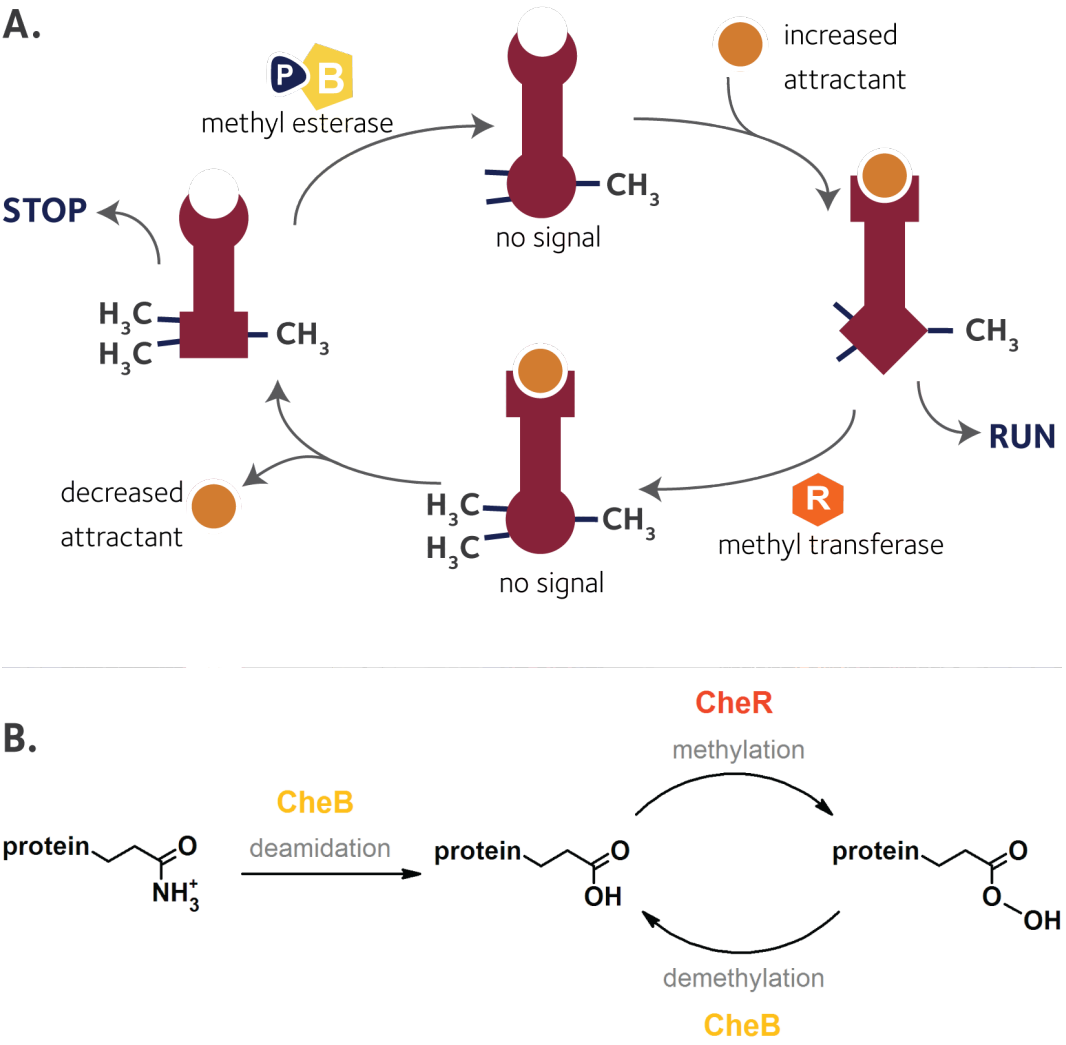

Figure 1.6: Methylation-based adaptation. A. Chemoreceptor sensitivity is tuned through methylation/demetylation; B. Post-translational modifications catalysed by CheB and CheR.

A subset of chemotactic bacteria – including some Firmicutes (*e.g. B. subtilis*), and β-, γ-, and ϵ-proteobacteria – contain CheV [3], a protein with a CheA phosphorylatable regulatory domain and a scaffolding domain homologous to CheW [84]. The role of CheV in adaptation has been studied in *B. subtilis*. It is believed that CheV provides an additional negative feedback loop. Upon CheA activation, CheV is phosphorylated, which disturbs the interaction between CheA and the MCP, causing a decrease in CheA activity [82].

1.2.6 Complexity in bacterial chemotaxis

Wuichet *et al.* have searched a nonredundant set of 450 prokaryotic genomes and found 245 (54%) containing chemotaxis genes [117]. Wuichet *et al.* found that multiple chemotaxis systems occur as frequently as single ones, which highlights the importance of studying bacterial models with multiple chemotaxis systems. *R. sphaeroides*, *Myxococcus xanthus*, and *Pseudomonas aeruginosa* have multiple sets of chemotaxis genes that constitute multiple chemotaxis systems with defined functions.

1.3 Chemotaxis in *Rhodobacter sphaeroides*

Almost 300 years after Antonie van Leeuwenhoek marvelled at what was likely *Selenomonas sp.* "very prettily a-swimming" under his microscope, the first strain of *Rhodobacter sphaeroides* was first described [111].

R. sphaeroides was initially known as *Rhodopseudomonas sphaeroides* and is an alphaproteobacterium within the order *Rhodobacterales* and the family *Rhodobacteraceae*. It is a photosynthetic purple non-sulfur bacterium preferring organic acids as carbon source during photoheterotrophic growth. *R. sphaeroides* can fix carbon dioxide during autotrophic growth and molecular nitrogen when other nitrogen sources are scarce.

Having such a diverse metabolism, it can be found in both soil and fresh water habitats, growing through fermentation or anaerobic respiration in the absence of light and oxygen, or growing as a chemoheterotroph in the absence of light and in the presence of oxygen, but preferably growing photosynthetically in the light and in the absence of oxygen. It has been used as a model system for studying

photosynthesis in bacteria as well as complex chemotaxis systems.

1.3.1 Genetic structure

A lot of the work on *R. sphaeroides* has focused on the strain identified by van Niel [111], *R. sphaeroides* 2.4.1, whose genome sequence has been available since 2001 [61].

However, the strain used in chemotaxis research is *R. sphaeroides* WS8N, a spontaneous nalidixic acid resistant variant of the naturally occuring WS8 published in 1972 [26]. *R. sphaeroides* WS8N exhibits enhanced swimming motility and chemotaxis compared to other strains of *R. sphaeroides* such as 2.4.1. For this reason, WS8N is used for the study of chemotaxis signaling [75, 80] and flagellum-based motility [70, 54]. This is also the strain which claimed the gold medal in the 100 μm swim at the 2012 Bacterial Olympics [118].

The WS8N genome sequence was published in 2011 and revealed a 4.42 Mbp genome organised into two chromosomes (3.14 and 0.97 Mbp) and two large plasmids (200 and 110 kbp), with a high GC content, 69.1% [80].

The smaller chromosome, ChrII, could be the remnant of a plasmid that has accumulated genes from the main chromosome over time. A more recent theory proposes that ChrII is an intermediate structure, termed chromid [40]. A chromid has a plasmid-type replication system, a nucleotide composition similar to that of a chromosome, and carries genes that are found in chromosomes in other species. Recombination between a chromosome and a plasmid, or an existing chromid and a plasmid could give rise to a chromid. It is also proposed that chromids are a mark of the evolution of a new genus, as they contain essential “lifestyle” genes, such as metabolism and motility genes [40].

1.3.2 Motility

R. sphaeroides is motile and encodes two sets of flagellar genes. Under laboratory conditions, only *fla1* is expressed, producing a single, randomly-positioned, unidirectional flagellum, which is typical of the family. The second set of flagellar genes, *fla2*, produces a tuft of flagella and has only been expressed in strains containing a *fla1* deletion. No laboratory conditions have been found to yield endogenous expression of *fla2*. Work presented in this thesis exclusively uses strains expressing *fla1*.

In contrast to the motility pattern encountered in *E. coli* and other flagellate species, the *R. sphaeroides* flagellar motor alternates between on and off [8]. Active forward motion is hence achieved when the flagellar motor is rotating, and reorientation occurs passively during periods when the motor is stopped and reorientation is a consequence of Brownian motion. It was found that when the motor is rotating, the flagellum adopts a helical conformation similar to the one seen in *E. coli*, but when the motor is off, the filament relaxes and becomes highly coiled hence potentially enhancing reorientation [9].

The swimming speed of *R. sphaeroides* in a homogenous environment has been measured to reach 80 μm/s. Direction change is achieved through transient stops which occur 0.31 ± 0.19 times per second. Rotation restarts after a mean stop time of 0.66 s and a median of 0.44 s. In comparison to *E. coli* which spends 0.6-0.7 of the time in runs [16], *R. sphaeroides* has a higher run bias, spending 0.80 ± 0.20 of the time in runs [22].

1.3.3 The chemotaxis operons

R. sphaeroides is chemotactic towards a range of attractants, including its preferred carbon sources succinate and propionate. The only known repellent is oxygen under anaerobic growth conditions.

The chemotaxis pathway of *R. sphaeroides* has multiple homologues of the simpler *Escherichia coli* pathway, encoded primarily in 3 operons, with the two expressed pathways localised to two discreet chemosensory clusters [90, 115] forming a complex signal integration pathway [107] that controls the single motor [12].

About 33 genes have been found that could encode chemotaxis signalling proteins (Table 1.1), all of which are found on the large chromosome, ChrI, with the exception of six of the MCP genes and $cheY_4$, which are found on ChrII. The chemotaxis genes are encoded by three operons (Figure 1.7) and multiple orphan genes. The first operon is not expressed under laboratory conditions. Expression of *che*Op$_1$ through mutational constraints has suggested that it expresses a chemotaxis pathway that functions similarly to the *E. coli* paradigm and to control the rotation of *fla2* motors [29, 30].

Both *che*Op$_2$ and *che*Op$_3$ are expressed under laboratory conditions and both are required for chemotaxis, as deleting either of them results in the loss of chemotaxis [94, 79]. The rest of this section will discuss the chemotaxis systems expressed by *che*Op$_2$ and *che*Op$_3$, with a focus on the latter.

Previous experiments showed that *che*Op$_1$ is not required for chemotaxis under laboratory conditions [37], but components of both *che*Op$_2$ and *che*Op$_3$ are essential for chemotaxis [93, 73]. The 2003 study by Wadhams *et al.* focuses on the localization of proteins encoded in *che*Op$_2$ and *che*Op$_3$ [115].

Table 1.1: Genomic and cellular location of the *R. sphaeroides* chemotaxis genes under laboratory conditions.

Gene	Genomic location	Cellular location	Reference
Histidine kinases			
$cheA_1$	ChrI *che*Op$_1$	not expressed	[38]
$cheA_2$	ChrI *che*Op$_2$	polar cluster	
$cheA_3$	ChrI *che*Op$_3$	cytoplasmic cluster	
$cheA_4$	ChrI *che*Op$_3$	cytoplasmic cluster	
Scaffold proteins			
$cheW_1$	ChrI *che*Op$_1$	not expressed	[38]
$cheW_2$	ChrI *che*Op$_2$	polar cluster	
$cheW_3$	ChrI *che*Op$_2$	polar cluster	
$cheW_4$	ChrI *che*Op$_3$	cytoplasmic cluster	
Response regulators			
$cheY_1$	ChrI *che*Op$_1$	not expressed	[38]
$cheY_2$	ChrI *che*Op$_1$	not expressed	[38]
$cheY_3$	ChrI *che*Op$_2$	partially cytoplasmic	[78]
$cheY_4$	ChrII separate locus	partially polar & cytoplasmic	[78]
$cheY_5$	ChrI *che*Op$_1$	not expressed	[38]
$cheY_6$	ChrI *che*Op$_3$	partially cytoplasmic	[78]
Methylesterases			
$cheB_1$	ChrI *che*Op$_2$	diffuse	
$cheB_2$	ChrI *che*Op$_3$	diffuse	
Methyltransferases			
$cheR_1$	ChrI *che*Op$_1$	not expressed	[38]
$cheR_2$	ChrI *che*Op$_2$	partially polar	
$cheR_3$	ChrI *che*Op$_3$	partially cytoplasmic	
Transmembrane chemoreceptors			
mcpA	ChrI *che*Op$_1$	not expressed	[38]
mcpB	ChrI *che*Op$_1$	not expressed	[38]
mcpE	ChrII separate locus	not tested	
mcpG	ChrII separate locus	polar	
mcpH	ChrII separate locus	polar	
mcpJ	ChrII separate locus	polar	
mcpM	ChrI separate locus	not tested	
mcpV	ChrII separate locus	not tested	
Soluble schemoreceptors			
tlpC	ChrI *che*Op$_2$	cytoplasmic	
tlpL	ChrI separate locus	diffuse	
tlpS	ChrI *che*Op$_1$	not expressed	[38]
tlpT	ChrI *che*Op$_3$	cytoplasmic	
Other genes			
cheD	ChrI *che*Op$_1$	not expressed	[38]

1.3.4 Cellular localisation of the chemotaxis proteins

Microscopy experiments using fluorescent protein fusions have shown that the proteins from *che*Op$_2$ and *che*Op$_3$ predominantly localise to two distinct foci: one at

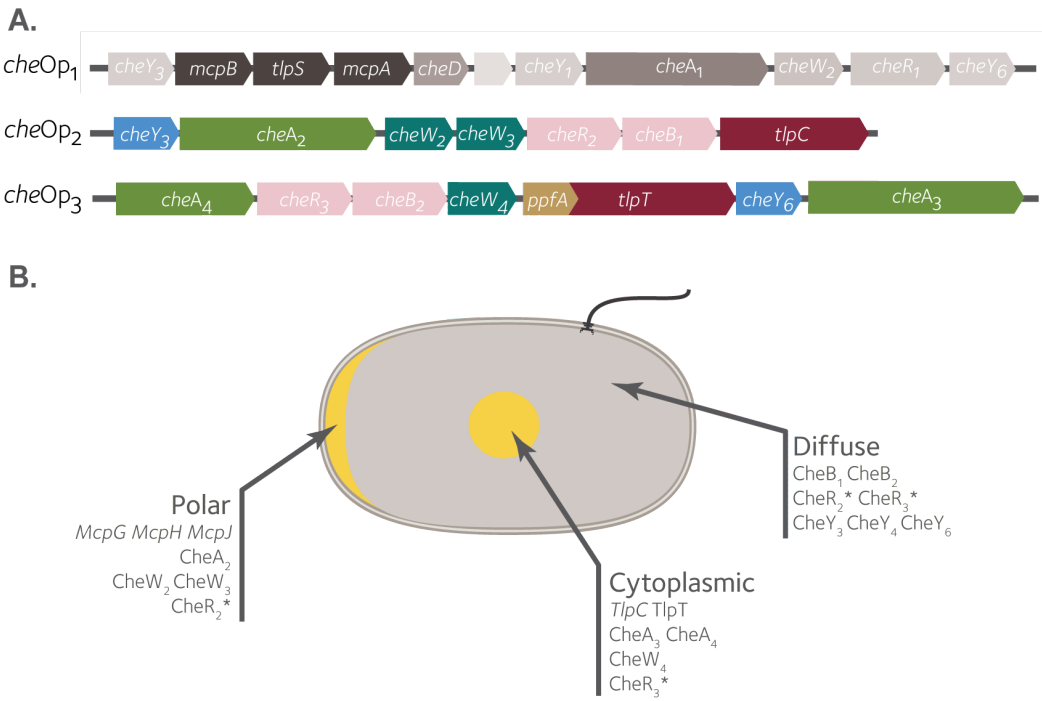

Figure 1.7: The *R. sphaeroides* chemotaxis operons and cellular localisation of the chemotaxis proteins expressed under laboratory conditions. **A.** The three chemotaxis operons encoding the majority of the *R. sphaeroides* chemotaxis proteins. **B.** Cellular localisation of the *R. sphaeroides* chemotaxis proteins, as shown by fluorescence microscopy [115].

the cell pole, and the other in the cytoplasm. The proteins expressed from *che*Op$_2$ generally cluster at the cell pole, while the *che*Op$_3$ proteins cluster in the cytoplasm. An exception to this rule is TlpC, which even though encoded by *che*Op$_2$, localises with the cytoplasmic cluster. The CheB proteins do not form foci and remain diffuse throughout the cytoplasm. CheR$_2$ and CheR$_3$ are partially localised with the polar and cytoplasmic clusters, respectively, but some proportion of the expressed fluorescent fusion was also found diffuse throughout the cytoplasm. Additional proteins shown to colocalise at the cell pole with the *che*Op$_2$ proteins are the transmembrane chemoreceptors McpH, McpJ, and McpG.

1.3.4.1 Polar chemosensory arrays

Electron cryotomography experiments showed that the *R. sphaeroides* polar chemosensory array adopts the conserved hexagonal lattice architecture [21]. During the cell cycle, polar chemosensory clusters form a number of dynamic unit-clusters which diffuse within the membrane, accumulating at the new poles after division and ensuring that each of the daughter cells will have a membrane-embedded chemosensory array [24].

1.3.4.2 Cytoplasmic chemosensory arrays

The cytoplasmic chemosensory cluster in *R. sphaeroides* is its distinguishing feature from other bacteria. Like the polar chemosensory cluster, the cytoplasmic one is essential for chemotaxis. While the polar chemosensory cluster senses the periplasmic environment, the cytoplasmic one senses the concentration of ligands inside the cell. Specific attractants for cytoplasmic chemoreceptors have not been identified. Nevertheless, chemotaxis towards many attractants *e.g.* ammonia, sugars, glutamate, alanine, requires their transport and metabolism. This suggests that the ligands sensed by the cytoplasmic cluster could include common metabolites. Such a scenario would be indicative of energy taxis.

Cryo-EM imaging of the *R. sphaeroides* and *Vibrio cholerae* cytoplasmic chemosensory arrays revealed trimers of receptor dimers organised in 12 nm hexagonal arrays, just like transmembrane ones (Figure 1.8). It was also revealed that the cytoplasmic fragments of transmembrane chemoreceptors form similar sandwich structures in the presence of molecular crowding agents, suggesting that molecular crowding and sandwiching can replace the stabilising effect of the membrane. It is also interesting to point out that the 12 nm hexagonal architecture seems universal among

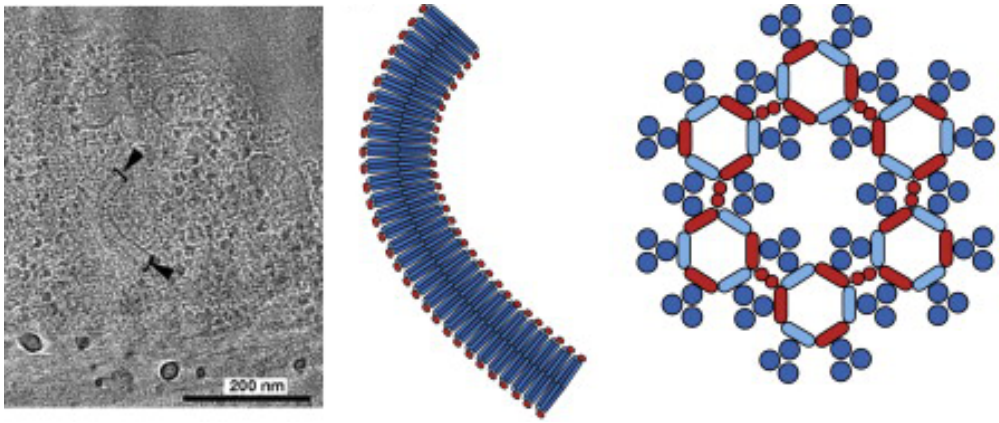

Figure 1.8: The *R. sphaeroides* cytoplasmic chemosensory cluster. From left to right: cryo-EM image showing a side view of the cytoplasmic chemosensory array, model model showing the cytoplasmic chemoreceptors sandwiched between two CheA/W baseplates, and model showing a top view of the cytoplasmic array. The images were taken from [49].

bacterial chemoreceptor arrays, whether membrane embedded or not.

In a new cell, the cytoplasmic array is positioned at mid-cell. As the cell cycle progresses, the array is split into two arrays that become positioned such that when the cell divides, each daughter cell inherits a single cluster positioned at mid-cell. Appropriate positioning of the cytoplasmic array during the cell cycle is ensured by a *che*Op$_3$ encoded ParA ATPase homologue, PpfA, using the chromosome as a scaffold.

1.3.5 Requirements for chemotaxis in *R. sphaeroides*

Both polar and cytoplasmic clusters have at least one homologue of each of the *E. coli* chemotaxis proteins: CheA, CheW, CheY, CheR, and CheB. A comprehensive study of individul chemotaxis protein deletions was carried out by Wadhams *et al.* to establish which of these proteins were indispensable for cluster formation, and what the order of cluster assembly might be [114].

While the integrity of the polar cluster was not affected by the deletion of any cy-

toplasmic cluster components, deleting $CheA_2$, $CheW_2$, or $CheW_3$, caused a reduction in polar clustering. The two polar CheW homologues appear to not be equivalent, with MCPs and $CheW_3$ forming a stable complex that likely interacts with $CheA_2$ and $CheW_2$ to drive the localisation of the entire cluster to the cell pole. The polar cluster formation is proposed to be initiated by MCPs targeting the cell membrane and forming a complex with $CheW_3$, which recruits $CheW_2$ and $CheA_2$, driving the formation of a tight polar cluster.

The cytoplasmic cluster could not form in the absence of TlpT or $CheW_4$, while deletion of *tlpC* caused partial delocalisation of the cluster. However, expression of *tlpT* from a plasmid in a Δ*tlpT* background causes the clusters to form. The cytoplasmic cluster CheAs localise to the cluster independently and neither are required for cluster formation. The cytoplasmic cluster CheA kinases will be further discussed in Chapter 4. Interestingly, deleting $CheW_3$, which localises to the polar cluster, caused some delocalisation of the cytoplasmic cluster. Hence it was suggested that cytoplasmic cluster formation starts with TlpT and $CheW_4$ coming together in the cytoplasm, and interacting with TlpC to form a tighter cluster. Lastly, $CheA_3$ and $CheA_4$ independently join the cytoplasmic cluster. This assembly model fits with the architecture of the cytoplasmic chemosensory array revealed by cryo-EM [18].

Both chemotaxis protein clusters are required for normal functioning chemotaxis in *R. sphaeroides*, deletion of either leading to loss of chemotaxis. Chemotaxis phenotypes as well as the ability to complement *E. coli* deletions are presented in Table 1.2.

Table 1.2: Effects of deleting expressed *R. sphaeroides* chemotaxis proteins.

Deletion	Chemotaxis	Reference	*E. coli* complementation	Reference
Polar cluster				
$\Delta cheA_2$	abolished	[38]	yes	[93]
$\Delta cheW_2$	abolished	[63]	no	[93]
$\Delta cheW_3$	abolished	[63]	yes	[93]
$\Delta cheR_2$	abolished	[64]	yes	[64]
Cytoplasmic cluster				
$\Delta cheA_3$	abolished	[79]	no	[79]
$\Delta cheA_4$	abolished	[79]	no	[79]
$\Delta cheW_4$	reduced	[79]	partial	[79]
$\Delta cheR_3$	abolished	[79]	no	[79]
Diffuse				
$\Delta cheB_1$	abolished	[64]	yes	[64]
$\Delta cheB_2$	reduced	[79]	partial	[79]
$\Delta cheY_3$	none	[78]	no	[93]
$\Delta cheY_4$	none	[78]	no	[93]
$\Delta cheY_6$	abolished	[79]	no	[79]

$\Delta cheY_{3,4}$ leads to loss of chemotaxis [78]

1.3.6 Current model for *R. sphaeroides* chemotaxis

The polar cluster is thought to respond to periplasmic stimuli while the cytoplasmic cluster to cytoplasmic ones and signal to the flagellar motor via CheY homologues in a sequence similar to the *E. coli* model. $CheR_2$ and $CheR_3$ increase the chemoreceptor sensitivity through methylation According to *in vitro* phosphorylation experiments, $CheA_2$ can phosphorylate $CheY_3$, $CheY_4$, $CheY_6$, $CheB_1$, and $CheB_2$ [73].

Upon encountering less favourable conditions, the CheA kinases are switched on. $CheA_2$ phosphorylates $CheY_3$, $CheY_4$, and $CheY_6$, while in the cytoplasmic cluster $CheA_4$ phosphorylates $CheA_3$, which then phosphorylates $CheY_6$. The flagellar motor can only be stopped by $CheY_6$ binding. The signal is terminated by $CheA_3$-catalysed dephosphorylation of $CheY_6$ [75].

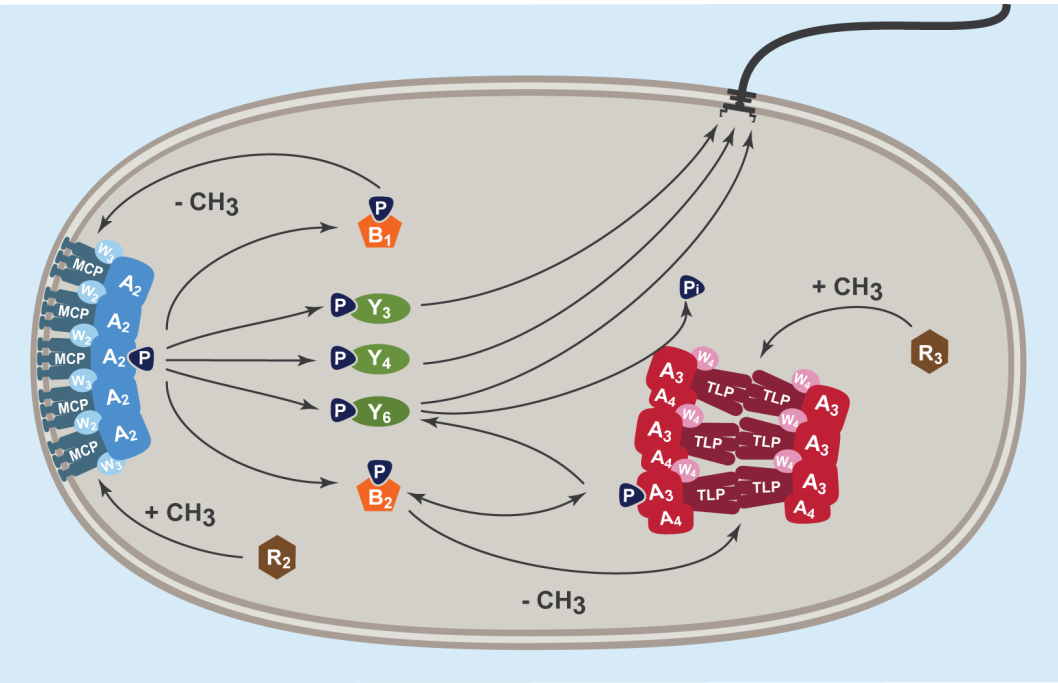

Figure 1.9: Current working model for *R. sphaeroides* chemotaxis. **Sensing:** External nutrient levels are detected by chemoreceptors. The CheR proteins methylate the MCPs and TLPs making them more sensitive. When nutrient levels drop, the chemoreceptors change conformation and activate their respective CheA kinases. **Signalling:** $CheA_2$-P phosphorylates $CheY_3$, $CheY_4$, and $CheY_6$, while $CheA_4$ phosphorylates $CheA_3$, which then phosphorylates $CheY_6$. The flagellar motor is stopped by $CheY_6$-P binding. $CheY_3$-P and $CheY_4$-P are proposed to act as a phosphate sink or interfere with $CheY_6$-P binding at the motor. **Adaptation:** $CheA_2$-P also phosphorylates $CheB_1$ and $CheB_2$, while $CheA_3$ phosphorylates just $CheB_2$. The CheB proteins demethylate the chemoreceptors, making them less sensitive and stopping their signalling. **Signal termination:** $CheY_6$-P loses its phosphate group by autodephosphorylation as well as through the phosphatase activity of $CheA_3$.

1.3.7 Modelling the signalling network of *R. sphaeroides*

1.3.7.1 Using an ODE model

Mathematical models based on ordinary differential equations (ODEs) have been built for the signalling reactions involving phosphorylation in the *R. sphaeroides* chemotaxis pathway [106, 107, 108]. The parameters of the model include rates of phosphosignalling reactions determined *in vitro* as well as *in vivo* protein concen-

trations.

$CheY_6$ acting as a phosphate sink for the other response regulators – thus allowing them to be rapidly dephosphorylated – and a reverse phosphorelay connecting the two clusters are the two major results from the ODE modelling work [7, 77].

1.3.7.2 Using control theory

Roberts *et al.* developed a model invalidation approach aimed at identifying potential interconnection structures of the *R. sphaeroides* chemotaxis network [83]. The method is based on manipulating and observing the entire system *in vivo*, and systematically invalidating a series of biologically viable models following successive model development/experiment design iterations. Control systems theory was used to design suitable experiments *in silico*.

The existence of multiple CheY homologs in *R. sphaeroides* does not allow for the FRET assay developed in *E. coli* to be applied [100, 99]. Roberts *et al.* used average cell traces from tethered cell experiments and *in vitro* reaction rates to establish the roles of $CheY_3$ and $CheY_4$ in the *R. sphaeroides* chemotaxis pathway.

Interestingly, the best fit model suggests that $CheY_3$-P and/or $CheY_4$-P do not bind cooperatively with $CheY_6$ to FliM. The model that could not be invalidated suggests $CheY_3$ and $CheY_4$ form a link between the polar and cytoplasmic signalling clusters, helping transmit the signal between the two clusters. The previously held hypothesis that $CheY_3$ and $CheY_4$ were phosphate sinks for the system was invalidated. The model of strictly cooperative binding of CheYs to the motor was also invalidated. Apart from fluorescent fusion protein localisation, there is no direct experimental proof of this.

Hamadeh *et al.* took the same model invalidation approach as Roberts *et al.* to ad-

dress the regulation of the two chemosensory clusters by $CheB_1$ and $CheB_2$ which are localised throughout the cytoplasm [36]. The assumptions of the models include $CheY_3$-P and $CheY_4$-P acting together to promote autophosphorylation of $CheA_{3,4}$ while $CheY_6$-P binds to FliM to increase the frequency of motor stopping [83].

The model that was not invalidated proposed that $CheB_1$-P and $CheB_2$-P both regulate the methylation state of the polar cluster, while $CheB_2$-P alone regulated the cytoplasmic cluster. The model corresponded to a cascade control system, a setup with one feedback loop nested within the other – in the case of *R. sphaeroides*, the cytoplasmic cluster feedback loop being nested within the polar cluster one. This organisation would enable robust chemotaxis in a noisy external environment as well as high sensitivity to fast changing inputs to the cytoplasmic cluster (perhaps from the metabolic state of the cell).

If $CheB_2$ interacts with both clusters, it remains unclear why there is a need for delocalised $CheB_1$ if it only regulates the polar cluster. It is also possible that the fluorescent tag used for cellular positioning experiments interferes with the localisation.

The network connectivity proposed by Hamadeh *et al.* suggests that the two chemotaxis pathways evolved independently and then became part of the same organism through horizontal gene transfer. From an evolutionary perspective, it is possible that the cascade control model has advantages in terms of robustness and chemotactic performance.

1.4 Project aims

Work in this thesis broadly focuses on the integration of the cytoplasmic chemosensory array into the *R. sphaeroides* chemotaxis signalling network. It addresses aspects of interaction specificity between array components as well as of array components and response regulators. The work also focuses on the proposal for methylation-based adaptation at the cytoplasmic array both from a chemoreceptor and an adaptation protein perspective. The specific aims of this thesis are:

- Structural studies on P5 domains of the cytoplasmic cluster
  - Purify P5-containing constructs to structural study standards
  - Perform biophysical characterisation of these constructs
  - Preliminary work on crystallisation
- Build an interaction map within the cytoplasmic cluster and to test some model predictions using a bacterial two-hybrid approach
  - Construct a library of chemotaxis gene-containing plasmids suitable for the method
  - Test interactions between cytoplasmic cluster components
  - Test interactions between *R. sphaeroides* adaptation proteins and potential interaction partners
  - Use this method to characterise CheA/CheW* interactions
- Investigate adaptation in *R. sphaeroides* adaptation – confirm methylation sites, analyse adaptation protein pair deletions.
  - Characterise the swimming behaviour of TlpT methylation site mutant

strains

  - Determine – based on the mutant work – whether methylation-based adaptation likely occurs in *R. sphaeroides*
  - Characterise the swimming behaviour of strains that have CheBCheR pairs deleted

- Confirm $CheB_2$ interactions with both chemosensory clusters *in vivo*
  - Adapt a single-molecule imaging technique to $CheB_2$
  - Obtain preliminary data on $CheB_2$ localisation based on diffusion coefficients.

- Integrate the data from above to improve the current model for *R. sphaeroides* chemotaxis.

CHAPTER 2

Materials and Methods

Recipes for all the media and buffer solutions mentioned in this chapter are listed in the Appendix.

2.1 Strains and plasmids

Escherichia coli and *Rhodobacter sphaeroides* strains were used in this study. Details on the strains used are presented in Tables 2.1 and 2.2. The generic plasmids used are listed in Table 2.3 while the library of bacterial two-hybrid plasmids is presented in Table 2.4.

Table 2.1: List of *E. coli* strains used.

Name	**Description**	**Source/Reference**
XL1Blue	Plasmid preparation; tetracycline resistant	Stratagene
BL21(DE3)pLysS	Protein overexpression; chloramphenicol resistant	Promega
S17-1 *λpir*	Used for conjugation	[68]
DHM1	Assay strain for BACTH	[50]

Table 2.2: List of *R. sphaeroides* strains used.

Name	Description	Source/Reference
WS8N	Nalidixic acid resistant derivative of wildtype WS8	[98]
JPA467	WS8NΔ*fliA*	
JPA1353	WS8N Δ*che*Op$_{1,2,3}$	
JPA1377	WS8N Δ*cheB*$_2$ Δ*cheR*$_3$	Elaine Byles (this laboratory)
JPA1378	WS8N Δ*cheB*$_1$ Δ*cheR*$_2$	Elaine Byles (this laboratory)
JPA1379	WS8N Δ*cheB*$_{1,2}$ Δ*cheR*$_{1,2,3}$	Elaine Byles (this laboratory)
JPA2332	WS8N *tlpT* (E296A)	Elaine Byles (this laboratory)
JPA2367	WS8N *tlpT* (E289A)	Elaine Byles (this laboratory)
JPA2368	WS8N *tlpT* (E289A, E296A)	Elaine Byles (this laboratory)
JPA2369	WS8N *tlpT* (E478A)	Elaine Byles (this laboratory)
JPA2370	WS8N *tlpT* (Q485A)	Elaine Byles (this laboratory)
JPA2371	WS8N *tlpT* (E289A, E296A, E478A, Q485A)	Elaine Byles (this laboratory)
JPA2332*	WS8N *tlpT* (E296A) *yfp-cheW*$_4$	This study
JPA2367*	WS8N *tlpT* (E289A) *yfp-cheW*$_4$	This study
JPA2368*	WS8N *tlpT* (E289A, E296A) *yfp cheW*$_4$	This study
JPA2369*	WS8N *tlpT* (E478A) *yfp-cheW*$_4$	This study
JPA2370*	WS8N *tlpT* (Q485A) *yfp-cheW*$_4$	This study
JPA2371*	WS8N *tlpT* (E289A, E296A, E478A, Q485A) *yfp-cheW*$_4$	This study

2.2 Microbiology techniques

2.2.1 *E. coli* growth conditions

Liquid cultures of *E. coli* were grown aerobically in Luria Bertani (LB) medium at 37°C with 225 rpm shaking. Stationary cultures were obtained after overnight incubation.

For selection purposes, the relevant *E. coli* strain was spread onto LB agar plates containing the desired antibiotic selection and grown overnight at 37°C to obtain colonies.

Table 2.3: List of plasmids used. All expression plasmids are IPTG-inducible

Name	Description	Source/Ref.
pIND4	IPTG-inducible expression vector for *E. coli* and *R. sphaeroides*, kanamycin resistant	[43]
pK18*mobsacB*	Allelic exchange suicide vector mobilised by *E. coli* S17-1 *λpir*	[89]
pQE80	IPTG-inducible *E. coli* expression vector, N-terminal His_6 tag	QUIAGEN
pKT25	IPTG-inducible *E. coli* expression vector for BACTH assay, kanamycin resistance, N-terminal T25 tag	[52]
pKNT25	IPTG-inducible *E. coli* expression vector for BACTH assay, kanamycin resistance C-terminal T25 tag	[50]
pUT18	IPTG-inducible *E. coli* expression vector for BACTH assay, ampicillin resistance C-terminal T18 tag	[52]
pUT18C	IPTG-inducible *E. coli* expression vector for BACTH assay, ampicillin resistance N-terminal T18 tag	[52]
pQE80-A4	pQE80 containing *R. sphaeroides* $cheA_4$	Kathryn Scott (this laboratory)
pK18-YFP-W4	500 bp upstream of *R. sphaeroides* $cheW_4$ - *yfp* - first 500 bp of $cheW_4$ in pK18*mobsacB*, for genomic tagging of $cheW_4$ with YFP	George Wadhams (this laboratory)

XL1-Blue (Promega) was used for plasmid overexpression. BL21(DE3)pLysS (Promega) was used for protein expression from plasmid. The S17-1 *λpir* strain was used for conjugation into *R. sphaeroides*.

2.2.2 *R. sphaeroides* growth conditions

Liquid cultures of *R. sphaeroides* were grown in succinate medium at 30°C, either aerobically in the dark shaking at 225 rpm, or anaerobically with illumination and

no shaking. Stationary cultures were typically obtained after two days.

Selection was carried out by streaking on LB agar plates containing the desired antibiotic selection and incubation at 30°C for two to three days.

2.2.3 Antibiotics

Antibiotics were added to the medium to the final concentration indicated in Table 2.5. Specific antibiotic resistance for strains and plasmids are detailed in Table 2.1 and Table 2.3. 1000x stock solutions of each antibiotic were stored at -20°C.

2.2.4 Preparation of competent *E. coli* strains

E. coli strains were made chemically competent using standard transformation buffer (TFB). A stationary liquid culture was diluted 1:50 in 50 mL LB media and grown aerobically until OD_{600} =0.4-0.5. The culture was then incubated on ice for 20 minutes, and centrifuged at 1000×g for 20 minutes at 4°C. The supernatant was discarded and the cell pellet was re-suspended in 16 mL of ice-cold buffer TFB I. The re-suspended cells were incubated on ice for 20 minutes, then centrifuged at 1000×g for 10 minutes at 4°C. The pellet was re-suspended in 2 mL ice-cold sterilised TFBII buffer and aliquoted into single-use volumes.

2.2.5 Transformation of *E. coli* cells

Plasmids were inserted into chemically competent *E. coli* cells (Section 2.2.4) by transformation. Ice-cold DNA was added to 50-100 μL thawed cells and incubated on ice for 30 minutes. The cells were heat-shocked by incubation at 42°C for 45 seconds, then immediately incubated on ice for 2 minutes. Following the recovery,

800 μL LB were added and the cells were incubated at 37 °C for an hour. Cells were then spread onto LBA plates containing the appropriate antibiotics and incubated overnight at 37°C.

2.2.5.1 Conjugation into *R. sphaeroides*

As *R. sphaeroides* cannot be made chemically competent, plasmids were transferred into *R. sphaeroides* by conjugation using the *E. coli* strain S17-1 λ*pir* (Table 2.1).

Plasmids were transferred to *E. coli* S17-1 λ*pir* cells by transformation (Section 2.2.5). A stationary culture obtained by growing a single colony of S17-1 λ*pir* transformant overnight at 37°C in LB with the appropriate antibiotic was diluted 1:100 and grown to OD_{600} =0.2-0.3.

1 mL of this S17-1 λ*pir* culture and 1 mL of a stationary *R. sphaeroides* culture (Section 2.2.2) were centrifuged individually at 6000×g for 3 minutes. The cell pellets were each gently re-suspended in 1 mL LB. The cells were centrifuged as before, then re-suspended in 100 μL LB. 100 μL *R. sphaeroides* suspension were mixed with 10 μL *E. coli* S17-1 λ*pir* suspension and pipetted onto a nitrocellulose filter placed on an LB plate.

After incubation overnight at 30°C, the filter paper was transferred to a microcentrifuge tube containing 800 μL LB and vortexed to re-suspend the cell pellet. The cell suspension was spread onto four LBA-Nal plates containing the selection antibiotic for the plasmid used, and the plates were incubated at 30°C for two to three days.

2.2.6 Genomic modification of *R. sphaeroides*

Recombination was used to fluorescently label genes on the *R. sphaeroides* genome. This was achieved using the suicide plasmid pK18*mobsacB*. The N-terminally tagged genomic insertion constructs contained the 500 bp immediately upstream of the gene of interest, the tag, and the first 500 bp of the gene of interest.

The suicide plasmid constructs were conjugated into the desired *R. sphaeroides* strain (Section 2.2.5.1) and spread onto LBA plates containing nalidixic acid and kanamycin. Since *R. sphaeroides* is unable to replicate pK18*mobsacB*, the only way that the kanamycin resistance of pK18*mobsacB* can be passed on is by genomic integration of the plasmid through homologous recombination. This is the first round of recombination.

Colonies from the first round of recombination were grown photoheterotrophically to stationary phase in SUX media with nalidixic acid. The absence of kanamycin allows the second round of recombination to occur, excising the vector backbone containing the *kan*^{r} and *sacB* genes from the genome. The second round of recombination is selected for by plating a series of stationary culture dilutions onto M22 10% sucrose agar plates. Replica plates of the colonies obtained here were also used to confirm the loss of the *kan*^{r} gene.

The second round of recombination also results in the excision of either the parental sequence or the mutation construct. Thus two outcomes are possible: the desired mutation is integrated into the genome or the strain reverts to wild-type genotype. To distinguish between these two possible outcomes, colonies that were confirmed to have lost the *kan*^{r} gene were further screened by colony PCR (Section 2.3.2.4).

2.2.7 Storage

Strains were stored long-term as glycerol stocks at -80°C. The stocks were kept in cryotubes and contained *E. coli* or *R. sphaeroides* stationary culture with 20% glycerol.

2.3 Molecular biology techniques

All standard genetic techniques were performed as described in [87].

2.3.1 Plasmid DNA extraction from *E. coli* cells

The QUIAGEN QIAprep Spin Miniprep Kit was used to extract sequencing-quality plasmid DNA from 5 mL stationary culture. The DNA was eluted using 50 μL MilliQ water and stored at -20°C

2.3.2 Cloning

2.3.2.1 Agarose gel DNA electrophoresis

DNA samples were separated using gel electrophoresis. The gel was made by mixing 1% agarose (w/v) with 0.5x TBE buffer. DNA loading dye was mixed with the DNA sample and loaded onto the agarose gel alongside 1 kb-plus DNA ladder (Invitrogen).

2.3.2.2 Purification of DNA from agarose gels

2.3.2.3 Polymerase chain reaction (PCR)

PCR was performed using *Phusion* DNA polymerase (Promega), according to the manufacturer's instructions. Primers were synthesized by Sigma-Genosys and are listed in Appendix **??**.

2.3.2.4 Colony PCR

Colony PCR was performed using GoTaq G2 DNA polymerase (Promega), according to the manufacturer's instructions. Primers were synthesized by Sigma-Genosys and are listed in Appendix **??**.

20 μL reactions were set up using the recipe and program described in Table 2.6.

2.4 Microscopy

2.4.1 Sample preparation

R. sphaeroides cells were grown aerobically, as described in Section 2.2.2, to OD_{700} =0.4-0.6.

2.4.2 Slide preparation

The cells were imaged on agarose pads. The pads were created by adding a thin layer of 1% agarose in succinate medium to a glass slide. 2 μL of culture were added to the pad and then a glass cover slip was placed on top.

For photoactivation localisation microscopy, it was important to remove background fluorescent particles. To this end, coverslips were previously burned in a furnace at 500°C for 1 hour. Burned coverslips were stored at room temperature wrapped in aluminium foil

2.4.3 Microscopes

2.4.3.1 Nikon Eclipse Ti

Fluorescence images of cells were acquired using an inverted Nikon Eclipse Ti microscope in phase contrast. 100× oil objectives (Nikon) were used. The microscope was controlled using NIS Elements software (Nikon). The microscope was used in epifluorescence mode using a mercury lamp and appropriate excitation and emission filters for YFP. Brightfield images used a halogen lamp for illumination. An Andor iXON CCD camera was used to capture images.

2.4.3.2 Photoactivation localisation (PALM) microscope

A custom built PALM microscope was used [109]. A Toptica Multi Laser Engine is used to provide laser excitation via a single-mode optical fibre. The fluorescence emission is recorded using an Andor iXon 879 ultra electron multiplying CCD (EM-CCD) camera. The resulting magnification is 96 nm/pixel. Sample positioning and focusing is achieved using an ASI controller. YFP fusion proteins were excited with a 473 nm laser, PAmCherry fusion proteins were photoactivated with the 405 nm laser, and excited with the 561 nm laser. Brightfield images were aquired using a LED light source (coolLED) and an Olympus condenser. The objective used was 100x oil immersion Olympus objective with 1.4 numerical aperture.

Table 2.4: Library of bacterial two-hybrid plasmids

Name	Description	Source
B2H-101	pKT25 containing *cheA*$_3$	This study
B2H-102	pKT25 containing *cheA*$_4$	This study
B2H-103	pKT25 containing *cheW*$_4$	This study
B2H-104	pKT25 containing *tlpC*	This study
B2H-105	pKT25 containing *tlpT*	This study
B2H-109	pKT25 containing *cheR*$_1$	Elaine Byles
B2H-110	pKT25 containing *cheR*$_2$	Elaine Byles
B2H-111	pKT25 containing *cheR*$_3$	Elaine Byles
B2H-112	pKT25 containing *cheB*$_1$	Elaine Byles
B2H-113	pKT25 containing *cheB*$_2$	Elaine Byles
B2H-116	pKT25 containing *cheW*$_4$ (I62A)	Elaine Byles
B2H-201	pKNT25 containing *cheA*$_3$	This study
B2H-202	pKNT25 containing *cheA*$_4$	This study
B2H-203	pKNT25 containing *cheW*$_4$	This study
B2H-204	pKNT25 containing *tlpC*	This study
B2H-205	pKNT25 containing *tlpT*	This study
B2H-209	pKNT25 containing *cheR*$_1$	Elaine Byles
B2H-210	pKNT25 containing *cheR*$_2$	Elaine Byles
B2H-211	pKNT25 containing *cheR*$_3$	Elaine Byles
B2H-212	pKNT25 containing *cheB*$_1$	Elaine Byles
B2H-213	pKNT25 containing *cheB*$_2$	Elaine Byles
B2H-216	pKNT25 containing *cheW*$_4$ (I62A)	Elaine Byles
B2H-301	pUT18 containing *cheA*$_3$	This study
B2H-302	pUT18 containing *cheA*$_4$	This study
B2H-303	pUT18 containing *cheW*$_4$	This study
B2H-304	pUT18 containing *tlpC*	This study
B2H-305	pUT18 containing *tlpT*	This study
B2H-309	pUT18 containing *cheR*$_1$	Elaine Byles
B2H-310	pUT18 containing *cheR*$_2$	Elaine Byles
B2H-311	pUT18 containing *cheR*$_3$	Elaine Byles
B2H-312	pUT18 containing *cheB*$_1$	Elaine Byles
B2H-313	pUT18 containing *cheB*$_2$	Elaine Byles
B2H-316	pUT18 containing *cheW*$_4$ (I62A)	Elaine Byles
B2H-401	pUT18C containing *cheA*$_3$	This study
B2H-402	pUT18C containing *cheA*$_4$	This study
B2H-403	pUT18C containing *cheW*$_4$	This study
B2H-404	pUT18C containing *tlpC*	This study
B2H-405	pUT18C containing *tlpT*	This study
B2H-409	pUT18C containing *cheR*$_1$	Elaine Byles
B2H-410	pUT18C containing *cheR*$_2$	Elaine Byles
B2H-411	pUT18C containing *cheR*$_3$	Elaine Byles
B2H-412	pUT18C containing *cheB*$_1$	Elaine Byles
B2H-413	pUT18C containing *cheB*$_2$	Elaine Byles
B2H-416	pUT18C containing *cheW*$_4$ (I62A)	Elaine Byles

Table 2.5: Final concentration used for antibiotics used in selection

Antibiotic	**Final concentration**	**Dissolved in**
Ampicillin (Amp)	100 μg/mL	water
Chloramphenicol (Cam)	30 μg/mL	ethanol
Kanamycin (Kan)	25 μg/mL	water
Nalidixic acid (Nal)	25 μg/mL	water
Tetracycline (Tet)	30 μg/mL	95% ethanol

Table 2.6: Colony PCR

Component	**Volume (μL)**
MilliQ water	100 μg/mL
5xGreen reaction buffer	9.225
50 mM $MgCl_2$	5
dNTPs (2.5 mM each)	2
DMSO	1.25
Forward primer	0.2
Reverse primer	0.2
GoTaq G2 polymerase	0.125
Template	*colony*
final volume	*20*

CHAPTER 3

Binary protein-protein interactions of the cytoplasmic array

The chemotaxis network of *R. sphaeroides* is complex, with two chemosensory systems controlling a single motor. The connectivity of this network has been explored through *in vitro*, *in vivo*, and modelling studies. The architectures of the polar and cytoplasmic chemosensory arrays have the same lattice structure which is highly conserved among bacterial species [21, 18]. While this arrangement predicts certain protein-protein interactions within the array, the specific relationships are less obvious in the case of the cytoplasmic array because of the non-cannonical chemoreceptors and trademark split CheA kinase, $CheA_{3,4}$. Work described in this chapter focuses on the cytoplasmic array and uses a two-hybrid system to identify interactions between its components.

3.1 The cytoplasmic chemosensory array

The products of the third *R. sphaeroides* chemotaxis operon, *che*Op_3, are essential for chemotaxis [79] and predominantly localise to a discrete region in the cytoplasm

[115]. Microscopy fluorescently-labelled deletion strains has revealed the protein requirements for this clustering to occur.

The chemoreceptor TlpT and the scaffold protein $CheW_4$ were shown to be essential for the localisation of the *che*Op_3 proteins to the specific region in the cytoplasm [114]. Neither of the two CheA kinases, were necessary for the localisation of the other proteins [114]. TlpC, expressed from *che*Op_2, localises to the cytoplasm like the products of *che*Op_3 [113, 115]. Fluorescence microscopy experiments have shown that deletion of *tlpC* reduces the fluorescence of the localised *che*Op_3 proteins and increases general cytoplasmic fluorescence [114]. Further, the fluorescence of YFP-$CheY_6$ predominantly localised to a cytoplasmic focus, with a proportion of the fluorescent protein also found diffuse throughout the cytoplasm. The methyltransferase $CheR_3$ had a localisation pattern similar to $CheY_6$, while the methylesterase $CheB_2$ was found diffuse throughout the cytoplasm.

Insight into the architecture of the phosphorelay network was obtained through extensive *in vitro* phosphotransfer experiments [73, 74, 78]. The cytoplasmic kinase activity requires both $CheA_3$ and $CheA_4$, as each has a subset of the domains present in a typical CheA kinase [74]. The *in vitro* kinetic studies suggested a mechanism in which $CheA_3$ associates with $CheA_4$, enabling the phosphorylation of the histidine residue in the $CheA_3$ P1 domain by the P4 kinase domain of $CheA_4$. The $CheA_{3,4}$ kinase can transfer a phosphoryl group to two response regulators: $CheY_6$, which in phosphorylated form can stop flagellar motor rotation, and $CheB_2$ a putative methylesterase response regulator, likely involved in the adaptation of the cytoplasmic chemosensory system.

Mathematical modelling predicted a bi-functional kinase-phosphatase as well as the ability of the cytoplasmic chemosensory system to signal to the polar one by reverse phosphotransfer through $CheB_2$–P [107]. Both of these results are supported by *in*

vitro phosphotransfer data [75, 76].

Modelling also suggests that $CheY_3$ and $CheY_4$ are likely to interact with the cytoplasmic chemosensory system, acting antagonistically to ligand in altering kinase activity [83]. Further modelling work builds on these results and proposes that $CheB_2$ interacts with both the cytoplasmic and the polar chemosensory systems [36]. At present, there is no direct experimental evidence for either of these two models.

Recent electron cryotomography experiments have revealed that the overall arrangement of the cytoplasmic chemotaxis proteins is identical to the highly-conserved chemosensory array structure across bacterial and archaeal species [21, 18, 20]. This model describes trimers of receptor dimers are networked by six-membered rings containing kinase P5 domains and scaffold proteins. In the case of cytoplasmic arrays, two such lattices are sandwiched together such that the kinase and scaffold protein layer faces the exterior.

The electron cryotomography data do not provide enough resolution to determine precise locations of the protein components. There is a lack of high-resolution structural information about the components of the chemosensory array that could be used to model the arrangement of the different proteins in the lattice. Complexity stemming from the presence of atypical soluble chemoreceptors and the split $CheA_{3,4}$ kinase greatly limits the extent to which findings from membrane embedded arrays can be applied here.

While the signalling network in *R. sphaeroides* has been well-characterised through *in vitro* phosphotransfer, none of the interactions between individual cytoplasmic array components have been verified experimentally. Work presented in this chapter aims to identify interactions between cytoplasmic cluster components *in vivo* using a bacterial two-hybrid approach.

3.2 Overview of the bacterial adenylate cyclase two-hybrid (BACTH) system

The concept behind a two-hybrid system is the detection of an interaction via reconstructing the activity of a modular protein. The most widely used such systems use yeast as host organisms and are based on the reconstruction of transcription factors [23]. A recent large-scale study publisehd in *Nature Biotechnology* used a yeast two-hybrid screen to map the binary protein-protein interaction landscape of *E. coli*, [81]. The identified interactions provided information on the internal topology of complexes and a draft reference for the basic physical wiring network in *E. coli*.

Analogous two-hybrid protocols have been developed in organisms other than yeast. As discussed in a review by Legrain and Selig [60], *E. coli* appears better suited for protein-protein interaction screens than *Saccharomyces cerevisiae* because it has a much lower generation time and the molecular biology techniques are much better adapted to bacteria than yeast. Similar to the yeast two-hybrid systems, there are bacterial two-hybrid systems based on transcriptional activation/repression of reporter genes, but there is also a separate class based on enzyme reconstitution.

The bacterial two-hybrid system used here is based on the reconstitution of adenylate cyclase. It is an alternative to the widely used yeast two-hybrid technology for identifying protein-protein interactions, as detailed in a protocol-focused review by Battesti and Bouveret [11]. The system was initially described in 1998 [51] and named BACTH (Figure3.1).

The adenylate cyclase in question is a toxin from *Bordetella petrussis*, the organism responsible for whooping cough. The enzyme becomes active within the eukary-

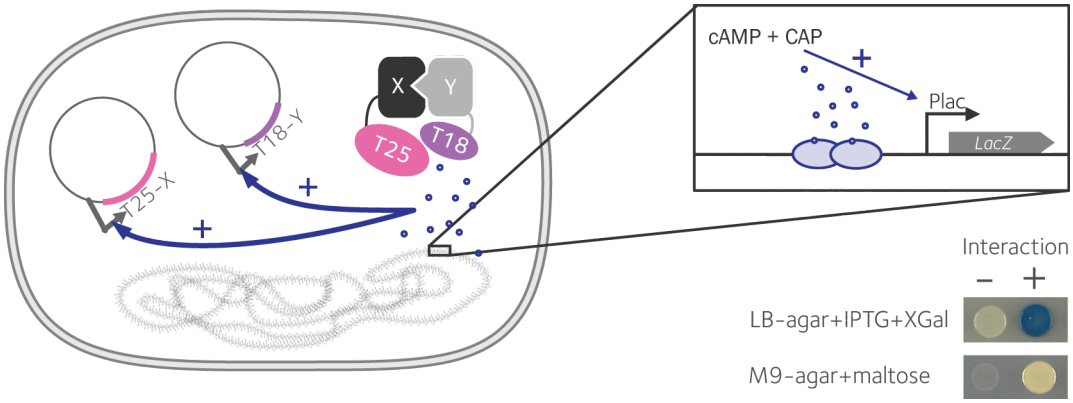

Figure 3.1: Description of the BACTH system. The BACTH system is based on the reconstruction of the cyclic AMP (cAMP) signalling cascade. If the T18 and T25 fragments of the *Bordetella pertussis* adenylate cyclase are physicaly brought together by the interaction of the proteins of interest coexpressed in a background lacking adenylate cyclase, the cAMP signalling cascade is activated leading to the expression of the *lac* operon.

otic host only when it binds to calmodulin, an exclusively eukaryotic protein. The catalytic domain can be divided into a 25 kDa and a 18 kDa subdomain which can produce cyclic AMP (cAMP) when expressed in the presence of calmodulin. This ability of the two sub-domains to recover adenylate cyclase activity when brought together by calmodulin forms the basis of the BATCH system.

The proteins of interest are expressed as fusions of the T25 or T18 domains in a *cya*$^{-}$ strain. If the proteins of interest do indeed interact, they bring together the two catalytic subdomains, fulfilling the role of calmodulin in restoring adenylate cyclase activity. The cAMP form a complex with the catabolite activator protein (CAP), acting as a transcriptional regulator. The lactose and maltose operons are activated by the cAMP/CAP complex, which can be easily detected in *E. coli*.

The BACTH system has been used to identify interactions in a number of bacterial protein networks. For example, *Chlamydiae*, intracellular eukaryotic pathogens, grow in cytoplasmic vesicles. The membranes of these vesicles are modified by chlamydial proteins whose interaction patterns have been investigated using BACTH

[35]. Zhang *et al.* have used the BACTH system to identify specific domains involved in the interactions that cause the elongation and contraction events of the Type VI secretion system [119].

The method also aided the construction of a protein-protein interaction network for the biosynthesis of the major cell wall component arabinogalactan in *Corynebacterium glutamicum*. In this case, twenty-four putative homotypic and heterotypic interactions were identified *in vivo* using BACTH [45].

The δ-proteobacterium *Myxococcus xanthus* has 21 chemoreceptors and gene clusters encoding for 8 separate chemosensory systems. To understand how the chemoreceptors are distributed among the chemosensory systems, the BACTH system was recently used to probe interactions suggested by phylogenetic and cellular localisation studies [67]. Indeed, 10 years ago, Lancero *et al.* used a yeast two-hybrid system to find interactions within the Dif chemosensory pathway proteins in *M. xanthus* [57].

Previous work in the laboratory has used the BACTH method to uncover interactions between proteins involved in *R. sphaeroides* cell division (N. Dubarry, unpublished), as well as to screen for interactions between *R. sphaeroides* MCPs and the polar array scaffold proteins $CheW_2$ and $CheW_3$ [5].

3.3 Experimental approach

Work described in this chapter focuses on the cytoplasmic array and aims to use the BACTH system to obtain *in vivo* information about its interactions. First, an interaction network is established between cytoplasmic array components. Then the interactions of the cytoplasmic cluster with the adaptation proteins is addressed. Finally, the applicability of the BACTH system to study the CheA/CheW interface

residues is investigated.

3.4 Results

3.4.1 Creation of a BATCH plasmid database for *R. sphaeroides* chemotaxis proteins

R. sphaeroides chemotaxis genes were cloned into all four BACTH plasmids, with the aim of obtainig both N- and C-terminal fusions of both the T18 and T25 tags [52, 50]. The successful plasmids are listed in Table 2.4.

Restriction cloning was employed for library creation, with the pair of restriction sites chosen to be compatible for all *R. sphaeroides* chemotaxis genes across the four BACTH plasmids. This means that the same pair of restriction enzymes can be used for all cloning into BACTH plasmids. In addition, one pair of amplification primers is enough to generate inserts for cloning into all T18 and T25 plasmids.

The pair of restriction enzymes are *XbaI* and *KpnI-HF*, for the 5' and 3' end of the cloned fragment, respectively. It should be kept in mind that the two restriction sites are not in frame and this should be taken into account when designing the amplification primers. The two restriction sites are located within the open reading frame so there is no need for the addition of a stop codon.

Amplification of the desired DNA fragment was carried out by PCR using an annealing temperature of 65°C, which was further optimised if required. The purified PCR products and a sample of each plasmid were digested for with XbaI and KpnI-HF at 37°C in CutSmart Buffer (New England Biolabs). The digests were purified,

ligated, and transformed into *E. coli* XL-1 Blue cells for plasmid production. A selection of the colonies obtained after transformation were verified for the presence of the insert. Successful colonies were grown to stationary phase and the plasmid DNA was extracted. The desired insertions into the plasmid were verified by sequencing.

Thus the cloning process for obtaining the BACTH constructs is straightforward and the strategy outlined here can easily be applied to obtain analogous constructs. For the work described below, 44 BACTH constructs were made in collaboration with Elaine Byles, including all of the cytoplasmic chemosensory array components and its putative interaction partners.

3.4.2 Detecting interactions using BACTH

In this work, signal detection was done on LB agar plates containing ampicillin and kanamycin for the selection of cells containing both T18 and T25 constructs. IPTG was added to induce expression, and X-Gal to detect *beta*-galactosidase activity.

An important aspect of detecting an interaction is the incubation time, which is dependent on the strength of the interaction between the two recombinant proteins. As pointed out by Battesti and Bouveret, incubation times shorter than 48 h might not be sufficient to accumulate enough cAMP to detect a signal [11]. In the protocol used here, the assay plates were incubated at room temperature for a week after the initial 48 h incubation at 30°C. The BACTH system presents a certain level of heterogeneity due to expression regulation by cAMP, hence assay cultures were inoculated with several co-transformation colonies. Furthermore, each interaction was tested in at least two independent assays performed on different days, form

different co-transformations.

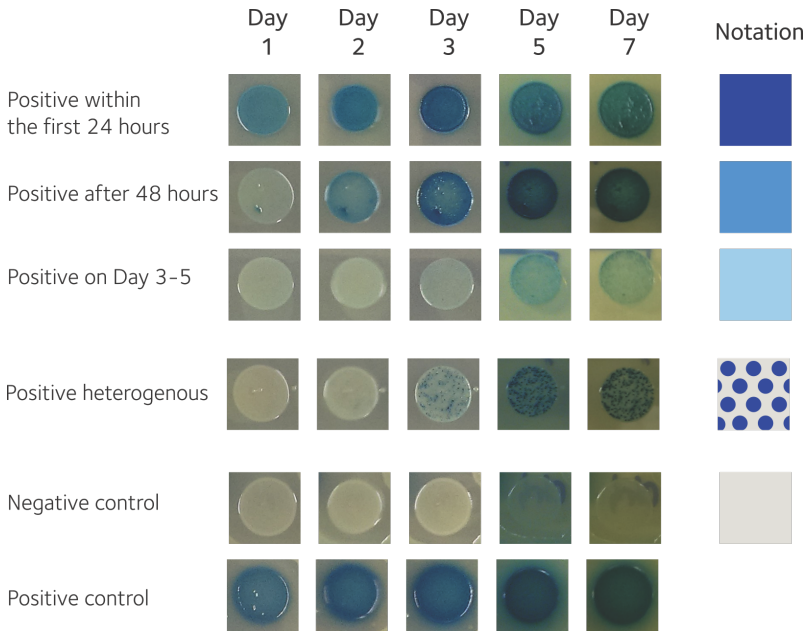

Figure 3.2: Signal detection in the BACTH assay. The incubation time required for the blue colour to appear is dependent on the strength of the interaction between the two recombinant proteins. The interactions observed in BACTH assays were colour-coded in relation to how long it took for the blue colour to appear, from which relative strength of interaction was inferred. A darker shade of blue indicates a stronger interaction. Grey indicates lack of interaction.

Depending on whether blue colouration developed and when this happened, a code was associated, as described in Figure 3.2. The intensity of the blue colour is proportional to the strength of the interaction, while grey represents a lack of colour after the one week room temperature incubation.

For every BACTH construct in the database, a control assay was run using co-transformants of the plasmid in question with an empty plasmid expressing just the complementary tag. All plasmids produced negative results upon testing against the complementary tag alone. This suggests that any interactions detected when testing pairs of constructs with complementary tags from the database are due to an interaction between the proteins themselves, not between the tags.

Alongside each assay, a negative and a positive control were also carried out. The

negative control was a co-transformant of two empty BACTH plasmids with complementary tags *e.g.* pKT25 and pUT18. The positive controls used were the pT18-zip and pKT25-zip plasmids coding for a strongly interacting leucine zipper motif [51].

3.4.3 Interactions between the cytoplasmic cluster proteins

The four constructs – N- and C-terminal fusions using the T18 and T25 tags – were successfully obtained for the components of the cytoplasmic chemosensory array. Pairs of complementary constructs were tested against each other according to the protocol described above.

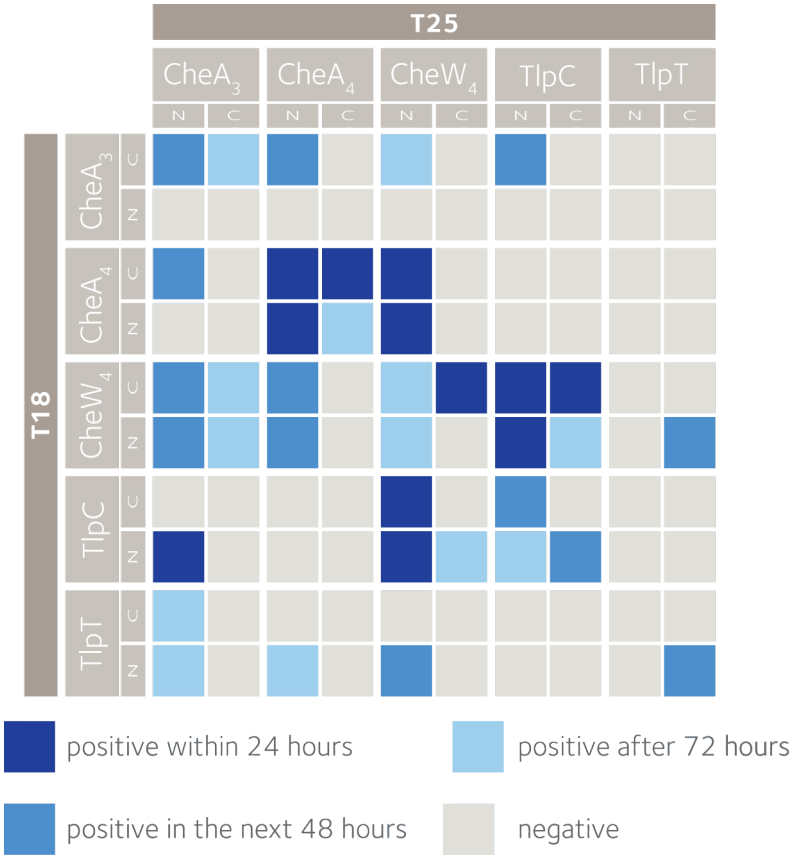

Figure 3.3: Interactions within the cytoplasmic cluster.

The identified interactions were summarised in Figure 3.3. Results for both N- and

C-terminal fusions are presented.

Interactions between most pairs of proteins can be observed. All proteins interacted with themselves, which was expected for $CheA_4$ which has a dimerisation domain (P3) and the chemoreceptors TlpT, and TlpC (transmembrane chemoreceptors are known to form dimers).

There were no interactions detected for the T18-tagged $CheA_3$ and T25-tagged TlpT constructs. Given that other fusions of $CheA_3$ and TlpT do show some interactions, it is possible that the constructs not showing any interactions are not functional.

No interactions at all have been detected between $CheA_4$ and TlpC. The $CheA_4$ and TlpC fusions report interactions of varying strength with other fusion constructs from the library, so lack of interaction due to misfolding is unlikely. Steric hindrance remains a plausible reason for the lack of interaction reported by this method.

The strongest interactions were observed for the formation of $CheA_4$ and $CheW_4$ dimers, as well as for $CheW_4$-$CheA_4$, $CheW_4$-TlpC, and $CheA_3$-TlpC interactions. While the strength of interaction between the various constructs does vary, and it may be influenced by the tags, the pairs mentioned above consistently displayed strong interactions.

3.4.4 Interactions of the adaptation proteins with the cytoplasmic cluster proteins

Interactions between the cytoplasmic chemoreceptors on the one hand, and the $CheB_{1,2}$ and $CheR_{2,3}$ adaptation proteins on the other, have also been investigated. Following the same protocol outlined above, no interactions were detected between

cytoplasmic chemoreceptors and adaptation proteins.

$CheB_2$ was also tested against the rest of the cytoplasmic proteins. Heterogeneous signals became visible after 7 days of incubation at room temperature. Interactions with $CheA_3$ could be observed. This suggests that the lack of signal obtained when testing $CheB_2$ against chemoreceptors previously was likely not due to incorrect production of the fusion protein. An interaction with $CheA_3$ was expected given the phosphotransfer results of Porter *et al.* [73].

3.4.5 Testing proposed interacting residues at the $CheA_{3,4}$/$CheW_4$ interface

As discussed in the Introduction, a typical CheA histidine kinase consists of five domains named P1 through to P5. No crystal structure for a full-length CheA is available, but due to its modular architecture, crystal structures of certain domains are available (Table 4.1).

Work in the laboratory used the existing structure of the P4 and P5 domains of *T. maritima* CheA in complex with CheW (PDB code 2CH4) to predict residues involved in the interaction between *R. sphaeroides* $CheA_4$ and $CheW_4$ (K.A. Scott, unpublished). The prediction was made using the protein contact prediction tool, i-Patch [39], and yielded 7 $CheW_4$ residues that were carried forward into experimental testing. Individual point mutations of these residues to alanine were created in strains with fluorescently-labelled cytoplasmic array. Since $CheW_4$ is required for array formation, it was proposed that disrupting the $CheA_4$-$CheW_4$ interface would lead to some level of fluorescence delocalisation. Most delocalisation was observed with the $CheW_4$ I62A mutant. To further investigate that this mutation disrupts the interaction with $CheA_4$, interactions were probed using BACTH.

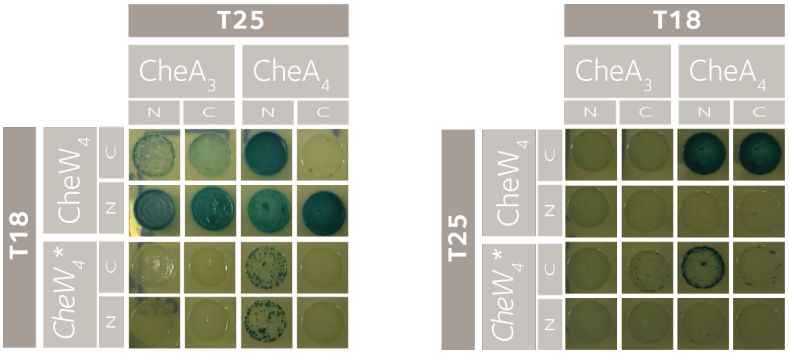

Figure 3.4: $CheW_4$ interactions with $CheA_3$ and $CheA_4$ are disrupted by an I62A mutation in $CheW_4$. The images shown were taken after 5 days. $CheW_4$ * is the short hand notation used in this figure for the $CheW_4$ I62A mutant.

Wild-type $CheW_4$ as well as the I62A mutant were tested against $CheA_3$ and $CheA_4$, in all combinations (Figure 3.4). While signals for wild-type $CheW_4$ became visible within the usual time frame (max 48 hours), it took 5 days of incubation at room temperature for signals belonging to $CheW_4$ I62A to be observed. The signal pattern for the mutant and wild-type was the same. That is to say, positive interaction for the I62 mutant were found with a subset of the wild-type interaction partners. It is interesting to note that while all $CheW_4$ I62A interactions appear weaker, interactions with $CheA_3$ seem to have been completely disrupted.

3.5 Discussion

While BACTH is a powerful method for detecting protein-protein inveractions *in vivo*, it is important to bear in mind that the method cannot be used to invalidate a proposed interaction. A negative result in the BACTH assay does not imply a lack of interaction.

There were no reports in the literature of proteins able to bind to the T18 or T25 domains and restore adenylate cyclase activity [11]. Hence any adenylate cyclase activity in the assay would be due to the interaction of the T18 and T25 domains,

which would have been brought together by the pair of proteins being investigated, and thus correlated with an interaction.

The cAMP-driven positive feedback loop means that only if the interation occurs, will the cAMP level rise and allow full expression of the hybrid proteins (Figure 3.1). This produces a threshold effect which keeps false positives due to sticky proteins or misfolding very low. It was shown that by using a *lac* promoter not inducible by cAMP increased the number of false positives [11]. However, this auto-activation is one of the potential reasons for the heterogeneity discussed in Section 3.4.2.

The fusion proteins are expressed from plasmids that have the same promoter (the wild-type *lac* operon promoter), but different replication origins: the T25 plasmids have a p15A origin (low copy number), while the T18 plasmids have a ColE1 origin (high copy number). This difference in plasmid copy number leads to different relative amounts of the two hybrid proteins being produced. For certain interactions, stoichiometry plays a crucial role, which is why both possible combinations of T18 and T25 fusions should be tested. It is not rare for an interaction to be observed in only one of the cases [11].

The BACTH assay is not a direct measure of an interaction. The outcome of the two proteins interacting is the restoration of adenylate cyclase activity, whereas the output of the assay is the level of expression of the lactose and maltose operons. Hence the plate assays are an indirect assessment of an interaction involving a complex signalling cascade which could be affected by a series of other signals, apart from cAMP.

Other proteins competing for binding to the assay proteins can reduce the likelihood of a T18 and a T25 fragment to meet and generate a signal. Cross-interaction with *E. coli* chemotaxis proteins should be kept in mind, particularly for interactions involv-

ing proteins that are able to complement the corresponding *E. coli* deletion strain. This is the case with $CheB_2$, which practically complements an *E. coli* CheB deletion, but not with $CheR_3$ which cannot complement an *E. coli cheR* strain. Hence the reason why the interactions between $CheB_2$ and $CheA_3$ only become visible after 7 days, could in part be due to $CheA_3$ interacting with the *E. coli* CheB.

It is important to take into consideration that some of the proteins investigated in this chapter are naturally part of an array. The lack of an analogous cytoplasmic array in the assay organism could have a significant influence on the three-dimensional structure of the protein and its ability to interact with other proteins.

Overall, the interaction map obtained for the components of the *R. sphaeroides* cytoplasmic chemosensory array reflects the connectivities observed in the *E. coli* paradigm. A notable exception is an interaction not being observed between the cytoplasmic chemoreceptor TlpC and the atypical histidine kinase $CheA_4$. The other cytoplasmic atypical histidine kinase showed interactions with TlpC. Both kinases showed interaction with TlpT.

The existence of an interaction between TlpC and $CheA_4$ should be tested through other methods too since BACTH cannot provide a verdict in this situation. If indeed, there is no interaction between TlpC and $CheA_4$, this could provide an explanation between the differences in localisation patterns observed for $CheA_3$ and $CheA_4$ [90]. Scott *et al.* found that the P5 domains of the two cytoplasmic CheAs, even though they are both responsible for localisation to the cytoplasmic array, are not equivalent. A domain-swapping approach revealed that if both proteins have the $CheA_3$ P5 domain, then the two CheAs colocalise to the cytoplasm as expected. However, if both CheAs have the $CheA_4$ P5 domain, $CheA_4$ preferentially localises to the cytoplasm. Further BACTH experiments could help identify the domains involved in the $CheA_3$-TlpC interaction and help build a better image of the interactions taking

place in this atypical chemosensory array.

An interaction was expected between $CheR_3$ and TlpT, as previous work from the laboratory suggests that TlpT provides the binding site for $CheR_3$. Fluorescence microscopy showed that $CheR_3$ was no longer localised with the cytoplasmic array if the last 5 residues of the C-terminal end of TlpT were deleted (Dr. M. Roberts, unpublished). Some form of steric hindrance is the most likely reason why this interaction could not be observed in the BACTH assay.

The BACTH system has also been successfully used to illustrate the effect of a mutation at the CheA/CheW binding interface. The approach could be extended to the other proposed inerface residues. It could also be used in other cases as a quick way to determine whether proposed contact residue mutations indeed affect the ability of the two proteins to interact.

3.6 Conclusions and future work

The BACTH method was successfully used to investigate protein-protein interactions involving the cytoplasmic array components. The interaction map obtained from the list of ineractions confirmed by this method reveals a highly connected network. All of the cytoplasmic array proteins interact with themselves and each other, with one notable exception: no interaction was found between $CheA_4$ and TlpC. To confirm the lack of interaction, other methods should be employed as a negative in the BACTH assay does not exclude the possibility of an interaction.

In the same manner that residues involved in the interaction between $CheA_3$ and $CheW_4$ were confirmed using BACTH, future work could be directed at revealing which domains are involved in the interactions identified in this chapter.

The method could be further applied to *che*Op$_1$ and *che*Op$_2$ proteins. The cAMP-based two-hybrid system is appropriate for studying molecular interactions involving membrane proteins [50], making the system easily applicable to the study interactions within the polar array. Moreover, the interactions between the *che*Op$_1$ proteins which are not expressed in a WS8N background under laboratory conditions could be revealed through this approach.

CHAPTER 4

Studies on the *R. sphaeroides* cytoplasmic CheA kinases

The two atypical histidine kinases from the *R. sphaeroides* cytoplasmic array have the P5 domain in common. While this domain is responsible for localisation with the cytoplasmic array, the P5 domains are not equivalent [90]. By swapping domains, it was shown that if both $CheA_3$ and $CheA_4$ have the $CheA_3$ P5 domain (A3P5), the two CheAs still localise to the cytoplasmic array. However, if both CheAs contained the $CheA_4$ P5 (A4P5) domain, only $CheA_4$ would localise to the array as normal, while most of $CheA_3$ was diffuse. This could be a consequence of structural differences between the two P5 domains. Work presented in this chapter focuses on the P5 domain of $CheA_3$ and $CheA_4$

4.1 CheA kinases

As discussed in the Introduction, bacterial chemotaxis relies on a two-component signal transduction pathway to integrate information from several chemoreceptors to tune swimming behaviour. The bacterial chemotaxis histidine kinase is CheA.

With a structure that is markedly different from ordinary histidine kinases, CheA proteins are placed in a category of their own. Typical CheA histidine kinases are modular and contain five domains, each having a dedicated role: histidine phosphotransfer (P1), response regulator binding (P2), dimerisation (P3), ATP binding (P4), and localisation (P5).

4.1.1 Atypical CheA kinases in *R. sphaeroides*

Atypical histidine kinases contain a subset of the domains shown for a typical CheA in Figure 4.1 or have additional ones. Organisms such as *R. sphaeroides*, *M. xanthus*, and *P. aeruginosa* encode atypical CheA kinases. Genomic analyses have revealed that up to 2.3% of all identified histidine kinases have the kinase (ATP-binding) and phosphotransfer activities on two distinct proteins that form a complex [117]. The *R. sphaeroides* genome encodes three chemotaxis systems, each locus encoding homologues of the *E. coli* chemotaxis system. *R. sphaeroides* does not have any CheZ homologues, which is typical of non-enteric bacteria. *che*Op$_1$ and *che*Op$_2$ encode for typical CheA kinases and are associated with membrane-embedded chemoreceptors, while *che*Op$_3$ encodes for two atypical CheA proteins – CheA$_3$ and CheA$_4$ – which are associated with cytoplasmic chemoreceptors. CheA$_3$ and CheA$_4$ each contain a subset of the domains found in a typical CheA. Neither can autophosphorylate, but CheA$_4$ can bind ATP and transfer a phosphoryl group onto the P1 domain of CheA$_3$. Hence the cytoplasmic cluster kinase is a *split* emzyme. In the genomic study by Wuichet *et al.*, 11 instances of split CheA sequences were found, compared to the 470 complete ones in their genome set.

CheA$_3$ only contains the P1 and P5 domains linked by a 794-amino acid sequence. Transfer of phosphoryl groups from CheA$_3$ to CheY$_1$, CheY$_6$, and CheB$_2$ was observed *in vitro* [74]. Signal termination in phosphorylation-based signalling net-

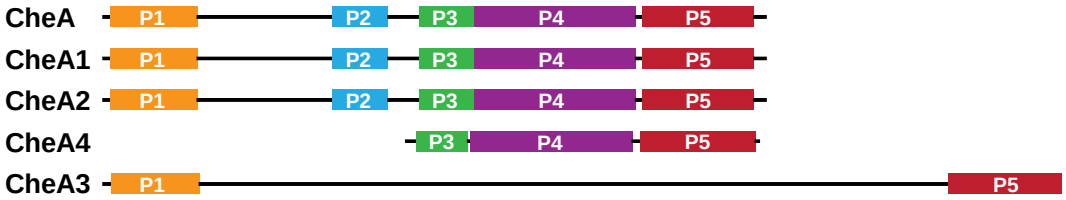

Figure 4.1: CheA domain structure.The *R. sphaeroides* $CheA_1$ (not expressed under laboratory conditions) and $CheA_2$ display the classical five-domain structure found in the *E. coli* CheA: P1 (yellow) contains the histidine substrate for phosphorylation; P2 (blue) interacts with response regulators; P3 (green) is involved in dimerization; P4 (purple) binds ATP; P5 (red) is the regulatory domain. Neither of the cytoplasmic *R. sphaeroides* CheAs contain all the P domains: $CheA_3$ only has domains P1 and P5 separated by a 794 amino acid sequence, while $CheA_4$ has domains P3, P4 and P5.

works occurs through some form of dephosphorylation. *R. sphaeroides* does not encode any homologues of known CheY-P phosphatases, but $CheA_3$ was shown to be a specific phosphatase for $CheY_6$-P. The phosphatase domain of $CheA_3$ was narrowed down to a 200 amino acid segment within the linker between the P1 and P5 domains [75]. Deletion of the phosphatase domain did not abolish phosphotransfer *in vitro*, but the strain containing this mutation was not chemotactic.

$CheA_4$ contains the P3, P4, and P5 domains and *in vitro* phosphorylation assays have shown that it can phosphorylate all other *R. sphaeroides* CheA proteins [74]. $CheA_3$ however cannot be phosphorylated by either $CheA_1$ or $CheA_2$ which have the required kinase domain. The reaction kinetics suggest that $CheA_3$ associates with a $CheA_4$ homodimer. It was suggested that the physical separation of $CheA_3$ and $CheA_4$ to the cytoplasm from $CheA_2$ at the polar cluster prevents undesired cross-talk within the network.

Both $CheA_3$ and $CheA_4$ have a P5 domain which most likely serves as a binding site for $CheW_4$. Scott *et al.* took a domain-swapping approach to investigate the specificity of the $CheA_3$ and $CheA_4$ P5 domains [90]. Results showed that each P5 domain was sufficient to determine the localisation of the two CheAs. How-

ever, even though the P5 domains are similar, they are not functionally equivalent. $CheA_3$ and $CheA_4$ containing the $CheA_3$ P5 domain, simultaneous localisation to the cytoplasmic cluster of both CheAs was observed. In contrast, if both proteins had the $CheA_4$ P5 domain, only $CheA_4$ localised with the cluster while $CheA_3$ was diffuse throughout the cytoplasm.

The CheA kinases play a key role in the integrity of the chemoreceptor arrays. Typical CheA kinases are indispensable for the structural integrity of their corresponding chemosensory arrays, as seen for the chemosensory array in *E. coli* or for the polar chemosensory array in *R. sphaeroides*. The cytoplasmic chemosensory array in *R. sphaeroides* has two atypical CheA kinases both of which contain a P5 domain and neither of which are crucial for cluster localisation [114]

Table 4.1: Currently available structures of CheA chemotaxis protein domains, deposited in the Protein Data Bank.

PDB ID	**CheA domain**					**Other proteins present**			**Organism**	**Method**
	P1	P2	P3	P4	P5	CheY	CheW	MCP		
1FWP	–	✓	–	–	–	–	–	–	*E. coli*	NMR
1A0O	–	✓	–	–	–	✓	–	–	*E. coli*	X-ray
1EAY	–	✓	–	–	–	✓	–	–	*E. coli*	X-ray
1FFG	–	✓	–	–	–	✓	–	–	*E. coli*	X-ray
1FFS	–	✓	–	–	–	✓	–	–	*E. coli*	X-ray
1FFW	–	✓	–	–	–	✓	–	–	*E. coli*	X-ray
2LP4	✓	✓	–	–	–	✓	–	–	*E. coli*	X-ray
3KYI	✓	–	–	–	–	✓	–	–	*R. sphaeroides*	X-ray
3KYJ	✓	–	–	–	–	✓	–	–	*R. sphaeroides*	X-ray
1TQG	✓	–	–	–	–	–	–	–	*T. maritima*	X-ray
2LD6	✓	–	–	–	–	–	–	–	*T. maritima*	NMR
1U0S	–	✓	–	–	–	✓	–	–	*T. maritima*	X-ray
1B3Q	–	–	✓	✓	✓	–	–	–	*T. maritima*	X-ray
1I58	–	–	–	✓	–	–	–	–	*T. maritima*	X-ray
1I59	–	–	–	✓	–	–	–	–	*T. maritima*	X-ray
1I5A	–	–	–	✓	–	–	–	–	*T. maritima*	X-ray
1I5B	–	–	–	✓	–	–	–	–	*T. maritima*	X-ray
1I5C	–	–	–	✓	–	–	–	–	*T. maritima*	X-ray
1I5D	–	–	–	✓	–	–	–	–	*T. maritima*	X-ray
2CH4	–	–	–	✓	✓	–	✓	–	*T. maritima*	X-ray
3UR1	–	–	–	✓	✓	–	✓	✓	*T. maritima*	X-ray
4JPB	–	–	–	✓	✓	–	✓	✓	*T. maritima*	X-ray
1I5N	✓	–	–	–	–	–	–	–	*S. typhimurium*	X-ray

The principal aims of this chapter are to:

- express and purify $CheA_4$ and the P5 domain from $CheA_3$
- perform biophysical characterisation of the P5 constructs
- perform initial work towards crystallisation of the P5 domains

4.2 Experimental approach

A separate approach was taken to the purification of the two P5 domains for crystallisation. The reason for this lies in their domain architecture. While $CheA_4$ contains the domains P3-5 close together in a highly ordered conformation, the P5 domain in $CheA_3$ is separated from the P1 domain by a disordered 794 amino acid domain. The disordered domain is likely to interfere with crystallisation – as suggested by the lack of success in crystallising full-length $CheA_3$ within the laboratory – hence it was decided to express the $CheA_3$ P5 domain (A3P5) separately with an N-terminal maltose-binding protein (MBP) tag.

The construct used to obtain the $CheA_4$ P5 domain (A4P5) for crystallography was wild-type $CheA_4$ itself as it contains domains P3, P4 and P5 very closely linked. The close vicinity with other structured domains should induce more stability to the P5 domain and make it suitable for structural studies. The fact that a crystal structure containing the *T. maritima* CheA P3-P5 domains already exists (1B3Q, see Table 4.1) is also indicative of the fact that having the three domains together might increase the chances of obtaining structural data through this method.

4.3 Results

4.3.1 Purification and characterisation of His_6-$CheA_4$

Wild-type $CheA_4$ was expressed with an N-terminal 6xHis tag from a pQE60 plasmid. Previous $CheA_4$ purification work [63, 73] was used as a starting point to develop a protocol that would yield protein suitable for structural studies.

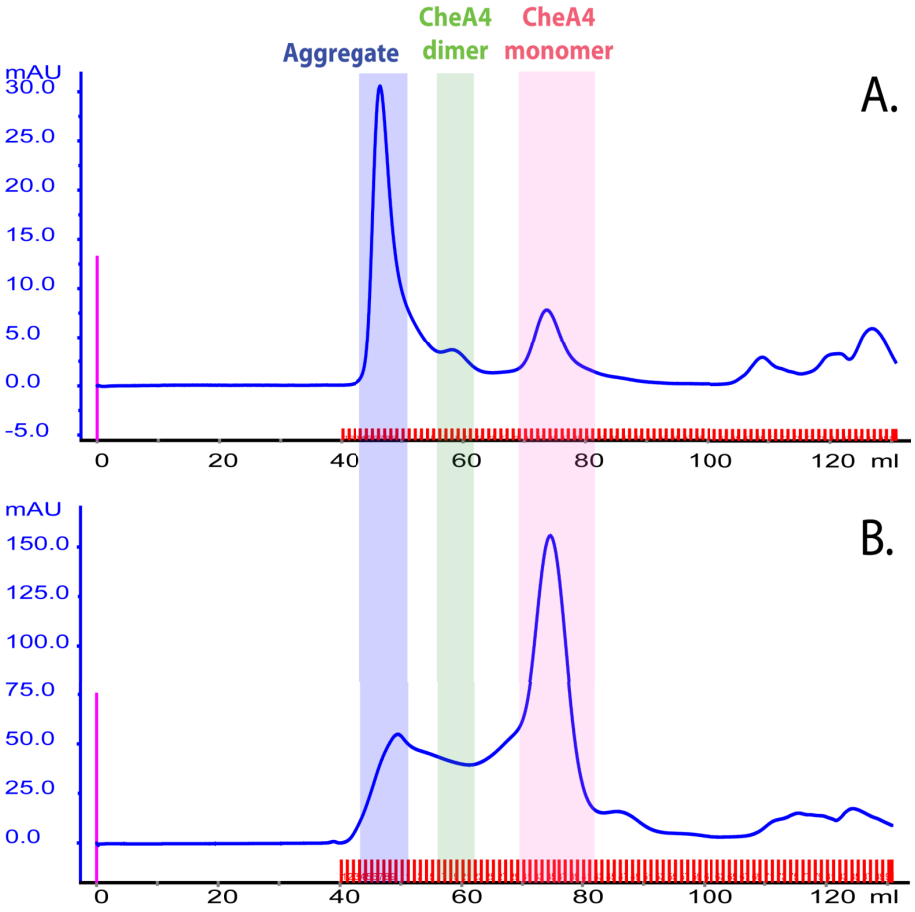

Figure 4.2: Chromatograms showing the size exclusion step of the His_6-$CheA_4$ purification. A. The chromatogram shows protein aggregate being predominantly formed when the preceeding affinity purification step was carried out by gravity column. B. The $CheA_4$ monomer is observed as the dominant species in the size exclusion chromatogram when the preceeding affinity step was carried out using the batch protocol.

It was found that performing the affinity purification step using a gravity column led to the elution of $CheA_4$ as an aggregate from the size exclusion column (Figure 4.2). The size exclusion step is necessary because the protein is not pure enough for structural studies after the affinity step. Switching to a batch protocol for the affinity step followed by a normal size exclusion step yielded protein of satisfatory purity for structural studies (above 90% as judged by SDS-PAGE, Figure 4.3).

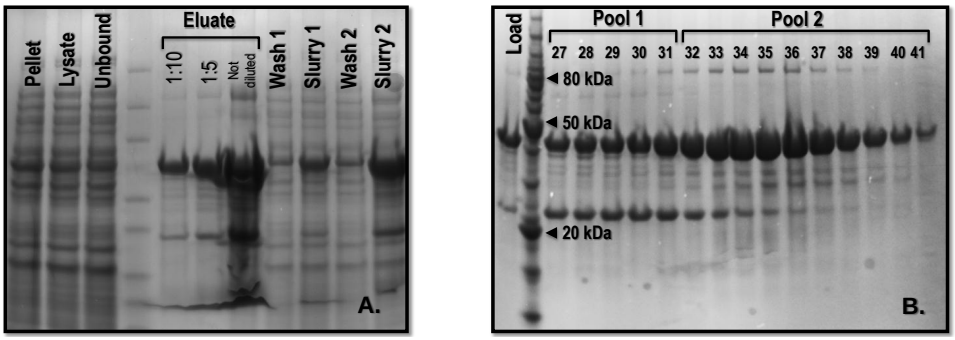

Figure 4.3: SDS-PAGE gels from the His_6-$CheA_4$ purification steps. A. SDS-PAGE gel containing samples from the affinity purification step done using the batch protocol. B. SDS-PAGE gel of samples from the fractions eluted in the size exclusion step. The fractions containing $CheA_4$ were collected into two pools according to their purity after the size exclusion step.

4.3.1.1 Biophysical characterisation

Electrospray ionisation mass spectrometry The purified His_6-$CheA_4$ was characterised by electrospray ionisation mass spectrometry (ESI-MS) to accurately determine the molecular weight of the protein and confirm its identity. The molecular weight measured by ESI-MS was 44508.70 ± 9.99 Da which is in good agreement with the one calculated using ProtParam (http://web.expasy.org/protparam/) from the amino acid sequence, 44483.1 Da.

Non-denaturing ESI-MS studies on $CheA_4$ Performing ESI-MS under non-denaturing conditions can provide information on the oligomerisation state of a protein [88,

62].

The non-denaturing conditions for the experiment are provided through the use of volatile protein buffers, such as ammonium acetate. The very specific buffer conditions required are not compatible with all proteins, and for some, non-denaturing ESI-MS is not a useful tool.

Initially met with skepticism, the method relies on the detection of dehydrated and ionised protein molecules in vacuum. Protein stability in the gas phase has been suggested to be due to the van de Waals attraction between the amino acids. This attractive force persists in the gas phase and is thought to provide enough stability to retain interactions [62].

Here, non-denaturing ESI-MS was performed to test the oligomeric state of $CheA_4$ proposed by analytical ultracentifugation (AUC). The protein spectrum obtained by non-denaturing ESI-MS contains multiple charged ions produced by protonation in solution. Protein oligomerisation can be distinguished from monomeric species, as the molecular weight of the species present can be calculated from the series of m/z values.

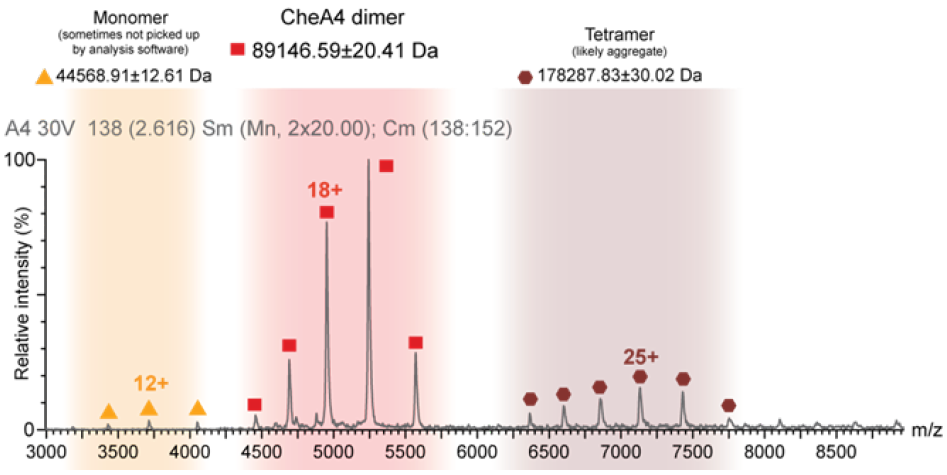

Figure 4.4: ESI-MS data for $CheA_4$. The spectrum contains three sets of peaks, corresponding to the monomer (yellow), dimer (red) and tetramer (brown).

The non-denaturing ESI-MS spectrum for $CheA_4$ is presented in Figure 4.4. Three sets of peaks (charge distributions) can be distinguished in the spectrum. The molecular weight of the species that gave rise to a distribution was calculated as described above. The dimer (89146.59±20.41) charge state distribution has the highest intesity and is therefore the dominant species in the gas phase. A tetramer and monomer species were also observed, but their relative peak intensities were much lower (under 20% of the dimer). The calculated molecular weight for the monomer was in good agreement with the value obtained through the denaturing protocol.

Differential scanning calorimetry (DSC) The thermal stability of a protein can be assessed using differential scanning calorimetry, a routine biophysical characterisation technique that measures the melting temperature of a protein sample. DSC was carried out to determine if there is scope for the thermofluor assay [59] to be further used to screen for improved crystallisation conditions. A measured melting temperature of 42°C indicated that the solution conditions could be improved to achieve higher stability [47].

4.3.1.2 Buffer optimisation

The measurement of the melting temperature for $CheA_4$ by DSC suggested that its stability could be improved. Important factors that influence protein stability are the buffer components. To find buffer conditions with a stabilising effect on $CheA_4$, a thermofluor assay was carried out [59]. It is a high-throughput method that uses a fluorescent dye that binds to molten globules and thermal denaturation intermediates. The experiment is performed in a real-time PCR machine, where the temperature is increased incrementally and the fluorescence signal is recorded. The

resulting fluorescence trace can be used to determine melting temperatures.

The experiment can be performed in 96-well plates, making it quick and relatively cheap to screen for additives that improve protein stability. Under the assumption that stability is proportional to the melting, additives that increase the melting temperature the most are used to formulate a new buffer.

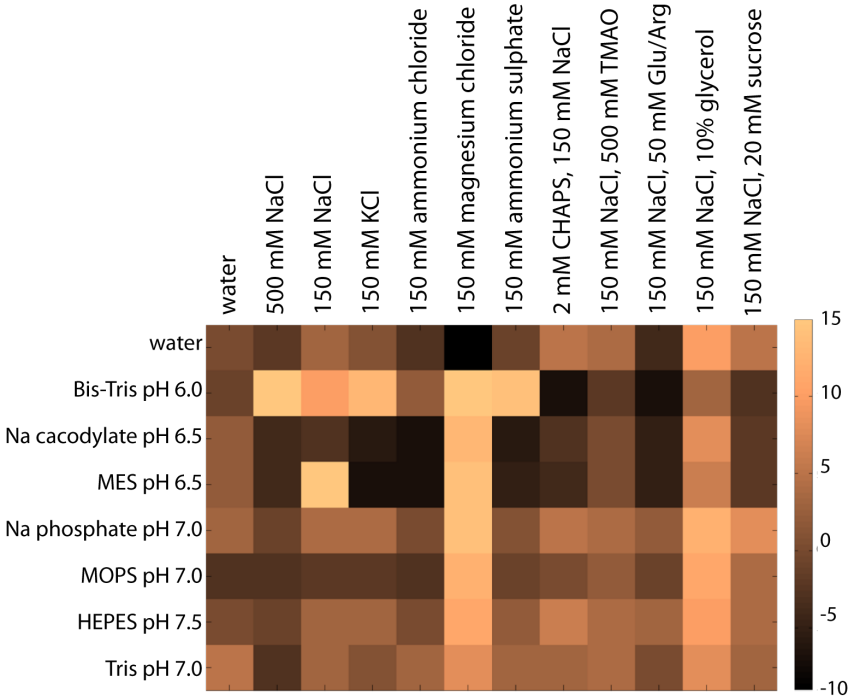

Figure 4.5: Results of the thermofluor assay. The plot is a representation of the variation in melting temperature as a result of using a certain additive combination. Light colours corelate with positive changes and dark colours with negative ones. The values are calculated relative to the protein being dissolved in water (well A1, top left hand corner).

Figure 4.5 displays the result of the thermofluor assay. Seven buffer conditions at various pH values were tested in combination with eleven buffer additives. Adding 150 mM $MgCl_2$ as well as using a combination of 150 mM NaCl and 10% glycerol have a pronounced positive influence regardless of buffer choice. Switching to Bis-Tris buffer could also be considered, except the addition of glycerol and NaCl to Bis-

Tris buffer loses its stabilising effect. Consequently, it was decided to use a 50 mM Tris buffer pH7.5, with 150 mM NaCl, 150 mM $MgCl_2$, and 10% glycerol.

4.3.1.3 Crystallisation trials

The crystallisation trials were performed at the Department of Biochemistry Crystallisation Facility. Protein from two different purification batches was used to set up 96-well crystallisation plates using the sitting drop method. Five crystallisation condition screens were used:

- ***JCSG+*** (Molecular Dimensions) is a broad sparse matrix covering PEG salt conditions, organics and polyalcohols, and a broad pH range.
- ***Morpheus***(Molecular Dimensions) contains a selection of low molecular weight ligands found to promote the formation of crystal contacts.
- ***PACT Premier*** (Molecular Dimensions) is a three-part multi-component screen against pH, cations, and anions. It is intended for use together with JCSG+ for initial screening.
- ***PEG/Ion HT*** (Hampton Research) is a systematic screen that uses PEG 3350as the precipitant against a variety of ions, ionic strength, and pH.
- ***JCSG Core Suite*** (Quiagen) provides a total of 384 crystallisation conditions that provided the highest hit rates in initial screening.

Precipitates and clear drops were obtained across all screening plates. Microcrystals and small irregular crystals (Figure 4.6A,B), both too small for harvesting, were obtained in a number of conditions, mostly across JCSG plates. Needle-shaped crystals were harvested from the E3.2 well from the JCSG+ plate. Unfortunately the harvested crystals did not diffract.

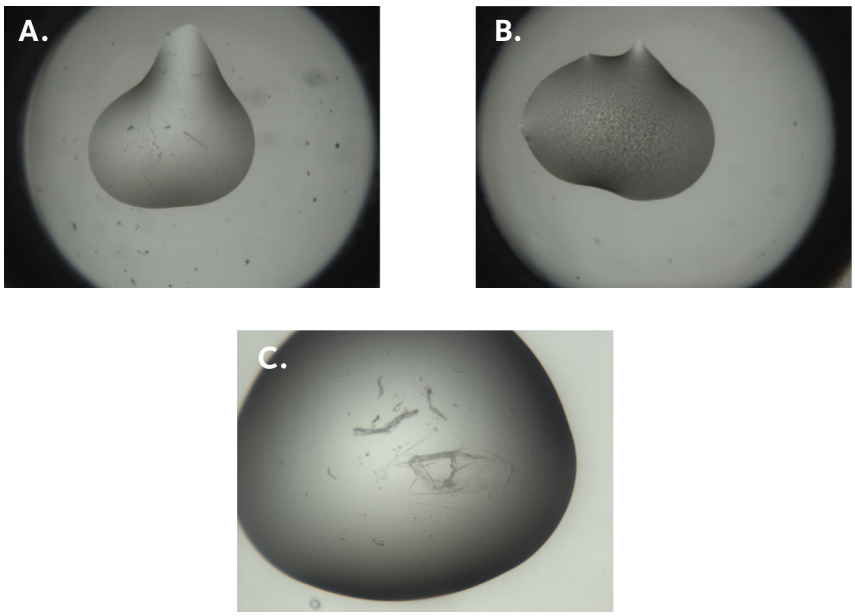

Figure 4.6: Crystallisation results. **A.** Small irregular crystals, representative image; **B.** Microcrystals, representative image; **C.** Harvested crystals from JCSG+ well E3.2.

4.3.2 Expression and purification of the MBP-A3P5 domain

This work addresses the role of the structural differences between the P5 domains of $CheA_3$ and $CheA_4$. It was hoped that structural data on the P5 domains would provide information on $CheA_3$ and $CheA_4$ association with the cytoplasmic cluster and the implications for stoichiometry and architecture of the array, as well as the function of the two proteins.

4.3.2.1 Genetic construct

A plasmid containing the P5 domain of $CheA_3$ cloned into the pMAL-p2G vector (New England Biolabs) was used. The construct encodes an N-terminal maltose binding protein (MBP) fusion. The pMAL vector is known to yield reliable expression of the fusion protein. Fusion to MBP has been shown to enhance the solubility of proteins and the elution process is gentle, not requiring detergents or harsh de-

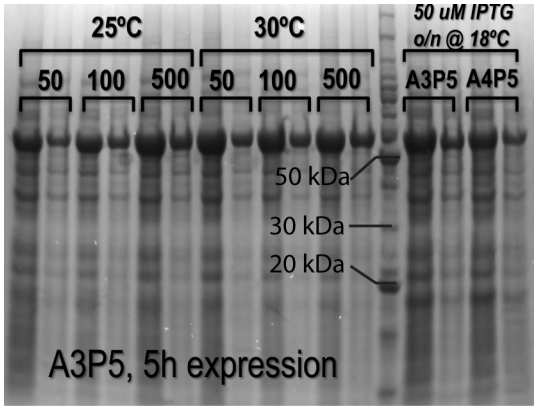

Figure 4.7: Expression trials of the P5 domain from $CheA_3$. Expression was tested for A3P5 at three IPTG concentrations (50 μm, 100 μm and 500 μm final concentration) and three temperatures (30°C, 25°C and 18°C). The length of expression was either 5 hours (for 30°C and 25°C expression) or over night (18°C, 25°C and 30°C). The overexpressed fusion proteins are 58 kDa.

naturants.

4.3.2.2 Expression trials

Since the pMAL constructs had not been expressed before, the first step taken was to conduct expression trials. A series of IPTG concentrations, expression temperatures as well as periods of expression were tested.

The SDS-PAGE gels used to assess expression of the P5 fusions are displayed in Figure 4.7. All conditions yielded overexpression of the MBP fusion proteins to some degree. The conditions with the highest level of overexpression which were then chosen for large scale expression are: 5 h expression at 25°C in the presence of 500 μm IPTG for the A4P5 domain and over night expression at 25°C in the presence of 500 μm IPTG for the A3P5 domain.

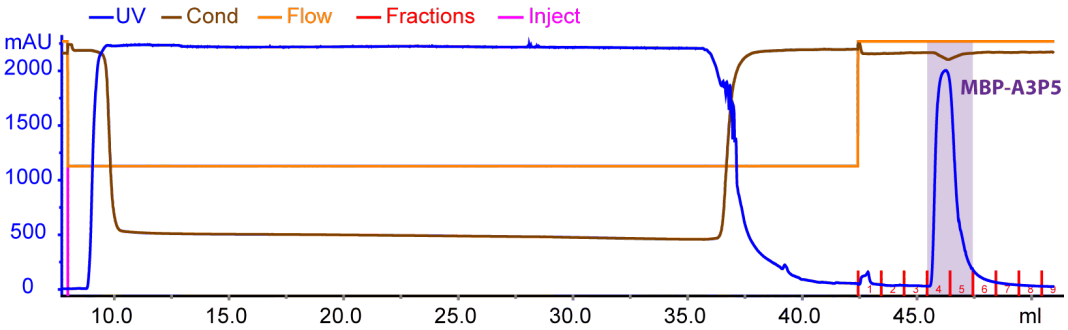

Figure 4.8: Purification of the overexpressed A3P5-domain fusion protein. The chromatogram shows a single peak in the elution step, corresponding to the purified MBP-A3P5 fusion protein.

4.3.2.3 Purification

Purification of the overexpressed P5-domain fusion protein was performed using a pre-packed MBP-Trap column (GE Healthcare), using the manufacturer's instructions. The chromatogram in Figure 4.8 shows the single peak that was eluted from the MBP column. Fractions 4 and 5 were pooled and used for subsequent optimisation of tag cleavage.

4.3.2.4 Tag cleavage

Between the A3P5 domain and the MBP tag, there is a GenenaseI cleavage site. Trial digestions were run, testing a range of parameters: the concentration of GenenaseI, incubation time, and incubation temperature Figure(4.9). Even though the digestion was complete even after the shortest digestion time and at the lowest temperature (2 hours at 4°C), there was a white precipitate formed in all cases. SDS-PAGE analysis of the supernatant showed that it contained just MBP.

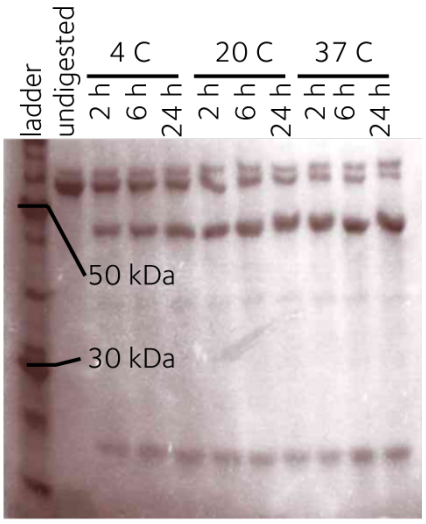

Figure 4.9: Digestion trials of the overexpressed A3P5-domain fusion protein shown on an SDS-PAGE gel. Temperature and digestion time variation are illustrated, using the same concentration of GenenaseI. The fusion protein is 58 kDa, the A3P5 domain is 15 kDa and the MBP tag is 43 kDa

4.4 Dicussion

$CheA_4$ was successfully purified to crystallography standard. Non-denaturing mass spectrometry revealed it to be a dimer, which is consistent with the presence of the dimerisation domain (P3). One round of buffer optimisation for crystallisation was carried out but no difracting crystals were obtained under the conditions tested. Further buffer optimisation as well as more crystallisation conditions could be tested in further attempts to crystallise $CheA_4$.

The A3P5 domain was successfully purified with an N-terminal MBP fusion. Tag cleavage caused the protein to precipitate, indicating that the domain is likely very unstable on its own. Future efforts could be directed at construct design level towards optimising the boundary of the domain.

CHAPTER 5

Studies on the adaptation pathway of the cytoplasmic chemosensory cluster

5.1 Introduction

The evolutionary conservation of CheB and CheR in most bacterial chemotaxis systems suggests that methylation-based adaptation is an effective fitness strategy for a variety of organisms and conditions.

R. sphaeroides has two CheB proteins, one in each of the two expressed chemotaxis operons, and three CheR proteins, one in each of the three existing chemotaxis operons. Despite the presence of the methyltransferases and methylesterases in *R. sphaeroides*, there is no evidence in the literature confirming that methylation-based adaptation occurs. Nevertheless, it is clear that some form of adaptation does occur in *R. sphaeroides* as cells are able to respond to step changes in attractant and then resume the initial swimming pattern.

It is generally assumed that the products of genes encoded in the same operon are functionally linked. However, the localisation of fluorescent fusions of the ex-

pressed adaptation proteins show that $CheR_2$ and $CheR_3$ are found with the polar and the cytoplasmic arrays, respectively, while $CheB_1$ and $CheB_2$ are found diffuse throughout the cytoplasm [115]. Deletion of $CheB_1$ causes loss of chemotaxis, but deletion of $CheB_2$ only causes a reduction [64, 28].

In addition, modelling and *in vitro* phosphotransfer suggest $CheB_2$ performs reverse phosphorelay linking the cytoplasmic and polar arrays, but it remains unclear what the purpose of having diffuse $CheB_1$ might be [76, 107]. Taken together, the genomic and cellular localisations of the adaptation proteins suggest that $CheB_1$ and $CheR_2$ are likely involved in polar chemosensory array adaptation, while $CheB_2$ and $CheR_3$ are likely involved in cytoplasmic chemosensory array adaptation.

Overall, there is evidence to suggest that methylation based adaptation could take place in *R. sphaeroides*. Each chemosensory array has available the basic requirements for such a mechanism to take place: chemoreceptors (transmembrane or cytoplasmic), a putative methyltransferase, and a putative methylesterase.

5.2 Experimental approach

Recent *in vitro* work in the laboratory has successfully identified three putative sites that could undergo methylation and deamidation in the cytoplasmic chemoreceptor TlpT, as well as predicted an additional site through bioinformatic analysis. The first part of this chapter aims to characterise the swimming behaviour strains containing genomic mutations to alanine of the methylation sites proposed for TlpT. The results will provide a clearer view, supported by *in vivo* evidence, of whether the proposed residues are indeed involved in methylation-based adaptation.

The second part of this chapter focuses on the adaptation proteins and aims to characterise the swimming behaviour of strains containing pairs of CheBR deletions.

These data will be analysed together with existing adaptation protein mutant strain data recently obtained in the laboratory, as well as the TlpT mutant data, to produce an updated model for adaptation in *R. sphaeroides*.

Swimming behaviour is assessed at population level under steady state conditions and at single cell level under dynamic conditions as tethered cells are challanged with increasing and decreasing gradients of attractant. Used together, these methods will provide detailed information about the functionality of the adaptation pathway in the various mutant strains.

5.3 Swimming behaviour under steady state conditions

5.3.1 Previous work

In a homogenous environment *R. sphaeroides* explores the environment using a random walk. Early work on characterising the swimming pattern of bacteria under steady state conditions (free swimming) *E. coli* includes manual tracking of single cells [14] and recording of whole swimming populations [6]. Tumbling events were then identified within the tracks extracted from the recordings. This was achieved either by hand, based on observable patterns [86, 72], or by applying heuristics such as defining speed and angle change limits, the results of which would be verified against manual classification [6].

The initial whole population tracking and heuristic classification approach applied to free swimming *R. sphaeroides* used VHS tape with a 24 frames per second recording framerate and software which is now obsolete [86, 72]. Because *R.*

sphaeroides uses stops as opposed to active tumbles as it is the case with *E. coli*, the angle change following a reorientation event is not as great as that observed in *E. coli*. For this reason, in combination with the low frame rate, this method was unable to detect the transient stops which *R. sphaeroides* uses to change swimming direction, hence was not able to discriminate between wild-type and for example, a $\Delta cheR_3$ strain which is visibly stoppier than wild-type.

A bead assay was developed to explore *R. sphaeroides* free swimming [22]. An optical bead was attached to the shortened flagellum using antibodies. The movement of individual beads was recorded and position change was used to calculate rotational velocity at each time point, which was then used to identify stopping events. Load on the motor is significantly less in this case compared to the tethered cell assay, but most importantly the rate of data capture was significantly improved from the past VHS recordings. The higher frame rate allowed for the observation of transient stopping events, but the analysis method is time consuming as only one cell can be analysed at a time.

Most recent work makes use of the same high frame rate video recordings (50 Hz) but has brought significant improvements in the analysis software. Analysis software development by Rosser *et al.* produced a simple and effective protocol for the removal of spurious tracks, followed by an analysis based on a two-state hidden Markov model [85]. The model uses mutant strains exhibiting only run events and only reorientating events as references, generating an empirical prior. This method is more robust to noise and introduces less systematic bias than the heuristic method.

The work described here performs 50 Hz video recordings of swimming bacteria, with tracking and analysis carried out using the protocol described by Rosser *et al.* [85]. The mutants used here focus on adaptation in *R. sphaeroides*, particularly

adaptation at the cytoplasmic chemosensory array, and adds to the extensive chemotaxis mutant characterisation performed by this method in our laboratory [28].

5.3.2 Data collection and processing

R. sphaeroides cultures were grown to OD_{700} =0.4-0.6 and used to fill rectangular capilary tubes. Free swimming cells were recorded at 50 Hz for 2 min, using 20x magnification. For each strain, data were acquired on at least three different days with at least three videos recorded for each day, giving at least nine repeats across a minimum of three biological replicates.

Video data processing was carried out in three stages: tracking, clean-up, and classification.

5.3.2.1 Tracking

The tracking software [116] identifies and records the positions of objects – image areas that are darker or lighter than the average surrounding intensity – in each frame of the video using the Niblack algorithm with user-defined parameters for each set of videos. This information is used to produce tracks by connecting objects from consecutive frames using probability densities. From the track information, the distance travelled between frames and the change in direction are calculated [85].

5.3.2.2 Data clean-up

The tracking software will produce tracks for the various objects that were identified. Not all of these tracks however belong to bacteria and not all of them are

useful for the subsequent analysis. Hence the data need to be cleaned up and this was achieved using the censoring protocol described by Rosser *et al.* [28, 85].

All strains were censored using the same parameters, except for the non-motile reference strain (JPA467) which was treated separately. Four types of erroneous tracks are typically found within each dataset, each of them with a potential for introducing errors during the subsequent track analysis. The types of tracks discussed in this chapter are exemplified in Figure 5.1. The categories of unusable tracks are discussed below, explaining how they are likely to arise and how the censoring is achieved.

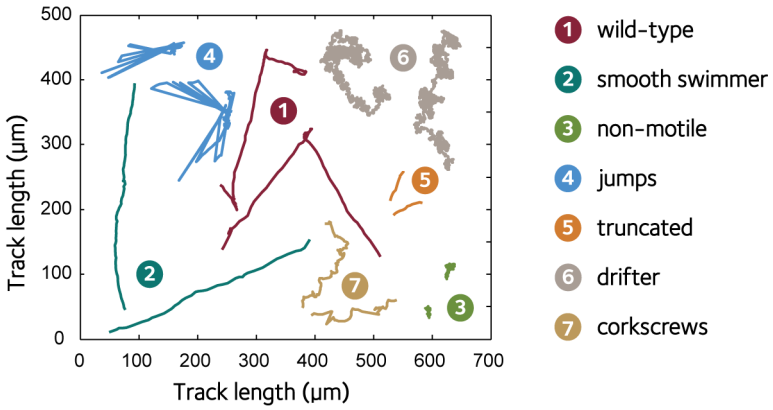

Figure 5.1: Types of free swimming tracks showing 1) wild-type, 2) smooth swimmer, 3) non-motile, and the types of unusable tracks 4) jumps, 5) truncated, 6) drifter, 7) corkscrews. Non-motile tracks belong to the unusable tracks category if the strain is motile.

- ***Jumps*** occur as a result of an error in the tracking program when adjacent tracks are mistakenly joined together. These tracks are removed by setting a top speed cutoff. The mean swimming speed of *R. sphaeroides* is 40 μm/s, so any frame-to frame transition that is much greater than this is likely due to a jump. The top speed cutoff parameter is hence set to double the mean swimming speed to account for variability.

- ***Truncated*** tracks refer to very short tracks that appear due to the bacterial cell

spending very little time in the focal plane (the focal depth for the objective used here is 5.8 μm). The presence of these tracks in the final dataset for analysis would bias the results towards runs, as during the short time the cell spends in the focal plane, it is typically swimming, not tumbling. Truncated tracks were removed from the dataset by setting a minimum number of frames for each track.

- ***Non-motile or drifters*** are cells that are dead or have lost their flagellum during the sample preparation process. The majority of these cells remain stationary throughout the video due to the absence of flow. Drifters arise as a non-motile cell is pushed around by motile cells and typically doesn't move very far. The persistence of these tracks in the final data set has the potential to skew summary statistics of track speeds and angle changes, as well as time a cell spends stopped. The minimum bounding radius (MBR) is defined as the radius of the smallest circle needed to include the whole track.

- ***Corkscrew*** tracks are formed when the moving object is off-centre in its movement, *e.g.* when the moving object is a dividing cell. In such cases, the resulting track has a characteristic zig-zag shape and could be erroneously classified as a sort series of runs and stops, thereby skewing the track analysis data. To remove these tracks, the angle change in each track is quantified by calculating the median absolute curvature (MAC). A track consisting of many stops or a drifting cell will have the highest MAC values, followed by corkscrew tracks, wild-type, and smooth swimmers, respectively. The tracks are sorted and the most tortuous are removed. Thus the MAC value can be used to censor for corkscrews, provided the majority of drifters have already been removed.

The role of the censoring steps is to remove enough of the polluting tracks without

removing significant numbers of useful tracks. Thus chosing the appropriate parameter for each censoring step is important, particularly in the case of the MBR when analysis tracks that are stoppier than wild-type. In this case, too high an MBR will remove the most stoppy tracks along with the drifters.

A tracks data set for a particular strain can be visualised as a 2D histogram of the tracks' mean absolute framewise angle changes (MAC) against their normalised effective means speed (NEMS) [66]. Such a plot will have the average normalised speeds increasing along the x axis and tortuosity increasing along the y axis. Hence, tracks from non-motile strains are expected to be concentrated in the top left hand region of the plot (slow and tortuous), while tracks from smooth swimmers are expected to be found in the bottom right hand corner (fast and not tortuous).

It was shown that a population consisting of self-propelled particles and non-motile diffusing particles has a well-separated bimodal distribution in the MAC-NEMS plot [66]. An example would be the MAC-NEMS distribution of wild-type tracks before MBR censoring (Figure 5.2C). MAC-NEMS plots are a useful tool for track population visualisation particularly while carrying out the censoring process.

Thus after removing jumps from a wild-type tracks dataset, the corresponding MAC-NEMS plot will have two clusters 1) high MAC and low NEMS (top left), and 2) low MAC and high NEMS (bottom right) (Figure 5.2C). In the case of the non-motile reference strain, only the cluster at high MAC and low NEMS (top left) is present (Figure 5.2A). Hence an MBR parameter that removes a significant proportion of the non-motile tracks from the non-motile strain is chosen (Figure 5.2B). When applied to a motile strain, this parameter should remove most tracks corresponding to drifting and non-motile cells (Figure 5.2D). MBR is chosen to exclude up to 99% of tracks from the non-motile reference strain [28]. Further removal of tracks less than 1 s and of the 10% most tortuous tracks tightens the tracks distribution in the

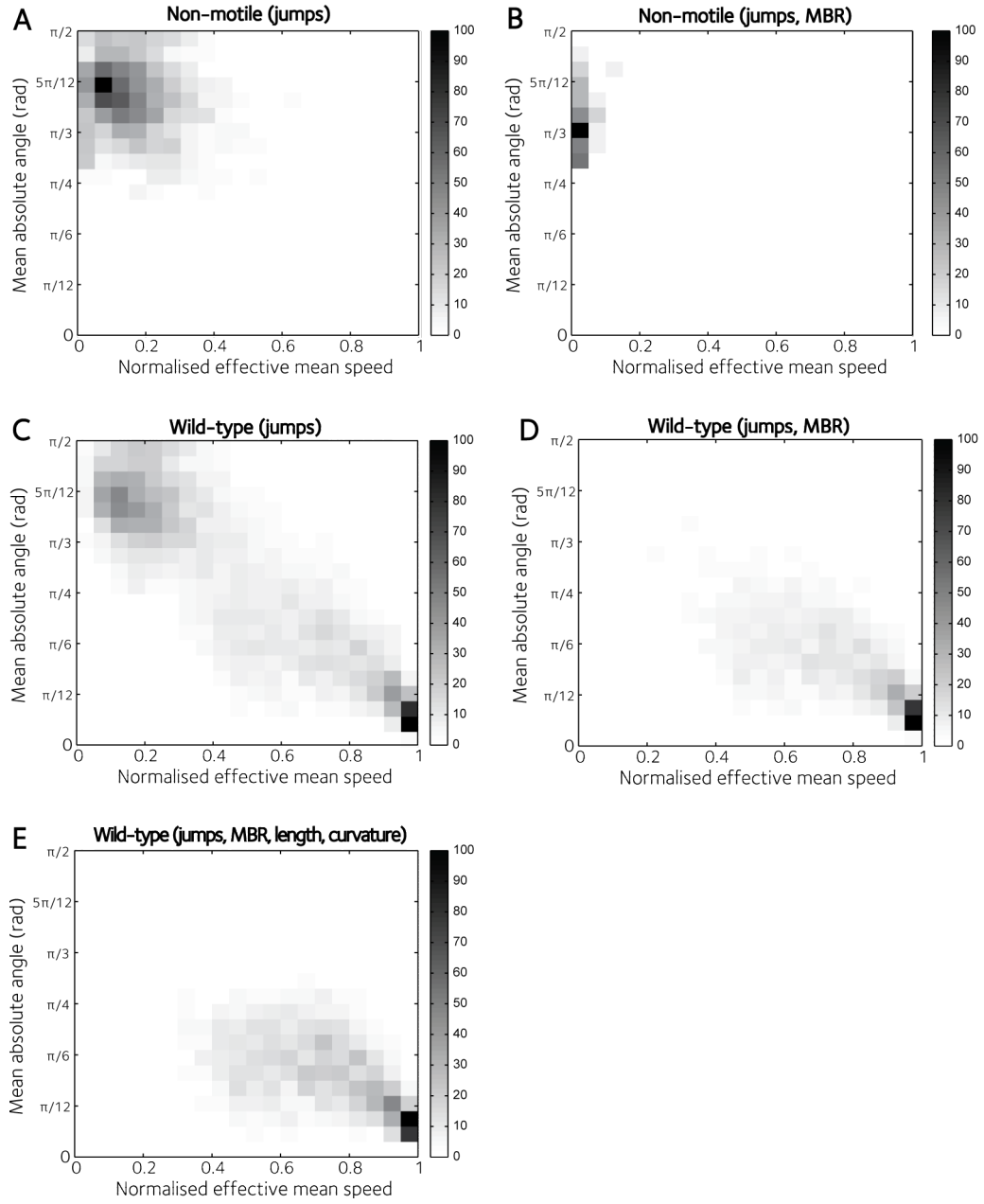

Figure 5.2: The MAC-NEMS space occupied by non-motile tracks **A.** with jumps removed and **B.** after censoring all remaining tracks with MBR=8 μm, and of wild-type tracks **C.** with only jumps removed, **D.** after censoring the remaining tracks with MBR=8 μm and **E.** after also removing tracks that persist for less than 1 s as well as the 10% most tortuous tracks. The plots are 400 bin two-dimensional histograms.

MAC-NEMS plot around the bottom right region (Figure 5.2E).

The parameters used for each censoring step are listed in Table 5.1 in the order in which they were applied. This censoring protocol was applied to all mutant strains studied in this chapter, as well as wild-type and smooth swimming reference strains. Lists detailing how many tracks have been removed at each censoring step can be found in Table 5.4 for the TlpT methylation site mutants and Table 5.7 for the adaptation protein mutants.

Table 5.1: Censoring parameters for sets of tracks from motile strains

Parameter	**Cells removed**	**Value**
Top speed cutoff	Jumpers	80 μm/s
Minimum track length	Truncated	50 frames (1 s)
MBR	Non-motile and drifters	8 μm
Percentage most tortuous	Corkscrews and remaining drifters	10%

The numbers in the tables detailing each censoring step show that censoring removes up to 90% of all tracks obtained for each strain, leaving a dataset of at least 400 tracks available for further analysis. It is interesting to note that there is a broad range of percentage of tracks left for analysis after censoring – form 12% for wild type to 4% for the TlpT double mutant analysed in this chapter. It has been proposed that this is related to the motility of each strain [28].

5.3.2.3 Classification

Following the protocol of de Beyer [28], censored data sets were classified in two ways. First, if the strains are sufficiently different, this will be obvious by eye. To this end, MAC-NEMS plots and representative sets of randomly-selected tracks for each strain were compared by eye. Generally speaking, due to the noise present in these data sets, if a difference cannot be picked up by eye, it is unlikely that further statistical tests will find them different [28].

Second, runs and stops were identified within each track from all data sets using published software by Rosser *et al.* [85]. Data sets from known non-motile and smooth swimming strains were used to determine whether each frame transition from a mutant strain data set is part of a run or a stop. All motile strains were classified using this protocol and then compared using summary statistics of the tracks. The distribution of the total fraction of time spent stopped was chosen to compare the free swimming behaviour of mutant strains.

5.4 Inferred motor behaviour under dynamic conditions

Bacteria propel themselves through liquids by individual rotation of their flagellar filaments, as argued by Berg and Anderson in 1973 [13]. Clear experimental proof of this was produced a year later by Silverman and Simon who managed to tether *E. coli* cells to a microscope slide using antiflagellin antibodies and record their rotational movement [95]. Larsen *et al.* further made the link between changes in flagellar rotation and the chemotactic response [58]. The rotating cells would be recorded onto video tape at around 20 frames per second and the rotation pattern would be analyseld in real time one cell at a time.

Since the 70s, the tethered cell method remained the most reliable and accurate way to measure the chemotactic response of bacteia at a single cell level. The tethered cell protocol has cetainly been refined over the decades, with the construction of flow chambers, use of high speed digital recording, and development of analysis software capable of clearly distinguishing the rotational behaviour. Current video processing and analysis software enables the simultaneous analysis of all tethered cells in a field of view, with rotation traces being used to characterise the swimming

behaviour of the cell throughout the experiment [56].

5.4.1 Data collection and processing

R. sphaeroides cultures grown to OD_{700} =0.4-0.6 were treated with chloramphenicol to prevent growth and resuspended in motility buffer. The cells were attached to a glass slide using anti-flagellin antibody. After an incubation period, the slide was fitted into a flow chamber. The cells equilibrated in motility buffer were exposed to an increasing gradient of attractant (propionate), followed by a decreasing gradient hence returning to the initial bufer conditions (Figure 5.3A). The behaviour of the cells was recorded using a high-speed video camera (100 Hz).

The videos were processed to extract information on the rotating cells. This was achieved using the BRAS software [56] which allows the user to select individual rotating cells for analysis. The raw data obtained contains a high level of noise. To get around this issue, the raw data processing software, Click&Mean [56], uses a Fourier transformation to smooth the data and displays a density map of the frequencies composing the rotational signal. Each density map thus corresponds to the rotational behaviour of a single bacterial cell, with the dominant components of the rotational signal coloured in increasingly dark shades of blue. Click&Mean uses a 128 frame smoothing window, meaning that at a 100 Hz acquisition frame rate, events under 1.28 seconds *i.e.* transient stops are lost through smoothing.

The output from Click&Mean does not classify cells, hence this is done by hand for each cell. The advantage of using the Fourier transform is that the intervals when the cell is rotating and when it is not are very clear and distinguishable. The cells are then classified according to previously described categories [8, 55, 71, 28], detailed below:

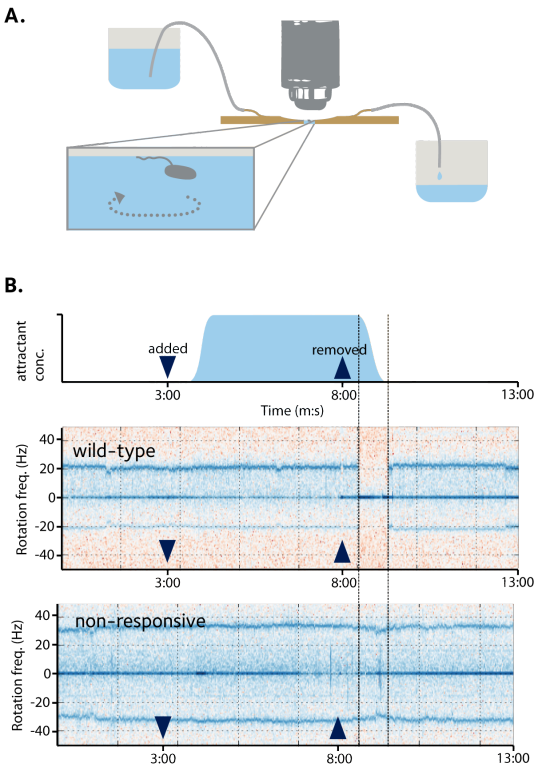

Figure 5.3: Overview of the tethered cell assay.**A.** Experimental setup for tethered cell analysis. Bacteria are tethered inside the flow cell (brown). Flow is created through the pressure difference between the source (higher) and the waste containers (lower). **B.** Examples of the wild-type and non-responsive tethering phenotypes (bottom two plots) shown in relation to the environment experienced by the tethered bacteria (top). For wild-type, when the decrease in nutrient is sensed, the chemotaxis pathway is activated and the flagellar motor is stopped to effect a change in swimming direction. Due to their ability to adapt, the wild-type cells only remain stationary while they are experiencing the unfavourable environmental conditions of a decresing attractant gradient, resuming rotation as they adapt. A non-responsive cell will not cease rotation while experiencing the decreasing gradient of attractant. The dotted lines highlight the adaptation stop interval characteristic of a wild-type cell.

- ***Adaptive*** cells cease rotating upon removal of attractant but adapt and restart rotating before the end of the experiment. Wild type cells are part of this category.

- ***Responsive*** cells stop rotating upon attractant removal and do not restart by the end of the experiment.

- ***Non-responsive*** cells show continuous rotation throughout the experiment, not responding to the addition or removal of attractant.
- ***Stoppy non-responsive*** cells do not react to addition or removal of attractant, but spend more time stopped, a behaviour visible in the Click&Mean output. The density maps show numerous clear 5-30 s gaps.
- ***Inhibited*** cells start rotating on addition of propionate, but some of these cells cease rotation before propionate is removed, some do not respond to propionate removal, and others are able to adapt.

Once the cells have been classified, the time it takes adaptive cells to resume rotation (adaptation time) can be measured using Click&Mean. The full adaptation time is defined by Kojadinovic *et al.* [56] as the time interval it takes the cell to resume full rotation. Thus, the tethered cell assay can provide information about the frequency of certain chemotaxis phenotypes and the duration of adaptation periods.

5.5 Investigation of putative methylation sites in TlpT

5.5.1 The discovery of putative methylation sites in TlpT

Previous work in the laboratory has identified four methylation sites in the *R. sphaeroides* cytoplasmic chemoreceptor TlpT: E296, E478, Q485, and E289 [28] (Figure 5.4). The first three of the sites were identified using tandem mass spectrometry (MS/MS) of purified TlpT in the presence of adaptation proteins, while the fourth, E289, was predicted using bioinformatics.

For the MS/MS experiments, *R. sphaeroides* was treated as a black box that could methylate and deamidate TlpT. The chemoreceptor was expressed in *R. sphaeroides*

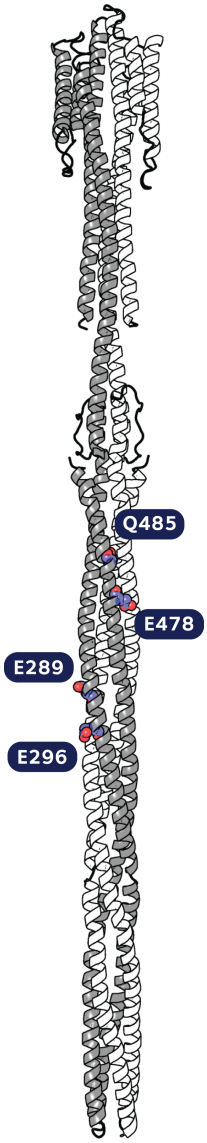

Figure 5.4: Putative methylation sites in TlpT. The proposed methylation sites (E296, E478, Q485, and E289) are highlighted mapped onto a model of TlpT by Dr. James R. Allen. The model was obtained by threading the TlpT sequence onto the *E. coli* serine chemoreceptor, Tsr. TlpT is represented as a dimer with the individual monomers coloured in white and grey. The methylation sites are labelled on the grey monomer.

and isolated using affinity chromatography. Three background strains were used for expression: 1) wild-type, which was expected to allow methylation and demethylation of any Q or E site, 2) a strain lacking all methyltransferases ($\Delta cheR_{1,2,3}$), which was expected produce no methylation on either E or Q sites, but full deamidation

of Q sites, and 3) a strain lacking all methylesterases ($\Delta cheB_{1,2}$), which was expected to methylate E sites but not the Q sites since they cannot be deamidated. The MS/MS results revealed E296 and E478 to be methylated. Q485 was the only residue found to be deamidated but no methylation was found at this site, even though it is correctly in frame with E478 to be modified by the adaptation pathway.

Using the locations of the sites identified by MS/MS and their relative position versus the known sites in *E. coli* Tsr, the fourth site, E289, was predicted (Figure 5.5). The corresponding in-frame methylation site in Tsr is E304, which showed the lowest levels of methylation in similar MS/MS experiments. It was thus suggested that the two sites are likely to have the same reactivity and were either not methylated to a great extent, or demethylated quickly during the isolation process.

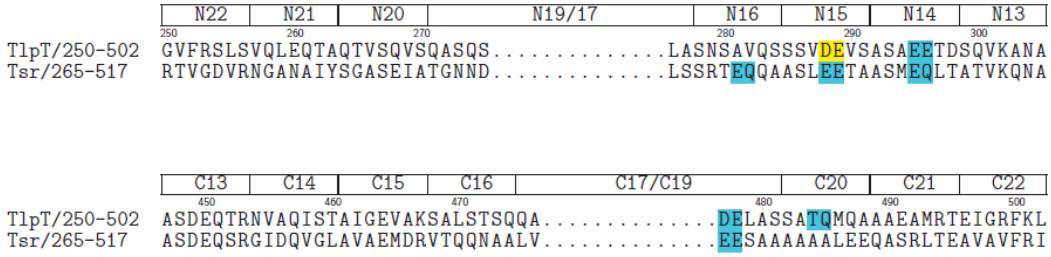

Figure 5.5: Alignment of *R. sphaeroides* TlpT and *E. coli* Tsr showing the positions of the identified methylation and deamidation sites (blue) and sites predicted from this information(yellow). Top ruler numbering is in relation to TlpT. This image was taken from [28].

Previous work in the laboratory characterised the free swimming and tethering phenotypes for the E296Q and E296D mutants [28]. Both mutants displayed wild-type phenotypes. Hence it was proposed that Q was post-translationally modified back to E and that a single mutation (E296D) might have too subtle an effect. Reasoning by the *E. coli* paradigm, the expected phenotype for for the E296D mutant was similar to a decrease in the methyltransferase CheR, hence showed smoother free swimming and less sensitivity to drops in attractant concentrations. By the same

reasoning, it was predicted that an E296A mutant would show stoppier free swimming and longer adaptation times.

Table 5.2: List of the strains that contain alanine mutations of proposed methylation sites in TlpT.

Strain number	Genotype
JPA2332	*tlpT* E296A
JPA2367	*tlpT* E289A
JPA2368	*tlpT* E289A E296A
JPA2369	*tlpT* E 478A
JPA2370	*tlpT* Q485A
JPA2371	*tlpT* E289A E296A Q485A E478A

To sum up, TlpT methylation by the *R. sphaeroides* adaptation proteins was identified at E296 and E478, while Q485 only showed deamidation. To directly show the involvement of these residues in *R. sphaeroides* adaptation, the dynamic and steady state swimming behaviour of strains containing the respective alanine mutants alone and in combination are characterised in this section. Additionally, an E296A mutation is characterised alone and in combination to assess whether methylation does indeed occur at this site as proposed by de Beyer [28]. Beyond the four strains containing single alanine mutations in TlpT, the number of possible combinations containing at least two sites mutated to A rises up to 11. The time necessary to produce strains containing each of these mutation combinations was prohibitively long for this study. Hence a double mutant and the quadruple mutant were selected and added to the list of strains analysed in this section. The TlpT alanine mutant strains were constructed by Elaine Byles.

5.5.2 Results

5.5.2.1 Integrity of the cytoplasmic cluster

TlpT is a component of the cytoplasmic chemosensory cluster whose absence prevents cluster formation and produces a non-chemotactic phenotype. It is crucial to probe whether the TlpT methylation site mutant strains contain clusters so that the observed phenotypes can be correlated with the methylation site mutation rather than the disruption of the cytoplasmic cluster. The presence of the cluster was verified using fluorescent fusions of the cytoplasmic cluster component, $CheW_4$.

An N-terminal YFP fusion of $CheW_4$ was previously used by Wadhams *et al.* to show its localisation to the cytoplasmic cluster. This genetic modification had no effect on cluster formation or functionality. The same genomically expressed fluorescent construct was employed to test for the presence of cytoplasmic clusters here.

The *yfp*-containing strains were obtained by homologous recombination using an existing pK18pK18*mobsacB* construct. The insertion of the *yfp* gene was verified by colony PCR using primers situated upstream and downstream from the site of insertion. Once confirmed by colony PCR, samples from cultures of OD_{700} =0.4-0.6 were analysed by phase-contrast fluorescence microscopy. The cells were immobilised on agar pads and both phase-contrast as well as fluorescent images were captured for every strain.

Fluorescent foci were observed in the YFP images, as presented in Figure 5.6. Upon overlay with the phase-contrast images, it became clear that each cell contained a variable number of foci. This is due to the mode of cytoplasmic cluster partition upon cell division, an event which is frequent at the optical density used here. For

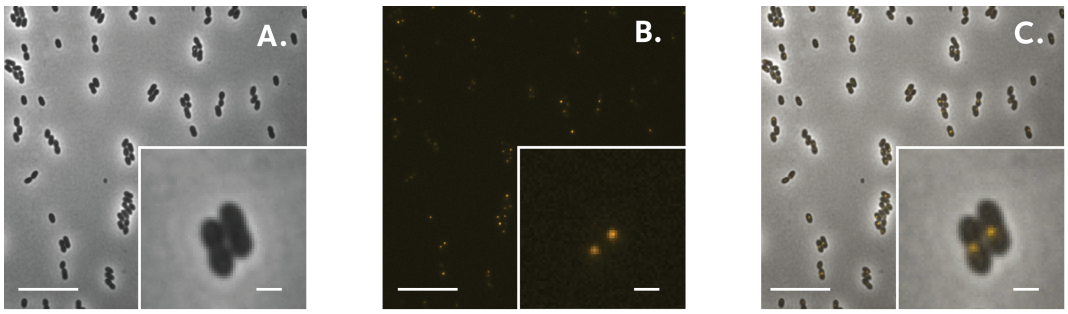

Figure 5.6: Representative image for cytoplasmic clusters in TlpT mutant strains. A: Phase-contrast image of the cells; B: fluorescence image of the same set of cells; C: Overlay of the first two images. The scale bars represent 10 μm in the big image and 1 μm in the inset.

the purpose of this work, it was only important to assess whether a cluster – and therefore a fluorescent focus – existed, so the cells in the images were sorted into two categories: cells with and cells without a cytoplasmic cluster. The sorting and counting was carried out by eye and the results are summarised in Table 5.3.

Table 5.3: Cytoplasmic cluster formation in TlpT mutant strains. The cytoplasmic cluster was observed by fluorescence phase-contrast microscopy.

Strain	Description	No. of cells	Cells with cytoplasmic cluster		Cells without cytoplasmic cluster	
			No.	%	No.	%
JPA2378	TlpT E478A YFP-CheW$_4$	190	182	95.7	8	4.3
JPA2379	TlpT E289A E296A YFP-CheW$_4$	184	178	96.7	6	3.3
JPA2380	TlpT Q485A YFP-CheW$_4$	159	151	94.9	8	5.1
JPA2381	TlpT E289A E296A Q485A E478A YFP-CheW$_4$	191	180	94.2	11	5.8

Fluorescently-labelled variants were obtained for all TlpT mutant strains except JPA2332 (TlpT E296A) and JPA2367 (TlpT E289A). It is important to point out that JPA2368 which contains both E296A and E289A mutations on TlpT was successfully tagged with YFP and foci were observed. In addition, the quadruple mutant strain, JPA2371, containing these two mutations along with two others displayed fluorescent clusters. It was therefore judged that the mutations present in

the JPA2332 and JPA2367 strains do not disturb the integrity of the cytoplasmic chemosensory cluster and the phenotypes observed hereafter are not due to the absence of the cluster.

5.5.2.2 Free swimming analysis

Data were collected and processed following the steps outlined in Section 5.3.2. The numbers of tracks left after each censoring step are listed in Table 5.4. It has been suggested that under the censoring protocol used, strains with a less motile phenotype *i.e.* ones that spend a longer time stopped, tend to be removed in excess at the minimum boundry radius (MBR) step, leading to a smaller number of cells left after censoring. This is the case with strains JPA2368 and JPA2370, for example which had the highest percentage of tracks removed during MBR censoring.

Table 5.4: Censoring table for TlpT mutant strains. Numbers of tracks removed after each censoring step are listed. Numbers between brackets show the numbers of tracks that would have been removed if that censoring step were used.

Strain	Description	Initial	Censoring				Final
			Speed	MBR	Length	MAC	
WS8N	wild-type	12350	3723	6345	579	171	1532
JPA467	non-motile	3879	291	(3490)	(35)	(7)	3588 (56)
JPA1353	non-chemotactic	8555	2357	4732	333	114	1019
JPA2332	*tlpT* E296A	7530	1785	4400	362	99	884
JPA2367	*tlpT* E289A	5134	1014	3290	247	59	524
JPA2368	*tlpT* E478A	12655	2362	9560	213	52	468
JPA2369	*tlpT* Q485A	3598	979	1873	184	57	505
JPA2370	*tlpT* E296A E289A	8791	1890	6173	190	54	484
JPA2371	*tlpT* E289A E296A Q485A E478A	4295	307	(3922)	(19)	(5)	3988 (42)

Figures 5.7 and 5.8 show MAC-NEMS plots and 25 randomly selected tracks from each strain, for the TlpT strains containing one site mutated to alanine and strains containing at multiple sites mutated to alanine, respectively. In both cases, clear distinctions between strains are immediately noticed by eye.

The single mutants E296A, E289A, and E478A (Figure 5.7) all display smooth

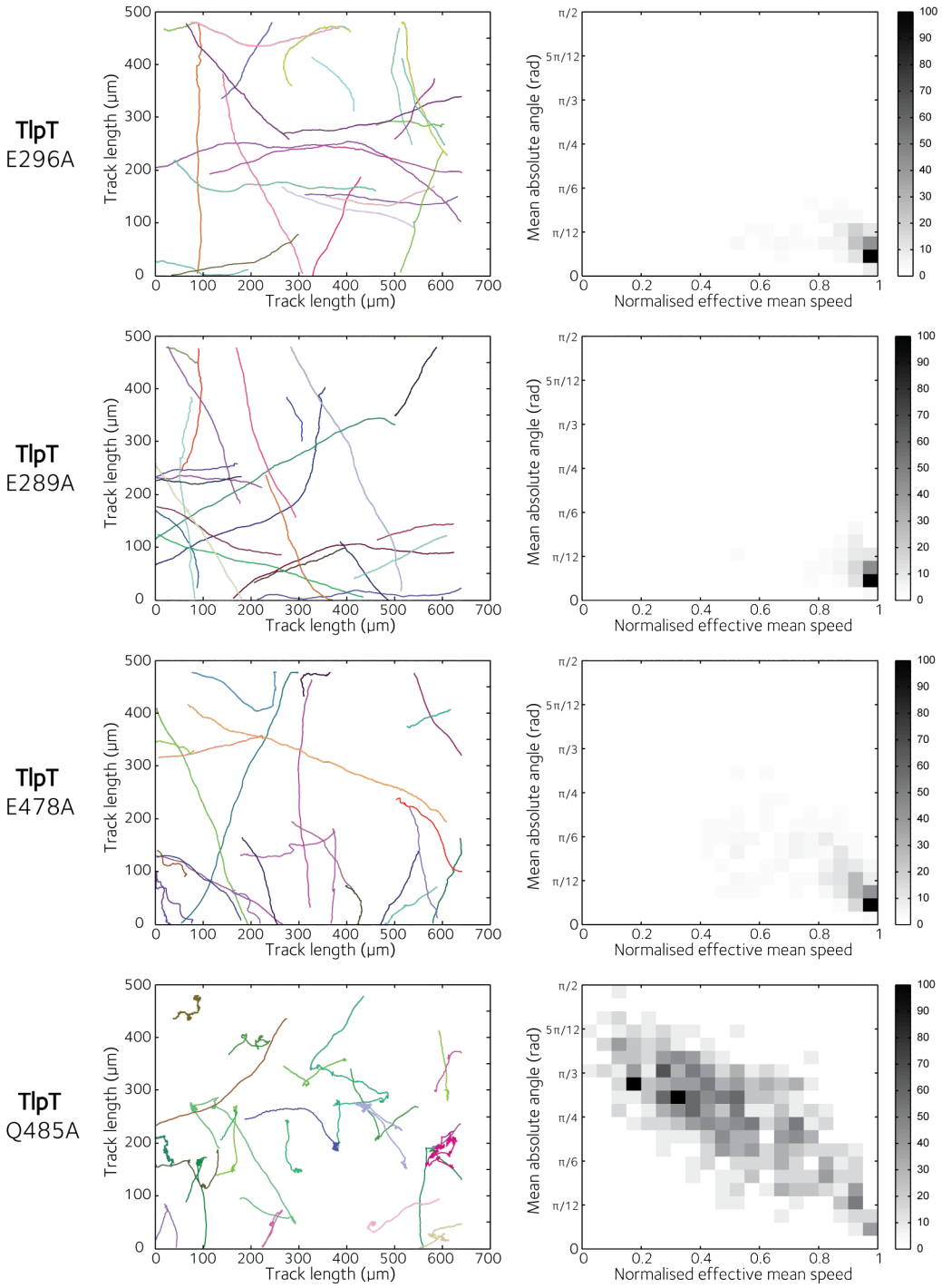

Figure 5.7: Free swimming tracks and MAC-NEMS plots for TlpT single mutants.

free swimming tracks, characterised by long, close to liniar tracks. This phenotype supported by MAC-NEMS plots showing the population of tracks concentrated in the bottom right region of the plot.

The single mutant Q485A (Figure 5.7), as well as the mutant containing both E296A and E289A (Figure 5.8) produced tracks that seem to have abnormally long transient stops. This is confirmed by the respective MAC-NEMS plots which show a broad distribution of swimming speeds and rotation angles, with the most intense signal located in the top left region, at lower mean speeds and higher angle change, indicative of a strain that is more stoppy than wild-type (Figure 5.2).

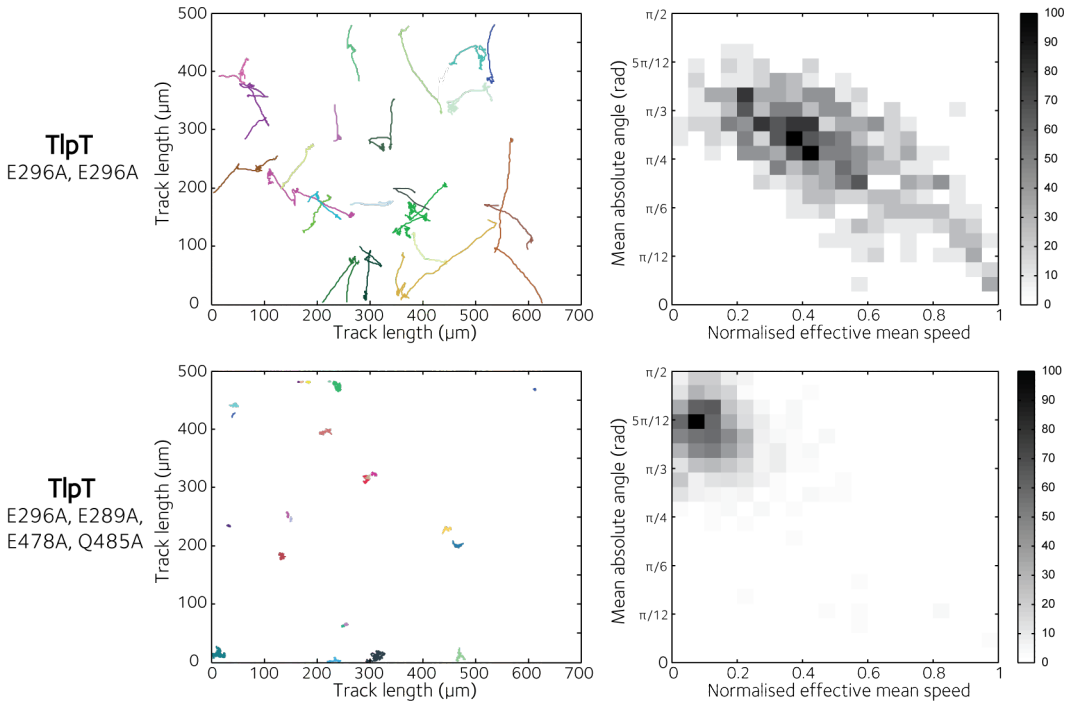

Figure 5.8: Free swimming tracks and MAC-NEMS plots for TlpT double and quadruple mutants.

The quadruple mutant (Figure 5.8) forms part of a separate category, with a free swimming phenotype similar to the non-motile reference strain. In the MAC-NEMS plot, the population of tracks is indeed found in the region of very low speed and large angle change. (Figure 5.2).

Free swimming data was also acquired for the Δ*tlpC* strain. This was to compare with existing free swimming data on Δ*tlpT* [28]. The behaviour of these two deletion strains has already been characterised under dynamic conditions, hence will

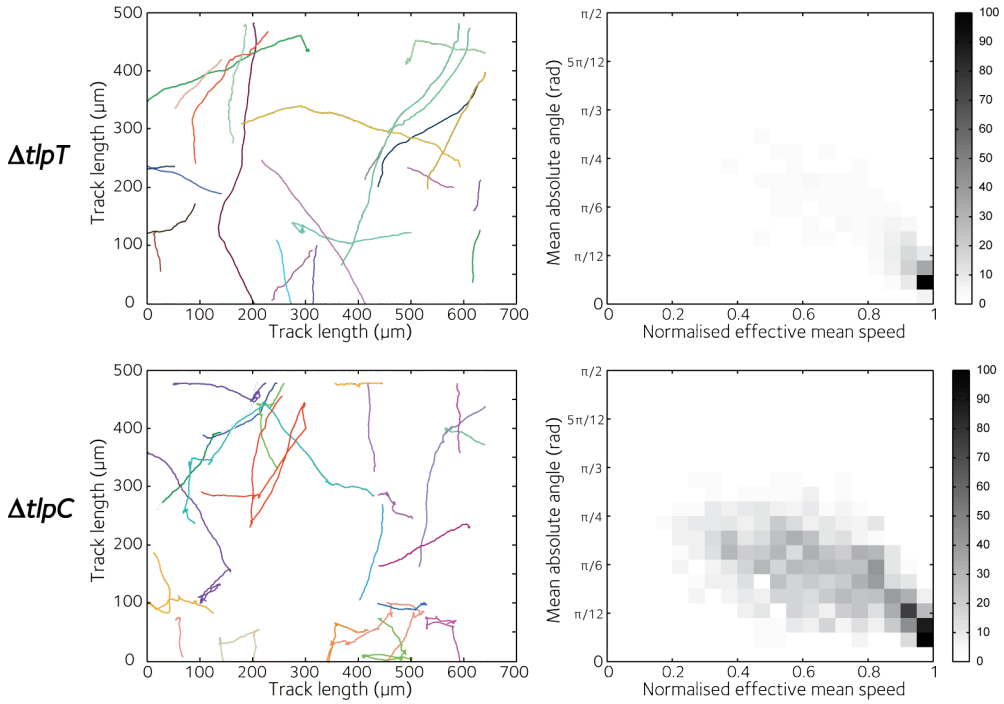

Figure 5.9: Free swimming tracks and MAC-NEMS plots for Δ*tlpT* and Δ*tlpC*.

not be repeated [28, 113]. The free swimming tracks and MAC-NEMS plots for Δ*tlpC* and Δ*tlpT* are presented in Figure 5.9. The free swimming phenotypes for the two mutants are visibly different, with the Δ*tlpC* phenotype appearing similar to wild-type.

Even though the data sets appear to be significantly different by eye, the complete track data analysis was undertaken. This is to allow the data on the TlpT mutants to be analysed alongside the adaptation protein mutants from the second part of this chapter. A summary statistic such as the mean speed per track used in *E. coli* was not considered because previous work has shown that the average swimming speed of *R. sphaeroides* wild-type cells is not statisticaly different from that of smooth swimming cells (lacking all chemotaxis proteins). This is likely because *R. sphaeroides* spends 80% of the time in runs, while *E. coli* only spends 60%, thus making the average swimming speeds more distinguishable for the two phenotypes in the latter case, but not the former.

The approach taken was to first separate the tracks into run and stop segments. From the new set of summary statistics obtained after this classification, the *time spent stopped* statistic was chosen over the *frequency of stops* statistic. This choice was made on the one hand because the classification algorithm has a bias towards assigning runs, thus introducing a level of noise into the data, and on the other hand because the censoring protocol is not perfect, leaving a number of non-motile and tortuous tracks in the data set, which would lead to an over-representation of the number of stops identified [28]. However, these errors stemming from various stages of the analysis only give rise to short-lived stops compared to true ones, hence using the total *time spent stopped* should be a lot less polluted by these errors and allow for discrimination between wild-type phenotypes and smoother or stoppier ones.

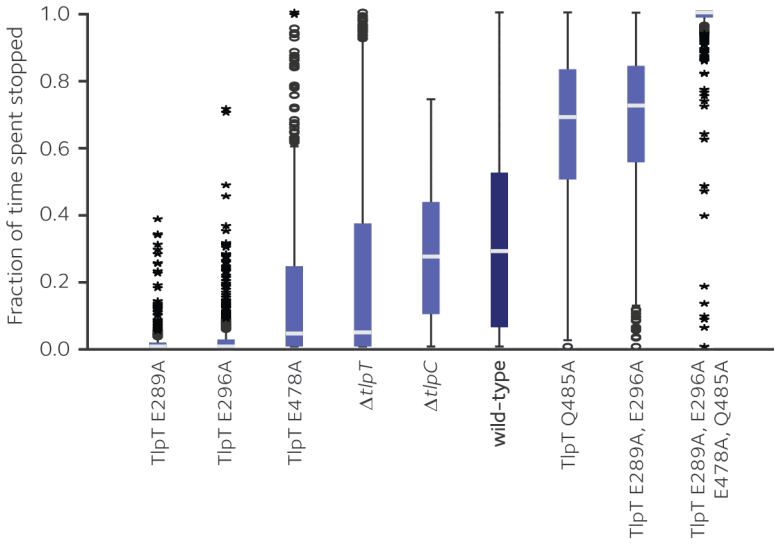

Figure 5.10: Distribution of the fraction of time each TLP mutant spends stopped. The wild-type is in a darker shade of blue. The boxplot edges indicate first and third quartile, the box centre indicates the median, whiskers indicate minimum and maximum, and circles and stars indicate outliers and extreme values, respectively.

The distributions of proportion of time spent stopped are presented in Figure 5.10, and the differences noticed by eye seem apparent here too. The distribution for the Δ*tlpC* strain indeed appears similar to wild-type and different to Δ*tlpT*. The data for

Δ*tlpT* was obtained by de Beyer [28], using the same censoring and track analysis parameters.

A one-way Kruskal-Wallis test rejects the null hypothesis that these distributions are from the same population ($p<0.05$). Post-hoc pairwise comparisons are made with the null hypothesis that each pair is from the same population. The corresponding probabilities for this hypothesis are shown in Figure 5.11A. The resulting groups are presented in Figure 5.11B, sorted by the amount of time spent stopped.

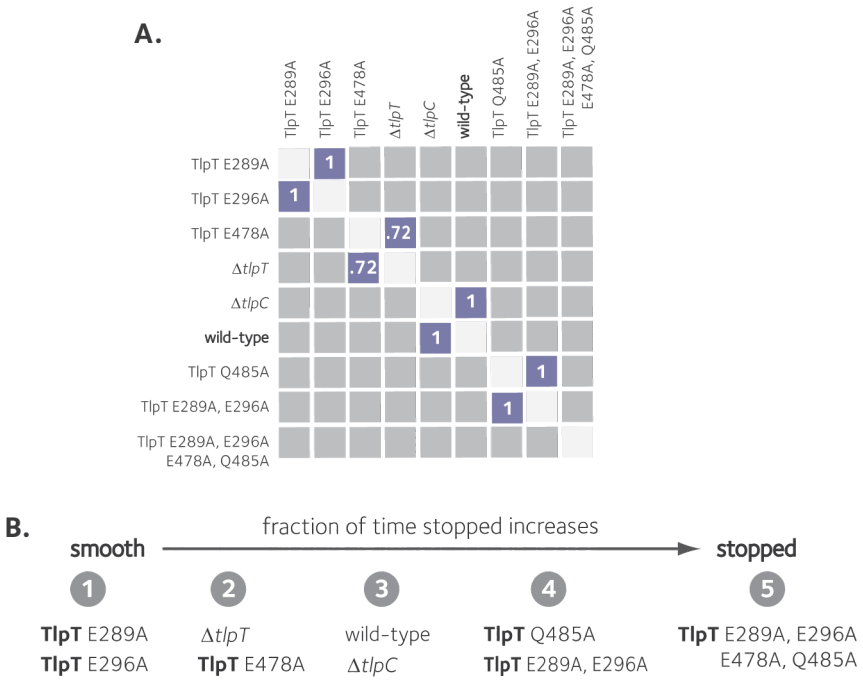

Figure 5.11: Analysis of TLP mutants. **A.** Pairwise comparisons between distributions of fraction of time spent stopped for each TLP mutant, giving the probability that a pair is from the same population. Blue squares indicate an accepted null hypothesis *i.e.* data sets from the same group. The numbers in the squares indicate the p-value for that comparison. Comparisons not shown are $p<0.001$. **B.** The mutants were grouped based on the proportion of time spent stopped, as judged by paiwise comparisons.

Five groups emerged after this classification. The quadruple mutant forms a group alone. There are two groups found to spend more time stopped than wild-type, and another two groups found to spend less time stopped than wild-type. As suspected,

Δ*tlpC* was grouped with wild-type. In the case of the strains more stoppy than wild-type, the difference was also noticed by eye: the quadruple mutant is practically non-motile while the E478A single mutant and the E289A, E296A double mutant also spend a large proportion of the time stopped, but are an obviously motile strain (Figures 5.7 and 5.8).

Interestingly, while by eye the strains containing single E mutations all appeared to be smooth swimming, statistical testing separated them into two groups: E478A was grouped with Δ*tlpT* spending less time time stopped than wild-type, while E289A and E296A were grouped separately with the least time spent stopped. The analysis was also carried out without Δ*tlpT*, in which case E478A remained grouped separately from the other two E mutants.

5.5.2.3 Tethered cell analysis

Data collection and analysis were performed as outlined in the first part of this chapter. The phenotypes identified for each strain are listed in Table 5.8.

Table 5.5: List of the tethering phenotypes obtained for the TlpT methylation site mutants.

Strain	Genotype	Tethering phenotype	# cells
JPA2332	*tlpT* E296A	Non-responsive	25
JPA2367	*tlpT* E289A	Non-responsive	14
JPA2368	*tlpT* E289A E296A	Responsive, early stop	14
JPA2369	*tlpT* E478A	Non-responsive	24
JPA2370	*tlpT* Q485A	Adaptive and stoppy	11
JPA2371	*tlpT* E289A E296A Q485A E478A	Responsive, early stop	10

Strains containing single E296A, E289A, E478A mutations in TlpT showed no response to a decrease in propionate. The TlpT Q485A mutant was the only one able to respond and adapt to the decrease in propionate concentration. A representative tethering signal density map for this strain is shown in Figure 5.12. The bacterium

appears to be rotating smoothly before the negative stimulus, when rotation stopped. Adaptation took longer than in wild-type, averaging 134±33 s, compared to the 40±5 s taken by wild-type. Rotation resumed in a stoppy fashion and it remains unclear whether it became smooth. Judging by the free swimming results, this was deemed unlikely, hence TlpT Q485A was classified as stoppy adaptive.

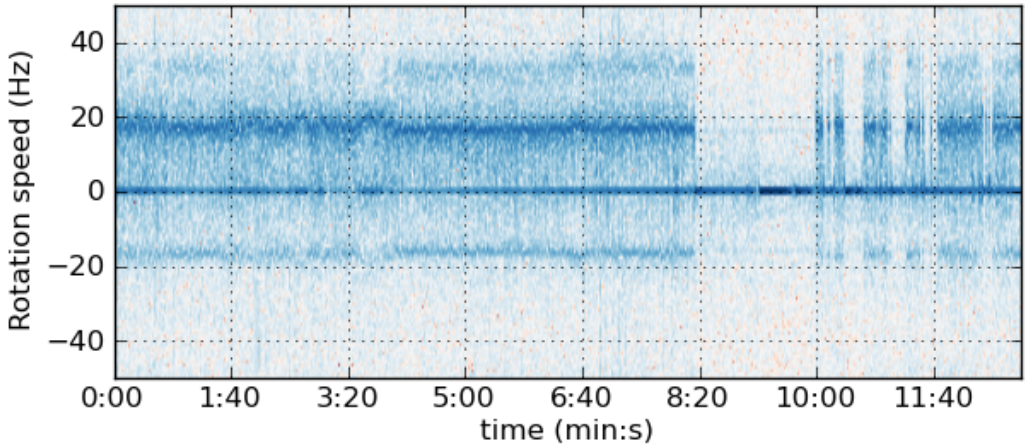

Figure 5.12: Representative rotational signal density map for TlpT Q485A. Cells are able to respond to a drop in nutrient concentration and take longer than wild-type to adapt. They resume spinning in a stoppy fashion.

The remaining two strains, each containing more than one mutation, have been classified as *Responsive, early stop*. The TlpT E289A E296A double mutant is motile but spends more time stopped than wild-type. This is visible in some cases in the tethering plots. In most cases, cells stopped swimming before the decrease in propionate, making it impossible to assess their response. Indeed, in some cases intermittent rotation was observed both before and after the negative stimulus, but not in response to any stimulus.

The quadruple mutant, also classified as *Responsive, early stop*, is a non-motile strain, as shown by the free swimming experiments. However, cells started rotating on the addition of propionate. Hence the tethering protocol was slightly modified for this strain, by adding a small amount (10 μM) of propionate to the initial buffer.

This allowed a field of view to be chosen with bacteria that were rotating. However, despite the protocol modification, most cells stopped rotating before a decreasing gradient of propionate. This suggests the bacteria have stopped motors, but on addition of an attractant the cells rotate, adapting to stopped.

5.6 The adaptation proteins of *R. sphaeroides*

5.6.1 Introduction

In *E. coli* there is a single CheB and a single CheR protein. Deleting CheB results in overmethylated chemoreceptors and that leads to a tumbling phenotype during steady state swimming. Cells can resume smooth swiming in the presence of a sufficiently large attractant stimulus, but cannot adapt quickly and recover tumbling swimming very slowly. Conversely, deleting CheR in *E. coli* results in unmethylated chemoreceptors, with smooth swimming cells under steady state. A sufficiently large repellent stimulus will cause the cells to tumble, but again, they cannot quickly adapt to the stimulus.

R. sphaeroides has two CheB and two CheR proteins expressed under laboratory conditions. Previous data from the laboratory shows that Δ*cheR*$_2$ and Δ*cheR*$_3$ strains are not chemotactic, being unable to respond to drops in nutrient concentrations. CheR$_2$ can complement an *E. coli* CheR deletion [64]. A Δ*cheB*$_1$ strain is not chemotactic, while a Δ*cheB*$_2$ strain shows reduced chemotaxis, with CheB$_1$ being able to fully complement an *E. coli* strain lacking CheB, while CheB$_2$ only being able to partially complement the same *E. coli* deletion strain. These differences in the effects of deleting either of the *R. sphaeroides* CheB proteins suggests different roles in adaptation [64].

Recent work from the laboratory performed an extensive characterisation of the swimming behaviour under both steady state and dynamic conditions for a series of *R. sphaeroides* chemotaxis protein mutants [28]. A summary of the data is presented in Table 5.6 and will be discussed in the context of the results obtained here, in Section 5.7.

Table 5.6: Summary of previous adaptation protein deletion swimming phenotypes obtained under steady state (free swimming) and dynamic (tethering) conditions.

Genotype	Free swimming phenotype	Tethering phenotype
$\Delta cheR_2$	more stoppy than wild-type	Non-responsive stoppy
$\Delta cheB_2$	more stoppy than wild-type	Adaptive; longer adaptation time
wild-type	*wild-type*	Adaptive
$\Delta cheR_3$	smooth swimming	Non-responsive
$\Delta cheR_{1,2,3}$	smooth swimming	not tested
$\Delta cheB_1$	smooth swimming	Non-responsive
$\Delta cheB_{1,2}$	smooth swimming	not tested

As mentioned in the Introduction and discussed at the beginning of this chapter, the $CheB_1$, $CheR_2$ pair is assumed to be responsible for polar chemosensory array adaptation, while the $CheB_2$, $CheR_3$ pair is assumed to fulfill the same function for the cytoplasmic chemosensory cluster. However, no mutant strains with the adaptation proteins deleted in these combinations have been studied. Strains containing the adaptation protein pairs associated with each of the arrays deleted in turn and together (mutagenesis performed by Elaine Byles) were characterised in this section under steady state as well as dynamic swimming conditions. Finally, the data obtained with the pair deletions is integrated with the data set obtained by de Beyer [28].

5.6.2 Results

5.6.2.1 Free swimming

The free swimming data was acquired and analysed as described in the method section of this chapter (Section 5.3.2). A summary of the censoring steps is presented in Table 5.7.

Table 5.7: Censoring table for Δ*cheBcheR* strains. Numbers of tracks removed after each censoring step are listed. Numbers between brackets show the numbers of tracks that would have been removed if that censoring step were used.

Strain	Description	Initial	Censoring				Final
			Speed	MBR	Length	MAC	
WS8N	wild-type	12350	3723	6345	579	171	1532
JPA467	non-motile (Δ*fli*A)	3879	291	(3490)	(35)	(7)	3588 (56)
JPA1353	Δ*che*Op$_{1,2,3}$	8555	2357	4732	333	114	1019
JPA1377	Δ*cheR*$_3$ Δ*cheB*$_2$	7274	1709	4385	368	82	730
JPA1378	Δ*cheR*$_2$ Δ*cheB*$_1$	8238	1588	5774	194	69	613
JPA1379	Δ*cheR*$_{1,2,3}$ Δ*cheB*$_{1,2}$	7493	1462	5101	308	63	559

Figure 5.13 shows 25 representative tracks and the MAC-NEMS plot for each of the Δ*cheBcheR* mutants. For all three mutants, the highest density of tracks is found at high NEMS and low MAC. Of the three strains, the Δ*cheB*$_1$*R*$_2$ strain shows the broadest distribution of speeds and angle changes. The MAC-NEMS plot suggests a mildly stoppy phenotype, with a small proportion of the tracks distribution extending towards the top left corner. The tracks themselves typically have one or more changes in angle, indicative of transient stops.

The Δ*cheB*$_2$*R*$_3$ tracks distribution is tightly placed in the bottom right corner of the plot, suggesting a smooth swimming phenotype. Indeed, the corresponding tracks are mostly straight lines, and hardly any angle changes can be seen in the random selection of 25 tracks plotted in Figure 5.13.

Judging by both the MAC-NEMS plot and the shape of the tracks, the Δ*cheB*$_{1,2}$*R*$_{1,2,3}$

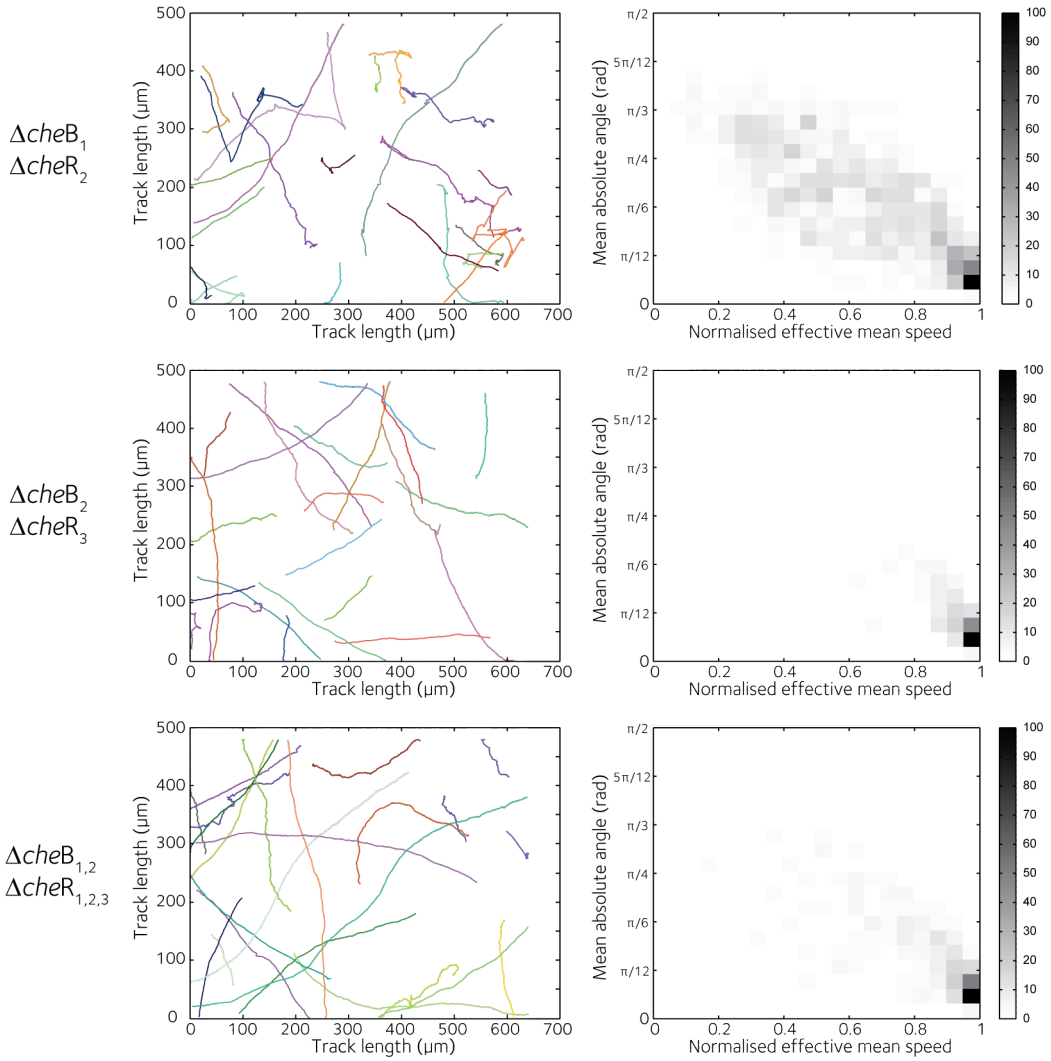

Figure 5.13: Free swimming tracks and MAC-NEMS plots for Δ*cheBcheR* mutants.

strain appears to have an intermediate free swimming phenotype. The broadness of the tracks distribution in the MAC-NEMS plot extends further than the Δ$cheB_2R_3$ strain, but not as far as the Δ$cheB_1R_2$ strain. The shape of most tracks, while not a straight line as in the case of the Δ$cheB_2R_3$ strain, does not have sharp angle changes immediately obvious to the naked eye, as it is the case with the Δ$cheB_1R_2$ strain. It is likely that the rounded shape of the tracks belonging to the Δ$cheB_{1,2}R_{1,2,3}$ strain is due to shorter, more frequent transient stops than in the case of the Δ$cheB_1R_2$ strain.

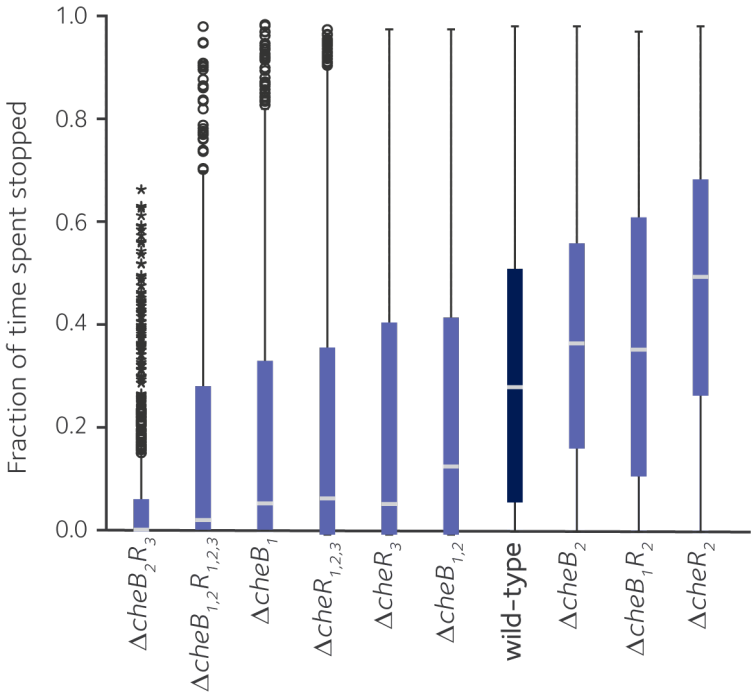

Figure 5.14: Distribution of the fraction of time each Δ*cheBcheR* mutant spends stopped. The wild-type is in a darker shade of blue. The boxplot edges indicate first and third quartile, the box centre indicates the median, whiskers indicate minimum and maximum, and circles and stars indicate outliers and extreme values, respectively.

While certain differences between these three strains could be noticed by eye, a statistical test would verify if the differences were indeed significant between the phenotypes of the three strains to consider them distinct. To provide more context to the discussion of the intricacies of *R. sphaeroides* adaptation at the end of this chapter, the free swimming data for these three mutants was analysed together with equivalent data obtained by de Beyer for other adaptation protein mutant strains [28] (Table 5.6). This merging of data sets was possible becuse the same acquisition, tracking, censoring and classification parameters were used for all strains.

The distributions of the proportion of time spent stopped by each of the strains considered in this analysis are presented as boxplots in Figure 5.14. A one-way Kruskal-Wallis test was used to check the null hypothesis that all these distributions

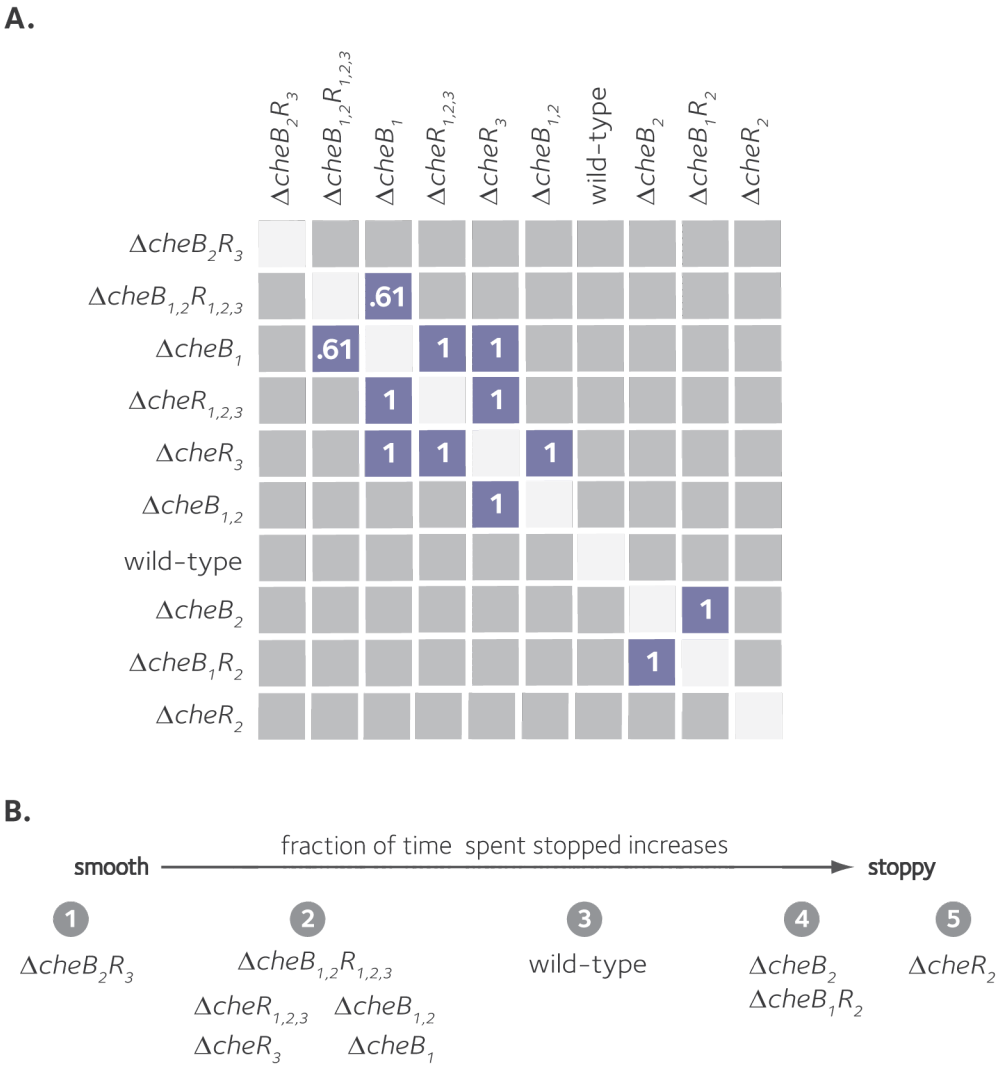

Figure 5.15: Analysis of Δ*cheBcheR* mutants. **A.** Pairwise comparisons between distributions of fraction of time spent stopped for each Δ*cheBcheR* mutant, giving the probability that a pair is from the same population. Blue squares indicate an accepted null hypothesis *i.e.* data sets from the same group. The numbers in the squares indicate the p-value for that comparison. Comparisons not shown are p<0.001. **B.** The mutants were grouped based on the proportion of time spent stopped, as judged by pairwise comparisons.

arise from the same population. The null hypothesis was rejected (p<0.001) so a post-hoc pairwise comparison was performed(Figure 5.15A). The groups resulting after considering all identities established by the pairwise comparison are presented in Figure 5.15B.

In this case too, five groups emerged. The first group only contains Δ*cheB*$_2$*R*$_3$ and is the mutant that spends the least time stopped. In the order of increasing amount of time spent stopped, next in line is a six-membered group that spends less time stopped than wild-type. This group contains the Δ*cheB*$_{1,2}$*R*$_{1,2,3}$ strain, confirming its intermediate phenotype between the Δ*cheB*$_2$*R*$_3$ and Δ*cheB*$_1$*R*$_2$ mutants. The middle group only contains the wild-type strain. The Δ*cheB*$_1$*R*$_2$ strain characterised as part of this work was grouped with the Δ*cheB*$_2$ strain, spending more time stopped than wild-type, but less time than the stoppiest strain of the extended data set, Δ*cheR*$_2$.

5.6.2.2 Tethered cell assay

The Δ*cheBcheR* mutants were tethered to the glass coverslip of a flow cell using anti-flagella antibodies and their rotational movement was recorded. A wild-type *R. sphaeroides* cell stops swimming when it experiences unfavourable conditions *e.g.* a decreasing propionate gradient. The Δ*cheBcheR* mutants analysed in this section reacted neither to an increasing propionate gradient, nor to a decreasing one.

Table 5.8: List of the tethering phenotypes obtained for the Δ*cheBcheR* mutants.

Strain	Genotype	Smooth	Stoppy	Total
JPA1377	Δ*cheB*$_2$ Δ*cheR*$_3$	8	9	17
JPA1378	Δ*cheB*$_1$ Δ*cheR*$_2$	8	21	29
JPA1379	Δ*cheB*$_{1,2}$ Δ*cheR*$_{1,2,3}$	9	15	24

The non-responsive phenotype however, could be separated by eye into smooth non-responsive (similar to the example in Figure 5.3B) and stoppy non-responsive, a phenotype not reported in literature and first characterised by de Beyer [28] (Figure 5.16A). The stoppy non-responsive phenotype was identified in tethered cell traces showing numerous clear, long (5-30s) gaps. Tethering phenotypes obtained for the Δ*cheBcheR* mutants here argue that the stoppy phenotype cannot be defined

discretely; Figure 5.16A shows an example of stoppy non-responsive phenotype with clear gaps 5-30 s in length, as described by de Beyer. In the same data set, tethering phenotypes similar to Figure 5.16B were also found to varying degrees. The areas of dense signal on this plot contain shorter (2 s) alternating intervals of rotation and pause. This explanation is supported by the video recording showing the spinning cell completing 1-2 rotations with short pauses inbetween. In the same data set still, cells combining the two phenotypic variations from panels A and B within the 13 minute experiment can be found.

Overall, both smooth and stoppy non-responsive phenotypes were observed for all three strains. The proportions between them varies, suggesting that some are probably more stoppy than others. The cells are able to start and stop rotation but not influenced by external chemosensory signals. This suggests fluctuations in $CheY_6$–P generated by means other than removal of attractant.

The tethered cell technique with its associated analysis software in its present setup is not optimised for quantifying short-lived events, so a pertinent comparison with the free swimming results cannot be made. It should also be emphasised that the clasification of tethering phenotypes is made by hand and is thus subjective and becomes problematic with less clear phenotypes.

5.7 Discussion

5.7.1 Analysis of TlpT methylation site mutants

The putative methylation sites from TlpT – E289, E296, E478, Q485 – were mutated to alanine. This type of mutation permanently removes the charge from those sites, causing the receptor to behave as if it were permanently methylated. Methy-

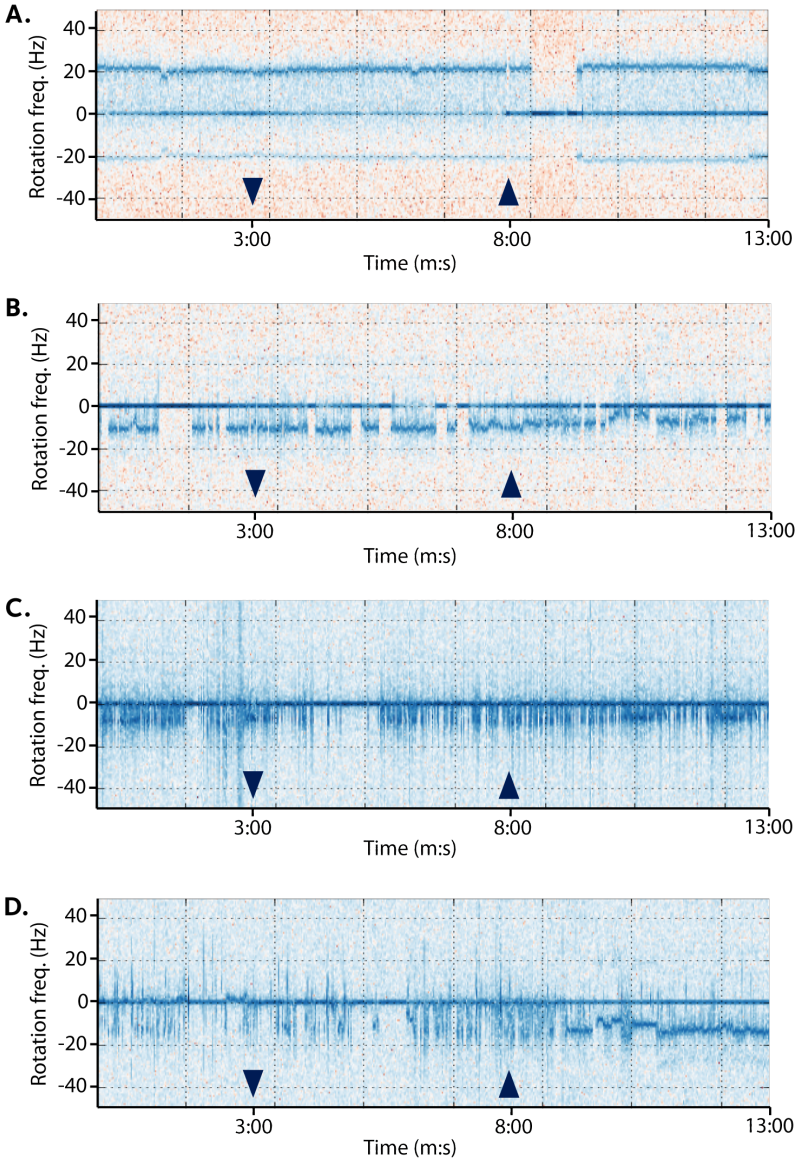

Figure 5.16: Variations of the stoppy non-responsive tethering phenotype. **A.** The stoppy non-responsive phenotype; **B.** A stoppy non-responsive phenotype with very short alternating pauses and swimming intervals; **C.** A cell displaying both short-interval and long-interval stoppy phenotypes.

lation sensitizes the chemosensory array by allowing tighter packing of the receptors.

Of the four single alanine mutations of proposed methylation sites in TlpT, E289A, E296A, and E478A produced a smooth free swimming phenotype and was non-responsive in tethering experiments. The fourth site, Q485A, which was only found

deamidated and not methylated in the MS/MS analysis, displayed adaptation on propionate removal. The adaptation times were longer than for wild-type and it also spent more time stopped than wild-type. Interestingly, the double mutant TlpT E289A, E296A is more stoppy than wild-type in free swimming experiments, and has a responsive phenotype in tethering experiments. Further, the quadruple mutant TlpT E289A, E296A, E478a, Q485A is completely stopped under steady state, but can respond to attractant and then stop after 3-4 minutes in the tethering experiments. This pattern of swimming on propionate addition, followed by a return to the stopped state is consistent with adaptation on propionate addition. As described above, there is a range of phenotypes obtained for different combinations of methylation site mutants. This diversity is similar to that seen in *B. subtilis* McpB [120].

The similarity between the tethering phenotypes of the three non-responsive single mutants and the non-chemotactic phenotype obtained through the deletion of TlpT *i.e.* the absence of a cytoplasmic array deserves a brief discussion. The presence of a cytoplasmic chemosensory array in the TlpT methylation site mutant strains was verified using fluorescence microscopy. Even though the array does form, as inferred from observing the fluorescent foci, the possibility that the assembled array is not functional remains. However, if a single mutation to alanine, like E296A, does disrupt the functionality of the array, then the same phenotype should be observed when E296A is present alongside other methylation site mutants. This is not what is observed. On the contrary, when two single non-resposnsive mutations are found in a double mutant strain, the phenotyoe become stoppy responsive. This is very likely a consequence of the methylation sites not being equivalent, as seen in *B. subtilis* and pointed out before, but to conclusively prove this, more double mutant combinations should be tested.

The methylation state of the cytoplasmic chemoreceptor TlpT has an influence on adaptation as seen in tethered cell experiments. Adaptation to a decreasing attractant gradient, as seen for TlpT Q485A, suggests that the cytoplasmic array is involved in the adaptive response to a negative stimulus. The ability to elicit an adaptive response to a positive stimulus was not impaired, as seen in the case of the quadruple mutant: the cell started rotating on propionate addition, and then stopped rotating as it adapted.

The Δ*tlpC* strain displays an intriguing behaviour. While during steady state it spends as much time stopped as wild-type, it is non-responsive in the tethered cell experiment [113]. This comes in spite of the slight cluster delocalisation observed by fluorescence microscopy in a Δ*tlpC yfp-cheW*$_4$ strain [114]. A lack of TlpC does not seem to change the wild-type ability to perform transient stops during swimming, but the ability to respond chemotactically is lost.

5.7.2 The adaptation proteins of *R. sphaeroides*

All of the free swimming data described in this chapter was analysed and presented together with previous data from the laboratory in Figure 5.17. The figure shows a ranking of groups of *R. sphaeroides* deletion strains based on the fraction of time they spend stopped. Strains from the same group were judged to have the same distribution of fraction of time spent stopped. The strains were sorted from smooth swimming to stopped, and also coloured according to their respective tethering phenotype (where available).

In Figure 5.17 the smooth end of the spectrum, contains strains that have a decreased level of CheY$_6$ in comparison to wild-type. As a consequence, the motor bias shifts towards smoother swimming (less time spent stopped). Similarly, the

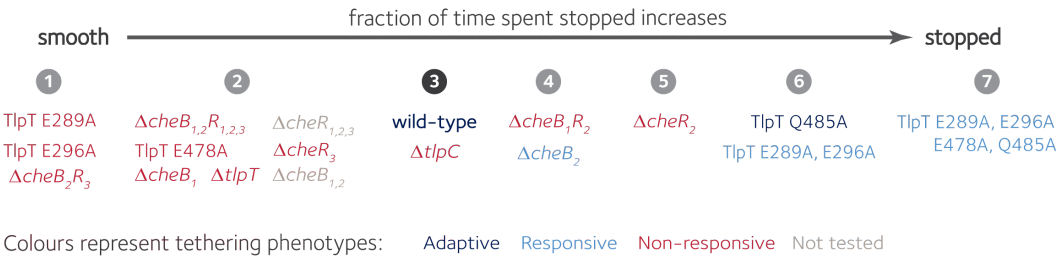

Figure 5.17: Overview of free swimming and tethering data. The groups of strains, numbered 1 to 7, are sorted from smooth to stopped, according to their steady-state behaviour. The group containing wild-type is emphasized with a darker shade. The tethering phenotypes are colour-coded red for non-repsonsive and blue for responsive, with darker blue for adaptive.

stoppy end contains strains that presumably have a relative excess of $CheY_6$–P. In both categories of steady state swimming phenotypes there are unexpected strains present.

It is surprising to see Δ*cheB*$_1$ classified as a smooth swimmer. Under the current model for chemotaxis, the expected phenotype for Δ*cheB*$_1$ is stoppy due to the oversensitized polar array producing $CheY_6$–P in excess.

It is also intriguing that Δ*cheR*$_2$ is classified as stoppy when, by the current model for chemotaxis, this strain should be smooth swimming due to lack of signaling through the polar array.

Taken together, an interesting trend emerges from the data presented in Figure 5.17. Judging by the phenotypes presented, it seems that an overstimulated polar array or an understimulated cytoplasmic array will produce smooth swimming. Conversely, an overstimulated cytoplasmic array or an understimulated polar array produces stoppy swimming

Results presented here have emphasized discrepancies between predictions made using the current model for chemotaxis and the phenotypes observed experimentally. A new model for chemotaxis in *R. sphaeroides* will be proposed in the *Con-*

cluding remarks chapter, taking into account the discrepancies pointed out here, as well as other results described in this thesis.

CHAPTER 6

Photoactivation localisation microscopy of $CheB_2$

R. sphaeroides expresses two CheB methylesterase response regulators. $CheB_2$ is encoded in *che*Op_3 with the components of the cytoplasmic chemosensory array, but fluorescence microscopy experiments show $CheB_2$ is found diffuse throughout the cytoplasm. Modelling and *in vivo* phosphotransfer work proposes that $CheB_2$ interacts with both the polar and cytoplasmic chemosensory arrays. As of yet, it remains unclear what the place of $CheB_2$ is within the *R. sphaeroides* chemotaxis signalling network. Work described in this chapter aims to provide more detailed qualitative *in vivo* information about the localisation of $CheB_2$ by employing single-molecule fluorescence microscopy.

6.1 $CheB_2$, a methylesterase response regulator

In *E. coli*, the methylesterase response regulator CheB localises with the membrane-embedded chemosensory array. In contrast, both CheB homologues of *R. sphaeroides* are found to be diffuse throughout the cytoplasm [115], making $CheB_1$ and $CheB_2$

available to both the membrane-embedded as well as the cytoplasmic chemosensory cluster (Figure 6.1 A, B). This further contrasts with the antagonistic adaptation proteins, the *R. sphaeroides* methyltransferases $CheR_2$ and $CheR_3$. The CheRs localise with the polar and cytoplasmic arrays, respectively (Figure 6.1 C, D).

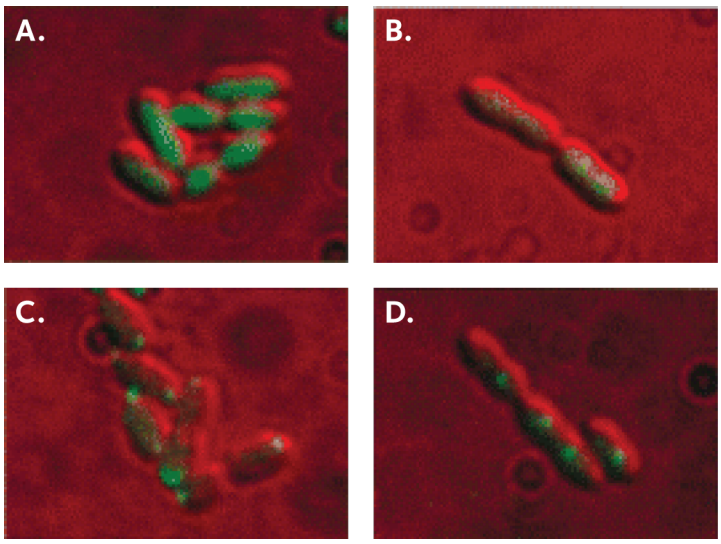

Figure 6.1: Localisation of *R. sphaeroides* adaptation proteins. **A.** YFP-$CheB_1$ is found diffuse throughout the cytoplasm. **B.** YFP-$CheB_2$ is found diffuse throughout the cytoplasm. **C.** YFP-$CheR_2$ is localised to foci at the cell poles. **D.** YFP-$CheB_1$ is localised to foci in the cytoplasm. The images were taken from [115].

Phosphotransfer experiments provided comprehensive *in vitro* information on the relationships between the *R. sphaeroides* CheA kinases and CheB response regulators. Among the CheAs expressed under laboratory conditions, $CheA_2$ can phosphorylate both $CheB_1$ and $CheB_2$ [73], while the $CheA_3$ $CheA_4$ split kinase can only phosphorylate $CheB_2$ [74]. It was also shown that $CheB_2$–P can phosphorylate $CheA_2$ [73]. Hence it is currently assumend that $CheB_2$ can be activated by both the polar and cytoplasmic arrays, but it is unclear in the adaptation of which chemoreceptors $CheB_2$ is involved.

Mathematical modelling by Tindall *et al.* revealed potential reverse phosphorelay from $CheA_3$–P to $CheA_2$, through $CheB_2$ [107]. Thus the cytoplasmic cluster could activate the polar cluster and indirectly activate its non-cognate response regulators.

The model did not include the sensory and adaptation pathways (chemoreceptors and CheR homologs), it only considered the CheA, CheB and CheY homologs. The question of reverse phosphorelay relevance under physiologucal conditions is further raised by the model being based on *in vitro* kinetic parameters.

Modelling work by Hamadeh *et al.* uses a model invalidation approach to establish the connectivity between the two CheB proteins and the two chemosensory clusters [36]. The model assumed a cascade control architecture for the two *R. sphaeroides* chemosensory pathways and considered the sensory and adaptation systems as well. All proposed connectivity models were invalidated, except for one proposing that $CheB_2$ interacts with both clusters, while $CheB_1$ only interacts with the polar cluster.

An *E. coli* CheB deletion strain can be complemented fully by expressing $CheB_1$, but only partially by expressing $CheB_2$ [64]. Deleting $CheB_1$ leads to a loss of chemotaxis, with the cells being unable to respond to a drop in nutrient. In the case of $CheB_2$, deletion only leads to a decrease in chemotaxis ability, with the cells being slower to adapt to changes in nutrient concentration.

Theoretical as well as *in vitro* work suggests that $CheB_2$ interacts with both chemosensory clusters, while $CheB_1$ only interacts with the polar cluster with *in vivo* data on these interactions being unavailable. Work in this chapter focuses on $CheB_2$ and makes use of photoactivation localisation microscopy (PALM) to track the movement of individual $CheB_2$ molecules in live cells.

6.2 Photoactivation localisation microscopy (PALM)

6.2.1 The method

Single-molecule fluorescene microscopy techniques allow protein visualisation with nanometer resolution. Photoactivated localisation microscopy (PALM) makes use of fluorescent proteins that can be activated from an initial dark state to a fluorescent state. By tuning the intensity of the activation laser, a subset of all labelled molecules can be activated at any time to determine their positions in a sequential manner. Tracks can be obtained by linking the positions of a single molecule fluorescence spot over time.

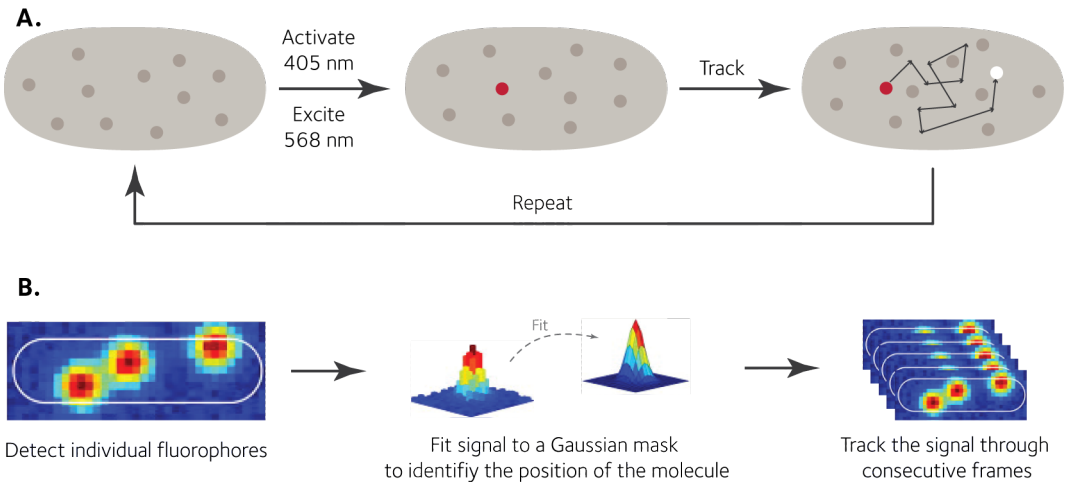

Figure 6.2: Photoactivation localisation microscopy and tracking. **A.** The protein of interest is expressed as a photoactivatable fusion (grey spots). A subset of fusion proteins are activated with low laser power (405 nm for PAmCherry) and then excited (568 nm for PAmCherry) with high laser power, allowing for individual molecules to be imaged (red sport). The excited molecules will remain visible throughout several consecutive frames (until photobleaching occurs, white spot) allowing for its tracking over time. **B.** Signal from individual fluorophores is detected and fitted to a Gaussian mask. The fitting allows for determination of the positon of the fluorescent molecule below the diffraction limit. The signal from an individual molecule can be tracked through a series of consecutive frames, allowing for movement parameters, such as diffusion coefficients, to be calculated.

By factoring in the time between acquisitions, diffusion coefficients can be calculated. Recent developments have increased the sensitivity and temporal resolution of PALM, such that it is now possible to track single diffusing fluorescent proteins in the bacterial cytoplasm of live cells.

6.2.2 Applications of PALM in live bacterial cells

Uphoff *et al.* have used photoactivation localisation and tracking to discriminate between bound and unbound DNA repair enzymes based on their diffusion coefficients [109]. The method was applied in live *E. coli* cells to identify transient DNA-binding events of single Pol1 molecules tagged with PAmCherry. Integration of their single-molecule observations generated a systems-level description of the DNA repair pathway *in vivo*.

Diepold *et al.* used PALM in their study of YscQ, the C-ring component of the *Yersenia enterolitica* type III secretion system [31]. They were able to discriminate *in vivo* between free-moving cytosolic and stable injectisome bound YscQ.

Work from our laboratory used a photoactivatable mCherry (PAmCherry) fusion in live *R. sphaeroides* cells to track the movement of individual PpfA molecules involved in cytoplasmic chemosensory array segregation upon cell division [48]. The method revealed a clear diversity of mobilities for PpfA.

6.3 Experimental approach

To gain more detailed insight into the localisation of $CheB_2$, the approach taken was to study the cellular localisation of individual molecules in correlation with their diffusion coefficient.

The *R. sphaeroides* inducible plasmid, pIND4 [43], was used to express $CheB_2$ tagged with a photoactivatable fluorophor, PAmChery, in a range of background strains (Table 6.1). The choice of fluorescent tag was due to its long fluorescence lifetime, minimal blinking on the exposure time scales used, and lack of oligomerisaion, making it well-suited for the study of single molecule dynamics [92]. The decision was reinforced by the previous successful use of PAmCherry in our laboratory [31], especially in *R. sphaeroides* [48].

Table 6.1: Background strains used for PALM imaging and tracking.

Strain	Genotype	Description
WS8N	wild-type	This is the wild-type background.
JPA1353	$\Delta che\text{Op}_{1,2,3}$ $\Delta cheY_{4,5}$	All chemotaxis genes are deleted.
JPA1301	$\Delta che\text{Op}_3$	the cytoplasmic cluster genes are deleted.
JPA1457	*yfp-cheW*$_4$	The cytoplasmic cluster is fluorescently labelled.
JPA1418	*yfp-cheW*$_3$	The polar cluster is fluorescently labelled.

The background strains listed in Table 6.1 would allow for $CheB_2$ tracking in a wild-type background strain (WS8N), as well as in strains lacking the cytoplasmic array ($\Delta che\text{Op}_3$), or both the polar and cytoplasmic arrays ($\Delta che\text{Op}_{1,2,3}$). Differences between diffusion coefficient distributions of PAmCherry-$CheB_2$ corresponding to these strains could be correlated with its interaction with the arrays. Additionally, colocalisation experiments using background strains with fluorescently tagged polar (YFP-$CheW_3$) or cytoplasmic clusters (YFP-$CheW_4$) would provide detailed information on the cellular localisation of the bound molecules (tracks with a low diffusion coefficient).

6.4 Creation of a photoactivatable fusion construct

PAmCherry was cloned upstream of *cheB*$_2$ within pIND4, the *R. sphaeroides* inducible expression plasmid [43]. An N-terminal fusion protein was chosen over a C-terminal one due to previous fluorescent protein work having been carried out with N-terminal fusions [115].

Creating the *PAmCherry-cheB*$_2$/pIND4 construct was not straightforward. The task was complicated through a) the GC abundance within the *cheB*$_2$ gene (75%) leading to a hairpin-rich secondary structure which interfered with PCR fragment assembly by overlap extension, and b) multiple instances of two out of the four restriction sites in the pIND4 multiple cloning site being present in both *cheB*$_2$ and *PAmCherry*.

The approach taken was to insert the *PAmCherry* gene upstream of *cheB*$_2$ in a pre-existing *cheB*$_2$/pIND4 plasmid. The *cheB*$_2$/pIND4 plasmid did not have any restriction sites upstream of *cheB*$_2$, so an *NdeI* restriction site was inserted through site-directed mutagenesis. The new plasmid was digested with *NdeI*, and treated with Antarctic Phosphatase (NEB) to remove 5′ phosphoryl termini required by ligases thereby preventing plasmid re-ligation. The linearised plasmid was then ligated with the *NdeI* double digested *PAmCherry* PCR fragment. The ligation products were transformed into XL-1 *E. coli* cells, which were grown on kanamycin selection LB-agar plates.

Because the insertion was performed using a single restriction site, *PAmCherry* could have been ligated in in both directions. A pair of primers was designed to identify colonies containing the correct insertion by colony PCR screen (see 2.3.2.4). The forward primer was designed to bind in the middle of the *PAmCherry* gene, while the reverse primer was designed to bind 800 bp downstream, within

cheB$_2$. Thus, a successful construct would yield a 800 bp PCR product, while an unsuccessful one would not produce a PCR product as the reaction would have been performed using two reverse primers. Two of the colonies identified as successful by PCR were grown up overnight in LB medium with kanamycin and the corresponding plasmids were isolated using the Quiagen MiniPrep kit. The two plasmid samples were all sent for sequencing and all were shown to contain the correct insert.

The successful plasmid was used to transform S17-1 *λpir* so that it could be conjugated into the desired *R. sphaeroides* strains (Table 6.1).

6.5 Results

6.5.1 Tracking and classification of PAmCherry-CheB$_2$ molecules in live *R. sphaeroides* cells

Cultures of *R. sphaeroides* cells containing PAmCherry-CheB$_2$ were grown in succinate medium to OD$_{700}$ of 0.4-0.6. Expression was not induced as it was shown in our laboratory that the leaky expression from pIND4 is enough to perform PALM experiments. Keeping a low number of photoactivatable molecules in the cell is important when doing PALM because the technique relies on the ability to activate and image single molecules [110].

The cells were the immobilised on 1% low-melting agarose pads prepared with filtered M9 medium. YFP fusion protein-containing strains were imaged using a 473 nm laser. Single molecules of PAmCherry-CheB$_2$ were imaged using simultaneous illumination at 405 nm (activation) and 561 nm (excitation), with a 15 ms exposure at 50 ms intervals for 15000 frames. The intensity of the 405 nm laser was con-

tinuously ajusted throughout the experiment to ensure that only single molecules of PAmCherry were activated within an individual cell at any time [110].

The data analysis package, STORMTRACKER, contains custom-built MATLAB (MathWorks) software, and was used to localise and track the molecules, as well as to analyse the diffusion characteristics of PAmCherry-$CheB_2$ [27, 42, 110]. The mean-squared displacement (MSD) between consecutive localisations making of a track of N steps is first calculated using Equation 6.1). Apparent diffusion coefficients were then calculated according to Equation 6.2, where the last term corrects for the localisation error. For the setup used in these experiments, σ_{loc}=40 nm and Δt=50 ms [110].

$$MSD = \frac{1}{N-1}\sum_{i=1}^{N-1}(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2 \tag{6.1}$$

$$D^* = \frac{MSD}{4\Delta t} - \frac{\sigma_{loc}^2}{\Delta t} \tag{6.2}$$

Only tracks of at least four steps were considered for the calculation of diffusion coefficients. It should be noted that D^* is an apparent diffusion coefficient which differs from a standard diffusion coefficient because of cell confinement, motion blurring, and localisation error [65, 104].

6.5.2 PAmCherry-$CheB_2$ in a wild-type background

PAmCherry-$CheB_2$ conjugated into a WS8N background was imaged as described above. The single molecule spots were localised for all the frames in the set (Figure 6.3B). The localisations were used to construct tracks, by connecting spots within a radius of 8 pixels in successive frames (Figure 6.3D).

A clearer way of visualising PAmCherry-$CheB_2$ cellular positioning is by constructing a density map of the localisations. Figure 6.3C presents a density map of the localisations in 6.3B, obtained by convolution using a Gaussian mask.

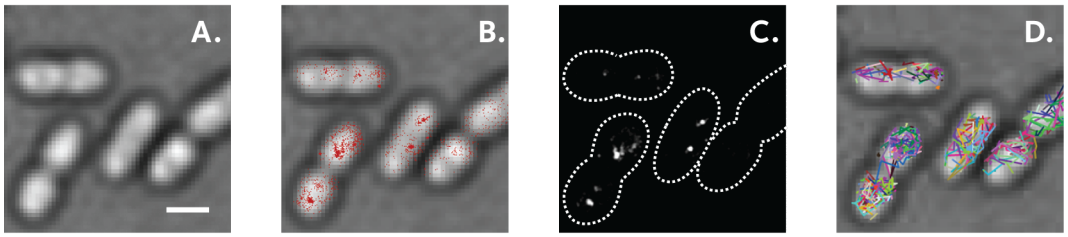

Figure 6.3: Localisation and tracking of PAmCherry-$CheB_2$ in WS8N. **A.** Transmitted light image of wild type cells containing PAmCherry-$CheB_2$. **B.** Localisations of individual fluorescent spots from an entire data set overlaid on to of the trasmitted light image. **C.** Density map of the localisations, obtained by convolution using a Gaussian mask. The colour map is grayscale with light areas indicating high density of localisation. **D.** Tracks of PAmCherry-$CheB_2$ overlaid on top of the transmitted light image. The scale bar in the first panel represents 1 μm.

Overlaying the localisations on top of the brightfield image provides a display of where PAmCherry-$CheB_2$ is most often located. By convoluting the localisations, a density map is produced, with the preferred localisations of PAmCherry-$CheB_2$ emerging as light spots (Figure 6.3C). The density map reveals a higher density of PAmCherry-$CheB_2$ localisations at specific points in the cytoplasm, as well as around the edge of the cell.

The tracks resulting from the localisations were plotted over the same field of view and revealed them to be distributed throughout the cell. (Figure 6.3D). This distribution was expected and is consistent with published data on $CheB_2$ appearing diffuse throughout the cytoplasm by fluorescence microscopy [115].

The distribution calculated D^* values for each of the PAmCherry-$CheB_2$ tracks in the WS8N data set was plotted in Figure 6.4A. A skewed distribution is obtained, with a large proportion of the tracks having a low D^*, close to zero. Without any prior knowledge of *in vivo* PAmCherry-$CheB_2$ diffusion coefficients, it is impossi-

ble to identify the distributions corresponding to bound and unbound PAmCherry-$CheB_2$.

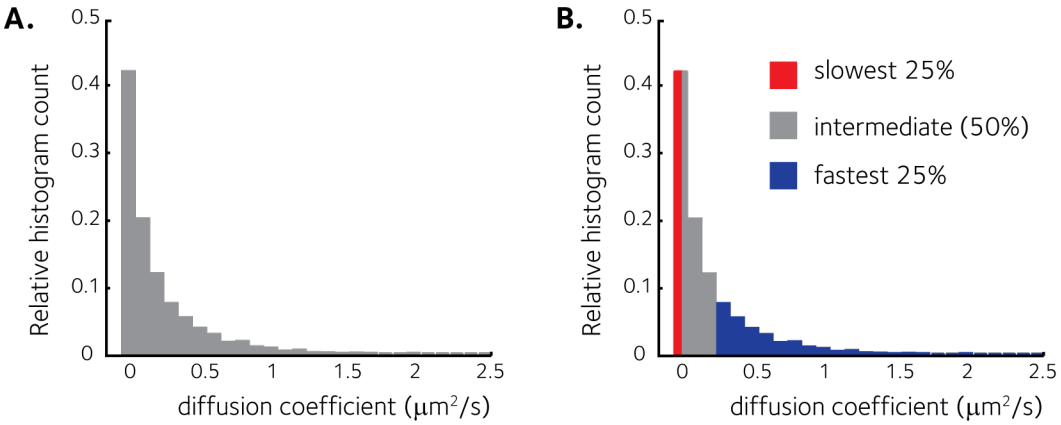

Figure 6.4: Tracks classification according to calculated diffusion coefficient. **A.** Histogram of the calculated D^* for PAmCherry-$CheB_2$ tracks in a wild-type background. **B.** Separation of the distribution into the slowest 25% tracks (D^* $\leq 0.0038 \mu m^2/s$) and fastest 25% tracks (D^* $\geq 0.2627 \mu m^2/s$).

Similar studies have benefited from well-separated distributions for slow and fast molecules, facilitating the classification of tracks into ones belonging to bound and unbound molecules [109, 31]. Alternatively, other studies used D^* values calculated for homologous proteins under similar experiment conditions [48].

To investigate if indeed, the population of tracks obtained here contains both slow-moving and fast-moving molecules with different cellular localisation patterns, the 25% slowest-moving tracks and the 25% fastest moving tracks were selected from the data set and carried forward for analysis. Figure 6.4B shows the slowest 25% with $D^* \leq 0.0038 \mu m^2/s$ coloured in red, and the fastest 25% with a $D^* \geq 0.2627 \mu m^2/s$. The tracks found in the intermediate 50% (grey) are likely to belong to both population of fast-moving and slow-moving PAmCherry-$CheB_2$.

By plotting these three categories of tracks separately, different cellular distributions of tracks emerge (Figure 6.5). The molecules with the lowest diffusion coefficient

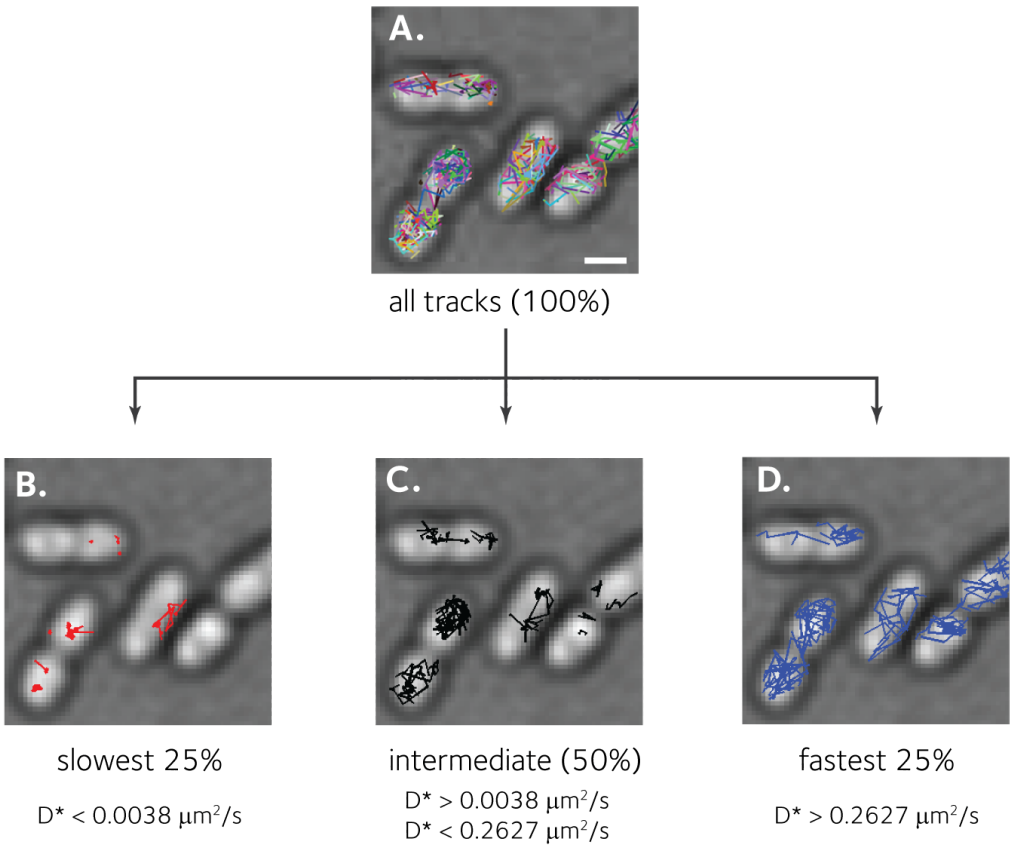

Figure 6.5: Tracks classification. The various subsets of tracks from the same field of view are plotted on top of the transmitted light image **A.** All tracks, each coloured in a different colour. **B.** The slowest 25% of tracks (red); **C.** The intermediate fraction of tracks (black); **D.** The fastest 25% of tracks (blue). The scale bar represents 1 μm.

(panel B) produced tracks that localise to discrete regions of the cell, either in the cytoplasm or towards the edge of the cell. In contrast, panel D shows the tracks with the highest diffusion coefficient covering the entire area of the cell. Panel C, displaying the tracks with intermediate diffusion coefficients shows trajectories zig-zagging throughout the area of the cell, as well as tracks localised to discrete regions.

Judging by the results displayed in Figure 6.5, the tracks have successfully been classified into slow, intermediate, and fast. Even though this approach only makes use of 50% of the data, it ensures that the two categories carried forward exclusively

contain tracks belonging to slow-moving and fast moving PAmCherry-CheB$_2$, respectively. The D^* limits used here to define the two categories of tracks in a wild-type background will be used in all subsequent tracks dataset classifications.

Overall, the tracks with the highest diffusion coefficients (blue) appear dispersed throughout the cytoplasm, while the tracks with the lowest diffusion coefficients (red) are concentrated to specific foci either in the cytoplasm or on the edge of the cell. The latter could be interpreted as the slower sub-population of PAmCherry-CheB$_2$ interacting with both chemosensory arrays.

6.5.3 Colocalisation with the chemosensory arrays

To further investigate whether the slower PAmCherry-CheB$_2$ molecules found to be concentrated to specific foci either in the cytoplasm or on the edge of the cell could indeed be interpreted as interacting with the chemosensory arrays, colocalisation experiments were carried out. To this end, PAmCherry-CheB$_2$ was conjugated into strains with fluorescently tagged polar and cytoplasmic arrays. Because the PALM microscope setup does not allow for the simultaneous imaging of a third fluorescent protein, separate strains were chosen to image the polar array (JPA1418 *yfp-cheW*$_3$) and the cytoplasmic array (JPA1457 *yfp-cheW*$_4$).

The same imaging, tracking, and track classification protocol as before was applied. As seen in a wild-type background, the slower tracks cluster in distinct areas of the cell (Figure 6.6C and 6.7C). The areas with an increased population of PAmCherry-CheB$_2$ are also highlighted in the density maps (Figure 6.6D and 6.7D) a similar distribution is seen.

Panel C from Figure 6.6 shows slow tracks located to the polar areas overlapping with the YFP-CheW$_3$ foci marking the polar chemosensory array. Similarly, panel

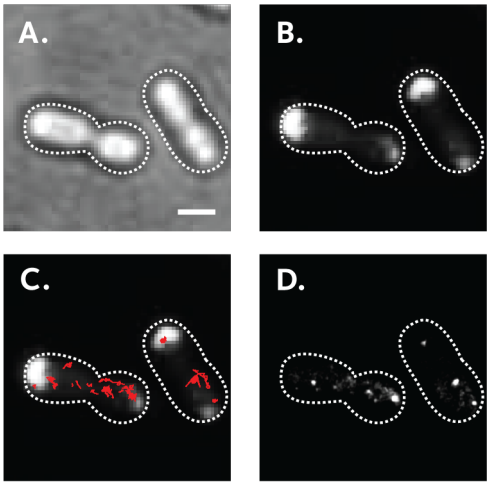

Figure 6.6: Colocalisation of PAmCherry-CheB$_2$ with the polar chemosensory array. The scale bar represents 1 μm and is the same for the whole series of images.

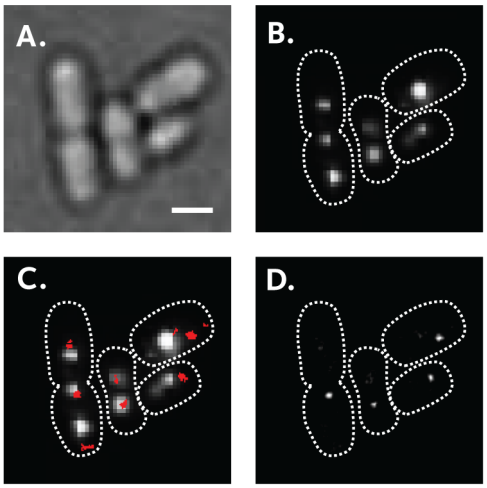

Figure 6.7: Colocalisation of PAmCherry-CheB$_2$ with the cytoplasmic chemosensory array. The scale bar represents 1 μm and is the same for the whole series of images.

C from Figure 6.7 shows a proportion of the slow PAmCherry-CheB$_2$ tracks on top of the YFP-CheW$_4$ foci marking the location of the cytoplasmic array. Both images have a number of slow PAmCherry-CheB$_2$ tracks not coinciding with a YFP spot, likely because this is where the untagged chemosensory array would be located.

Overall, the positions of the slowest 25% of the tracks colocalise with both the polar

and the cytoplasmic cluster.

6.5.4 PAmCherry-CheB$_2$ in the absence of chemosensory arrays

To confirm PAmCherry-CheB$_2$ interaction with the cytoplasmic chemosensory array, the same experiment was performed in a Δ*che*Op$_3$ background, ensuring that none of the array components – apart from TlpC which cannot form the array on its own [114] – are present. Under these conditions, a decrease in the number of slow tracks is expected. To further investigate whether the slow-moving PAmCherry-CheB$_2$ molecules were due to binding to a chemosensory array, the experiment was also performed in a Δ*che*Op$_{1,2,3}$ background. These cells would have no chemosensory arrays and a significant decrease – even depletion – of the tracks population with a $D^* \leq 0.0038$ μm^2/s is expected.

PALM data for these strains was acquired on at least two different days for at least two fields of view each, adding up to at least 3000 usable tracks per strain. The tracking, calculation of diffusion coefficients, and track classification was performed as before. For the boxplots representing diffusion coefficient distributions in Figure 6.8, all data available for that particular strain was used. The boxplots reveal diffusion coefficient distributions that appear very similar, apart from the Δ*che*Op$_{1,2,3}$ background which visibly shows a broader distribution.

A Kruskal-Wallis test with a significance of 0.05 was applied to judge whether these distribuitions are the same or not. The null hypothesis of the three distributions being the same was rejected, with pairwise comparisons showing that the three distributions are all different from each other.

As expected, the population of tracks for the Δ*che*Op$_{1,2,3}$ strain shows an increase in mean diffusion coefficient, but the effect is not as dramatic as predicted. The

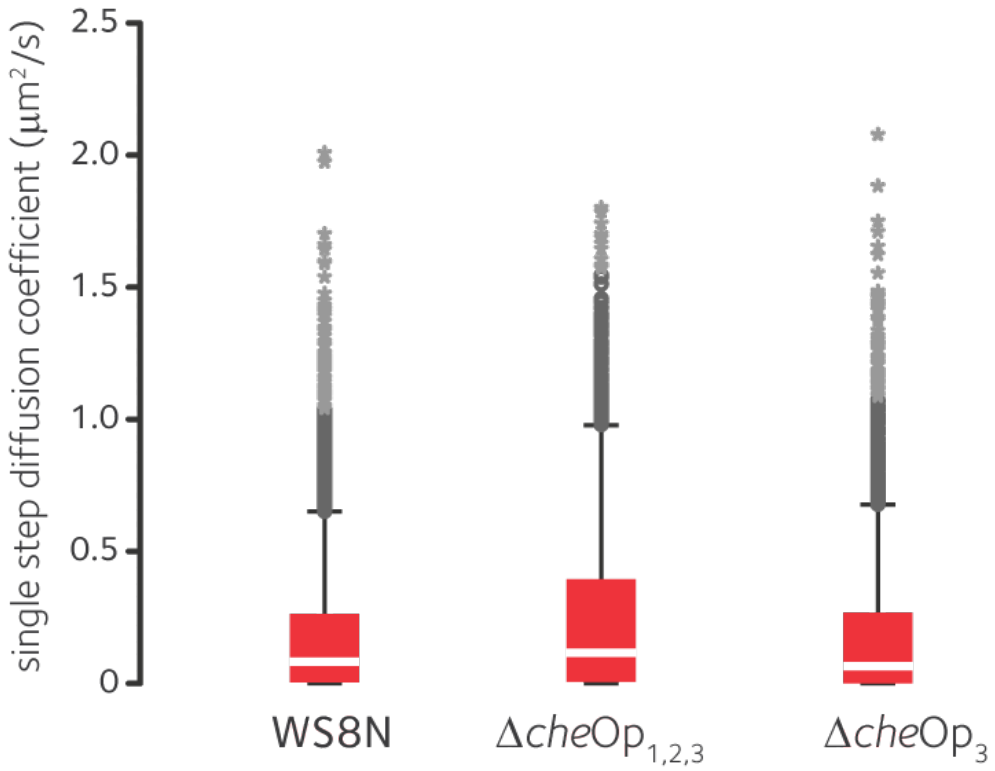

Figure 6.8: Boxplot showing diffusion coefficient distributions of D^* for PAmCherry-CheB$_2$ in a wild-type background (n=4249 tracks), a Δ*che*Op$_{1,2,3}$ background (n=4124 tracks), and a Δ*che*Op$_3$ background (n=3109 tracks).

corresponding histogram (Figure 6.9) confirms the persistence of a population with low diffusion coefficient.

Even though the histograms for the wild-type, Δ*che*Op$_3$, and Δ*che*Op$_{1,2,3}$ strains look very similar, differences emerge when single PAmCherry-CheB$_2$ molecule localisations and tracks are plotted in top of the transmitted light images (Figure 6.9). The general pattern that is observed shows the wild-type strain as described before, with slower tracks localising to discrete regions in the cytoplasm and at the edge of the cell. The density map highlights the cytoplasmic foci, suggesting a preference for this localisation. The tracks corresponding to fast molecules cover the whole area of the cell.

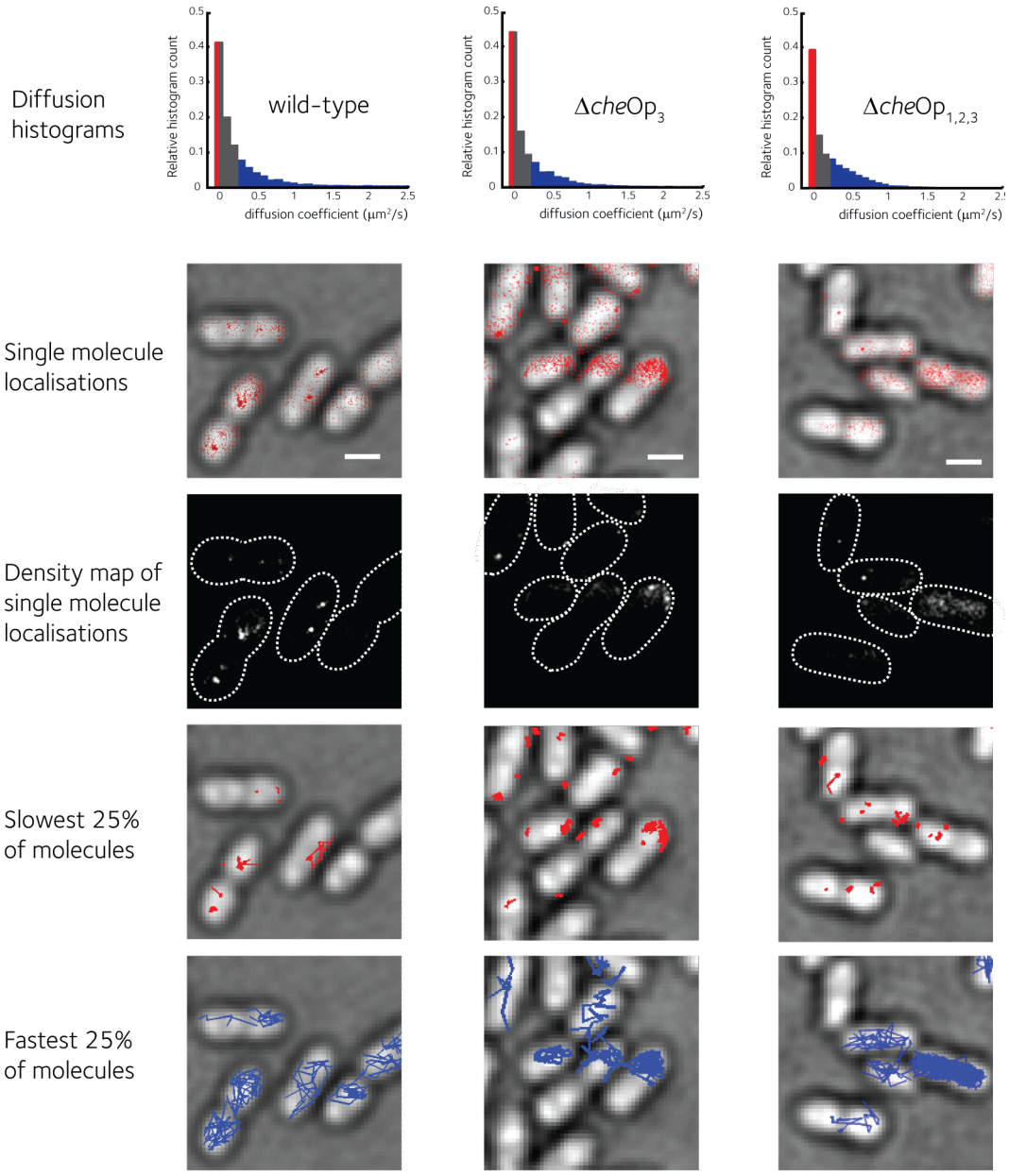

Figure 6.9: PAmCherry-CheB$_2$ tracks in deletion background strains. The data for wild-type, Δ*che*Op$_3$, and Δ*che*Op$_{1,2,3}$ is organised into columns. The histogram showing the distribution of diffusion coefficients for each strain is presented at the top of each column. Next, images showing single molecule localisations overlaid on top of the transmitted light image are presented. The second row of images shows the density maps of single molecule localisations for each strain. The last two rows of images show the slowest 25% and the fastest 25% of each population of tracks, overlaid on top of the corresponding transmitted light image The scale bar is 1μm and displayed in the first panel of the series.

In the case of the Δ*che*Op$_3$ strain, the localisation densities from the cytoplasm are no longer present. Indeed, the tracks corresponding to the slowest PAmCherry-CheB$_2$ molecules are localised at the cell poles and cell edges. The fast tracks fill the area of the cell in this case too.

When all the chemotaxis operons are deleted, in Δ*che*Op$_{1,2,3}$, a population of slow-moving cells persists and its localisation does not appear specific to any region of the cell. The slowest tracks, while present in discrete groups, are scattered throughout the cell, with no obvious preference for the cytoplasm or the edge of the cell. In this case too, the fast tracks fill the area of the cell.

6.6 Discussion

The current chapter describes the successful application of PALM imaging and tracking to provide qualitative insights into the cellular localisation of CheB$_2$. The ability to see below the diffraction limit has brought new *in vivo* information regarding the potential interaction partners of CheB$_2$.

First, PAmCherry-CheB$_2$ molecules could be classified into slow-moving and fast-moving, based on the calculated D^*. Because two distinct populations could not be identified from the distribution of D^*, subsets of the slowest and fastest 25% of tracks were used. These subsets of tracks were shown to have different and homogeneous localisation patterns. The slowest tracks according to the calculated diffusion coefficient, accumulated to distinct regions of the cell, either in the cytoplasm or at the edge of the cell or cell poles. The fastest 25% of tracks from all strains studied here were found to cover the whole area of the cell.

Second, the slow tracks were found to colocalise with both the polar and cytoplasmic arrays. Because the PALM microscope setup did not allow imaging of a *R.*

sphaeroides strain having both the polar and the cytoplasmic arrays tagged, it cannot be concluded that the slow-diffusing PAmCherry-$CheB_2$ molecules colocalise exclusively with the chemosensory arrays. This result is in agreement with existing *in vitro* and modelling data.

Third, work with deletion strains aimed to confirm interactions with the clusters by a process of elimination. In a Δ*che*Op_3 background lacking the cytoplasmic cluster and most of its components – apart from TlpC which is encoded in *che*Op_2 – the cellular distribution pattern changed to having concentrated spots of PAmCherry-$CheB_2$ at the cell pole. In some cases, slow tracks were localised around the edge of the cell, not only at the poles. This can be explained through the formation of mini-clusters that migrate from the old pole to the new pole during the cell division process [24].

Even though the images displaying PAmCherry-$CheB_2$ localisations did not show preference for any region of the cell in a Δ*che*Op_3 background, a high population of slow molecules was still present in a Δ*che*$Op_{1,2,3}$ background which does not have any chemosensory arrays. This was unexpected under our assumption of $CheB_2$ interacting with $CheA_2$ from the polar array (encoded in *che*Op_2). Deleting *che*$Op_{1,2,3}$ leaves the orphan chemoreceptors found mostly on Chromosome II (Table 1.1, which would still be expressed in the strain used here.

Studies on *E. coli* and *S. typhimurium* which contain a single CheR or CheB show that, apart from binding at the methylation site, some chemoreceptors bind CheR or CheB through additional high-affinity sites at the C terminus, known as the pentapetide tether [10]. Such a high affinity site has not yet been identified in *R. sphaeroides*, but their existence has been predicted in McpA, McpB, McpM, McpG, McpR, TlpC, TlpS, and TlpT using a relaxed consensus pentapeptide sequence [28]. The first three and TlpS are not expressed under laboratory conditions so will not

be considered. TlpT was not expressed in either of the deletion strains, while TlpC being expressed in a Δ*che*Op$_3$ background does not form cytoplasmic arrays and is unlikely to interact with PAmCherry-CheB$_2$ given the bacterial two-hybrid results (Section 3.4.4). The localisation of McpR has not yet been tested [28], hence McpG remains the only candidate from a high affinity pentapetide tether perspective.

Indeed, McpG was shown to localise at the cell poles and in the absence of *che*Op$_2$ proteins, McpG became delocalised throughout the membrane [112]. While a direct interaction between CheB$_2$ and McpG has not been observed, one could speculate that it is the reason for the high population of slow-diffusing PAmCherry-CheB$_2$ in a Δ*che*Op$_{1,2,3}$ background. This proposal could be tested by carrying out the same experiment in a background strain with an additional McpG deletion. In adition to that, the CheB$_2$-McpG interaction could also be probed unsing the BACTH method.

However, not just McpG, but any polarly localised chemoreceptor expressed under laboratory conditions (Table 1.1) is a good candidate for binding CheB$_2$, since none of the orphan chemoreceptor genes were deleted in the Δ*che*Op$_3$ and Δ*che*Op$_{1,2,3}$ strains used in the study.

At this stage, an interaction between CheB$_2$ and chemoreceptors in the polar array is mere speculation. However, entertaining this possibility could mean that CheB$_2$ is involved in the adaptation of the polar array, playing a role in deactivating CheA$_2$ when the cytoplasmic array is signalling.

While the predicted reverse phosphotransfer through CheB$_2$ may well be a way in which the two chemosensory arrays are connected, work in this chapter suggests it may not be the only way.

6.7 Conclusions and future work

To conclude, it was possible to identify two types of PAmCherry-$CheB_2$ based on apparent diffusion coefficient. It was also shown that PAmCherry-$CheB_2$ colocalises with both the polar and cytoplasmic clusters, as predicted by *in vitro* and modelling work. Interestingly, in the absence of all the chemotaxis operons, PAmCherry-$CheB_2$ interaction partners were still present, likely candidates being membrane-embedded chemoreceptors, such as McpG.

The results obtained here can be further developed by investigating potential interactions between $CheB_2$ and the transmembrane chemoreceptors through both PALM imaging and tracking, BACTH interactions assays, as well as other methods. The qualitative data described here would benefit from quantitative measurements of protein localisation in relation to cellular landmarks [103]. Lastly, the methods discussed here could be successfully applied to study the behaviour of other chemotaxis proteins such as the similarly diffuse $CheB_1$, or to establish whether the prediction of Robers *et al.* of $CheY_3$ and $CheY_4$ regulating the cytoplasmic cluster [83] hold true *in vivo*.

Moreover, regardless of the background strain used, the localisation density maps reveal high-density regions which correspond to the positioning of the chemosensory arrays at certain points during the cell cycle.

CHAPTER 7

Concluding remarks

The current model for chemotaxis proposes periplasmic sensing of decreasing propionate gradients by the polar chemosensory array and cytoplasmic metabolic sensing by the cytoplasmic array. Signalling by the polar array is assumed to occur through $CheY_{3,4,6}$–P and signalling through the cytoplasmic array only through $CheY_6$–P. Adaptation at the polar and cytoplasmic arrays is assumed to occur through $CheB_1$, $CheR_2$, and $CheB_1$, $CheR_2$, respectively.

Results described in this thesis have showed that this model does not correctly predict the phenotypes for Δ*cheB*$_1$ and Δ*cheR*$_2$. It also does not predict adaptive response to addition of propionate. The current model does not take into account the methanol release experiments that have suggested that adaptation occurs both upon addition and removal of propionate [64].

A new model for chemotaxis in *R. sphaeroides* is proposed (Figure 7.1) in an attempt to incorporate the results described in this thesis. The proposed model senses positive stimuli *e.g.* increase in propionate through the polar array and signals only through $CheY_{3,4}$. $CheY_{3,4}$ is proposed to introduce bias towards runs, potentially by displacing $CheY_6$–P from the motor. The cytoplasmic cluster senses the metabolic

state of the cell which could reflect negative stimuli *e.g.* decrease in propionate, and effect a stop by signalling only through $CheY_6$–P, the only CheY capable of stopping the motor. It also includes the possibility of $CheB_2$ desensitizing the polar array as well (red arrow) and the putative role of TlpC in regulating the kinase/phosphatase equilibrium (blue arrow)

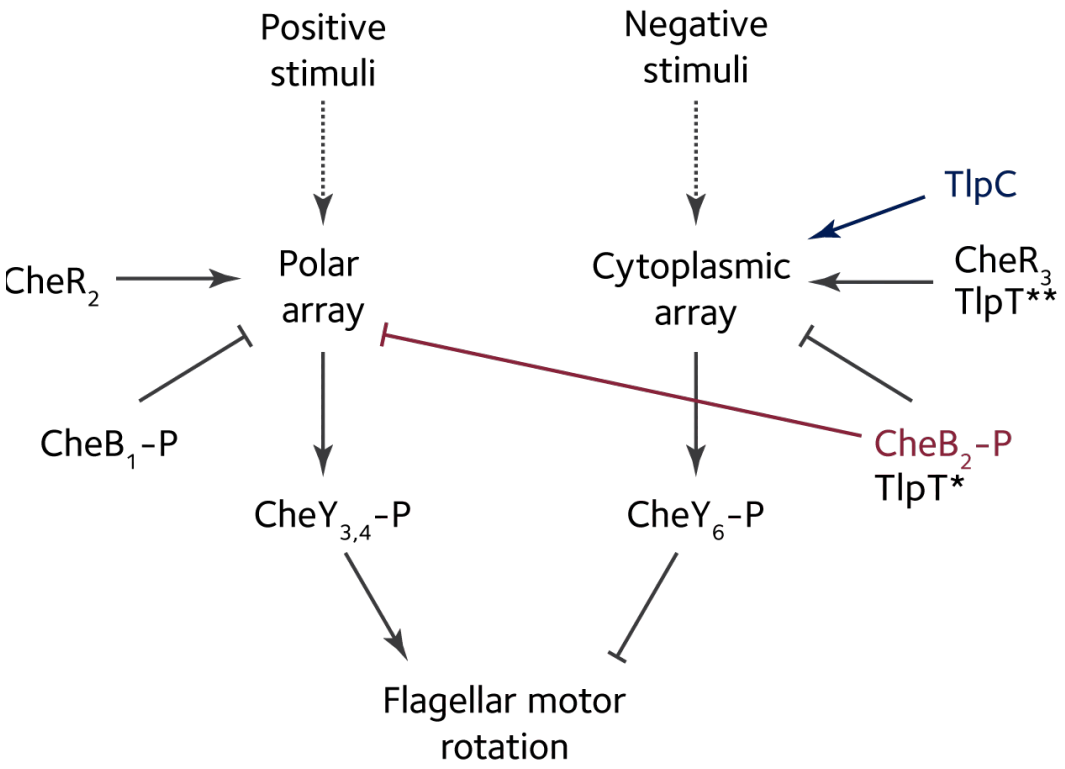

Figure 7.1: Outline of the new model proposed for *R. sphaeroides* chemotaxis.

Building the interaction network between the components of the cytoplasmic cluster revealed a network architecture that mirrors the one of the well-studied polar arrays. $CheA_4$ was shown to interact with itself, and further biophysical studies revealed it forms dimers. This suggests that the role of $CheA_4$ within the cytoplasmic array might be to connect the hexagonal rings in the CheA/CheW baseplates through its P3 dimerisation domain. Similar studies could be carried out to find out whether $CheA_3$ could perform a different function.

The $CheA_4$-TlpC pair stood out by not showing any interactions in the two-hybrid

assays. This could mean that the two cytoplasmic chemoreceptors respond to or regulate different responses. It could be the case, for example, that TlpT – which interacts with both $CheA_3$ and $CheA_4$ – regulates kinase activity at the cytoplasmic array, while TlpC regulates the phosphatase activity of $CheA_3$.

These putative interactions propose an antagonistic action of the signalling output from the two cytoplasmic chemoreceptors, threby regulating the level of $CheY_6$–P in the cell. This model is consistent with partial $CheY_6$ localisation at the cytoplasmic array.

Further, steady state swimming experiments performed with combinations of adaptation protein mutants suggested distinct roles for the two arrays. Smooth swimming was found to be a result of increased signalling through the polar array, while stoppy swimming occurs when the cytoplasmic array signalling is increased. This correlates the polar cluster with producing runs, and the cytoplasmic cluster with producing stops. It is thus unlikely that $CheY_6$ gets phosphorylated to physiologically relevant levels by $CheA_2$.

Super-sensitized TlpT mutants are non-motile due to the flagellar motor being saturated with $CheY_6$–P. This can be temporarily displaced through addition of propionate which causes the cells to start rotating. Rotation ceases typically within 3-4 minutes.

Δ*tlpC* suggested that TlpC is involved in the response to dynamic changes, by having a wild-type free swimming phenotype and a non-responsive tethering phenotype. The proposed role of TlpC in maintaining the balance between kinase and phosphatase activities could be applicable to this scenario.

Despite unsuccessful attempts to map the interactions of the cytoplasmic array proteins with the adaptation proteins, single molecule microscopy has provided in-

sight into the cellular localisation of $CheB_2$. It was revealed to predominantly colocalise with the cytoplasmic array, though a Δ*che*Op_3 strain showed it accumulating to the cell poles. It was also speculated that $CheB_2$ might interact with MCPs, as suggested by localisation experiments with both the Δ*che*Op_3 and Δ*che*$Op_{1,2,3}$ strains.

Overall, the results obtained here fit with a chemotaxis model where the polar array responds and adapts to positive changes in periplasmic propionate concentrations and the cytoplasmic array elicits and adaptive response to metabolic cues produced as a result of a decrease in propionate. Evidence for methylation-based adaptation occurring at the cytoplasmic array was brought through *in vivo* study of methylation site mutants. Additionally, the data suggest separate roles for the two cytoplasmic chemoreceptors, and propose TlpC to be involved – directly or indirectly – in regulating phosphatase activity. Experimental proof for the latter could validate the model proposed here.

CHAPTER 8

Appendices

Table 8.1: Primers used for BACTH cloning. Restriction sites are highlighted in bold.

Primer name	Sequence
B2H-cheA3-F	CATCAT**TCTAGA**ATGTCCGACGCATTCGACG
B2H-cheA3-R	ATGATGGGTACCGCCGGCTTCGGCAGGCAGC
B2H-cheA4-F	CATCAT**TCTAGA**ATGAGCGCGCCTGTCGAGGCC
B2H-cheA4-R	ATGATGGGTACCTGCCGCACCTCCCGCCG
B2H-cheB2-F	CATCAT**TCTAGA**AATGAAGGCTGCCGCCGACCG
B2H-cheB2-R	ATGATGGGTACCCTGTCGTTTGACGCGAGGCTG
B2H-cheR1-F	CATCAT**TCTAGA**AGTGAAATACTGGTCGCG
B2H-cheR1-R	ATGATGGGTACCCTATGCTGTCCCTGAGTGCC
B2H-cheR2-F	CATCAT**TCTAGA**AATGAGTCAGGCCTTGCGTCC
B2H-cheR2-R	ATGATGGGTACCCTGGTTTTCCGAAAGACTGAGG
B2H-cheR3-F	CATCAT**TCTAGA**AATGACGCTCCGCGAGGAGGCG
B2H-cheR3-R	ATGATGGGTACCCTGCGGGCCTCCTTCTGATGG
B2H-cheW4-F	CATATG**TCTAGA**ATGCCCGAAGACCTGACC
B2H-cheW4-R	ATGTATGGGTACCGGCCTCCAGCTCGGCGGG
B2H-tlpC-F	CATCAT**TCTAGA**ATGAAACTGGAGCCCGATCTGC
B2H-tlpC-R	ATGATGGGTACCGAAGAAGATGTCTGCCAGATCC
B2H-tlpT-F	CATCAT**TCTAGA**ATGACACGCCGCCCCAAGACG
B2H-tlpT-R	ATGATGGGTACCGAAGTCGCCGAAGCCCCGC

Table 8.2: Extra primers designed for BACTH construct sequencing.

Primer name	Sequence
T18-F	GGAAAAGCCTGTTCGACGATGG
T18-R	CCGATATTCATGTCGCCGTCG
T25-F	CGATTTCGAGGCGGTCAAGG
T25-R	CCAGGCGGAACATCAATGTGG

Table 8.3: Primers used for PALM construct cloning. Restriction sites or modified sites are highlighted in bold.

Primer name	Sequence
B2PALM-PAmCh-F	CATACT**CATATG**GTGAGCAAGGGCGAGGAGG
B2PALM-PAmCh-R	CATACT**CATATG**CTTGTACAGCTCGTCCATGC
B2PALM-check-F	CGTATCACGAGGCCCTTTCGTCTTCACCTC
B2PALM-check-R	CGTTCGTACTGTTCCACGATGGTGTAGTCC
QC-B2(D70N)-F	CGACGTGGTGACGCTC**A**ATCTCGAGATGCC
QC-B2(D70N)-R	GGCATCTCGAGAT**T**GAGCGTCACCACGTCG
QC-B2PALM-F	AAAGAGGAGAAATTA**CAT**ATGAAGGCTGCCGCCG
QC-B2PALM-R	CGGCGGCAGCCTTCAT**ATG**TAATTTCTCCTCTTT

Table 8.4: Primers used for colony PCR. These primers were used to screen for the insertion of *yfp* upstream of $cheW_4$ after the second recombination step.

Primer name	Sequence
rec-X-W4-F	CGCCCAGCCTCGCGTCAAACGACTGAACCAGAGGAGACG
rec-X-W4-R	GGTCTTGCCCACGCCCCCCTTCTGGTTGCAGATCATCG

Recipe 8.5: Luria Bertatni (LB) broth. Ingredients were dissolved in MilliQ water, pH adjusted to 7.0 and autoclaved.

Ingredient	Quantity per litre
Yeast extract	5 g
Bacto-tryptone	10 g
NaCl	5 g

Recipe 8.6: 2TY media. Ingredients were dissolved in MilliQ water, pH adjusted to 7.0 and autoclaved.

Ingredient	Quantity per litre
Yeast extract	10 g
Bacto-tryptone	16 g
NaCl	5 g

Recipe 8.7: Succinate media (Sux) Ingredients were dissolved in MilliQ water, pH adjusted to 7.2 with KOH and autoclaved.

Ingredient	Quantity per litre
1M Phosphate buffer pH 7.0	20 mL
Concentrated base (see below)	20 mL
Growth factors (see below)	2 mL
$(NH_4)_2SO_4$	0.5 g
NaCl	0.5 g
Sodium succinate	2 g
Casamino acids	1 g

Recipe 8.8: LB agar. Ingredients were dissolved in MilliQ water, pH adjusted to 7.0 and autoclaved.

Ingredient	Quantity per litre
Agar	20
Yeast extract	5 g
Bacto-tryptone	5 g
NaCl	5 g

Recipe 8.9: 1 M phosphate buffer pH 7.0 Buffer pH was adjusted with KOH, the solution autoclaved and stored at 4°C.

Ingredient	Quantity per litre
K_2HPO_4	174.18 g
KH_2PO_4	136.09 g

Recipe 8.10: Concentrated baseNTA was dissolved first and pH adjusted to 5.0 with KOH. This pH allows $FeSO_4$ to dissolve. After the addition of all the other ingredients, the pH was adjusted to 6.8 with KOH. The solution was autoclaved and stored a t 4°C.

Ingredient	Quantity per litre
Nitrilotriacetic acid (NTA)	5.94 g
Metals 44 solution	25 mL
$MgSO_4 \cdot 7H_2O$	14.5 g
$CaCl_2 \cdot 2H_2O$	1.67 g
$FeSO_4 \cdot 7H_2O$	50 mg
Ammonium molybdate $\cdot 4\ H_2O$	4.6 mg

Recipe 8.11: Metals 44 solution The solution was filter sterilised and stored at 4°C.

Ingredient	Quantity per litre
EDTA	2.5 g
$ZnSO_4 \cdot 7H_2O$	11 g
$FeSO_4 \cdot 7H_2O$	5 g
$CuSO_4 \cdot 5H_2O$	0.39 g
$MnSO_4 \cdot 4H_2O$	3 g
$CoCl_2 \cdot 6H_2O$	0.2 g
3M H_2SO_4	1.5 mL
H_3BO_3 0.12 g	

Recipe 8.12: Growth factor The solution was autoclaved and stored at 4°C.

Ingredient	Quantity per litre
Biotin	20 mg
$NaHCO_3$	0.5 g
Niacin	1g
Thiamine hydrochloride	0.5 g

Recipe 8.13: M22 minimal media The solution was made in MilliQ water, the pH adjusted to 7.2 with KOH, and autoclaved.

Ingredient	Quantity per litre
1 M phosphate buffer	20 mL
Concentrated base	20 mL
Growth factor	2 mL
$(NH_4)_2SO_4$	0.5 g
NaCl	0.5 g

Recipe 8.14: 10x M9 salts The salts were dissolved in MilliQ water.

Ingredient	Quantity per litre
Na_2HPO_4	75.2 g
KH_2PO_4	30 g
NH_4Cl	10 g
NaCl	5 g

Recipe 8.15: SOC media

Ingredient	Concentration
Glucose	20 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
NaCl	10 mM
Tryptone	20 g/L
Yeast extract	5 g/L

Recipe 8.16: 5x DNA loading dye

Ingredient	Quantity w/v
Glycerol	70%
Bromophenol blue	0.1%
Xylene cyanol	0.1 %
TBE buffer	29.8%

Recipe 8.17: 10x TBE Ingredients were dissolved into MilliQ and pH adjusted to 8.3.

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

Recipe 8.18: TFB I The pH was adjusted to 5.8 with 0.2 M acetic acid and the final solution was filter sterilised.

Ingredient	Concentration
KCH_3COO	30 mM
RbCl	100 mM
$CaCl_2 \cdot 2H_2O$	10 mM
$MnCl_2 \cdot 4H_2O$	50 mM
Glycerol	15% (v/v)

Recipe 8.19: TFB II The pH was adjusted to 6.8 with 0.2 M HCl and the final solution was filter sterilised.

Ingredient	Concentration
PIPES	10 mM
RbCl	10 mM
$CaCl_2 \cdot 2H_2O$	75 mM
Glycerol	15% (v/v)

Recipe 8.20: PIPES buffer The pH was adjusted to 7.2 with HCl and the final solution was filter sterilised.

Ingredient	Concentration
Piperazine-N,N‘-bis(2-ethanesulfonic acid)	10 mM

Recipe 8.21: Phosphate buffer saline (PBS)

Ingredient	Concentration
Dulbecco A PBS tablets	10 in 1 L

Bibliography

[1] Adler, J. (1966). Chemotaxis in bacteria. *Science*, 153(3737):708.

[2] Adler, J. (1983). Bacterial chemotaxis and molecular neurobiology. *Cold Spring Harb Symp Quant Biol*, 48 Pt 2:803.

[3] Alexander, R.P., Lowenthal, A.C., Harshey, R.M., and Ottemann, K.M. (2010). CheV: CheW-like coupling proteins at the core of the chemotaxis signaling network. *Trends Microbiol*, 18(11):494.

[4] Alexander, R.P. and Zhulin, I.B. (2007). Evolutionary genomics reveals conserved structural determinants of signaling and adaptation in microbial chemoreceptors. *Proc Natl Acad Sci U S A*, 104(8):2885.

[5] Allen, J.R. (2014). *The structure, function and specificity of the Rhodobacter sphaeroides membrane-associated chemotaxis array*. Ph.D. thesis, Department of Biochemistry, University of Oxford.

[6] Alon, U., Camarena, L., Surette, M.G., y Arcas, B.A., Liu, Y., Leibler, S., and Stock, J.B. (1998). Response regulator output in bacterial chemotaxis. *EMBO J*, 17(15):4238.

[7] Amin, M., Porter, S.L., and Soyer, O.S. (2013). Split histidine kinases enable ultrasensitivity and bistability in two-component signaling networks. *PLoS Comput Biol*, 9(3):e1002949.

[8] Armitage, J.P. and Macnab, R.M. (1987). Unidirectional, intermittent rotation of the flagellum of *Rhodobacter sphaeroides*. *J Bacteriol*, 169(2):514.

[9] Armitage, J.P., Pitta, T.P., Vigeant, M.A., Packer, H.L., and Ford, R.M. (1999). Transformations in flagellar structure of *Rhodobacter sphaeroides* and possible relationship to changes in swimming speed. *J Bacteriol*, 181(16):4825.

[10] Bartelli, N.L. and Hazelbauer, G.L. (2011). Direct evidence that the carboxyl-terminal sequence of a bacterial chemoreceptor is an unstructured linker and enzyme tether. *Protein Sci*, 20(11):1856.

[11] Battesti, A. and Bouveret, E. (2012). The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. *Methods*, 58(4):325.

[12] Bell, C.H., Porter, S.L., Strawson, A., Stuart, D.I., and Armitage, J.P. (2010). Using structural information to change the phosphotransfer specificity of a two-component chemotaxis signalling complex. *PLoS Biol*, 8(2):e1000306.

[13] Berg, H.C. and Anderson, R.A. (1973). Bacteria swim by rotating their flagellar filaments. *Nature*.

[14] Berg, H.C., Brown, D.A., *et al.* (1972). Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*, 239(5374):500.

[15] Bilwes, A.M., Alex, L.A., Crane, B.R., and Simon, M.I. (1999). Structure of CheA, a signal-transducing histidine kinase. *Cell*, 96(1):131.

[16] Block, S.M., Segall, J.E., and Berg, H.C. (1983). Adaptation kinetics in bacterial chemotaxis. *J Bacteriol*, 154(1):312.

[17] Bourret, R.B., Davagnino, J., and Simon, M.I. (1993). The carboxy-terminal portion of the CheA kinase mediates regulation of autophosphorylation by transducer and CheW. *J Bacteriol*, 175(7):2097.

[18] Briegel, A., Ladinsky, M.S., Oikonomou, C., Jones, C.W., Harris, M.J., Fowler, D.J., Chang, Y.W., Thompson, L.K., Armitage, J.P., and Jensen, G.J. (2014). Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling. *Elife*, 3:e02151.

[19] Briegel, A., Li, X., Bilwes, A.M., Hughes, K.T., Jensen, G.J., and Crane, B.R. (2012). Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. *Proc Natl Acad Sci U S A*, 109(10):3766.

[20] Briegel, A., Ortega, D.R., Huang, A.N., Oikonomou, C.M., Gunsalus, R.P., and Jensen, G.J. (2015). Structural conservation of chemotaxis machinery across Archaea and Bacteria. *Environ Microbiol Rep*, 7(3):414.

[21] Briegel, A., Ortega, D.R., Tocheva, E.I., Wuichet, K., Li, Z., Chen, S., Müller, A., Iancu, C.V., Murphy, G.E., Dobro, M.J., Zhulin, I.B., and Jensen, G.J. (2009). Universal architecture of bacterial chemoreceptor arrays. *Proc Natl Acad Sci U S A*, 106(40):17181.

[22] Brown, M.T. (2009). *Control of the unidirectional motor in Rhodobacter sphaeroides*. Ph.D. thesis, Department of Biochemistry, University of Oxford.

[23] Brückner, A., Polge, C., Lentze, N., Auerbach, D., and Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci*, 10(6):2763.

[24] Chiu, S.W., Roberts, M.A., Leake, M.C., and Armitage, J.P. (2013). Positioning of chemosensory proteins and FtsZ through the Rhodobacter sphaeroides cell cycle. *Mol Microbiol*, 90(2):322.

[25] Clarke, S. and Koshland, Jr, D. (1979). Membrane receptors for aspartate and serine in bacterial chemotaxis. *J Biol Chem*, 254(19):9695.

[26] Clayton, R.K. and Clayton, B.J. (1972). Relations between pigments and proteins in the photosynthetic membranes of *Rhodopseudomonas spheroides*. *Biochim Biophys Acta*, 283(3):492.

[27] Crocker, J.C. and Grier, D.G. (1996). Methods of digital video microscopy for colloidal studies. *J Colloid Interface Sci*, 179(1):298.

[28] de Beyer, J.A. (2013). *Roles of the two chemotaxis clusters in Rhodobacter sphaeroides*. Ph.D. thesis, Department of Biochemistry, University of Oxford.

[29] del Campo, A., Ballado, T., Camarena, L., and Dreyfus, G. (2011). In *Rhodobacter sphaeroides*, chemotactic operon 1 regulates rotation of the flagellar system 2. *J Bacteriol*, 193(23):6781.

[30] del Campo, A.M., Ballado, T., de la Mora, J., Poggio, S., Camarena, L., and Dreyfus, G. (2007). Chemotactic control of the two flagellar systems of *Rhodobacter sphaeroides* is mediated by different sets of CheY and FliM proteins. *J Bacteriol*, 189(22):8397.

[31] Diepold, A., Kudryashev, M., Delalez, N.J., Berry, R.M., and Armitage, J.P. (2015). Composition, formation, and regulation of the cytosolic C-ring, a dynamic component of the type III secretion injectisome. *PLoS Biol*, 13(1).

[32] Erbse, A.H. and Falke, J.J. (2009). The core signaling proteins of bacterial chemotaxis assemble to form an ultrastable complex. *Biochemistry*, 48(29):6975.

[33] Falke, J.J. and Piasta, K.N. (2014). Architecture and signal transduction mechanism of the bacterial chemosensory array: progress, controversies, and challenges. *Curr Opin Struct Biol*, 29:85.

[34] García-Fontana, C., Corral Lugo, A., and Krell, T. (2014). Specificity of the $CheR_2$ methyltransferase in *Pseudomonas aeruginosa* is directed by a C-terminal pentapeptide in the McpB chemoreceptor. *Sci Signal*, 7(320):ra34.

[35] Gauliard, E., Ouellette, S.P., Rueden, K.J., and Ladant, D. (2015). Characterization of interactions between inclusion membrane proteins from *Chlamydia trachomatis. Front Cell Infect Microbiol*, 5:13.

[36] Hamadeh, A., Roberts, M.A., August, E., McSharry, P.E., Maini, P.K., Armitage, J.P., and Papachristodoulou, A. (2011). Feedback control architecture and the bacterial chemotaxis network. *PLoS Comput Biol*, 7(5):e1001130.

[37] Hamblin, P.A., Bourne, N.A., and Armitage, J.P. (1997). Characterization of the chemotaxis protein CheW from *Rhodobacter sphaeroides* and its effect on the behaviour of *Escherichia coli*. *Mol Microbiol*, 24(1):41.

[38] Hamblin, P.A., Maguire, B.A., Grishanin, R.N., and Armitage, J.P. (1997). Evidence for two chemosensory pathways in Rhodobacter sphaeroides. *Mol Microbiol*, 26(5):1083.

[39] Hamer, R., Luo, Q., Armitage, J.P., Reinert, G., and Deane, C.M. (2010). i-Patch: Interprotein contact prediction using local network information. *Proteins: Struct , Funct , Bioinf*, 78(13):2781.

[40] Harrison, P.W., Lower, R.P.J., Kim, N.K.D., and Young, J.P.W. (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol*, 18(4):141.

[41] Hazelbauer, G.L., Engström, P., and Harayama, S. (1981). Methyl-accepting chemotaxis protein III and transducer gene *trg*. *J Bacteriol*, 145(1):43.

[42] Holden, S.J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O.J., and Kapanidis, A.N. (2010). Defining the limits of single-molecule FRET resolution in TIRF microscopy. *Biophys J*, 99(9):3102.

[43] Ind, A.C., Porter, S.L., Brown, M.T., Byles, E.D., de Beyer, J.A., Godfrey, S.A., and Armitage, J.P. (2009). Inducible-expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*. *Appl Environ Microbiol*, 75(20):6613.

[44] Jahreis, K., Morrison, T.B., Garzón, A., and Parkinson, J.S. (2004). Chemotactic signaling by an *Escherichia coli* CheA mutant that lacks the binding domain for phosphoacceptor partners. *J Bacteriol*, 186(9):2664.

[45] Jankute, M., Byng, C.V., Alderwick, L.J., and Besra, G.S. (2014). Elucidation of a protein-protein interaction network involved in *Corynebacterium glutamicum* cell wall biosynthesis as determined by bacterial two-hybrid analysis. *Glycoconj J*, 31(6-7):475.

[46] Jarrell, K.F. and McBride, M.J. (2008). The surprisingly diverse ways that prokaryotes move. *Nat Rev Microbiol*, 6(6):466.

[47] Johnson, C.M. (2013). Differential scanning calorimetry as a tool for protein folding and stability. *Arch Biochem Biophys*, 531(1-2):100.

[48] Jones, C.W. (2013). *Dynamics, formation and segregation of the cytoplasmic chemoreceptor cluster in Rhodobacter sphaeroides*. Ph.D. thesis, University of Oxford, Medical Sciences Division, Department of Biochemistry.

[49] Jones, C.W. and Armitage, J.P. (2015). Positioning of bacterial chemoreceptors. *Trends Microbiol*, 23(5):247.

[50] Karimova, G., Dautin, N., and Ladant, D. (2005). Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol*, 187(7):2233.

[51] Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A*, 95(10):5752.

[52] Karimova, G., Ullmann, A., and Ladant, D. (2001). Protein-protein interaction between *Bacillus stearothermophilus* tyrosyl-tRNA synthetase subdomains revealed by a bacterial two-hybrid system. *J Mol Microbiol Biotechnol*, 3(1):73.

[53] Kim, K.K., Yokota, H., and Kim, S.H. (1999). Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature*, 400(6746):787.

[54] Kobayashi, K., Saitoh, T., Shah, D.S.H., Ohnishi, K., Goodfellow, I.G., Sockett, R.E., and Aizawa, S.I. (2003). Purification and characterization of the flagellar basal body of *Rhodobacter sphaeroides*. *J Bacteriol*, 185(17):5295.

[55] Kojadinovic, M., Armitage, J.P., Tindall, M.J., and Wadhams, G.H. (2013). Response kinetics in the complex chemotaxis signalling pathway of *Rhodobacter sphaeroides*. *J R Soc Interface*, 10(81):20121001.

[56] Kojadinovic, M., Sirinelli, A., Wadhams, G.H., and Armitage, J.P. (2011). New motion analysis system for characterization of the chemosensory response kinetics of *Rhodobacter sphaeroides* under different growth conditions. *Appl Environ Microbiol*, 77(12):4082.

[57] Lancero, H.L., Castaneda, S., Caberoy, N.B., Ma, X., Garza, A.G., and Shi, W. (2005). Analysing protein-protein interactions of the *Myxococcus xanthus* Dif signalling pathway using the yeast two-hybrid system. *Microbiology*, 151(Pt 5):1535.

[58] Larsen, S.H., Reader, R.W., Kort, E.N., Tso, W.W., and Adler, J. (1974). Change in direction of flagellar rotation is the basis of the chemotactic response in *Escherichia coli*. *Nature*.

[59] Lavinder, J.J., Hari, S.B., Sullivan, B.J., and Magliery, T.J. (2009). High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. *J Am Chem Soc*, 131(11):3794.

[60] Legrain, P. and Selig, L. (2000). Genome-wide protein interaction maps using two-hybrid systems. *FEBS Lett*, 480(1):32.

[61] Mackenzie, C., Choudhary, M., Larimer, F.W., Predki, P.F., Stilwagen, S., Armitage, J.P., Barber, R.D., Donohue, T.J., Hosler, J.P., Newman, J.E., Shapleigh, J.P., Sockett, R.E., Zeilstra-Ryalls, J., and Kaplan, S. (2001). The home stretch, a first analysis of the nearly completed genome of *Rhodobacter sphaeroides* 2.4.1. *Photosynth Res*, 70(1):19.

[62] Marcoux, J. and Robinson, C.V. (2013). Twenty years of gas phase structural biology. *Structure*, 21(9):1541.

[63] Martin, A.C., Wadhams, G.H., and Armitage, J.P. (2001). The roles of the multiple CheW and CheA homologues in chemotaxis and in chemoreceptor localization in *Rhodobacter sphaeroides*. *Mol Microbiol*, 40(6):1261.

[64] Martin, A.C., Wadhams, G.H., Shah, D.S., Porter, S.L., Mantotta, J.C., Craig, T.J., Verdult, P.H., Jones, H., and Armitage, J.P. (2001). CheR- and CheB-dependent chemosensory adaptation system of *Rhodobacter sphaeroides*. *J Bacteriol*, 183(24):7135.

[65] Michalet, X. and Berglund, A.J. (2012). Optimal diffusion coefficient estimation in single-particle tracking. *Phys Rev E Stat Nonlin Soft Matter Phys*, 85(6 Pt 1):061916.

[66] Miño, G., Mallouk, T.E., Darnige, T., Hoyos, M., Dauchet, J., Dunstan, J., Soto, R., Wang, Y., Rousselet, A., and Clement, E. (2011). Enhanced diffusion due to active swimmers at a solid surface. *Phys Rev Lett*, 106(4):048102.

[67] Moine, A., Agrebi, R., Espinosa, L., Kirby, J.R., Zusman, D.R., Mignot, T., and Mauriello, E.M.F. (2014). Functional organization of a multimodular bacterial chemosensory apparatus. *PLoS Genet*, 10(3):e1004164.

[68] Penfold, R.J. and Pemberton, J.M. (1992). An improved suicide vector for construction of chromosomal insertion mutations in bacteria. *Gene*, 118(1):145.

[69] Piasta, K.N. and Falke, J.J. (2014). Increasing and decreasing the ultrastability of bacterial chemotaxis core signaling complexes by modifying protein-protein contacts. *Biochemistry*, 53(35):5592.

[70] Poggio, S., Abreu-Goodger, C., Fabela, S., Osorio, A., Dreyfus, G., Vinuesa, P., and Camarena, L. (2007). A complete set of flagellar genes acquired by horizontal transfer coexists with the endogenous flagellar system in *Rhodobacter sphaeroides*. *J Bacteriol*, 189(8):3208.

[71] Poole, P.S. and Armitage, J.P. (1988). Motility response of Rhodobacter sphaeroides to chemotactic stimulation. *J Bacteriol*, 170(12):5673.

[72] Poole, P.S., Sinclair, D.R., and Armitage, J.P. (1988). Real time computer tracking of free-swimming and tethered rotating cells. *Anal Biochem*, 175(1):52.

[73] Porter, S.L. and Armitage, J.P. (2002). Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. *J Mol Biol*, 324(1):35.

[74] Porter, S.L. and Armitage, J.P. (2004). Chemotaxis in *Rhodobacter sphaeroides* requires an atypical histidine protein kinase. *J Biol Chem*, 279(52):54573.

[75] Porter, S.L., Roberts, M.A.J., Manning, C.S., and Armitage, J.P. (2008). A bifunctional kinase-phosphatase in bacterial chemotaxis. *Proc Natl Acad Sci U S A*, 105(47):18531.

[76] Porter, S.L., Wadhams, G.H., and Armitage, J.P. (2008). *Rhodobacter sphaeroides*: complexity in chemotactic signalling. *Trends Microbiol*, 16(6):251.

[77] Porter, S.L., Wadhams, G.H., and Armitage, J.P. (2011). Signal processing in complex chemotaxis pathways. *Nat Rev Microbiol*, 9(3):153.

[78] Porter, S.L., Wadhams, G.H., Martin, A.C., Byles, E.D., Lancaster, D.E., and Armitage, J.P. (2006). The CheYs of *Rhodobacter sphaeroides*. *J Biol Chem*, 281(43):32694.

[79] Porter, S.L., Warren, A.V., Martin, A.C., and Armitage, J.P. (2002). The third chemotaxis locus of *Rhodobacter sphaeroides* is essential for chemotaxis. *Mol Microbiol*, 46(4):1081.

[80] Porter, S.L., Wilkinson, D.A., Byles, E.D., Wadhams, G.H., Taylor, S., Saunders, N.J., and Armitage, J.P. (2011). Genome sequence of *Rhodobacter sphaeroides* Strain WS8N. *J Bacteriol*, 193(15):4027.

[81] Rajagopala, S.V., Sikorski, P., Kumar, A., Mosca, R., Vlasblom, J., Arnold, R., Franca-Koh, J., Pakala, S.B., Phanse, S., Ceol, A., *et al.* (2014). The binary protein-protein interaction landscape of *Escherichia coli*. *Nat Biotechnol*, 32(3):285.

[82] Rao, C.V., Glekas, G.D., and Ordal, G.W. (2008). The three adaptation systems of *Bacillus subtilis chemotaxis. Trends Microbiol*, 16(10):480.

[83] Roberts, M.A., August, E., Hamadeh, A., Maini, P.K., McSharry, P.E., Armitage, J.P., and Papachristodoulou, A. (2009). A model invalidation-based approach for elucidating biological signalling pathways, applied to the chemotaxis pathway in *Rhodobacter sphaeroides*. *BMC Syst Biol*, 3:105.

[84] Rosario, M.M., Fredrick, K.L., Ordal, G.W., and Helmann, J.D. (1994). Chemotaxis in *Bacillus subtilis* requires either of two functionally redundant CheW homologs. *J Bacteriol*, 176(9):2736.

[85] Rosser, G., Fletcher, A.G., Wilkinson, D.A., de Beyer, J.A., Yates, C.A., Armitage, J.P., Maini, P.K., and Baker, R.E. (2013). Novel methods for analysing bacterial tracks reveal persistence in *Rhodobacter sphaeroides*. *PLoS Comput Biol*, 9(10):e1003276.

[86] Sager, B.M., Sekelsky, J.J., Matsumura, P., and Adler, J. (1988). Use of a computer to assay motility in bacteria. *Anal Biochem*, 173(2):271.

[87] Sambrook, J. and Russell, J. (2001). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[88] Sanglier, S., Atmanene, C., Chevreux, G., and Dorsselaer, A.V. (2008). Nondenaturing mass spectrometry to study noncovalent protein/protein and protein/ligand complexes: technical aspects and application to the determination of binding stoichiometries. *Methods Mol Biol*, 484:217.

[89] Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994). Small mobilizable multipurpose cloning vectors derived from the *Escherichia coli* plasmids pK18

and pK19 – selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene*, 145:169.

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

[91] Segall, J.E., Manson, M.D., and Berg, H.C. (1982). Signal processing times in bacterial chemotaxis. *Nature*, 296(5860):855.

[92] Sengupta, P., van Engelenburg, S.B., and Lippincott-Schwartz, J. (2014). Superresolution imaging of biological systems using photoactivated localization microscopy. *Chem Rev*, 114(6):3189.

[93] Shah, D.S., Porter, S.L., Harris, D.C., Wadhams, G.H., Hamblin, P.A., and Armitage, J.P. (2000). Identification of a fourth cheY gene in *Rhodobacter sphaeroides* and interspecies interaction within the bacterial chemotaxis signal transduction pathway. *Mol Microbiol*, 35(1):101.

[94] Shah, D.S., Porter, S.L., Martin, A.C., Hamblin, P.A., and Armitage, J.P. (2000). Fine tuning bacterial chemotaxis: analysis of *Rhodobacter sphaeroides* behaviour under aerobic and anaerobic conditions by mutation of the major chemotaxis operons and cheY genes. *EMBO J*, 19(17):4601.

[95] Silverman, M. and Simon, M. (1974). Flagellar rotation and the mechanism of bacterial motility. *Nature*, 249:73.

[96] Silversmith, R.E., Levin, M.D., Schilling, E., and Bourret, R.B. (2008). Kinetic characterization of catalysis by the chemotaxis phosphatase CheZ. Modulation of activity by the phosphorylated CheY substrate. *J Biol Chem*, 283(2):756.

[97] Slivka, P.F. and Falke, J.J. (2012). Isolated bacterial chemosensory array possesses quasi- and ultrastable components: functional links between array stability, cooperativity, and order. *Biochemistry*, 51(51):10218.

[98] Sockett, R.E., Foster, J.C.A., and Armitage, J.P. (1990). Molecular biology of the *Rhodobacter sphaeroides* flagellum. *FEMS Symp*, 53:473.

[99] Sourjik, V. and Berg, H.C. (2002). Binding of the *Escherichia coli* response regulator CheY to its target measured *in vivo* by fluorescence resonance energy transfer. *Proc Natl Acad Sci U S A*, 99(20):12669.

[100] Sourjik, V., Vaknin, A., Shimizu, T.S., and Berg, H.C. (2007). *In vivo* measurement by FRET of pathway activity in bacterial chemotaxis. *Methods Enzymol*, 423:365.

[101] Stewart, R.C., Jahreis, K., and Parkinson, J.S. (2000). Rapid phosphotransfer to CheY from a CheA protein lacking the CheY-binding domain. *Biochemistry*, 39(43):13157.

[102] Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000). Two-component signal transduction. *Annu Rev Biochem*, 69:183.

[103] Stracy, M., Lesterlin, C., Garza de Leon, F., Uphoff, S., Zawadzki, P., and Kapanidis, A.N. (2015). Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proc Natl Acad Sci U S A*, 112(32):E4390.

[104] Stracy, M., Uphoff, S., Garza de Leon, F., and Kapanidis, A.N. (2014). *In vivo* single-molecule imaging of bacterial DNA replication, transcription, and repair. *FEBS Lett*, 588(19):3585.

[105] Swanson, R.V., Schuster, S.C., and Simon, M.I. (1993). Expression of CheA fragments which define domains encoding kinase, phosphotransfer, and CheY binding activities. *Biochemistry*, 32(30):7623.

[106] Tindall, M.J., Maini, P.K., Porter, S.L., and Armitage, J.P. (2008). Overview of mathematical approaches used to model bacterial chemotaxis II: bacterial populations. *Bull Math Biol*, 70(6):1570.

[107] Tindall, M.J., Porter, S.L., Maini, P.K., and Armitage, J.P. (2010). Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. *PLoS Comput Biol*, 6(8).

[108] Tindall, M.J., Porter, S.L., Maini, P.K., Gaglia, G., and Armitage, J.P. (2008). Overview of mathematical approaches used to model bacterial chemotaxis I: the single cell. *Bull Math Biol*, 70(6):1525.

[109] Uphoff, S., Reyes-Lamothe, R., Garza de Leon, F., Sherratt, D.J., and Kapanidis, A.N. (2013). Single-molecule dna repair in live bacteria. *Proc Natl Acad Sci U S A*, 110(20):8063.

[110] Uphoff, S., Sherratt, D.J., and Kapanidis, A.N. (2014). Visualizing protein-dna interactions in live bacterial cells using photoactivated single-molecule tracking. *J Vis Exp*, (85).

[111] Van Niel, C. (1944). The culture, general physiology, morphology, and classification of the non-sulfur purple and brown bacteria. *Bacteriol Rev*, 8(1):1.

[112] Wadhams, G.H., Martin, A.C., and Armitage, J.P. (2000). Identification and localization of a methyl-accepting chemotaxis protein in *Rhodobacter sphaeroides*. *Mol Microbiol*, 36(6):1222.

[113] Wadhams, G.H., Martin, A.C., Porter, S.L., Maddock, J.R., Mantotta, J.C., King, H.M., and Armitage, J.P. (2002). TlpC, a novel chemotaxis protein in *Rhodobacter sphaeroides*, localizes to a discrete region in the cytoplasm. *Mol Microbiol*, 46(5):1211.

[114] Wadhams, G.H., Martin, A.C., Warren, A.V., and Armitage, J.P. (2005). Requirements for chemotaxis protein localization in *Rhodobacter sphaeroides*. *Mol Microbiol*, 58(3):895.

[115] Wadhams, G.H., Warren, A.V., Martin, A.C., and Armitage, J.P. (2003). Targeting of two signal transduction pathways to different regions of the bacterial cell. *Mol Microbiol*, 50(3):763.

[116] Wood, T.M., Yates, C., Wilkinson, D., Rosser, G., *et al.* (2012). Simplified multitarget tracking using the PHD filter for microscopic video data. *IEEE T Circ Syst Vid*, 22(5):702.

[117] Wuichet, K. and Zhulin, I.B. (2010). Origins and diversification of a complex signal transduction system in prokaryotes. *Sci Signal*, 3(128):ra50.

[118] Youle, M., Rohwer, F., Stacy, A., Whiteley, M., Steel, B.C., Delalez, N.J., Nord, A.L., Berry, R.M., Armitage, J.P., Kamoun, S., Hogenhout, S., Diggle, S.P., Gurney, J., Pollitt, E.J.G., Boetius, A., and Cary, S.C. (2012). The microbial olympics. *Nat Rev Microbiol*, 10(8):583.

[119] Zhang, X.Y., Brunet, Y.R., Logger, L., Douzi, B., Cambillau, C., Journet, L., and Cascales, E. (2013). Dissection of the TssB-TssC interface during type VI secretion sheath complex formation. *PLoS One*, 8(11):e81074.

[120] Zimmer, M.A., Tiu, J., Collins, M.A., and Ordal, G.W. (2000). Selective methylation changes on the *Bacillus subtilis* chemotaxis receptor McpB promote adaptation. *J Biol Chem*, 275(32):24264.