Specificity of localization and phosphotransfer in the CheA Proteins of *Rhodobacter sphaeroides*.

Kathryn A. Scott,[1,2,†] *Steven L. Porter,*[1,2,‡,†] *Eleanor A. L. Bagg,*[2,§] *Rebecca Hamer,*[1,3] *Jennifer L. Hill,*[2] *David A. Wilkinson*[2] *and Judith P. Armitage*[1,2]*

1. Oxford Centre for Integrative Systems Biology, University of Oxford, Oxford, OX1 3QU, UK

2. Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK

3. Department of Statistics, University of Oxford, Oxford, OX1 3TG, UK

[†]These authors contributed equally to this work

*Corresponding author

judith.armitage@bioch.ox.ac.uk

tel: +44 (0)1865 613293

fax: +44 (0)1865 613338

Additional Footnotes

[§]Current address: Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA

[‡]Current address: School of Biosciences, University of Exeter, Exeter, Devon, EX4 4QD

Running Title

Specificity of *R. sphaeroides* CheAs

Keywords

Rhodobacter sphaeroides; Chemotaxis; Two-component; Response regulator; Histidine protein kinase

Summary

Specificity of protein-protein interactions plays a vital role in signal transduction. The chemosensory pathway of *Rhodobacter sphaeroides* comprises multiple homologues of chemotaxis proteins characterised in organisms such as *Escherichia coli*. Three CheA homologues are essential for chemotaxis in *R. sphaeroides* under laboratory conditions. These CheAs are differentially localized to two chemosensory clusters, one at the cell-pole and one in the cytoplasm. The polar CheA, $CheA_2$, has the same domain structure as *E. coli* CheA and can phosphorylate all *R. sphaeroides* chemotaxis response regulators. $CheA_3$ and $CheA_4$ independently localise to the cytoplasmic cluster; each protein has a subset of the CheA domains and together they make a functional CheA protein. $CheA_4$ phosphorylates $CheA_3$ which then phosphorylates two response regulators, $CheY_6$ and $CheB_2$, *in vivo*. *R. sphaeroides* CheAs exhibit two interesting differences in specificity, in the response regulators that they phosphorylate and the chemosensory cluster to which they localize. Using a domain-swapping approach we have investigated the role of the P1 and P5 CheA domains in determining these specificities. We show that the P1 domain is sufficient to determine which response regulators will be phosphorylated *in vitro* while the P5 domain is sufficient to localise the CheAs to a specific chemosensory cluster.

Introduction

Specificity of interaction is essential for faithful transmission of information in signal transduction pathways. Two-component pathways, comprising histidine protein kinases (HPKs) and response regulators (RRs), are widely used by Archaea, bacteria and some eukaryotes to control processes as diverse as gene expression, chemotaxis, virulence and development. A single bacterial cell can have over 150 different two-component systems, which demands a high degree of specificity in order to prevent unwanted cross-talk.

A recent study has analyzed the phosphotransfer specificity determinants in canonical two-component systems. In the canonical systems the HPKs autophosphorylate a histidine residue within their Dhp (dimerization and histidine phosphorylation) domains and then transfer the phosphoryl group to an aspartate residue in their cognate RRs. Residues within the Dhp domain were shown to be critical for phosphotransfer specificity, and by changing just a few of these residues it was possible to reengineer phosphosignalling (Skerker *et al.*, 2008). In this study, we focus on the chemotaxis pathway (reviewed in (Wadhams and Armitage, 2004; Sourjik, 2004)). The chemotaxis HPK, CheA, differs from canonical HPKs in that its autophosphorylateable histidine residue is contained within a separate histidine containing phosphotransfer (Hpt or P1) domain found at the N-terminus of the protein rather than within the Dhp domain.

The architecture of CheA proteins shows some variation between species. The majority of CheAs, referred to henceforth as classical CheAs, form a homodimer whose monomers have five structural domains designated P1 to P5. P4 is the kinase domain, which binds ATP and transfers a phosphoryl group to a histidine in the P1 domain (Swanson *et al.*, 1993). The P3 domain is responsible for dimerization, but unlike the equivalent Dhp domain in canonical HPKs does not contain a phosphorylatable histidine residue. In *E. coli*, the P5 domain of CheA has been shown to bind to both CheW and the transmembrane chemoreceptors forming a chemoreceptor complex

localized at the poles of the cell (Maddock and Shapiro, 1993; Kim *et al.*, 1999; Shimizu *et al.*, 2000; Skidmore *et al.*, 2000; Homma *et al.*, 2004; Parkinson *et al.*, 2005; Briegel *et al.*, 2009). The P5 domain also couples chemoreceptor signalling to the CheA autophosphorylation rate, thereby transducing signals from the chemoreceptors about changing chemoeffector concentration into an intracellular signal (Bourret *et al.*, 1993; Morrison and Parkinson, 1994). The P2 domain binds CheY and CheB increasing their concentration in the vicinity of the P1 domain and hence increasing the rate of phosphotransfer from the P1 domain to these RRs (Stewart *et al.*, 2000; Jahreis *et al.*, 2004).

The purple non-sulfur bacterium *Rhodobacter sphaeroides* has three chemotaxis operons, *che*Op1-3, encoding multiple homologues of many of the chemosensory proteins found in *E. coli* (Ward *et al.*, 1995; Hamblin *et al.*, 1997; Porter *et al.*, 2002; Porter *et al.*, 2008b). The genes encoded by *che*Op1 are not expressed under laboratory conditions (Shah *et al.*, 2000b; Poggio *et al.*, 2007; del Campo *et al.*, 2007). The genes encoded by *che*Op2 and *che*Op3 are expressed and control the rotation of a single, sub-polar, flagellum encoded by the fla1 genes (Porter *et al.*, 2002). The proteins encoded by *che*Op2 and *che*Op3 form two signalling pathways, both of which are necessary for chemotaxis (Figure 1). The components of one pathway localise to a cluster at the pole of the cell whilst the components of the other pathway form a discrete cluster in the cytoplasm (Wadhams *et al.*, 2002; Wadhams *et al.*, 2003; Wadhams *et al.*, 2005). Along with the transmembrane chemoreceptors, the polar cluster includes one CheA homologue, $CheA_2$, and two CheW homologues, $CheW_2$ and $CheW_3$. The cytoplasmic cluster contains putative cytoplasmic chemoreceptors in addition to two CheA homologues, $CheA_3$ and $CheA_4$, and a single CheW homologue, $CheW_4$. $CheA_2$ shows the classical five domain homodimeric structure found in *E.coli*, however neither $CheA_3$ nor $CheA_4$ contain all of the domains necessary for signal transduction. $CheA_3$ contains only the P1 and P5 domains separated by a 794 amino acid sequence that does not contain any identifiable domains but which shows specific phosphatase activity for $CheY_6$-P (Porter

and Armitage, 2004; Porter *et al.*, 2008a). $CheA_4$ is a homodimer of domains P3-P5. Neither $CheA_3$ nor $CheA_4$ can autophosphorylate, however $CheA_4$ can phosphorylate the P1 domain of $CheA_3$.

In *R. sphaeroides*, the CheAs localised to the polar and cytoplasmic clusters show different phosphotransfer specificity. $CheA_2$-P can phosphorylate all chemotaxis response regulators (Porter and Armitage, 2002), however $CheA_3$-P can only phosphorylate $CheB_2$, and $CheY_6$ (it also phosphorylates $CheY_1$ *in vitro*, but this is not expressed under lab conditions) (Porter and Armitage, 2004). Either or both of the P1 and P2 domains of $CheA_2$ may play a role in determining phosphotransfer specificity. For example, it is possible that the P2 domain of $CheA_2$, by binding all RRs, could enable the P1 domain to phosphorylate them. Alternatively, the specificity of the phosphotransfer reaction may be due entirely to the interaction between the P1 domain and the RRs. The existence of two chemotaxis clusters in *R. sphaeroides* raises the additional question of which region of CheA determines the cluster to which it is targeted. Localization specificity could be due to the P5 domain, which in *E. coli* has been shown to bind to CheW and the chemoreceptors. However this does not preclude the involvement of other regions, particularly as neither of the *R. sphaeroides* chemosensory clusters observed under laboratory growth conditions contains only one CheA and one CheW. At the polar cluster there is a single CheA but two different CheWs, either or both of which could interact with $CheA_2$. In contrast, in the cytoplasmic cluster two atypical CheA homologues, each with a different P5 domain, are required for chemosensing, however only a single CheW is present. To address the question of specificity we investigated the roles of the P1 and P5 domains in determining the phosphotransfer and localization specificity of *R. sphaeroides* CheAs respectively using a domain-swapping approach. We show that the P1 domains of both $CheA_2$ and $CheA_3$ are responsible for determining the specificity of phosphotransfer to the RRs. We also show that the P5 domain of each protein is sufficient to determine localization. Domain-swapped proteins with the P5 domain of $CheA_2$ localise to the polar cluster, whilst those with the P5 domain of $CheA_3$ or $CheA_4$ localise to the cytoplasmic cluster.

Results

P1 domain-swapped CheAs can be phosphorylated in vitro

To investigate the role of the *R. sphaeroides* P1 domain in phosphotransfer specificity *in vitro*, P1 domain-swapped $CheA_2$ and $CheA_3$, designated (A_3P1)-$CheA_2$ and (A_2P1)-$CheA_3$, were overexpressed and purified from *E. coli*. Before investigating potential phosphotransfer to the RRs, it was important to address two questions. Firstly, to determine whether (A_3P1)-$CheA_2$ could autophosphorylate, (A_3P1)-$CheA_2$ was incubated in the presence of [γ–^{32}P]-ATP and the products analysed by SDS-PAGE and phosphorimaging (Figure 2). Secondly, to test whether (A_2P1)-$CheA_3$ could be phosphorylated by $CheA_4$, a mixture of $CheA_4$ and (A_2P1)-$CheA_3$ was incubated with [γ–^{32}P]-ATP and the products analysed similarly. The results showed that (A_3P1)-$CheA_2$ could autophosphorylate, although the initial rate for the reaction was ~13 fold slower than that for $CheA_2$. Similarly, $CheA_4$ was able to phosphorylate (A_2P1)-$CheA_3$ but with a ~12 fold lower initial rate than that for phosphorylation of $CheA_3$ by $CheA_4$.

The P1 domain determines specificity of phosphotransfer in vitro

To determine the phosphotransfer specificity of the domain-swapped proteins ^{32}P labelled (A_3P1)-$CheA_2$-P and (A_2P1)-$CheA_3$-P were incubated with each of the RRs, $CheY_1$ to $CheY_6$, $CheB_1$ and $CheB_2$, for 30 s and the resulting products analysed by SDS-PAGE and phosphorimaging (Figure 3). The data showed that both $CheA_2$-P and (A_2P1)-$CheA_3$-P could phosphotransfer to all 8 RRs, although in each case the extent of phosphotransfer to $CheY_6$ was small after 30 s. Similarly, $CheA_3$-P and (A_3P1)-$CheA_2$-P could both only phosphotransfer to $CheY_1$, $CheY_6$ and $CheB_2$. These data indicate that the phosphotransfer specificity of $CheA_2$ and $CheA_3$ is determined by the interactions of their P1 domains with the cognate RRs.

R. sphaeroides *strains P1 domain-swapped CheAs are non-chemotactic*

The components of both the polar and cytoplasmic chemosensory clusters are required for a wild-type chemotactic response in *R. sphaeroides* (Porter *et al.*, 2008b). The *in vitro* phosphotransfer assays demonstrate that the P1 domains of $CheA_2$ and $CheA_3$ each specify which RRs they phosphorylate. These results suggest that domain-swapped proteins can be used to investigate whether the specific localization of the phosphotransfer activity is important for chemotaxis *in vivo*. *R. sphaeroides* strains with just $cheA_2$, just $cheA_3$ or both $cheA_2$ and $cheA_3$ replaced with the domain-swapped gene in the genome were constructed. Strain JPA926 has $cheA_2$ replaced with (A_3P1)-$cheA_2$, JPA927 has $cheA_3$ replaced with (A_2P1)-$cheA_3$ and JPA1103 has both $cheA_2$ replaced with (A_3P1)-$cheA_2$ and $cheA_3$ replaced with (A_2P1)-$cheA_3$. Soft agar swim assays were performed to analyse chemotactic behaviour in response to propionate under aerobic conditions (Figure 4). For each of the three domain-swapped strains the colony diameter was within experimental error of that seen for a motile, but non-chemotactic strain. The interpretation of these data is, however, complicated by the reduction in the autophosphorylation rate of the domain-swapped proteins relative to the wild-type *in vitro*. However, the lack of chemotaxis in strain JPA1103, where the P1 domains of the polar kinase $CheA_2$ and the cytoplasmic kinase $CheA_3$ have been exchanged, is presumably due to the inability of the domain swapped proteins to produce the right balance of phosphorylated RR for chemotaxis.

The P5 domain determines specificity of protein localization

The P5 domains of *R. sphaeroides* show pairwise sequence identities ranging from 17 to 34 % (Table 1). P5 domain-swapped proteins were used to investigate whether the P5 domain was sufficient for localization of CheAs to a specific chemosensory cluster. In contrast to the *in vivo* P1 domain-swap experiments, P5 domain-swapped proteins were expressed from an IPTG-inducible expression vector in the appropriate *cheA* deletion background. This avoided any difficulties that could arise due to the expression of these chemosensory proteins in operons. Swapping P5 domains in the genome is likely

to affect downstream gene expression, since in *cheA*$_2$ and *cheA*$_4$ the Shine-Dalgarno sequence for the downstream gene is contained within the region encoding the P5 domain.

The DNA coding for CheA$_2$, CheA$_3$, CheA$_4$ and the P5 domain-swapped proteins, designated CheA$_x$-(A$_y$P5), were cloned into the *R. sphaeroides* expression vector pIND4 both with and without an N-terminal fusion to yellow fluorescent protein (YFP) or cyan fluorescent protein (CFP) (Ind *et al.*, 2009). The plasmids coding for YFP-CheA$_2$, YFP-CheA$_3$ and YFP-CheA$_4$, were also introduced into Δ*cheA*$_2$, Δ*cheA*$_3$ and Δ*cheA*$_4$ deletion strains respectively. All plasmids coding for domain-swapped YFP fusion proteins were introduced into a strain with all *cheA* genes deleted to ensure that domain-swapped proteins did not have to compete with wild type proteins for localization. Plasmids coding for domain-swapped YFP fusion proteins ending with the P5 domain of CheA$_2$ were also introduced into a Δ*cheA*$_2$ strain whilst those with the P5 domain of CheA$_3$ or CheA$_4$ were also introduced into a Δ*cheA*$_3$Δ*cheA*$_4$ strain.

Localization of the YFP/CFP fusion domain-swapped proteins was visualised using fluorescence microscopy; representative fluorescence images of strains grown in the presence of 1 μM IPTG are shown (Figure 5). Proteins containing the P5 domain of CheA$_2$ (Figure 5 A-C) localised to the polar cluster whilst proteins containing the P5 domain of CheA$_3$ or CheA$_4$ (Figure 5 D-I) localised to the cytoplasmic cluster. Even in the presence of higher concentrations of IPTG when the level of diffuse fluorescent protein increased, CheA$_x$-(A$_2$P5) proteins localised exclusively to the polar cluster whilst CheA$_x$-(A$_3$P5) and CheA$_x$-(A$_4$P5) localised exclusively to the cytoplasmic cluster. These results show that the P5 domain determines the cluster to which the CheA localizes.

The P5 domains of CheA$_3$ and CheA$_4$ are not equivalent

CheA$_3$ and CheA$_4$ both localise independently to the cytoplasmic chemosensory cluster in wild type *R. sphaeroides* (Wadhams *et al.*, 2003). Their P5 domains share 32 % sequence identity. These proteins may compete for the same binding site at the cytoplasmic cluster, alternatively there may be

distinct binding sites for $CheA_3$ and $CheA_4$. This raises the question of whether localization to the cytoplasmic cluster is possible if both CheAs have the same P5 domain.

To determine whether localization to the cytoplasmic cluster is possible when both proteins have the P5 domain of $CheA_4$ the plasmid coding for YFP-$CheA_3$-(A_4P5) was introduced into Δ*cheA*$_3$, and Δ*cheA*$_3$,*cfp-cheA*$_4$ strains. Localization of the fluorescently tagged proteins was visualised using fluorescence microscopy (Figure 6). When YFP-$CheA_3$-(A_4P5) was expressed in the Δ*cheA*$_3$ and Δ*cheA*$_3$,*cfp-cheA*$_4$ backgrounds the majority of the fluorescent protein was diffuse throughout the cytoplasm (Figure 6 B, C). In some cells, however, the fluorescence at the cytoplasmic cluster was visible above the fluorescence from the diffuse protein in the cytoplasm. CFP fluorescence images from the strain expressing YFP-$CheA_3$-(A_4P5) in a Δ*cheA*$_3$,*cfp-cheA*$_4$ background showed that CFP-$CheA_4$ was localised to the cytoplasmic cluster (Figure 6 D). These data indicate that CFP-$CheA_4$ localizes to the cytoplasmic cluster preferentially over YFP-$CheA_3$-(A_4P5).

To determine whether localization to the cytoplasmic cluster is possible when both proteins have the P5 domain of $CheA_3$ the plasmid coding for CFP-$CheA_4$-(A_3P5) was introduced into Δ*cheA*$_4$, and *yfp-cheA*$_3$Δ*cheA*$_4$ strains. When CFP-$CheA_4$-(A_3P5) was expressed in Δ*cheA*$_4$ and *yfp-cheA*$_3$Δ*cheA*$_4$ background strains, CFP fluorescence images show that the majority of the CFP-$CheA_4$-(A_3P5) is localised to the cytoplasmic cluster, with the remainder diffuse throughout the cytoplasm (Figure 6 F, G). YFP fluorescence images of CFP-$CheA_4$-(A_3P5) expressed in *yfp-cheA*$_3$Δ*cheA*$_4$ showed that YFP-$CheA_3$ was also localised to the cytoplasmic cluster in this case (Figure 6 H). These data indicate that YFP-$CheA_3$ and CFP-CheA4-(A_3P5) can localize to the cytoplasmic cluster together without apparent interference.

The P5 domains of $CheA_3$ and $CheA_4$ are not equivalent; when both $CheA_3$ and $CheA_4$ have A_3P5 they co-localize to the cytoplasmic cluster whereas when they both have A_4P5 there is

competition and $CheA_4$ preferentially localizes to the cytoplasmic cluster. This could be explained if the number of A_4P5 binding sites within the cytoplasmic cluster were limiting. In that case the apparent inability of $CheA_3$-(A_4P5) to compete with $CheA_4$ for binding, could be due to steric constraints, since $CheA_3$-(A_4P5) (1095 amino acids) is much larger than $CheA_4$ (399 amino acids).

The P5 domains of $CheA_3$ and $CheA_4$ differ in their ability to substitute for one another in chemotaxis assays

We next addressed the question of whether chemotaxis is observed in strains in which both cytoplasmic cluster CheAs have the same P5 domain. We have previously shown that strains in which $CheA_3$ or $CheA_4$ are fused to YFP or CFP have reduced chemotactic ability (Wadhams *et al.*, 2003), therefore chemotactic ability was measured in non-fusion strains. The genes coding for $CheA_3$, $CheA_3$-(A_4P5), $CheA_4$ and $CheA_4$-(A_3P5) were cloned into pIND4 (Ind *et al.*, 2009). The plasmids coding for $CheA_3$ and $CheA_3$-(A_4P5) were introduced into a Δ*cheA*$_3$ background strain whilst those coding for $CheA_4$ and $CheA_4$-(A_3P5) were introduced into a Δ*cheA*$_4$ background strain. Soft agar swim assays were performed under aerobic conditions with propionate as attractant and 0, 1, 10, 100 or 1000 μM IPTG (Figure 7). The data show that pIND4-*cheA*$_4$ can fully complement a Δ*cheA*$_4$ strain in the presence of 100 μM and 1 mM IPTG. However, pIND4-*cheA*$_4$*-(A*$_3$*P5)* only partially complements a Δ*cheA*$_4$ strain even in the presence of 1 mM IPTG. The behaviour of the $CheA_3$ proteins was more complex. pIND4-*cheA*$_3$ fully complemented a Δ*cheA*$_3$ strain induced with 10 μM IPTG, but Δ*cheA*$_3$ cells containing pIND4-*cheA*$_3$*-(A*$_4$*P5)* did not show a significant increase in colony diameter over a non-chemotactic strain at any induction level. These results are consistent with the localization results; when both $CheA_3$ and $CheA_4$ have A_3P5, they co-localize to the cytoplasmic cluster and give a partially functional chemotactic response whereas when both proteins have the P5 domain of $CheA_4$ then $CheA_3$-(A_4P5) fails to localize and the strain is non-chemotactic. Together these results indicate that while the A_3P5 can partially substitute for A_4P5, A_4P5 cannot

substitute for A_3P5, suggesting that the different cytoplasmic cluster CheAs have different P5 domains because they play different roles in chemosensory signalling.

Comparison of CheA surface residues

Falke and co-workers showed that residues important for interactions of the *Salmonella enterica* serovar Typhimurium CheA P5 domain with CheW lie within patches of conserved residues on the CheA surface (Miller *et al.*, 2006). In *R. sphaeroides*, the P5 domains determine not only localization of the CheA to a chemosensory cluster, but also discriminate between the polar and cytoplasmic clusters. Unfortunately, the architecture of the cytoplasmic cluster is not known, but it may involve different interaction patterns between the constituent CheA and CheW homologues than identified for the polar cluster with only one CheW and one CheA. It is therefore interesting to compare the sequence of the *R. sphaeroides* P5 domains with those from other organisms to investigate whether the *R. sphaeroides* P5 domains share the conserved surface residues. If $CheA_3$ and/or $CheA_4$ do not share the conserved residues it may indicate a substantially different binding interaction to that seen between classical CheAs and CheWs.

A non-redundant set of 367 CheA homologues from genomes where both CheA and CheW homologues were present were iteratively aligned using MUSCLE and MaxAlign to give a final alignment of 341 sequences (Edgar, 2004; Gouveia-Oliveira *et al.*, 2007). CheA homologues were required to include both a P4 and P5 domain, thus *R. sphaeroides* $CheA_3$ does not appear in the alignment. Conservation scores ranging from 0 (no conservation) to 11 (identity) were calculated for each sequence position using Jalview (Waterhouse *et al.*, 2009).

Residues on the *Thermotoga maritima* CheA P5 domain that interact with CheW were compared with the corresponding residues in *R. sphaeroides* CheAs (Figure 8) (Bilwes *et al.*, 1999; Park *et al.*, 2006). The most highly conserved residues in the *T. maritima* CheA-CheW binding site, leucine 640, glycine 659, isoleucine 661 and leucine 663, cluster together at the centre of the binding

site and the conservation of residues decreases towards the periphery. For *R. sphaeroides* CheAs the residues corresponding to the four most conserved positions in the *T. maritima* CheA-CheW binding site either matched the consensus sequence or contained a residue of the same amino acid class. Drawing conclusions from sequence comparison at less well conserved positions in the binding site was more difficult, however one or two positions in $CheA_3$ and $CheA_4$ clearly showed amino acids of a different class to the consensus sequence. For example, $CheA_4$ has a serine residue at the equivalent position to phenyalanine 650; at this position the consensus sequence is isoleucine and only five other sequences in the alignment showed a serine or threonine residue in this position. Similarly, $CheA_3$ has histidine residues at equivalent positions to leucine 599 and glutamate 649; histidine appears in these positions in only 1 and 2 other sequences respectively. These results suggest that it is possible that $CheA_3$ and $CheA_4$ interact with the chemoreceptor through the same binding interface as in a classical CheA-CheW interaction.

Discussion

The P1 domain determines specificity of phosphotransfer to the RRs

In a classical CheA dimer, the P4 domain binds ATP and transphosphorylates the P1 domain. The P2 domain binds RRs allowing phosphotransfer to occur between the P1 domain and the RR. This raises the question as to whether the specificity of transfer to the RRs is determined by either or both of the P1 or P2 domains. It has been shown in *E. coli* that the P2 domain of CheA binds to CheY with an apparent K_D of 3.7×10^{-7} M (Swanson *et al.*, 1993). When the P2 domain of *E. coli* CheA is deleted, however, phosphotransfer to CheY remains possible albeit with a significantly slower rate and increased K_m (Stewart *et al.*, 2000). Further, *in vivo* experiments showed that a *cheAΔP2* strain is chemotactic in soft agar swim assays (although at a reduced level compared with wild type). This suggests that in *E. coli* the role of the P2 domain is to increase the local concentration of RRs and thereby increase the rate of phosphotransfer to that necessary for a rapid chemotactic response

(Jahreis *et al.*, 2004). In *R. sphaeroides* one of the classical CheAs, $CheA_2$, can phosphorylate all chemotaxis RRs whilst the atypical $CheA_3$, which lacks the P2 domain, can only phosphorylate $CheY_1$, $CheY_6$ and $CheB_2$ (Porter *et al.*, 2008b). In this case it is conceivable that the P2 domain of $CheA_2$ is important for binding and recognizing the additional RRs. This hypothesis was tested using a P1 domain-swapping approach.

Phosphotransfer assays on the P1 domain-swapped proteins showed that (A_3P1)-$CheA_2$-P can phosphorylate $CheY_1$, $CheY_6$ and $CheB_2$ and that (A_2P1)-$CheA_3$-P can phosphorylate all chemotaxis RRs *in vitro*. This demonstrates that the interactions made between the RRs and the P1 domain determine the specificity of the phosphotransfer reaction and that while the P2 domain may catalyze the phosphotransfer reaction by increasing the local concentration of the RRs, it does not determine the specificity of the reactions.

Interestingly, none of the three strains with P1 domain-swapped CheAs were chemotactic. For the strains where a single domain-swapped protein was introduced this is perhaps unsurprising as phosphotransfer to $CheY_6$, $CheB_1$, $CheB_2$ and to either of $CheY_3$ or $CheY_4$ is essential for chemotaxis in *R. sphaeroides* (Martin *et al.*, 2001; Porter *et al.*, 2002; Porter and Armitage, 2004). In the strain where $CheA_2$ has the $CheA_3$ P1 domain then phosphoryl groups cannot be transferred to $CheY_3$, $CheY_4$ or $CheB_1$, and thus the resulting strain would be expected to be non-chemotactic. In the strain where $CheA_3$ has the $CheA_2$ P1 domain however, all RRs can be phosphorylated but the strain was not chemotactic. Previous work, however, showed that $CheA_3$-P transfers a phosphoryl group to $CheY_6$ at a significantly higher rate than $CheA_2$-P (Porter and Armitage, 2002; Porter and Armitage, 2004). It is therefore possible that in the strain where $CheA_3$ has the $CheA_2$ P1 domain the level of $CheY_6$-P formed may be insufficient to elicit a normal chemotactic response.

For the strain in which $CheA_2$ has the $CheA_3$ P1 domain and $CheA_3$ has the $CheA_2$ P1 domain, phosphotransfer should be able to occur to all of the RRs necessary for chemotaxis. Despite

this, the resulting strain did not show a chemotactic response. There are a number of possible explanations for this result. The autophosphorylation rate of (A_3P1)-CheA_2 and the rate of phosphorylation of (A_2P1)-CheA_3 by CheA_4 are lower than for the corresponding wild-type proteins measured *in vitro*. As a result the levels of (A_3P1)-CheA_2-P and (A_2P1)-CheA_3-P formed in response to activation of the chemoreceptors may be lower than the levels of CheA_2-P and CheA_3-P formed in the wild-type strain. Alternatively, it may be essential for specific RRs to be phosphorylated in response to activation of the polar and cytoplasmic chemosensory clusters i.e. the cross-wired signalling pathway produced by domain-swapping may not produce the right balance of phosphorylated RR to produce a functional signal at the motor.

The P5 domain determines specificity of protein localization

In *E. coli*, CheA localises to the chemosensory cluster through interactions with both CheW and the chemoreceptors (Maddock and Shapiro, 1993), with CheW being essential for changes in the autophosphorylation rate of CheA in response to activation of the chemoreceptors (Levit *et al.*, 2002). The P5 domain of CheA has been shown to interact with CheW and the chemoreceptors in *E. coli*, *S. enterica* serovar Typhimurium and *T. maritima* using a variety of different methods including chemical mapping, spin-labelling combined with ESR and x-ray crystallography (Miller *et al.*, 2006; Park *et al.*, 2006; Zhao and Parkinson, 2006a; Zhao and Parkinson, 2006b). Chemical mapping of *S. enterica* serovar Typhimurium CheA also defined regions of the P3, P4 and P5 domains which interact with the chemoreceptor (Miller *et al.*, 2006), whilst for *T. maritima* the K_D for the interaction between P3-P4-P5 and CheW is ~3-fold lower than that for P5 and CheW (Park *et al.*, 2006). Collectively these data show that although the interaction between the P5 domain and CheW is vital for stimulation of autokinase activity by the chemoreceptors other regions of CheA play an important role in the assembly of the chemosensory cluster.

In *R. sphaeroides* the presence of two chemosensory clusters to which different CheA and CheW homologues are localised raised the question of which domains are important in determining the specificity of localization. Using domain-swapping we have shown that the P5 domain contains the determinants for localization of CheA to either the polar or cytoplasmic cluster. At low levels of induction, CFP/YFP tagged domain-swapped CheAs containing the P5 domain of $CheA_2$ localised to the poles of the cells whilst those containing the P5 domain of $CheA_3$ or $CheA_4$ localised to the cytoplasmic chemosensory cluster. At higher levels of induction, the proportion of the fluorescence signal diffuse throughout the cytoplasm was increased but localization to the 'wrong' chemosensory cluster was not observed. This result is particularly interesting for $CheA_2$ and $CheA_4$ P5 domain-swapped proteins, as both contain P3 and P4 domains. The P3 and P4 domains have been reported to be important for interaction with the chemoreceptors in other organisms and may therefore be expected to play some role in cluster localization (Miller *et al.*, 2006; Park *et al.*, 2006). It is possible that the interaction between the P5 domain and CheW may be strong enough to overcome any competing interactions between the rest of CheA and the chemoreceptors. An alternative possibility is that the protein-protein interactions in the chemosensory clusters of *R. sphaeroides* differ markedly from those in other organisms. For the polar chemosensory cluster at least, this last hypothesis seems unlikely given the ability of $CheA_2$ to partially complement an *E. coli* Δ*cheA* mutant (Shah *et al.*, 2000a) and in light of recent work showing that the architecture of the chemoreceptor arrays is widely conserved amongst different bacteria (Briegel *et al.*, 2009).

P5 domain-swapped $CheA_3$ and $CheA_4$ were also used to investigate whether two proteins with the same P5 domain could bind to the cytoplasmic cluster simultaneously. The results differed depending upon which of the P5 domains was present. When both proteins had the P5 domain of $CheA_3$ simultaneous localization to the cytoplasmic cluster was observed. In contrast, when both proteins had the P5 domain of $CheA_4$ only CFP-$CheA_4$ localised to the cytoplasmic cluster whilst much of the YFP-$CheA_3$-(A_4P5) was diffuse throughout the cytoplasm. Further, soft agar swim

assays showed that $CheA_4$-(A_3P5) could partially complement a *cheA*$_4$ deletion strain, while $CheA_3$-(A_4P5) failed to complement a Δ*cheA*$_3$ deletion. Thus, whilst the P5 domains of $CheA_3$ and $CheA_4$ localise proteins to the cytoplasmic cluster they are not functionally equivalent. These results open a number of intriguing possibilities regarding the architecture and activity of the cytoplasmic cluster. One possibility is simply that there are more binding sites for $CheA_3$ than for $CheA_4$. Another possibility is that whilst the $CheA_3$ and $CheA_4$ P5 domains are likely to share a similar structure, they have different binding sites at the cytoplasmic cluster. In this case, the results suggest that the binding site for $CheA_4$ may be able to accommodate the P5 domain of $CheA_3$ sufficiently well to support chemotaxis, but that the converse is not true. It has been shown that key surface residues in the CheA-CheW binding interface are conserved across different genera (Miller *et al.*, 2006). The architecture of the *R. sphaeroides* cytoplasmic cluster is not known and may show different interactions between CheA and CheW than in a typical polar cluster. We therefore considered the possibility that the conserved surfaces residues may not be present in $CheA_3$ and/or $CheA_4$. Sequence analysis showed that for *R. sphaeroides* CheAs the residues corresponding to the four most conserved positions in the *T. maritima* CheA-CheW binding site either matched the consensus sequence or contained a residue of the same amino acid class. The presence of the conserved residues suggests that it may be possible for $CheA_3$ and $CheA_4$ to interact with the chemoreceptor through the same binding interface as in a classical CheA-CheW interaction.

R. sphaeroides has a complex chemosensory network requiring specificity in both protein localization and phosphorylation of the correct RRs. This study shows that the P1 domains of both $CheA_2$ and $CheA_3$ are sufficient to determine the specificity of phosphotransfer to the RRs. Similarly, the P5 domain contains the determinants for the specificity of localization of the CheAs.

Experimental Procedures

Construction of P1 domain-swapped CheAs

Overlap extension PCR was used to generate P1 domain-swapped proteins where the P1 domains of $CheA_2$ and $CheA_3$ were exchanged. The fragment of *E. coli* CheA comprising residues 1-149 and the corresponding fragment of $CheA_3$ (residues 1-182) have both been shown to be functional in phosphotransfer reactions *in vitro* (Garzon and Parkinson, 1996; Bell *et al.*, 2010). The boundaries for the P1 domain-swapped proteins were therefore chosen to correspond to residues 1-149 of *E. coli* CheA (i.e. residues 1-138 of $CheA_2$ and 1-182 of $CheA_3$). The domain-swapped fragment included the entire P1 domain plus most of the linker connecting it to the next domain. A construct containing (A_2P1)-$CheA_3$ (residues 1-138 of $CheA_2$ joined to residues 183-1095 of $CheA_3$) with ~500 bp of upstream and downstream flanking sequence from the *cheA*$_3$ genomic region was produced. A second construct was produced containing (A_3P1)-$CheA_2$ (residues 1-183 of $CheA_3$ joined to residues 139-516 of $CheA_2$) with ~500 bp of upstream and downstream flanking sequence from the *cheA*$_2$ genomic region. These constructs were cloned into the allelic-exchange suicide vector, pK18*mobsacB* (Schäfer *et al.*, 1994). The resulting plasmids were used to replace the wild-type *cheA*$_2$ and *cheA*$_3$ genes with the domain-swapped CheAs in the *R. sphaeroides* genome (Porter *et al.*, 2007).

Protein overexpression and purification

Wild-type His-tagged *R. sphaeroides* CheA, CheY and CheB proteins were overexpressed and purified as described previously (Martin *et al.*, 2001; Porter and Armitage, 2002; Porter and Armitage, 2004; Porter *et al.*, 2006). The coding sequences of (A_2P1)-$CheA_3$ and (A_3P1)-$CheA_2$ were amplified by PCR and cloned into C-terminal His-tag expression vector pQE60 (Qiagen). The domain-swapped CheAs were purified using the same method as wild-type $CheA_2$ and $CheA_3$ (Porter and Armitage, 2004).

CheA phosphorylation reactions

Assays were performed at 20 ^{o}C in TGMNKD buffer (50 mM Tris HCl, 10% (v/v) glycerol, 5 mM $MgCl_2$, 150 mM NaCl, 50 mM KCl, 1 mM DTT, pH 8.0). Reaction mixtures contained 5 μM CheA protein (either $CheA_2$, $CheA_3$, (A_2P1)-$CheA_3$ or (A_3P1)-$CheA_2$). Due to the absence of kinase domains in $CheA_3$ and (A_2P1)-$CheA_3$, reactions containing these proteins were supplemented with 20 μM $CheA_4$. Reactions were initiated by the addition of 0.5 mM [γ-^{32}P] ATP (specific activity 14.8 GBq mmol^{-1}; PerkinElmer). Reaction aliquots of 10 μl were taken at the specified time points and quenched immediately in 10 μl of 3 x SDS-PAGE loading dye (7.5% (w/v) SDS, 90 mM EDTA, 37.5 mM Tris HCl, 37.5% glycerol, 3% (v/v) β-mercaptoethanol, pH 6.8). Quenched samples were analyzed using SDS-PAGE and phosphorimaging as described previously (Porter and Armitage, 2002).

Phosphotransfer from the domain-swapped CheAs to the response regulators

Phosphotransfer assays were performed at 20 ^{o}C in TGMNKD buffer. Reaction mixtures contained 5 μM CheA (either $CheA_2$, $CheA_3$, (A_2P1)-$CheA_3$ or (A_3P1)-$CheA_2$) and 0.5 mM [γ-^{32}P] ATP (specific activity 14.8 GBq mmol^{-1}; PerkinElmer) plus 20 μM $CheA_4$ for reactions containing either $CheA_3$ or (A_2P1)-$CheA_3$. The ATP dependent phosphorylation of the CheAs was allowed to proceed for 30 minutes and then the phosphotransfer reactions were initiated by the addition of 10 μM response regulator. After 30 s a 10 μl aliquot of the reaction mixture was taken, quenched and analyzed by SDS-PAGE as described above.

Phenotypic analysis of R. sphaeroides strains

The soft agar swim responses of the *R. sphaeroides* strains were characterised as described previously (Porter *et al.*, 2002). Briefly, strains were grown for 48 hours under photoheterotrophic conditions at 30 ^{o}C using succinate medium in the presence of appropriate antibiotics (Sistrom, 1960). M22 soft agar swim plates containing 0.25 % agar, nalidixic acid, 100 μM sodium propionate

and varying concentrations of IPTG if appropriate, were then inoculated with 5 μl of each strain. Plates were incubated at 30 ^{o}C under aerobic conditions and the colony diameter measured after 48 hours. Nine data sets (three repeats for three independent cultures of each strain) were collected in each case.

Construction of P5 domain-swapped proteins

Overlap extension PCR was used to construct the DNA coding for $CheA_4$ and $CheA_2$ P5 domain-swapped proteins with and without an N-terminal fusion to YFP or CFP. The P5 domain of $CheA_2$ comprises residues 515-654, $CheA_3$ 954-1095 residues and $CheA_4$ residues 263-399, so for example $CheA_2$-(A_4P5) consists of residues 1-514 of $CheA_2$ joined to residues 263-399 of $CheA_4$. Domain boundaries were defined with reference to an alignment of the *R. sphaeroides* CheA sequences against that of CheA-289 from *T. maritima* for which a structure of the P3 - P5 domains is available (Bilwes *et al.*, 1999). In order to be consistent with previous fluorescent protein constructs used in the laboratory, the DNA coding for an XbaI restriction site was included between the coding sequence for YFP/CFP and that of CheA in the N-terminal fusion proteins. Due to their size $CheA_3$ domain-swapped proteins were constructed in two pieces, utilising a BstBI restriction enzyme site occurring at base 1517 of *cheA*$_3$. Overlap extension PCR was used to generate the DNA coding for bases 1-1527 and 1507-3285 which was then cloned into pIND4 using the appropriate restriction enzymes. DNA sequencing verified that each of the constructs had the expected sequence.

Fluorescence Analysis

Log-phase cultures were embedded in 1.2 % agarose on microscope slides as described previously (Wadhams *et al.*, 2000). DIC and fluorescence images were acquired using a Nikon eclipse TE-2000-E microscope with YFP/CFP filters (Chroma) and recorded with a cooled CCD camera (ANDOR iXon^{+}) at subsaturating intensities. All images for each strain were collected on the same day using the same microscope settings and images from three independent cultures, comprising

approximately 100 cells per culture, were analysed. Cells were scored according to the number and location of fluorescent clusters present in each case using in-house software and the results were also confirmed by visual inspection.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council.

Reference List

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**: e1000306.

Bilwes, A.M., Alex, L.A., Crane, B.R., and Simon, M.I. (1999) Structure of CheA, a signal-transducing histidine kinase. *Cell* **96**: 131-141.

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**: 2097-2101.

Briegel, A., Ortega, D.R., Tocheva, E.I., Wuichet, K., Li, Z., Chen, S. *et al.* (2009) Universal architecture of bacterial chemoreceptor arrays. *Proc Natl Acad Sci USA* **106**: 17181-17186.

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**: 8397-8401.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* **32**: 1792-1797.

Garzon, A., and Parkinson, J.S. (1996) Chemotactic signaling by the P1 phosphorylation domain liberated from the CheA histidine kinase of Escherichia coli. *J Bacteriol* **178**: 6752-6758.

Gouveia-Oliveira, R., Sackett, P., and Pedersen, A. (2007) MaxAlign: maximizing usable data in an alignment. *BMC Bioinformatics* **8**: 312.

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**: 1083-1096.

Homma, M., Shiomi, D., Homma, M., and Kawagishi, I. (2004) Attractant binding alters arrangement of chemoreceptor dimers within its cluster at a cell pole. *Proc Natl Acad Sci USA* **101**: 3462-3467.

Ind, A.C., Porter, S.L., Brown, M.T., Byles, E.D., de Beyer, J.A., Godfrey, S.A., and Armitage, J.P. (2009) An inducible expression plasmid for *Rhodobacter sphaeroides* and *Paracoccus denitrificans*. *Appl Environ Microbiol*: doi:10.1128/AEM.01587-09.

Jahreis, K., Morrison, T.B., Garzon, 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**: 2664-2672.

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**: 787-792.

Levit, M.N., Grebe, T.W., and Stock, J.B. (2002) Organization of the receptor-kinase signaling array that regulates *Escherichia coli* chemotaxis. *J Biol Chem* **277**: 36748-36754.

Maddock, J.R., and Shapiro, L. (1993) Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science* **259**: 1717-1723.

Martin, A.C., Wadhams, G.H., Shah, D.S.H., Porter, S.L., Mantotta, J.C., Craig, T.J. *et al.* (2001) CheR- and CheB-dependent chemosensory adaptation system of *Rhodobacter sphaeroides*. *J Bacteriol* **183**: 7135-7144.

Miller, A.S., Kohout, S.C., Gilman, K.A., and Falke, J.J. (2006) CheA kinase of bacterial chemotaxis: Chemical mapping of four essential docking sites. *Biochemistry* **45**: 8699-8711.

Morrison, T.B., and Parkinson, J.S. (1994) Liberation of an interaction domain from the phosphotransfer region of CheA, a signaling kinase of *Escherichia coli*. *Proc Natl Acad Sci USA* **91**: 5485-5489.

Park, S.Y., Borbat, P.P., Gonzalez-Bonet, G., Bhatnagar, J., Pollard, A.M., Freed, J.H. *et al.* (2006) Reconstruction of the chemotaxis receptor-kinase assembly. *Nature Structural & Molecular Biology* **13**: 400-407.

Parkinson, J.S., Ames, P., and Studdert, C.A. (2005) Collaborative signaling by bacterial chemoreceptors. *Curr Opin Microbiol* **8**: 116-121.

Penfold, R.J., and Pemberton, J.M. (1992) An improved suicide vector for construction of chromosomal insertion mutations in bacteria. *Gene* **118**: 145-146.

Poggio, S., breu-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**: 3208-3216.

Porter, S.L., Wadhams, G.H., and Armitage, J.P. (2007) *In vivo* and *in vitro* analysis of the *Rhodobacter sphaeroides* chemotaxis signaling complexes. *Method Enzymol* **423**: 392-413.

Porter, S.L., and Armitage, J.P. (2002) Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. *J Mol Biol* **324**: 35-45.

Porter, S.L., and Armitage, J.P. (2004) Chemotaxis in *Rhodobacter sphaeroides* requires an atypical histidine protein kinase. *J Biol Chem* **279**: 54573-54580.

Porter, S.L., Roberts, M.A.J., Manning C.S., and Armitage, J.P. (2008a) A bifunctional kinase-phosphatase in bacterial chemotaxis. *Proc Natl Acad Sci USA* **105**: 18531-18536.

Porter, S.L., Wadhams, G.H., and Armitage, J.P. (2008b) *Rhodobacter sphaeroides*: Complexity in chemotactic signalling. *Trends Microbiol* **16**: 251-260.

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**: 1081-1094.

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**: 32694-32704.

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**: 69-73.

Shah, D.S., Porter, S.L., Harris, D.C., Wadhams, G.H., Hamblin, P.A., and Armitage, J.P. (2000a) Identification of a fourth *cheY* gene in *Rhodobacter sphaeroides* and interspecies interaction within the bacterial chemotaxis signal transduction pathway. *Mol Microbiol* **35**: 101-112.

Shah, D.S.H., Porter, S.L., Martin, A.C., Hamblin, P.A., and Armitage, J.P. (2000b) 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**: 4601-4613.

Shimizu, T.S., Le Novere, N., Levin, M.D., Beavil, A.J., Sutton, B.J., and Bray, D. (2000) Molecular model of a lattice of signalling proteins involved in bacterial chemotaxis. *Nature Cell Biology* **2**: 792-796.

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

Skerker, J.M., Perchuk, B.S., Siryaporn, A., Lubin, E.A., Ashenberg, O., Goulian, M., and Laub, M.T. (2008) Rewiring the specificity of two-component signal transduction systems. *Cell* **133**: 1043-1054.

Skidmore, J.M., Ellefson, D.D., McNamara, B.P., Couto, M.M.P., Wolfe, A.J., and Maddock, J.R. (2000) Polar clustering of the chemoreceptor complex in *Escherichia coli* occurs in the absence of complete CheA function. *J Bacteriol* **182**: 967-973.

Sockett, R.E., Foster, J.C.A., and Armitage, J.P. (1990) Molecular biology of the *Rhodobacter sphaeroides* flagellum. *FEMS Symp* **53**: 473-479.

Sourjik, V. (2004) Receptor clustering and signal processing in E. coli chemotaxis. *Trends Microbiol* **12**: 569-576.

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**: 13157-13165.

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**: 7623-7629.

Wadhams, G.H., and Armitage, J.P. (2004) Making sense of it all: Bacterial chemotaxis. *Nat Rev Mol Cell Bio* **5**: 1024-1037.

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**: 1222-1233.

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**: 1211-1221.

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**: 895-902.

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**: 763-770.

Ward, M.J., Bell, A.W., Hamblin, P.A., Packer, H.L., and Armitage, J.P. (1995) Identification of a chemotaxis operon with 2 *cheY* genes in *Rhodobacter sphaeroides*. *Mol Microbiol* **17**: 357-366.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189-1191.

Zhao, J.H., and Parkinson, J.S. (2006a) Mutational analysis of the chemoreceptor-coupling domain of the Escherichia coli chemotaxis signaling kinase CheA. *J Bacteriol* **188**: 3299-3307.

Zhao, J.S., and Parkinson, J.S. (2006b) Cysteine-scanning analysis of the chemoreceptor-coupling domain of the Escherichia coli chemotaxis signaling kinase CheA. *J Bacteriol* **188**: 4321-4330.

Table 1. Sequence identity of domains

	$CheA_2$ P1	$CheA_2$ P5	$CheA_3$ P5	$CheA_4$ P5
$CheA_3$ P1	24 %	n/a	n/a	n/a
$CheA_2$ P5	n/a	100 %	-	-
$CheA_3$ P5	n/a	17 %	100 %	-
$CheA_4$ P5	n/a	27 %	32 %	100 %

Table 2. Plasmids and bacterial strains used in this study

Plasmid/Strain	Description	Source/reference
E. coli strains		
M15pREP4	Expression host containing pREP4; kanamycin resistance.	Qiagen
S17-1λpir	Strain capable of mobilizing the suicide vector pK18mobsacB into *R. sphaeroides*; streptomycin resistant.	(Penfold and Pemberton, 1992)
XL1 Blue	General cloning strain and expression host. *lacI*^{q}; tetracycline resistant	Stratagene
R. sphaeroides strains		
WS8N	Spontaneous nalidixic acid resistant mutant of wild type WS8	(Sockett *et al.*, 1990)
JPA1213	WS8N containing *cheY*$_{6}$*(D56N)* in place of wild type *cheY*$_{6}$ in the chromosome. Non motile.	(Porter *et al.*, 2006)
JPA1315	Δ*cheA*$_{1}$, Δ*cheA*$_{2}$ & Δ*cheA*$_{3}$ derivative of WS8N. Non-chemotactic.	(Porter *et al.*, 2002)
JPA926	WS8N containing *(A*$_{3}$*P1)-cheA*$_{2}$ in place of *cheA*$_{2}$ in the chromosome.	This study
JPA927	WS8N containing *(A*$_{2}$*P1)-cheA*$_{3}$ in place of *cheA*$_{3}$ in the chromosome.	This study
JPA1103	WS8N containing *(A*$_{3}$*P1)-cheA*$_{2}$ in place of *cheA*$_{2}$ and *(A*$_{2}$*P1)-cheA*$_{3}$ in place of *cheA*$_{3}$ in the chromosome.	This study
JPA1314	Δ*cheA*$_{3}$ derivative of WS8N.	(Porter *et al.*, 2002)

JPA1308	Δ*cheA*$_4$ derivative of WS8N.	(Porter *et al.*, 2002)
JPA1345	Δ*cheA*$_1$, Δ*cheA*$_2$, Δ*cheA*$_3$ & Δ*cheA*$_4$ derivative of WS8N.	(Porter *et al.*, 2002)
JPA1902	Δ*cheA*$_3$ & Δ*cheA*$_4$ derivative of WS8N.	This study
JPA1436	Δ*cheA*$_3$ derivative of WS8N containing a *cfp-cheA*$_4$ fusion in place of the wild-type *cheA*$_4$ in the chromosome.	Previously created by George Wadhams
JPA1535	Δ*cheA*$_4$ derivative of WS8N containing a *yfp-cheA*$_3$ fusion in place of the wild-type *cheA*$_3$ in the chromosome.	(Wadhams *et al.*, 2005)
Plasmids		
pQE60	IPTG-inducible expression vector for *E. coli*. Introduces RGS(H)6 at the C-terminus of the protein. Confers ampicillin resistance.	Qiagen
pREP4	Plasmid containing the *lacIq* gene and conferring kanamycin resistance. Compatible with pQE60.	Qiagen
pk18*mobsac*B	Allelic-exchange suicide vector mobilized by *E. coli* S17-1λpir. Confers kanamycin resistance and sucrose sensitivity.	(Schäfer *et al.*, 1994)
pIND4	IPTG-inducible expression vector for *R. sphaeroides*. Confers kanamycin resistance.	(Ind *et al.*, 2009)
pQE60(A_3P1)-CheA_2	Plasmid for expression of (A_3-P1)-CheA_2 in *E. coli*	This study
pQE60(A_2P1)-CheA_3	Plasmid for expression of (A_3-P1)-CheA_2 in *E. coli*	This study
pINDCheA_x-(A_YP5)	Plasmids for expression of CheA_X-(A_yP5) in	This study

	R. sphaeroides. Combinations with X = 3 or 4 and Y = 3 or 4 were constructed.	
pINDYFPCheA$_x$-(A$_Y$P5)	Plasmids for expression of a YFP-CheA$_x$-(A$_y$P5) fusion protein in *R. sphaeroides*. Combinations with X = 2, 3 or 4 and Y = 2, 3 or 4 were constructed.	This study
pINDCFPCheA$_4$-(A$_3$P5)	Plasmid for expression of a CFP-CheA$_4$-(A$_3$P5) fusion protein in *R. sphaeroides.*	This study

1

2

3

4

Figure Legends

Figure 1.The chemosensory network of R. sphaeroides

Schematic diagram showing the current working model of the *R. sphaeroides* fla1 signal transduction pathway. There are two clusters of chemosensory proteins, one at the cell pole and one in the cytoplasm. The polar chemosensory cluster is thought to respond to external signals whilst the cytoplasmic chemosensory cluster is thought to respond to the metabolic state of the cell. Chemosensory signal transduction requires three CheA kinase proteins ($CheA_2$, $CheA_3$ and $CheA_4$) and five response regulator proteins ($CheB_1$, $CheB_2$, $CheY_3$, $CheY_4$ and $CheY_6$). In the diagram, red arrows indicate processes involved in controlling rotation of the flagellar motor and blue arrows those involved in adaptation. Figure adapted from (Porter *et al.*, 2008b).

Figure 2. Phosphorylation of the P1 domain-swapped CheAs

Phosphorimages of SDS-PAGE gels measuring the rates of: A. autophosphorylation of $CheA_2$, B. phosphorylation of $CheA_3$ by $CheA_4$, C. phosphorylation of (A_2P1)-$CheA_3$ by $CheA_4$, D. autophosphorylation of (A_3P1)-$CheA_2$, All proteins except $CheA_4$ were present at a final concentration of 5 μM; the concentration of $CheA_4$ was 20 μM. Reactions were initiated by addition of 0.5 mM [γ-^{32}P] ATP. 10 μl reaction samples were taken at the time points indicated and quenched in 10 μl of 3 x SDS/EDTA loading dye. The quenched samples were analyzed by SDS-PAGE and detected by phosphorimaging.

Figure 3. Phosphotransfer from the P1 domain-swapped CheAs to the response regulators

Phosphorimages of SDS-PAGE gels measuring phosphotransfer to the chemotaxis response regulators. A schematic diagram illustrating the domain structure of the CheAs involved in the kinase reactions is shown on the left of each panel and the SDS-PAGE gel on the right for: A. $CheA_2$-P, B. $CheA_3$-P, C. (A_3P1)-$CheA_2$-P, D. (A_2P1)-$CheA_3$-P. 5 μM of each protein was

preincubated with 0.5 mM [γ-^{32}P] ATP for 30 minutes. Reactions B&C also contained 20 μM $CheA_4$. Phosphotransfer reactions were initiated by addition of response regulators (10 μM) to the reaction mix. 10 μl samples were removed after 30 s and quenched immediately by addition of 10 μl of 3 x SDS/EDTA loading dye. The quenched samples were analyzed by SDS-PAGE and detected by phosphorimaging. The lane labeled N shows a control reaction in which an equal volume of buffer was added instead of the response regulators. The remaining lanes are labeled according to which response regulator was used. Phosphotransfer is indicated by the appearance of phosphorylated response regulator and/or a reduction in the amount of CheA-P.

Figure 4. Soft agar swim chemotaxis assay of P1 domain-swapped CheAs

A histogram comparing the chemotactic ability of the P1 domain-swapped mutants, *(A_3P1)-cheA_2* replacing *cheA_2* (JPA926), *(A_2P1)-cheA_3* replacing *cheA_3* (JPA927) and the double domain-swapped mutant with *(A_3P1)-cheA_2* replacing *cheA_2* and *(A_2P1)-cheA_3* replacing *cheA_3* (JPA1103) with wild-type (WS8N), non-chemotactic (JPA1315) and non-motile (JPA1213) strains. The soft agar swim plates contained 100 μM sodium propionate and were incubated for 48 hours under aerobic conditions. The error bars indicate the standard error of the mean from 9 experiments.

Figure 5. Cellular localization of the P5 domain-swapped CheAs

Representative fluorescence images showing the localization of YFP tagged P5 domain-swapped proteins. The YFP fusion protein was expressed from an IPTG-inducible expression vector introduced into the appropriate *cheA* deletion strain by conjugation; a concentration of 1 μM IPTG was used to induce expression in each case. In the strains shown here constructs containing A_2P5 were introduced into a Δ*cheA_2 R. sphaeroides* strain and those containing A_3P5 or A_4P5 into a Δ*cheA_3*Δ*cheA_4* strain.

Figure 6. Effect of background strain on the localization of P5 domain-swapped mutants to the cytoplasmic cluster

Representative fluorescence images showing the degree of localization of YFP-CheA$_3$-(A$_4$P5) and CFP-CheA$_4$-(A$_3$P5) to the cytoplasmic cluster in different background strains. The fluorescently labelled protein was expressed from an IPTG-inducible expression vector introduced into the appropriate *cheA* deletion strain by conjugation; a concentration of 1 μM IPTG was used to induce expression in each case.

Figure 7. Soft agar swim chemotaxis assay of P5 domain-swapped CheA$_3$ and CheA$_4$.

A histogram comparing the chemotactic ability of P5 domain-swapped CheA$_3$ and CheA$_4$. pIND-*cheA$_3$* and pIND-*cheA$_3$-(A$_4$P5)* were introduced into a Δ*cheA$_3$* strain whilst pIND-*cheA$_4$* and pIND-*cheA$_4$-(A$_3$P5)* were introduced into a Δ*cheA$_4$* strain. Soft agar swim plates containing 100 μM sodium propionate and variable concentrations of IPTG were inoculated with 5 μl of each strain and grown at 30 ^{o}C for 48 hours. The wild type strain (WS8N) a non-motile strain (*cheY$_6$(D56N)*) a non-chemotactic strain (Δ*cheA$_1$*, Δ*cheA$_2$* & Δ*cheA$_3$*), in addition to Δ*cheA$_3$* and Δ*cheA$_4$* strains were included as controls. The error bars indicate the standard error of the mean from 9 experiments.

Figure 8. Comparison of residues comprising the T. maritima *CheA-CheW binding site with equivalent residues in* R. sphaeroides *CheAs*.

A) Surface representation of the P5 domain of *T. maritima* CheA coloured according to sequence conservation; colouring is from white (no sequence conservation) to blue (identity). B) Ribbon diagram of the P5 domain of *T. maritima* CheA with residues within 4.5 Å of CheW in the crystal structure of the CheA-CheW complex shown as sticks and coloured according to sequence conservation. The orientation of the protein is identical to that in panel A). It should be noted that G659 is not apparent in this figure due to the lack of sidechain. C) Sequence alignment of the CheA P5 domains from *T. maritima*, *E. coli*, *S. typhimurium* and *R. sphaeroides*. The sequence of the

CheA$_3$ P5 domain was not present in the alignment used for analysis but has been added here for comparison. The *T. maritima* sequence is coloured according to sequence conservation using the same scale as in A) and B); residues forming part of the CheW binding site are indicated with an asterix.

Figure 1

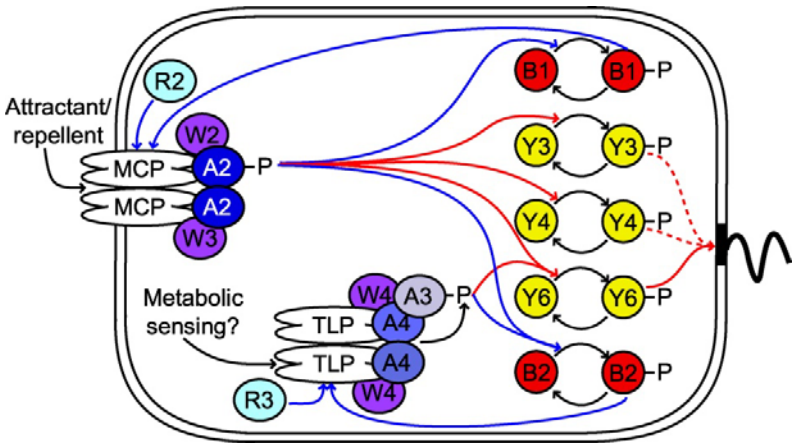

Attractant/
repellent
R2
W2
MCP
A2
P
MCP
A2
W3
Metabolic
sensing?
W4
A3
P
TLP
A4
TLP
A4
R3
W4
B1
B1
P
Y3
Y3
P
Y4
Y4
P
Y6
Y6
P
B2
B2
P

Figure 2

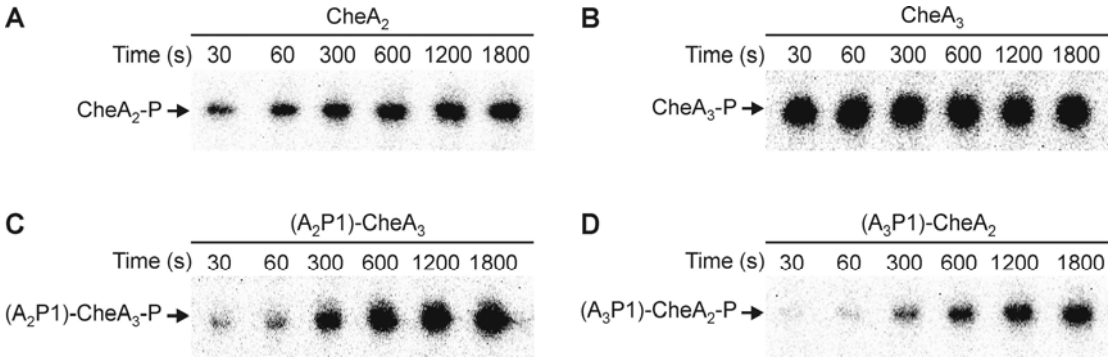

A
CheA2
Time (s) 30 60 300 600 1200 1800
CheA2-P
B
CheA3
Time (s) 30 60 300 600 1200 1800
CheA3-P
C
(A2P1)-CheA3
Time (s) 30 60 300 600 1200 1800
(A2P1)-CheA3-P
D
(A3P1)-CheA2
Time (s) 30 60 300 600 1200 1800
(A3P1)-CheA2-P

Figure 3

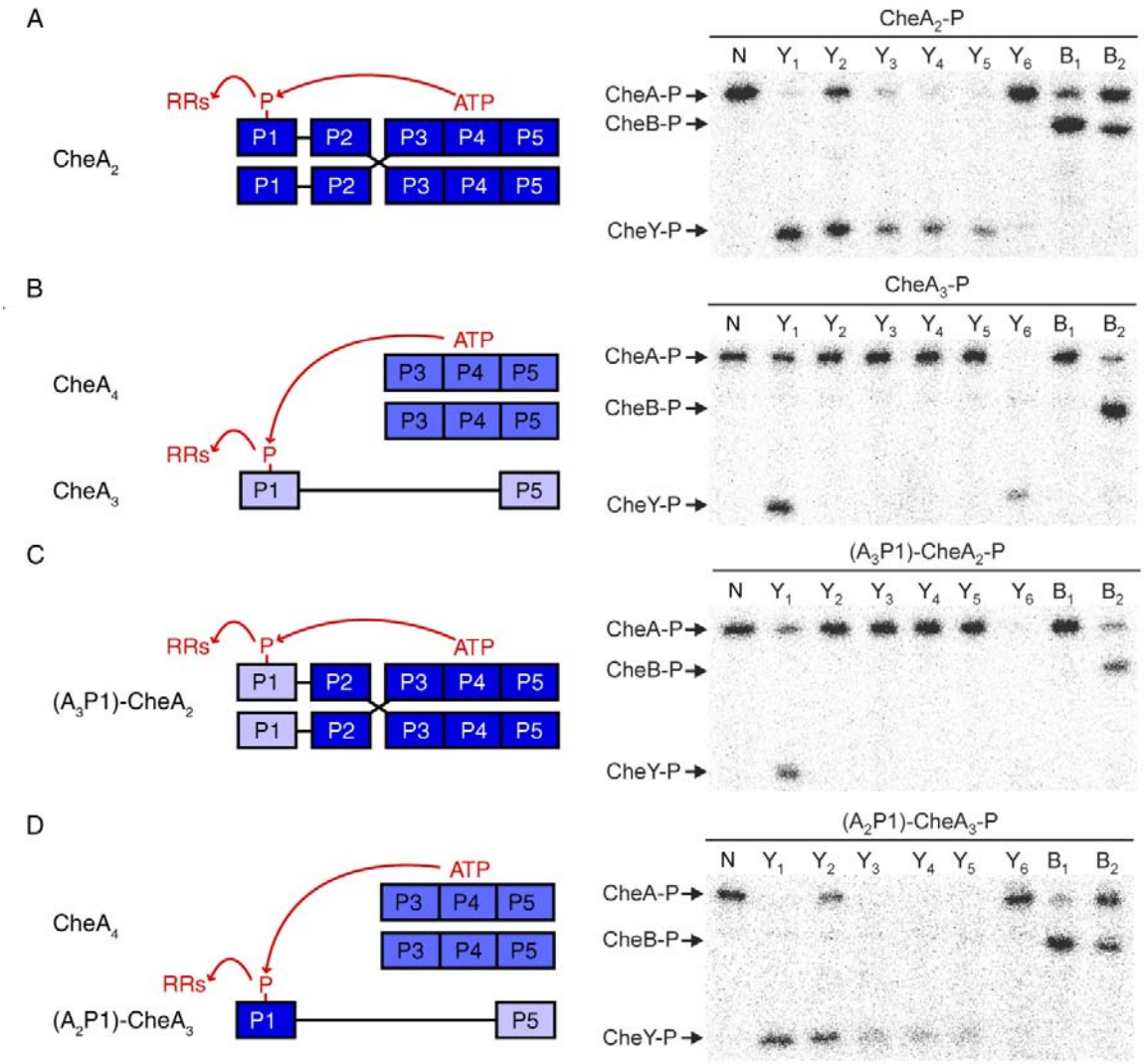

A
RRs
P
ATP
P1
P2
P3
P4
P5
CheA2
CheA2-P
N
Y1
Y2
Y3
Y4
Y5
Y6
B1
B2
CheA-P
CheB-P
CheY-P
B
CheA4
CheA3
CheA3-P
C
(A3P1)-CheA2
(A3P1)-CheA2-P
D
(A2P1)-CheA3
(A2P1)-CheA3-P

Figure 4

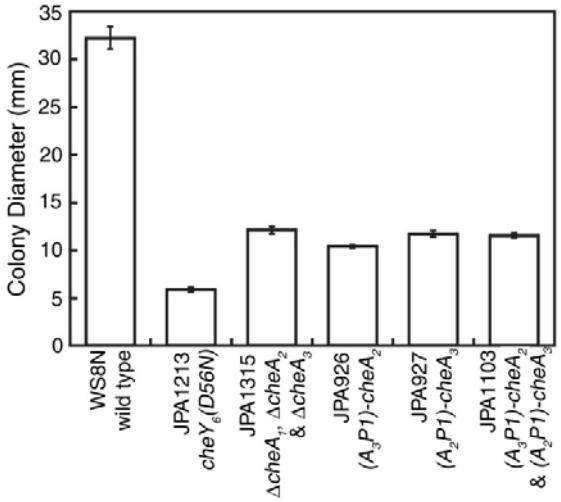

Colony Diameter (mm)
35
30
25
20
15
10
5
0
WS8N wild type
JPA1213 cheY6(D56N)
JPA1315 ΔcheA1, ΔcheA2 & ΔcheA3
JPA926 (A3P1)-cheA2
JPA927 (A2P1)-cheA3
JPA1103 (A3P1)-cheA2 & (A2P1)-cheA3

Figure 5

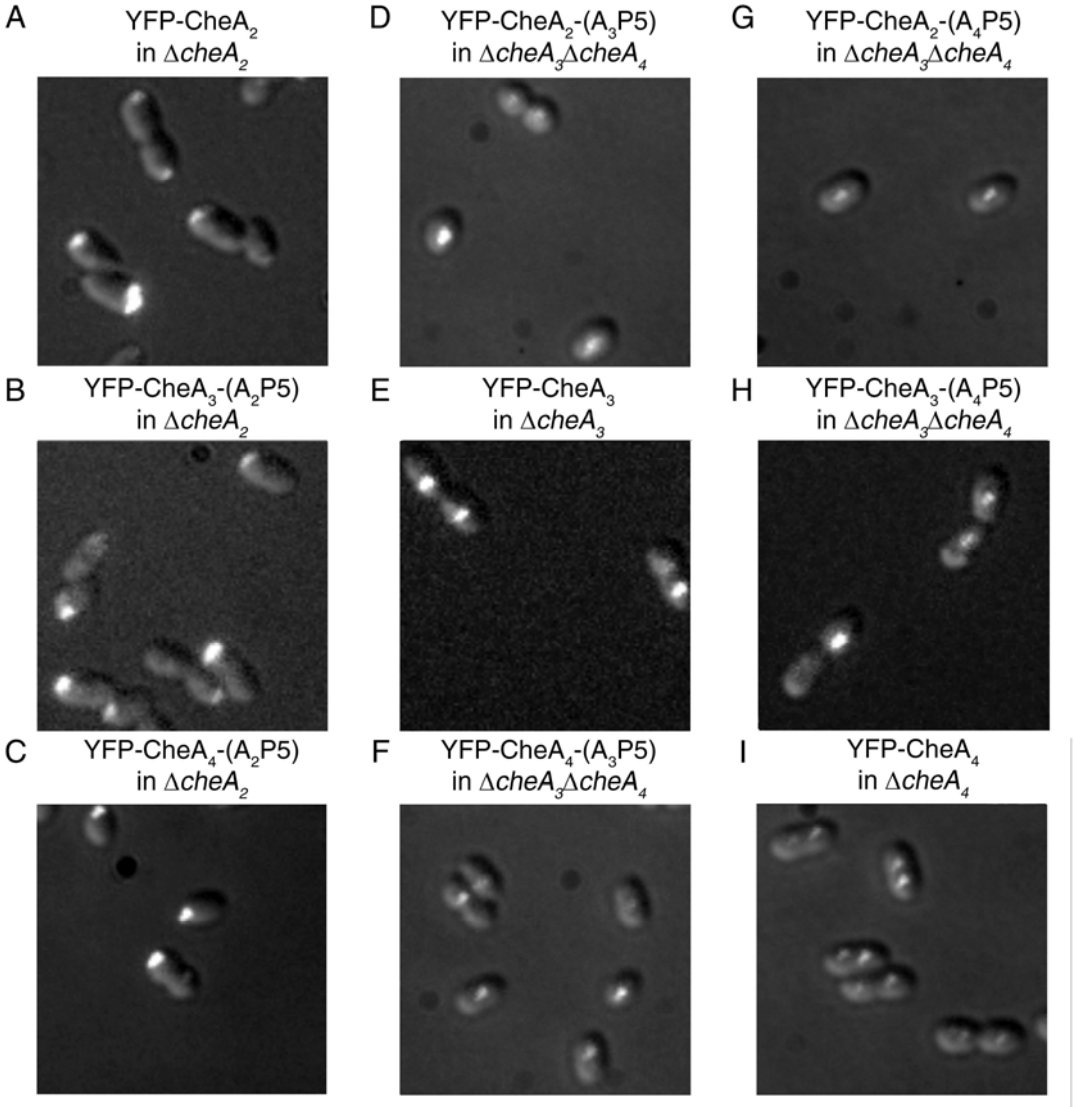

A
YFP-CheA2
in ΔcheA2
D
YFP-CheA2-(A3P5)
in ΔcheA3ΔcheA4
G
YFP-CheA2-(A4P5)
in ΔcheA3ΔcheA4
B
YFP-CheA3-(A2P5)
in ΔcheA2
E
YFP-CheA3
in ΔcheA3
H
YFP-CheA3-(A4P5)
in ΔcheA3ΔcheA4
C
YFP-CheA4-(A2P5)
in ΔcheA2
F
YFP-CheA4-(A3P5)
in ΔcheA3ΔcheA4
I
YFP-CheA4
in ΔcheA4

Figure 6

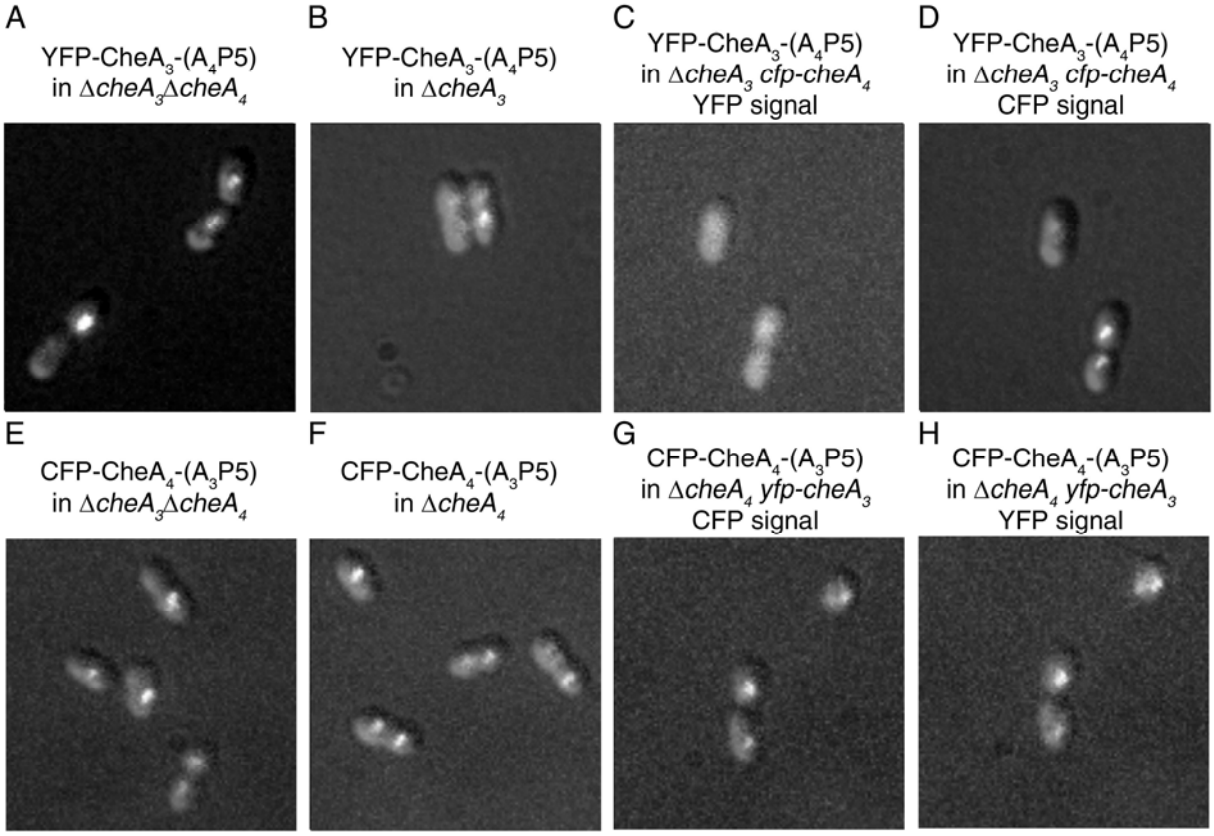

A
YFP-CheA3-(A4P5)
in ΔcheA3ΔcheA4
B
YFP-CheA3-(A4P5)
in ΔcheA3
C
YFP-CheA3-(A4P5)
in ΔcheA3 cfp-cheA4
YFP signal
D
YFP-CheA3-(A4P5)
in ΔcheA3 cfp-cheA4
CFP signal
E
CFP-CheA4-(A3P5)
in ΔcheA3ΔcheA4
F
CFP-CheA4-(A3P5)
in ΔcheA4
G
CFP-CheA4-(A3P5)
in ΔcheA4 yfp-cheA3
CFP signal
H
CFP-CheA4-(A3P5)
in ΔcheA4 yfp-cheA3
YFP signal

Figure 7

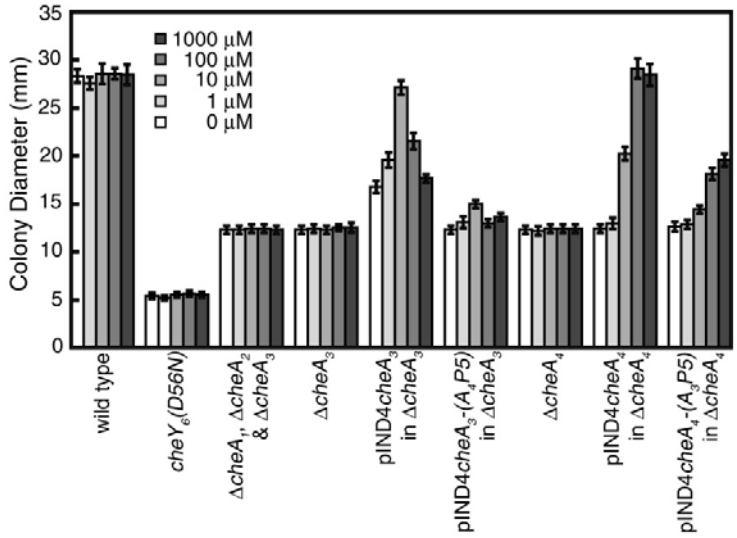
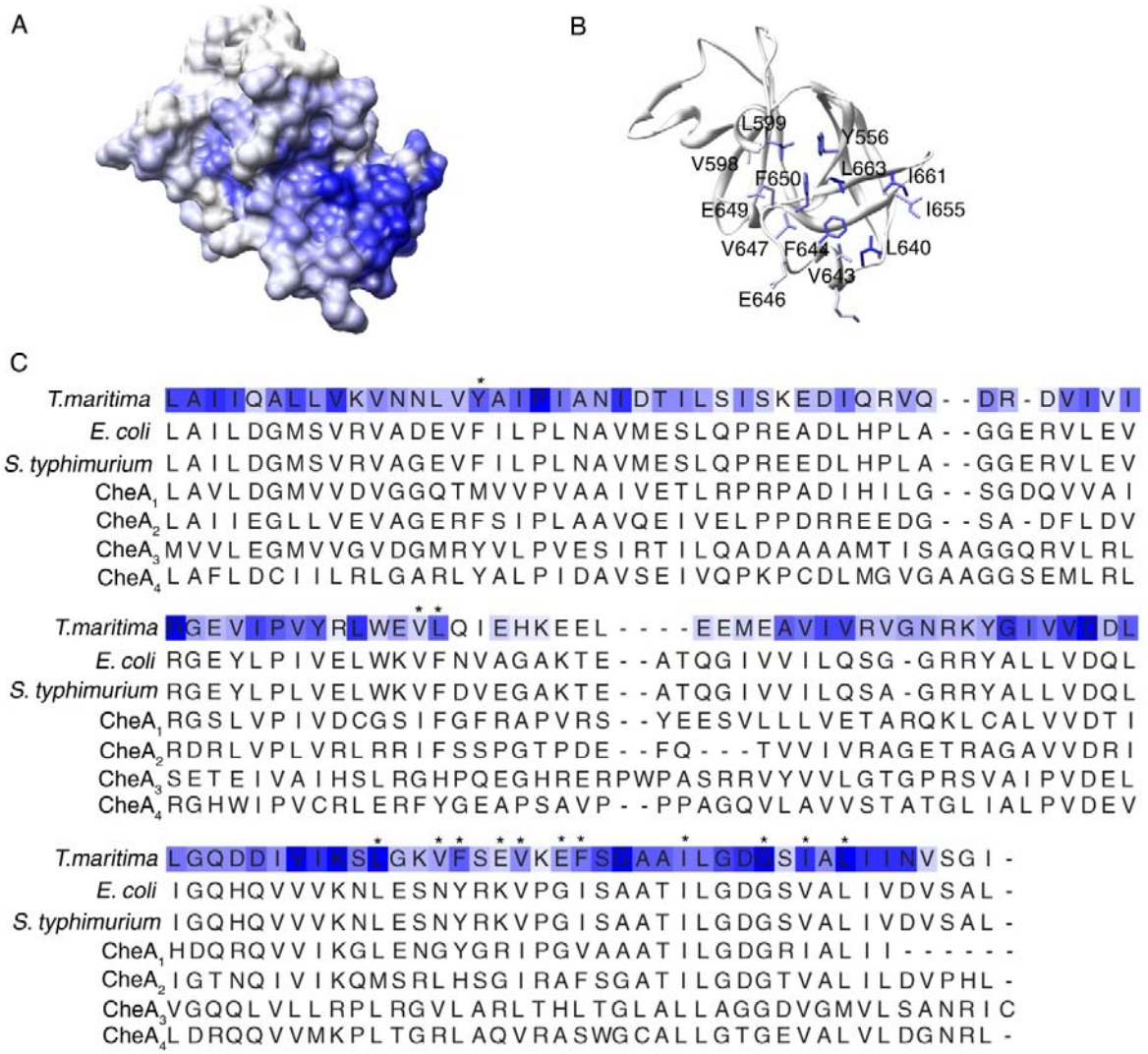

Colony Diameter (mm)
35
30
25
20
15
10
5
0
1000 µM
100 µM
10 µM
1 µM
0 µM
wild type
cheY6(D56N)
ΔcheA1, ΔcheA2 & ΔcheA3
ΔcheA3
pIND4cheA3 in ΔcheA3
pIND4cheA3-(A4P5) in ΔcheA3
ΔcheA4
pIND4cheA4 in ΔcheA4
pIND4cheA4-(A3P5) in ΔcheA4

Figure 8


A
B
L599
Y556
V598
L663
I661
F650
E649
I655
V647
F644
L640
V643
E646
C
T.maritima LATIQALLVKVNNLVYAIEIANIDTILSISKEDIQRVQ--DR-DVIVI
E. coli LAILDGMSVRVADEVFILPLNAVMESLQPREADLHPLA--GGERVLEV
S. typhimurium LAILDGMSVRVAGEVFILPLNAVMESLQPREEDLHPLA--GGERVLEV
CheA1 LAVLDGMVVDVGGQTMVVPVAAIVETLRPRPADIHILG--SGDQVVAI
CheA2 LAIIEGLLVEVAGERFSIPLAAVQEIVELPPDRREEDG--SA-DFLDV
CheA3 MVVLEGMVVGVDGMRYVLPVESIRTILQADAAAAMTISAAGGQRVLRL
CheA4 LAFLDCIILRLGARLYALPIDAVSEIVQPKPCDLMGVGAAGGSEMLRL
T.maritima RGEVIPVYRLWEVLQIEHKEEL----EEMEAVIVRVGNRKYGIVVEDL
E. coli RGEYLPIVELWKVFNVAGAKTE--ATQGIVVILQSG-GRRYALLVDQL
S. typhimurium RGEYLPLVELWKVFDVEGAKTE--ATQGIVVILQSA-GRRYALLVDQL
CheA1 RGSLVPIVDCGSIFGFRAPVRS--YEESVLLLVETARQKLCALVVDTI
CheA2 RDRLVPLVRLRRIFSSPGTPDE--FQ---TVVIVRAGETRAGAVVDRI
CheA3 SETEIVAIHSLRGHPQEGHRERPWPASRRVYVVLGTGPRSVAIPVDEL
CheA4 RGHWIPVCRLERFYGEAPSAVP--PPAGQVLAVVSTATGLIALPVDEV
T.maritima LGQDDIVIKSLGKVFSEVKEFSLAAILGDFSIAIIINVSGI-
E. coli IGQHQVVVKNLESNYRKVPGISAATILGDGSVALIVDVSAL-
S. typhimurium IGQHQVVVKNLESNYRKVPGISAATILGDGSVALIVDVSAL-
CheA1 HDQRQVVIKGLENGYGRIPGVAAATILGDGRIALII------
CheA2 IGTNQIVIKQMSRLHSGIRAFSGATILGDGTVALILDVPHL-
CheA3 VGQQLVLLRPLRGVLARLTHLTGLALLAGGDVGMVLSANRIC
CheA4 LDRQQVVMKPLTGRLAQVRASWGCALLGTGEVALVLDGNRL-