Spirochetes produce ordered chemoreceptor arrays of unusual composition, arrangement, and symmetry to compensate for a highly curved membrane

Alise R. Muok[1], Davi R. Ortega[2,3], Kurni Kurniyati[4], Wen Yang[1], Adam Sidi Mabrouk[1], Chunhao Li[4], Brian R. Crane[5], Ariane Briegel[1*]
[1]Institute for Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden
[2]Department of Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[3]Laboratory for Research in Complex Systems, 425 California St., Suite 1200, San Francisco, CA 94104, USA
[4]Department of Oral and Craniofacial Molecular Biology, Philips Research Institute for Oral Health, Virginia Commonwealth University, Richmond, VA 23298
[5]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14850

Author contributions: A.R.M., D.R.O, K.K., C.L., B.R.C, and A.B. designed research; A.R.M., D.R.O., and K.K. performed research; A.R.M., D.R.O, W.Y., K.K., and A.S. analyzed data; and A.R.M., D.R.O, and A.B. wrote the paper with input from all authors.

* Corresponding author: A.briegel@biology.leidenuniv.nl

Abstract
The prokaryotic chemotaxis system is arguably the best-understood signaling pathway in biology, but most insights have been obtained from only a few model organisms. In all previously described species, chemoreceptors organize with the histidine kinase (CheA) and coupling protein (CheW) into a hexagonal (P6 symmetry) extended array that is considered universal among archaea and bacteria. Here, for the first time, we apply cryo-electron tomography to whole *Treponema denticola* (*Td*) cells to investigate the structure of a spirochete (F2) chemotaxis system. The *Td* chemoreceptor arrays assume a truly unusual arrangement of the supra-molecular protein assembly that has likely evolved to accommodate the high membrane curvature present in spirochetes. A two-fold (P2) symmetry of the chemotaxis apparatus in *Td* emerges from a strict linear organization of the kinase CheA, which generates arrays that run parallel to the cell axis. The arrays have several additional atypical features, such as an extended dimerization domain of CheA and a variant CheW-CheR-like fusion protein that is critical for maintaining an ordered, functional chemosensory apparatus in an extremely curved cell. Furthermore, the previously characterized *Td* oxygen sensor ODP influences array integrity and its loss substantially orders CheA. These results demonstrate the importance of examining chemotaxis structures of non-model organisms *in vivo* and suggest a greater diversity of this signaling system than previously thought.

Introduction
Chemotaxis is a behavior most motile bacteria employ to sense their chemical environment and navigate toward favorable conditions. The cellular machinery that underlies this system is a supra-molecular assembly of proteins that is relatively conserved across motile bacteria and chemotactic archaea. The main components of the system are transmembrane chemotaxis receptors called methyl-accepting chemotaxis proteins (MCPs), the histidine kinase CheA, and the adaptor protein CheW. The intracellular tips of MCPs bind CheA and CheW and communicate changes from the external chemical environment into the cell by modulating CheA kinase activity[1–3]. Activation of CheA initiates an intracellular phosphorelay that ultimately controls flagellar rotation and cell movement. CheA functions as a dimer and possesses five domains (P1-P5) with distinct roles in autophosphorylation and array integration. The P1 domain is the phosphate substrate domain, P2 interacts with response regulators, P3 is the dimerization domain, P4 binds ATP, and P5 interacts with CheW. In the model species *Escherichia coli* (*Ec*), CheA P5 and CheW are paralogs that interact pseudo-symmetrically to form six-subunit rings. In all bacterial and archael species examined thus far, the MCPs are arranged in a trimer-of-dimer oligomeric state and further organize into a hexagonal lattice. In *Ec*, the receptors are connected by the highly ordered rings of CheA and CheW bound to the cytoplasmic tips of the receptors[4,5]. These insights have established a widely accepted central

model of the chemotaxis array[4,6,7]. However, emerging research has recently revealed divergent components and arrangements of the chemotaxis apparatus in non-canonical organisms. For example, many bacteria contain multiple copies of CheW, with some fused to other protein modules (i.e. CheV), and *in vivo* analyses of *Vibrio cholerae* (*Vc*) chemotaxis arrays have revealed that CheA and CheW in the rings apparently lack an ordered arrangement[8,9].

The chemotaxis systems in prokaryotes have been classified into 19 systems based on phylogenomic markers[10]. These classes include 17 'Flagellar' systems (F), one 'Alternative Cellular Function' system (ACF), and one 'Type Four Pilus system' (TFP). The spirochete chemotaxis system belongs to the 'Flagella class 2' (F2) category, which has not been investigated with structural methods[10]. Herein, we examine the *in vivo* structure of an F2 system in the spirochete *Treponema denticola* (*Td*) by cryo-electron tomography (cryo-ET). We imaged chemotaxis arrays and demonstrate the presence of a novel array architecture, that likely owes to the high curvature of the cells. Genetics experiments, bioinformatics analyses, structural investigations, and molecular modeling of *Td* chemotaxis proteins reveals adaptations that have likely evolved to accommodate formation of an extended chemotaxis array in a highly curved membrane. We demonstrate that a CheR-like fusion domain in a variant CheW is key for maintaining the structural integrity of the arrays. Furthermore, cryo-ET analysis of *Td* cells lacking the previously characterized oxygen sensor ODP demonstrates significant changes in CheA mobility[11].

Results

Conservation of F2 chemotaxis proteins

F2 systems are the signature systems of Spirochaetia and Brachyspirae

Based on their evolutionary history and gene cluster organization, the spirochete *T. denticola* has a single F2 chemotaxis system encoded in its genome[10]. In the Microbial Signal Transduction Database version 3 (MiST3)[12], 306 genomes contain at least 1 CheA of the F2 class (CheA-F2) (Dataset 1). Only three of these genomes were not members of the Spirochaetota phylum and are likely the consequence of lateral gene transfer (see Supplementary Methods).

However, there are 1096 genomes classified as Spirochaetota in GTDB[13] and 804 of these are present in the MiST3 database (Dataset 2)[12]. With these 804 genomes, we mapped the different classes of CheA kinases to the Spirochaetota taxonomy tree (Figure S1). Based on the topology of this tree, it appears that the major chemosensory systems in the genomes from the Spirochaetota phylum are: F1 (Leptospirae), a transitional hybrid F1/F2 system (Brachyspirae), and F2 (Spirochaetia) (Fig. S1).

CheW-CheR$_{like}$ is only present in complete F2 systems

Intriguingly, the *Td* genome contains two CheW proteins: a classical CheW protein (TDE1589) and a CheW that possesses a C-terminal CheR-like domain separated by a 28-residue linker (TDE1492, CheW-CheR$_{like}$). Typically, CheR is a methyltransferase that, together with the methylesterase CheB, controls the methylation state of the receptors and thus provides an adaptation system[14]. In the CheW-CheR$_{like}$ fusion protein, the linker residues are predicted to form a single alpha helix with flanking disordered regions (Jpred, Fig. S4D). The F2 system's main architectural difference from other systems is the presence of this unusual scaffold protein. Of the 306 genomes with at least one CheA-F2, all of them also contain CheW-CheR$_{like}$ with a few exceptions (See Supplementary Methods). However, all Spirochaetia genomes contain the CheW-CheR$_{like}$ gene (Fig. S1).

The CheW-CheR$_{like}$ has not been found in other chemosensory system classes. In the MiST3 database, there are 274 CheW-CheR$_{like}$ proteins in 273 genomes. Of those, 15 do not contain an F2 chemosensory system but these genomes are not fully assembled (14 at contig and 1 at scaffold level, see Supplementary Methods). Based on these results, we conclude that CheW-CheR$_{like}$ is likely only present in complete F2 systems.

The three F2 CheW domains evolved different sequence patterns

F2 systems contain three proteins with a CheW domain: the classical scaffold CheW, CheW-CheR$_{like}$, and the P5 domain from the histidine kinase CheA. To investigate sequence patterns of these CheW domains, we built a pipeline to produce non-redundant sequence datasets of CheW, CheW-CheR$_{like}$, and CheA of the class F2 from all genomes with at least one CheA-F2. The final alignment of each group contained the CheW domain portion of 74 CheW proteins, 59 CheW-CheR$_{like}$ proteins, and 73 CheA proteins. The sequences of each group are summarized in sequence logos (Fig. S3A). The

logos reveal conserved regions at the ring interfaces.

The F2 CheR domains evolved different sequence patterns

There are two proteins in the F2 system with a CheR domain: the CheR-F2 methyltransferase and the CheW-CheR$_{like}$ protein. To investigate sequence patterns in these proteins, we produced a final sequence dataset with the CheR domain of 83 CheR-F2 and 88 CheW-CheR$_{like}$ proteins and summarized them in sequence logos (Fig. S3B). The CheR$_{like}$ domain has a 20% identity to the *Td* CheR methyl-transferase (TDE0647, LALIGN). The two residues that are responsible for catalytic function to change chemoreceptor methylation levels, R79 and Y218 in *Td* CheR, are different in the two CheR domains[14]. Both residues are highly conserved in CheR-F2 proteins, but are modified in the CheW-CheR$_{like}$ protein (R79W and Y218F). Furthermore, the conserved region at the C-terminus of CheR is not conserved in CheW-CheR$_{like}$. Based on these results we speculate that the CheR$_{like}$ domain does not possess methyltransferase activity. Collectively, our analyses suggest that CheR and CheW-CheR$_{like}$ have different biological functions.

The structure of the *Treponema denticola* (*Td*) chemotaxis array in wild-type cells

Cell poles of intact *Td* cells were imaged by cryo-electron tomography (cryo-ET) and used for three-dimensional reconstructions. Top views (cross-sections through the array) and side views (visualizing the long axes of the receptors) of membrane-associated arrays were clearly visible (Fig. 1A, S4A). Sub-volume averaging revealed the conserved receptor trimer-of-dimers in the typical hexagonal arrangement. Remarkably, several novel features of the chemotaxis arrays are apparent (Fig. 1B). Specifically, a density of unknown origin is located in the center of the receptor hexagons and slightly above the plane of the CheA:CheW rings. This density, which will hereafter be referred to as the middle density, extends from two subunits in the rings (Fig. 2A). Additionally, there are small but distinct puncta of density in between some of the trimer-of-dimer modules (Fig. 1B). However, averages of the arrays at the CheA:CheW layer did not reveal discernible CheA density, indicating either a sparse or disordered distribution of CheA or a highly mobile kinase (Fig. 1C). Lysis of *Td* cells via lysozyme treatment followed by cryo-ET reveals intact arrays (Fig. S4B).

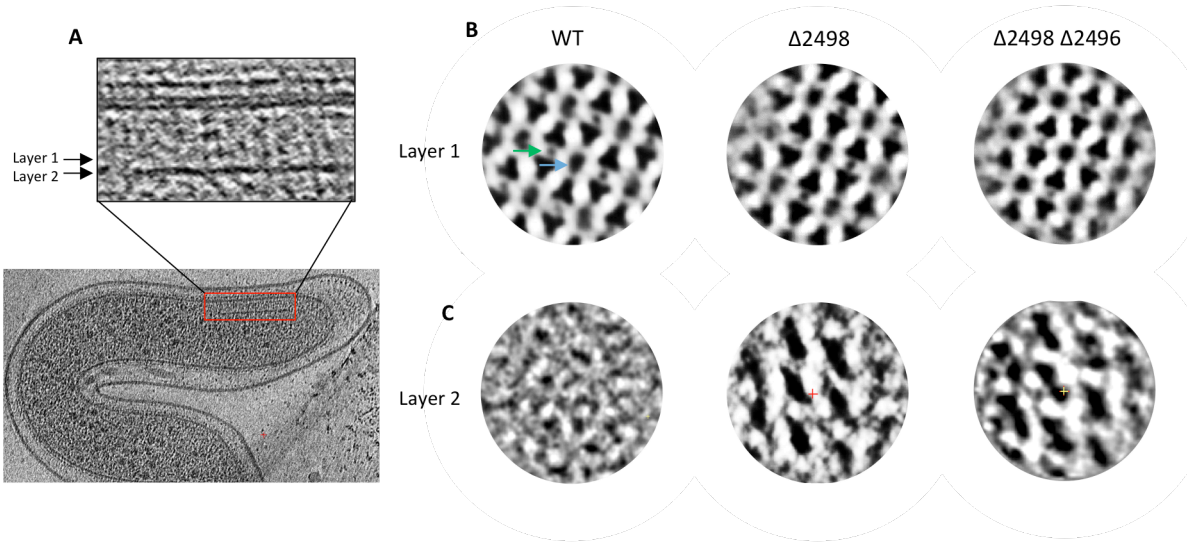


Fig. 1 Cryo-electron tomography of whole *T. denticola* (*Td*) cells reveals the protein arrangement of chemotaxis machinery. (A) Side-views of the membrane-associated chemotaxis apparatus illustrate the location of the receptor layer (Layer 1) and CheA:CheW baseplate (Layer 2). These layers are spaced ~90 Å from one another. (B) Sub-volume averaging of three *Td* strains reveals the universally-conserved receptor trimer-of-dimer arrangement with 12 nm spacing between opposing trimer-of-dimer modules. Notably, density is apparent in the center of the receptor hexagons (blue arrow) and between some receptor trimer-of-dimer modules (green arrow). In general, the densities in Layer 1 of the wild-type (WT) strain are better resolved. (C) Sub-volume averages at Layer 2 reveal the organization of CheA. Density corresponding to CheA is only apparent in the two *Td* deletion mutants and CheA is arranged in a linear fashion. In this arrangement, the density between the trimer-of-dimer modules (green arrow) corresponds to the CheA P3 domain.

Arrays in *T. denticola* deletion mutants

Previous experiments demonstrate that a protein called oxygen-binding diiron protein (ODP) functions as an oxygen sensor for chemotaxis in *Td*, however the study did not determine if ODP is an integral component of the array[11]. This protein is genetically coupled to a soluble receptor, TDE2496, with unclear signaling capabilities. TDE2496 likely integrates into the membrane-bound arrays based on the observation that no cytoplasmic arrays were observed in the tomograms. Moreover, the *Td* genome encodes only one CheA homolog, and cytoplasmic receptors often associate with distinct kinases[5,15]. TDE2496 forms trimers-of-dimers *in vitro* and can modulate CheA activity (Fig. S5A,B). To determine if the presence of ODP (TDE2498) and its cognate receptor (TDE2496) impacts array architecture or integrity, we conducted cryo-ET with two *Td* gene knock-out strains, Δ2498 and Δ2498Δ2496[11]. Importantly, reverse-transcription PCR confirms that deletion of ODP does not impact transcription of TDE2496[11]. The sub-volume averages of these strains reveal distinct differences in array densities compared to the wild-type (WT) strain (Fig. 1B,C, Fig. S4A). Namely, the location of CheA at the CheA:CheW layer (Layer 2, Fig 1C) is now clearly visible. Interestingly, CheA arranges in well-ordered linear rows. Placement of CheA necessarily positions the P3 domain in between two of the trimer-of-dimer modules in each hexagon. This position exclusively corresponds to the location of the puncta between receptor trimer-of-dimer modules observed in the WT array, indicating that this density corresponds to the P3 domain (Layer 1, Fig. 1B). Additionally, the receptor densities are less resolved than in the WT strain (Fig. 1B). This is further illustrated by the fact that when the selected cryo-ET particles for WT and mutant strains are averaged using the same size mask, the WT strain generates averages that resolve a larger portion of the array at Layer 1. Thus, the receptors and P3 domains are more ordered across the array in the WT strain whereas the P5-CheW layer is more ordered in the ODP deletion strain (Fig. S5C).

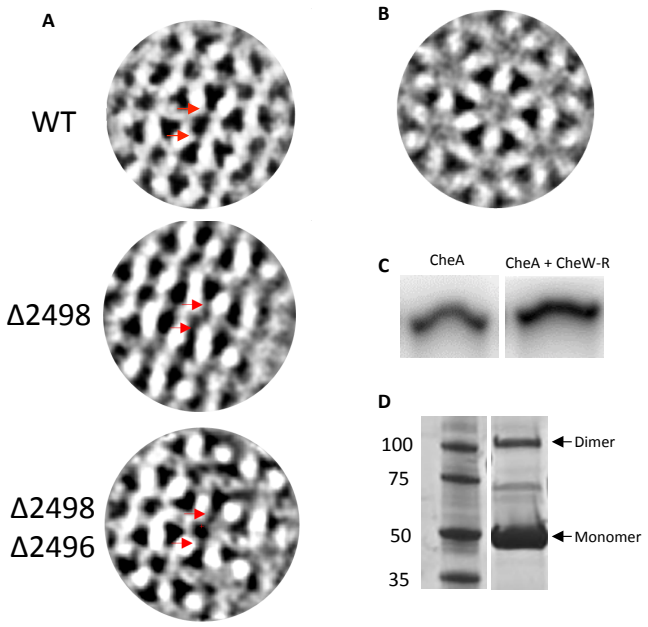


Fig. 2 Investigations into the identity of the density located in the center of the receptor hexagons, referred to as the middle density. (A) In Layer 1, the middle density extends to two positions in the CheA:CheW rings for all three *Td* strains (red arrows). (B) Deletion of a CheR-like domain that is covalently attached to the C-terminus of a CheW domain generates sub-tomogram averages that lack the middle density feature. (C) Radioisotope assays that monitor CheA autophosphorylation with [γ-^{32}P]ATP demonstrate that CheW-CheR$_{like}$ influences CheA activity *in vitro*. The proteins were incubated for 30 minutes at ambient temperatures and autophosphorylation was initiated for 12 minutes before quenching. Bands are from non-adjacent lanes on the same gel. (D) Purification of CheW-CheR$_{like}$ followed by Native PAGE gel analysis reveals the presence of a protein monomer (49 kDa) and protein dimer (98 kDa). Bands are from non-adjacent lanes on the same gel.

Analyses of the CheW-CheR$_{like}$ protein in *T. denticola*

To determine the composition of the middle density in the *Td* arrays, we examined spirochete genomes for unique chemotaxis proteins. The bioinformatics analyses identify CheW-CheR$_{like}$ as a conserved component of the F2 Spirochaetia chemotaxis system. In *Td,* the CheW-CheR$_{like}$ gene is co-transcribed with the only CheA, CheX, and CheY proteins in the genome (Fig. S4C)[15]. Furthermore, native gel electrophoresis of the purified CheW-CheR$_{like}$ protein demonstrates the presence of a protein dimer, and radioisotope assays that monitor CheA autophosphorylation demonstrate that the CheW-CheR$_{like}$ protein activates CheA kinase activity, similarly to *Tm* chemotaxis proteins (Fig. 2C,D, Fig. S4E,F). Therefore, we postulated that the middle density may be comprised of two CheR$_{like}$ domains that extend from two CheW-CheR$_{like}$ proteins in the CheA:CheW rings .

A *Td* strain lacking the $CheR_{like}$ domain ($\Delta CheR_{like}$) reveals a significant decrease in the prevalence and size of the arrays (Fig. S4A). Due to the small size of the arrays in $\Delta CheR_{like}$, only 194 particles were available for sub-tomogram averaging, but the resulting averages clearly demonstrate that the middle density is no longer present (Fig. 2B, S4A). Like the WT strain, the CheA density below the rings in $\Delta CheR_{like}$ is not apparent.

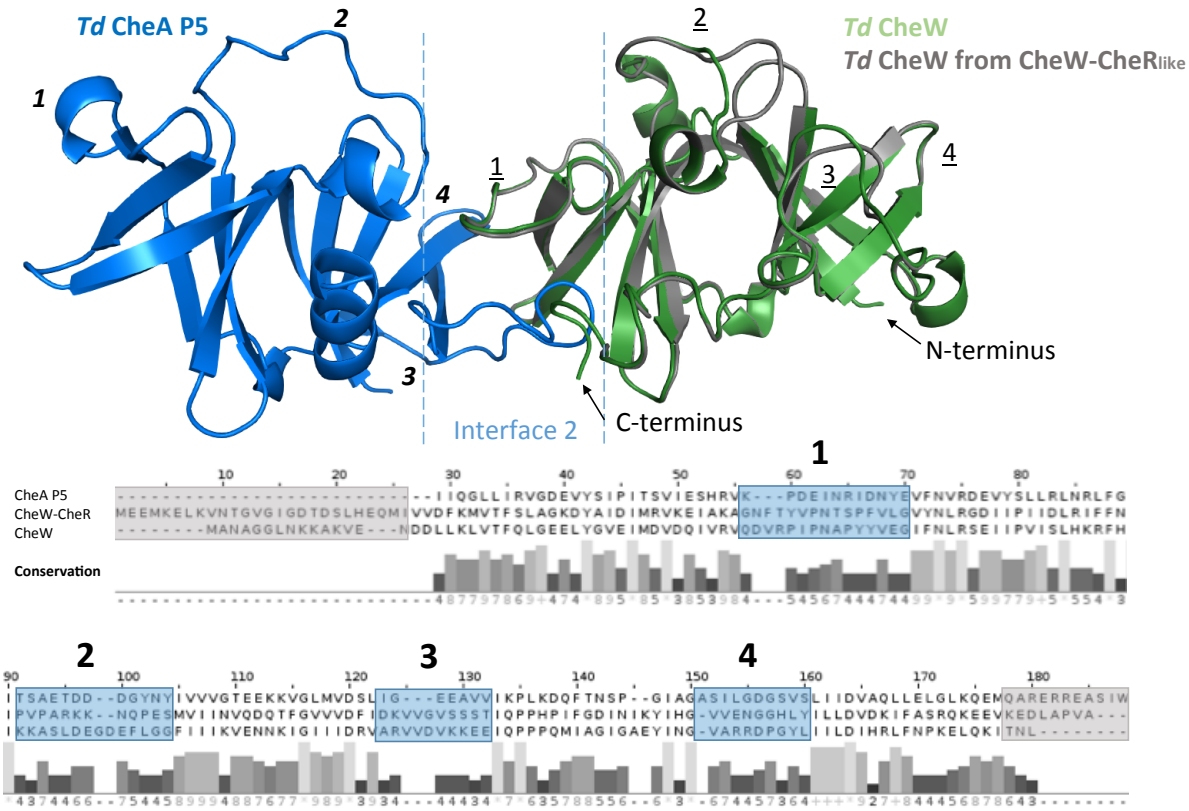


Fig. 3 Homology models of *Td* CheA P5, CheW, and the CheW domain of CheW-$CheR_{like}$ mapped onto a previously determined crystal structure of *Thermotoga maritima* CheA P5 and CheW (PDB ID: 3UR1). Variable regions between the three domains (blue boxes) were determined by sequence alignments of the three domains followed by conservation analysis. These variable regions are located at the rings interface regions. The CheW domains also have variable N-terminal and C-terminal regions that are not complimented in P5 (grey boxes). Variable regions for P5 and the CheW domains are denoted on the homology models by italicized numbers and underlined numbers, respectively.

Protein interfaces in the CheA:CheW:CheW-$CheR_{like}$ rings

Bioinformatics analyses demonstrate that all functional Spirochaetota F2 chemotaxis systems possess a CheW-$CheR_{like}$ homolog and at least one classical CheW protein. As only two of the CheW subunits in the hexagonal rings extend to the middle density (which probably arises from the CheW-$CheR_{like}$ protein), and two of the ring positions are occupied by CheA P5, it is likely that the other two positions are occupied by the classical CheW protein. (Fig. 2A, 5A). Within this arrangement three unique interaction interfaces are possible, interface 1 occurs between CheA P5 and the classical CheW (as seen in canonical systems), interface 2 occurs between CheA P5 and the CheW domain of CheW-$CheR_{like}$, and a third interface (interface 3) occurs between the classical CheW and the CheW domain of CheW-$CheR_{like}$ (Fig. 5A). To explore the binding interfaces within the *Td* rings, we analyzed homology models of the classical CheW, the CheW domain of CheW-$CheR_{like}$, and the CheA P5 domain. The CheW models were generated using a crystal structure of *Thermoanaerobacter tengcongensis* (*Tt*) CheW as the template (PDB ID: 2QDL), and the CheA P5 model was generated using a cryo-EM structure of *E. coli* CheA P5 (PDB ID:6S1K) (Fig. S6A-C)[16,17]. Three of the four regions with lowest sequence conservation among the three domains are located at interfaces 1-3 (Fig. 3). Alignment of the *Td* CheW and CheA P5 models to a crystal structure of *Tm* CheW in complex with *Tm* CheA P5 (PDB ID: 3UR1) further illustrates that these regions are located at the CheW:P5 ring interfaces (Fig 3)[18]. Mapping the variable regions onto the sequence logos of the F2

CheW domains demonstrates that they evolved different sequence patterns in these regions, with the exception of the variable region that is not located at the interaction interface (region 2, Fig. 3, S3).

CheA arrangement and array curvature in *T. denticola*

Sub-tomogram averaging reveals that *Td* CheA forms a linear arrangement across the chemotaxis array, linking the CheA:CheW rings into extended 'strands' that are held together by receptor:receptor interactions (Fig. 4A,B). These strands are apparent in the cryo-ET reconstructions and run relatively parallel to the axis of the cells, effectively allowing the strands to remain straight rather than bending to the cell curvature (Fig. 4D). Indeed, the angle between the cell axis and the extended strands is 10.4 +/- 8.6°, (*n* = 26 cells) and no significant difference was found among the three *Td* strains measured (Table S1C).

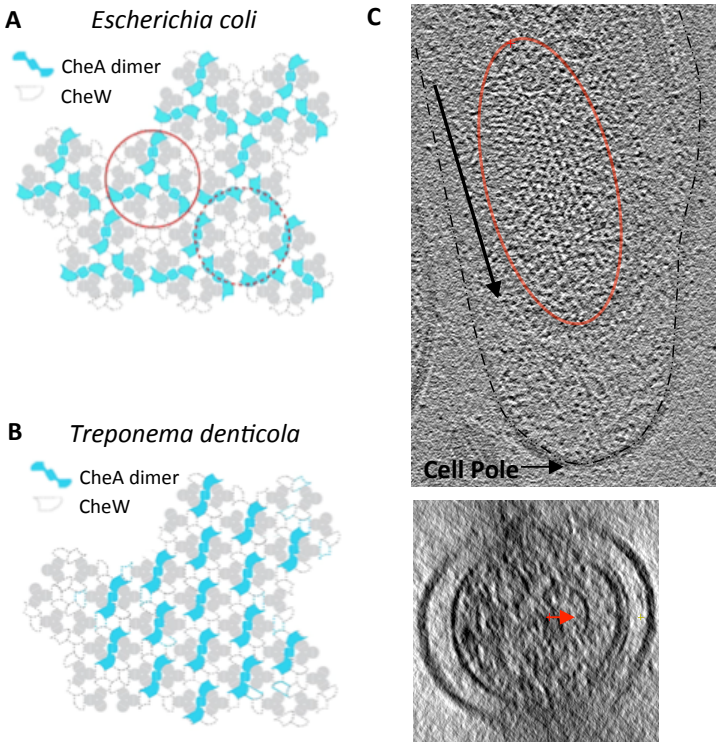


Fig. 4 The organization of CheA in *Td* differs from all previously reported arrangements. (A) The arrangement of CheA in *E. coli* and other canonical chemotaxis systems. CheA is arranged hexagonally to link adjacent CheA:CheW rings into an extended network across the cell membrane. A baseplate ring where three CheA subunits are incorporated is circled in red. A CheW-only baseplate ring is in a dashed circle. (B) In *Td,* CheA is arranged with a strict linear organization, producing 'strands' of CheA:CheW rings linked by the CheA P3 domain. (C) Top: The strands of CheA:CheW rings (red oval) follow the axis of *Td* cells (black arrow) with an average offset angle of 10.4 +/- 8.6°. The outer membrane of the cell is emphasized by a black dashed line. Bottom: The curvature of the inner membrane of *Td* at chemotaxis arrays in the reconstructions is 35.8 +/- 6.6 /µm (270 Å radius). The curvature of the CheA:CheW baseplate (red arrow) is 65.6 +/- 19 /µm (152 Å radius).

Previous studies in *Ec* and *Vc* demonstrate that the chemotaxis arrays can accommodate vastly different curvatures between lysed and artificially small mini-cell strains[8,19,20]. However, *Td* cells demonstrate a significantly higher curvature of the cell membrane than organisms investigated for their chemotaxis arrangement thus far [4,8,19]. As a measure of comparison, the *Vc* mini-cells used in a previous study have an inner membrane curvature of 9.15 +/- 4.5 /µm (radius 1092 Å, *n* = 6 cells), and the *Td* cells have an inner membrane curvature of 35.8 +/- 6.6 /µm (radius: 279 Å, *n* = 10 cells) (Fig. 4, Fig. S7, Table S1A,B). Additionally, the measured curvature of the *Td* CheA:CheW baseplate is 65.6 +/- 19 /µm (radius: 152 Å, *n* = 10 cells) (Fig. 4C, Table S1A).

The CheA:CheW rings present in a crystal structure (PDB ID: 3UR1) are flat and ~95 Å in diameter[18]. The length across two rings connected by a dimeric CheA is 224 Å (Fig. S8). To determine the extent of buckling that would need to occur in the two CheA:CheW rings if they ran perpendicular to the cell axis, the 224 Å rings were modeled as a chord in a circle with radius 152 Å. Using the equation $h = r - \sqrt{r^2 - L^2}$ (where h is the height of the circular segment, r is the circle radius, and L is half the chord length (224 Å /2)), the height of the circular segment is 49.2 Å (Fig. S8A). Therefore, the center of the two rings (the P3 domain) would need to buckle by an average of 49.2 Å toward the cell membrane to accommodate the cell curvature. Even with the CheA strands arranged perfectly to the cell axis, this arrangement still necessitates that each single CheA:CheW ring must bend to follow the baseplate curvature. Using the same equation above (where L is 95 Å /2), the height of the circular segment is 7.6 Å (Fig. S8B). Therefore, in *Td,* the center of a single ring must buckle toward the

membrane by an average of 7.6 Å to align to the measured baseplate curvature.

Spirochetes possess an atypical dimerization domain

The cryo-ET results reveal density corresponding to the P3 domain, which has not been previously reported in *in vivo* arrays. Sequence alignments of *Td* CheA with CheA homologs from a variety of model bacteria with previously characterized chemotaxis systems reveal that *Td* CheA possesses an additional ~50 residues located between the canonical dimerization domain (P3) helices (Figure S9A)[7]. CheA homologs from other spirochete genera including Borrelia and Brachyspira also possess additional residues in this region (Fig. S9B). Analysis of non-redundant P3 domains from all CheA classes reveal general sequence conservation in the canonical helices but highly divergent sequences at these additional residues (Fig. S10A). Furthermore, CheA-F2 proteins possess the most residues in this non-conserved region (Fig. S10B,C). X-ray crystallography experiments were used to determine the structure of the isolated *Td* P3 domain to 1.5 Å (PDB ID: 6Y1Y, Fig. 5B, Table S2). These experiments revealed that the additional residues adopt the coil-coiled motif of the classic dimerization domain with the exception of a break in one of the helices, producing a discontinuous coiled-coil (Fig. 5B). Interestingly, aromatic residues (Phe, Tyr) cluster near the helix breakages but it's unclear if this arrangement is functionally relevant (Fig. S11A). Asymmetry between the subunits is located exclusively at the additional residues following residue Tyr83, due to differing orientations of Tyr83 in the subunits (Fig. S11B,C). Fitting the new P3 domain into an all-atom chemotaxis array that was generated for previous molecular dynamics simulations (PDB ID: 3JA6) shows that these additional helices are within a ~15 Å from receptors (Fig. S11D)[21]. Additionally, the *Td* P3 domain possesses a different handedness than *Tm* P3 observed in previous crystal structures[22,23].

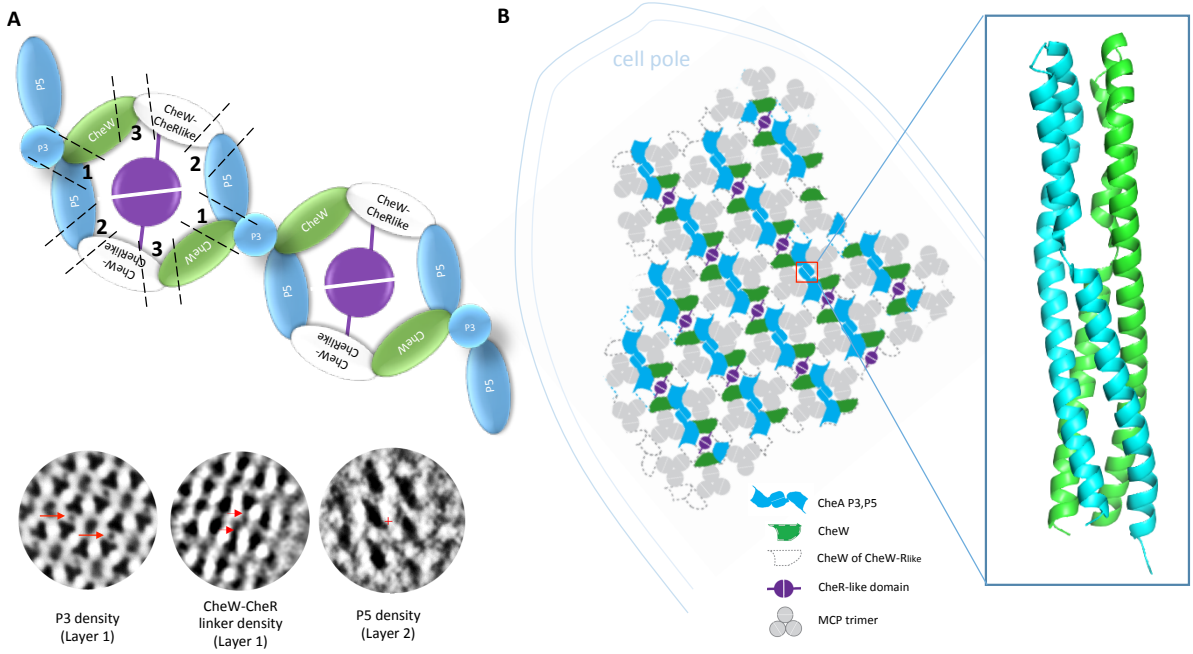


Fig. 5 The arrangement of chemotaxis proteins in *Td.* (A) Top: The CheA:CheW rings consists of three domains: CheA P5 (blue), a classical CheW protein (green), and the CheW domain of CheW-CheR$_{like}$ (white). Three unique interfaces are produced and the CheR$_{like}$ domain (purple) extends to the center of the rings to interact with one another. Bottom: The location of the CheW-CheR$_{like}$ protein in the rings is evident by examining the location of the CheA domains with respect to the CheW-CheR$_{like}$ linker. (B) The linear organization of CheA produces 'strands' of CheA:CheW rings that follow the cell axis and are linked by an atypical P3 domain.

Discussion

Here, we reveal the protein arrangement of F2 chemotaxis arrays through cryo-ET of intact *T. denticola* (*Td*) cells. In this system, three proteins comprise the rings at the receptor tips: CheA, CheW, and a CheW-CheR$_{like}$ protein. Like the *Ec* system, these proteins in *Td* are integrated into the array with strict organization[8,19]. However, a new linear arrangement of CheA is present that generates 'strands' of rings interlinked by the CheA dimerization domain (P3). The strands follow the axis of the *Td* cells, effectively allowing linked rings to align along the path of least curvature and avoid substantial buckling that would otherwise occur. Therefore, it is unlikely that a hexagonal CheA

arrangement present in canonical systems like *Ec* could form under such curvature constraints. We hypothesize that two of the $CheR_{like}$ domains in CheW-$CheR_{like}$ interact in the center of the rings for additional stability. The deleterious effect on array integrity observed in the $\Delta CheR_{like}$ strain is consistent with the expectation that $CheR_{like}$ dimerization plays a key role in array assembly and stabilization. The strict linear arrangement of CheA could be facilitated by the composition of the *Td* rings; three unique protein interfaces are present in the rings and restrict CheA P5 geometry (i.e. CheA P5 can only occupy these two positions in the six-member ring). Furthermore, the *Td* CheA P3 domain is clearly discernible in the averaged tomograms, which has not been previously observed *in vivo*[4,5,19]. As P3 has been previously implicated to directly engage receptors, the elongated P3 domain in spirochetes may have evolved to allow P3:receptor interactions in a highly curved array[21]. Due to the high cell curvature in *Td*, the receptors are expected to be more splayed compared to cells with less curvature, such as *Ec* and *Vc*. The elongated P3 may compensate for this increased distance between the receptors and the P3 domains. As *Td* cells have the smallest average diameter (0.1 – 0.4 µm) of all bacteria with determined chemotaxis architectures thus far, and the novel linear arrangement of CheA produces arrays that can better accommodate extreme cell curvature, we surmise that the array architecture in *Td* is an adaptation that evolved to produce an extended chemoreceptor apparatus in a highly curved membrane[4,5,7].

Bioinformatics analyses indicate that the unique protein features seen in *Td* are largely conserved in Spirochaetia F2 systems. Indeed, i*n vivo* genetics experiments in the *Bb* have shown that two CheW proteins, a classical CheW (CheW*1*) and a CheW-$CheR_{like}$ (CheW*3*), are essential for array formation and chemotactic behavior, and also possess variable regions at the protein interfaces[9]. *Bb* also possesses two CheA homologs (CheA*1*-F8 and CheA*2*-F2) but only one of the homologs (CheA*2*) contains an elongated P3 domain and is essential for chemotaxis and pathogenicity[24]. Therefore, we predict that a similar chemotaxis arrangement is present in *Bb*, which has a similar diameter as *Td*[9].

Unexpectedly, the placement of CheA in WT *Td* arrays could not be discerned (with the exception of the P3 domain), but was clearly visible in two *Td* mutants: one that lacks the previously characterized ODP sensor (Δ2498) and one that lacks both ODP and its cognate receptor (Δ2498Δ2496). As the density corresponding to the P3 domain in the WT strain is clearly discernible, the sparse density corresponding to all other CheA domains (P1, P2, P4, P5) is not attributed to low incorporation of CheA in these arrays. These results indicate that the kinase is highly mobile or more disordered in the WT strain, but is more constrained when ODP is deleted, suggesting that ODP directly affects array structure. However, densities in the three strains do not designate an obvious position for ODP, indicating that ODP may not be an integral component of the array, but rather peripherally interacts with the chemotaxis machinery.

In summary, we illustrate a novel chemotaxis arrangement that has evolved to compliment the spirochetes' high membrane curvature. Therefore, it is likely that the behavior and characteristics of chemoreceptor arrays in general can be influenced by perturbing the shape of the cell membrane. Importantly, previous cryo-ET chemotaxis studies have relied on artificial systems for higher resolution data, but these methods generate arrays with non-native curvature[8,17,20,21]. Collectively, these data exemplify the importance of examining biological structures in native *in vivo* conditions, as essential cellular features may not be recapitulated under *in vitro, ex vivo*, and artificial systems. While the use of model organism systems (such as *Ec, Bs, Tm)* has provided excellent insight into chemotaxis, examining the systems in non-model organisms can lead to new, unexpected advances for understanding the remarkable signaling system of bacterial chemotaxis.

Acknowledgements
This work is part of the research programme National Roadmap for Large-Scale Research Infrastructure 2017 – 2018 with project number 184.034.014, which is financed in part by the Dutch Research Council (NWO). This work was funded by a grant from the National Institutes of Health: R35GM122535 awarded to BRC, R01AI078958 and R01DE023080 to CLI, and by the European Union under a Marie-Sklodowska-Curie COFUND LEaDing fellowship to ARM. We thank the Netherlands Centre for Electron Nanoscopy (NeCEN) for access to cryo-ET data collection facilities, and NE-CAT at the Advanced Photon Source for access to X-ray crystallography data collection facilities. NE-CAT is supported by NIH/NIGMS awards P30 GM124165 and S10 RR029205.

1. Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem. Sci.* **33**, 9–19 (2008).
2. Pedetta, A., Parkinson, J. S. & Studdert, C. A. Signalling-Dependent Interactions Between the Kinase-Coupling Protein CheW and Chemoreceptors in Living Cells. **93**, 1144–1155 (2015).
3. Muok, A. R., Briegel, A. & Crane, B. R. Regulation of the chemotaxis histidine kinase CheA: A structural perspective. *Biochim. Biophys. Acta - Biomembr.* **1862**, 183030 (2019).
4. Briegel, A. *et al.* Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. *Proc. Natl. Acad. Sci.* **109**, 3766–3771 (2012).
5. Briegel, A. *et al.* Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling. *Elife* **2014**, 1–16 (2014).
6. Bhatnagar, J. *et al.* Structure of the ternary complex formed by a chemotaxis receptor signaling domain, the CheA histidine kinase, and the coupling protein CheW as determined by pulsed dipolar ESR spectroscopy. *Biochemistry* **49**, 3824–41 (2010).
7. Briegel, A. *et al.* Universal architecture of bacterial chemoreceptor arrays. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17181–17186 (2009).
8. Yang, W., Alvarado, A., Glatter, T., Ringgaard, S. & Briegel, A. Baseplate variability of *Vibrio cholerae* chemoreceptor arrays. *Proc. Natl. Acad. Sci.* **115**, 201811931 (2018).
9. Zhang, K. *et al.* Two CheW coupling proteins are essential in a chemosensory pathway of Borrelia burgdorferi. *Mol. Microbiol.* **85**, 782–794 (2012).
10. Wuichet, K. & Zhulin, I. Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. *Sci Signal* **128**, 1–27 (2010).
11. Muok, A. R. *et al.* A di-iron protein recruited as an Fe[II] and oxygen sensor for bacterial chemotaxis functions by stabilizing an iron-peroxy species. *Proc. Natl. Acad. Sci.* **116**, 14955–14960 (2019).
12. Gumerov, V. M., Ortega, D. R., Adebali, O., Ulrich, L. E. & Zhulin, I. B. MiST 3.0: an updated microbial signal transduction database with an emphasis on chemosensory systems. *Nucleic Acids Res.* **48**, D459–D464 (2020).
13. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996 (2018).
14. García-Fontana, C. *et al.* High specificity in CheR methyltransferase function: CheR2 of pseudomonas putida is essential for chemotaxis, whereas CheR1 is involved in biofilm formation. *J. Biol. Chem.* **288**, 18987–18999 (2013).
15. Greene, S. R. & Stamm, L. V. Molecular characterization of a chemotaxis operon in the oral spirochete, Treponema denticola. *Gene* **232**, 59–68 (1999).
16. Yao, W., Shi, L. & Liang, D. C. Crystal structure of scaffolding protein CheW from thermoanaerobacter tengcongensis. *Biochem. Biophys. Res. Commun.* **361**, 1027–1032 (2007).
17. Cassidy, C. K. *et al.* Structure and dynamics of the E. coli chemotaxis core signaling complex by cryo-electron tomography and molecular simulations. *Commun. Biol.* **3**, 1–10 (2020).
18. Li, X. *et al.* The 3.2 Å resolution structure of a receptor:CheA:CheW signaling complex defines overlapping binding sites and key residue interactions within bacterial chemosensory arrays. *Biochemistry* **52**, 3852–3865 (2013).
19. Yang, W. *et al.* In situ conformational changes of the escherichia coli serine chemoreceptor in different signaling states. *MBio* **10**, 1–14 (2019).
20. Burt, A. *et al.* Complete structure of the core signalling unit of the E. coli chemosensory array in an optimised minicell strain. *Nat. Commun.* 1–9 (2019). doi:10.1038/s41467-020-14350-9
21. Cassidy, C. K. *et al.* CryoEM and computer simulations reveal a novel kinase conformational switch in bacterial chemotaxis signaling. *Elife* **4**, 1–20 (2015).
22. Greenswag, A., Muok, A., Li, X. & Crane, B. Conformational Transitions that Enable Histidine Kinase Autophosphorylation and Receptor Array Integration. *J. Mol. Biol.* **427**, 87–92 (2016).
23. Bilwes, A. M., Alex, L. A., Crane, B. R. & Simon, M. I. Structure of CheA, a signal-transducing histidine kinase. *Cell* **96**, 131–141 (1999).
24. Sze, C. W., Zhang, K., Kariu, T., Pal, U. & Li, C. Borrelia burgdorferi needs chemotaxis to establish infection in mammals and to accomplish its enzootic cycle. *Infect. Immun.* **80**, 2485–2492 (2012).

Methods and Materials

Bacterial strains, culture conditions, and oligonucleotide primers

Treponema denticola (*Td*) ATCC 35405 (wild-type) was used in this study. The *Td* deletion mutants, Δ2498 and Δ2498Δ2986, were generated in a previous study[2]. Cells were grown in tryptone-yeast extract-gelatin-volatile fatty acids-serum (TYGVS) medium at 37°C in an anaerobic chamber in presence of 85% nitrogen, 10% carbon dioxide, and 5% hydrogen[1]. *Td* mutants were grown with an appropriate antibiotic for selective pressure as needed: erythromycin (50 µg/ml) and gentamycin (20 µg/ml). *Escherichia coli* 5α strain (New England Biolabs, Ipswich, MA) was used for DNA cloning. The *E. coli* strains were cultivated in lysogeny broth (LB) supplemented with appropriate concentrations of antibiotics. The oligonucleotide primers for PCR amplifications used in this study are listed in Table S3. These primers were synthesized by IDT (Integrated DNA Technologies, Coralville, IA).

Construction of a CheR truncated mutant (ΔCheR$_{like}$). *TDE1492::ermB* (Fig. S12) was constructed to replace the CheR-like domain (8781-1,308 nt) in *TDE1492* with a previously documented erythromycin B resistant cassette (*ermB*)[3]. The *TDE1492::ermB* vector was constructed by two-step PCR and DNA cloning. To construct this vector, the 5` end of *TDE1492* region and the downstream flanking region were PCR amplified with primers P_1/P_2 and P_3/P_4, respectively, and then fused together with primers P_1/P_4, generating Fragment 1. The Fragment 1 was cloned into the pMD19 T-vector (Takara Bio USA, Inc, Mountain View, CA). The *ermB* cassette was PCR amplified with primers P_5/P_6, generating Fragment 2. The Fragment 2 was cloned into the pGEM-T easy vector (Promega, Madison, WI). The Fragment 1 and 2 were digested using *Not*I and ligated, generating the *TDE1492::ermB* plasmid. The primers used here are listed in Table S3. To delete *TDE1492*, the plasmid of *TDE1492::ermB* was transformed into *Td* wild-type competent cells via heat shock, as previously described[4]. Erythromycin-resistance colonies that appeared on the plates were screened by PCR for the presence of *ermB* and absence of *TDE1492 (781-1,308 nt)* gene**.** The PCR results showed that the *TDE1492 (781-1,308 nt)* gene was replaced by *ermB* cassette as expected (Fig. S12). One positive clone (*ΔTDE1492*) was selected for further study.

Bioinformatics software and resources

The datasets used in the bioinformatics analysis were built using data from Microbial Signal Transduction Database v3 (MiST3) accessed February 2020[5] and the Genome Taxonomy Database v89 (GTDB) [6]. We built custom scripts using TypeScript-3.7.5 and NodeJS-12.13. To make these scripts, we also used packages publicly available at the node package manager repository (npm): we used RegArch-1.0.1[7] to separate CheW-CheR$_{like}$ from other CheWs, gtdb-local-0.0.12 (https://npmjs.com/package/gtdb-local) to use the GTDB taxonomy, Phylogician-TS-0.10.1-4 (https://npmjs.com/package/phylogician-ts) to visualize and manipulate the phylogenetic trees, BioSeq-TS-0.2.4 (https://npmjs.com/package/bioseq-ts) to handle protein sequences and MiST3-TS-0.7.6 (https://npmjs.com/package/mist3-ts) to access MiST3 API. Multiple sequence alignments were produced using L-INS-I algorithm from the MAFFT package. To reduce redundancy in sequence datasets we used CD-HIT v4.6[8] with unaligned sequences and Jalview[9] with aligned sequences. RAxML v8.2.10[10] was used to perform phylogenetic reconstructions, and low support branches in the phylogenetic trees were collapsed with TreeCollapseCL4[11]. Sequence logos were built using Weblogo 3.7[12].

Bioinformatics scripts and pipelines

We collected all information on the proteins classified as CheA (96,434) and CheW (134,165). To process this dataset we built several scripts and pipelines to produce the tables, figures and datasets used in this analysis (Fig. S13). The scripts are found in Supplementary File 1.

Chemosensory profile of Spirochaetotas

The “spiro-pipeline” selects all genomes from Spirochaetota phylum using gtdb-local package to access GTDB v89, then it filters only the genomes that are also present in MiST3 database. It collects the information on MiST for each genome and appends the complete taxonomy information from GTDB and signal transduction profiles. Finally, the pipeline builds the table with the information in markdown (Dataset 2).

Chemosensory profile of genomes with at least one CheA-F2

The “chea-pipeline” starts from the raw dataset taken from MiST3 database with information on 96,434 CheA genes. Based on MiST3 classification it selects the genomes with at least one CheA-F2 sequences and fetches information about these genomes. At this step, it also checks with the list

generated by the "wr-pipeline" of the genomes containing CheW-CheR$_{like}$. It proceeds to append chemosensory information for each genome and GTDB taxonomy. At this step the pipeline splits into four pathways, where one parses the information to build the Dataset 1 and the other three build FASTA formatted files with sequences from: CheA-F2, CheR-F2, CheW-CheR$_{like}$ and all CheWs belonging to genomes with at least 1 CheA-F2. We noticed that MiST3 currently misclassifies some CheR proteins, so we used RegArch to filter out false positives. We also used RegArch to separate CheW and CheW-CheR$_{like}$ sequences as MiST3 classifies both as CheW (scaffold). The RegArch definitions can be found in the script 'regArchDefinitions.ts' of the source code in Supplementary File 1.

In MiST3[12], 306 genomes contain at least 1 CheA-F2 (Dataset 1) with three exceptions: two of them belong to Acidobacteriota phylum,and one from Planctomycetota, which suggest that the presence of an F2 system in these genomes (outside of the Spirochaetota phylum) is the consequence of lateral gene transfer. Of the 306 genomes with at least one CheA-F2, only the following lack the CheW-CheR$_{like}$ protein: the three non-Spirochaetota genomes mentioned previously, 40 genomes of the Brachyspirae class, and three genomes from the order Borreliales (Fig. S1). Interestingly, the Borreliales genomes do appear to have the CheW-CheR$_{like}$ gene, except there is no gene product associated with them in the MiST3 database.

Classification of CheW

CheW proteins are not classified in MiST3. In order to make comparisons between the sequences of CheW domains in F2 systems, we must first select only CheW-F2. We first assign to the F2 class all the canonical CheW found in genomes with a single CheA-F2. Contrary to the F2 systems where the canonical CheW is not present in the chemosensory gene clusters, other classes do contain their canonical CheW within the rest of the gene cluster. CheW found within 5 genes from a classified CheA were assigned to the same class as CheA. To perform this classification, we selected the 598 full-length CheW sequences generated by the chea-pipeline. We used CD-HIT to remove 405 redundant sequences (-c 1). Next, we ran the "classify-w" pipeline on the remaining sequences (193). The pipeline reads the identifiers of the sequences and fetches the chemotaxis profile from MiST3 for each genome. It classifies CheWs as F2 classes if there is only 1 CheA of the class F2 in the profile. Next, it fetches the gene neighborhood (5 genes up and downstream) of the remaining CheWs and assigns a matching class if there is a classified CheA within these genes. We also aligned the sequences (193) with the L-INS-I algorithm of the MAFFT package, produced a phylogeny with 1000 rapid bootstrap using RAxML (-f a -m PROTGAMMAIAUTO -N 1000) and collapsed the phylogeny using TreeCollapse4 at 50% bootstrap. We mapped the CheW classification to the CheW tree in Figure S2. We expanded the F2 classification to the 74 sequences within the branch with only CheW-F2 sequences.

Comparison of the CheW domains in CheA, CheW-CheR$_{like}$ and CheW

We put together the sequences from CheA-F2 and CheW-CheR$_{like}$ (both trimmed by the PFAM model for the CheW domain, already annotated in MiST3) and the full sequence of the CheW-F2 selected in the previous step. We then use L-INS-I algorithm to align the sequences and Jalview to manually inspect and eliminate identical redundant sequences. Finally, we trimmed the whole alignment based on the boundaries of CheA-F2 and CheW-CheR$_{like}$ and eliminated one incomplete sequence: GCF_000413015.1-HMPREF1221_RS07250. The final alignment had a total of 206 sequences: 73 CheA, 59 CheW-CheR$_{like}$ and 74 CheW. We separated the alignments into individual files and built sequence logos to summarize the amino-acid diversity in each position for each group (Fig. S3A).

Comparison of CheR domains

First, we added together the trimmed part matching the CheR domain of the CheW-CheR$_{like}$ sequences and the 292 sequences of the CheR protein. We then aligned the 550 sequences using L-INS-I. We used Jalview to inspect the alignment and remove identical redundant sequences. The final alignment had the CheR domain of 83 CheR and 83 CheW-CheR$_{like}$ proteins. We separated the alignments into individual files and generated independent sequence logos (Fig. S3B).

Analyses of CheA P3 domains

The "p3-pipeline" pipeline processes the data for the analysis of the length of P3 domains of all CheA homologs in the MiST3 database. It reads the information for all 96,434 CheAs in the MiST3 database, trims the sequence matching the Pfam mode H-kinase_dim and builds FASTA formatted datasets for each chemotaxis class. For each dataset, we used CD-HIT with 75% identity cutoff and

aligned them using the L-INS-I algorithm from MAFFT. Using Jalview we manually inspect and edited the alignment to remove divergent sequences that opens major gaps in the alignment. We removed 6 F1 sequences, 37 F7 sequences, 20 F8 sequences and 5 F9 sequences. Then we merged the alignment using mafft-profile with each dataset as a seed alignment in a single shot (Fig. S10A). We selected the non-conserved central region of the alignment and measured the number of amino-acids in each sequence (Fig. S10B).

Cryo-ET and sub-tomogram averaging of *T. denticola* chemotaxis arrays

Cells were concentrated by centrifugation, and a 1/10 dilution of protein A- treated 10-nm colloidal gold solution (Cell Microscopy Core, Utrecht University, Utrecht, The Netherlands) was added to the cells and mixed by pipeting. 3 µL aliquots of the cell suspension were applied to glow-discharged R2/2 200 mesh copper Quanti-foil grids (Quantifoil Micro Tools, GmbH), the sample was pre-blotted for 30 seconds, and then blotted for 2 seconds. Grids were pre-blotted and blotted at 20 °C and at 95% humidity. The grids were plunge-frozen in liquid ethane using an automated Leica EM GP system (Leica Microsystems).

Data collection was achieved on a Titan Krios transmission electron microscope (Thermo Fisher Scientific) operating at 300 kV. Images for three strains (WT, Δ2498, Δ2498Δ2986) were recorded with a Gatan K2 Summit direct electron detector with a GIF Quantum energy filter (Gatan) operating with a slit width of 20 eV. Images were taken at a magnification of 42,000×, which corresponds to a pixel size of 3.513 Å. Tilt series were collected using SerialEM with a modified bidirectional tilt scheme (-20° to 60°, followed by -22° to -60°) with a 2° increment. Images for the $\Delta CheR_{like}$ strain were recorded with a Gatan K3 Summit direct electron detector equipped with a GIF Quantum energy filter (Gatan) operating with a slit width of 20 eV. Images were taken at a magnification of 26,000X, which corresponds to a pixel size of 3.27 Å. Tilt series were collected using SerialEM with a bidirectional dose-symmetric tilt scheme (-60° to 60°, starting from 0°) with a 2° increment. For all strains, the defocus was set to − 6µm and the cumulative exposure per cell was 100 e-/A^2.

Bead tracking-based tilt series alignment and drift correcting were done using IMOD[13]. CTFplotter was used for contrast transfer function determination and correction[14]. Tomograms were reconstructed using simultaneous iterative reconstruction with iteration number set to 6. Dynamo was used for manual particle picking and sub-tomogram averaging[15,16].

Purification of CheA, CheA P3, CheW-$CheR_{like}$, and TDE2496 proteins of *T. denticola*

DNA segments encoding the CheA P3 domain, CheW-$CheR_{like}$ protein, and TDE2496 in *T. denticola* were PCR amplified from *Td* genomic DNA using a forward oligonucleotide encoding an NdeI restriction site and a reverse primer encoding an EcoRI restriction site. The CheA protein was amplified using a forward oligonucleotide encoding an NdeI restriction site and a reverse primer encoding an BamHI restriction site. The PCR products were treated with the appropriate restriction enzymes, purified, and ligated into a pet28a plasmid with a poly-Histidine tag and kanamycin resistance marker. The plasmids were transformed into *Escherichia coli* BL21-DE3 cells and 4-8 L of cell culture were grown at 37°C until an O.D. of 0.6 was reached. The flasks were cooled to 21°C and 1 mM of IPTG was added to the culture. The cells were harvested after 16 hours of growth. The cells were lysed via sonication in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM Imidazole) while cooled on ice. The lysate was centrifuged at 20,000 X G for 1 hour at 4°C. The lysate was then run over a gravity-flow purification column containing 3 ml of Nickel-NTA resin. The resin was washed with 10 ml wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 20 mM Imidazole) and the protein was eluted with 10 ml elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 200 mM Imidazole) and collected in 1 ml fractions. The fractions were assessed for protein concentration via Bradford reagent and the fractions containing protein were run on a size-exclusion s75 and s200 column systems that monitored absorbance at 280 nm and collected 6 ml fractions. Fractions that contain CheA were concentrated to ~20 mg/ml via centrifugation in a protein concentrator containing a regenerated cellulose filter with a 50 kDa molecular-weight cut-off (MWCO). Fractions that contain CheA P3, CheW-$CheR_{like}$ and TDE2496 were concentrated with a 10 kDa MWCO filter to 32 mg/ml, 11 mg/ml and 7 mg/ml, respectively. The protein solutions were aliquoted, flash frozen in liquid nitrogen, and stored at -80°C. For CheW-$CheR_{like}$, CheA, and CheA P3 domain, the purifications were prepared at ambient temperatures. For TDE2496, the purification was prepared at 4°C.

Radioisotope assays

23 µl samples containing 2 µM CheA alone, or 2 µM CheA and 2 µM CheW-$CheR_{like}$, or 2 µM CheA and 2 µM CheW-$CheR_{like}$ with TDE2496 (2 µM or 12 µM), were incubated in 50 mM MOPS pH 7.5, 150 mM KCl, 10 mM $MgCl_2$ for 30 minutes at ambient temperatures. Phosphorylation of CheA was

initiated by the addition of 2 µl of a solution containing 1 mM ATP mixed and 2-10 µl radiolabeled γ-P32 ATP (3000 Ci/mmol, 10 µCi/µL; Perkin Elmer) and quenched with 25 µl of 3X LDS buffer containing 100 mM EDTA pH 8.0 after 1-12 minutes. The samples were run on a native Tris-glycine gel for 2 hours at 120 volts. The gels were dried, placed in a radiocassette for 24 hours, and then imaged with a Typhoon phosphor-imager (GE Healthcare). The intensity of the radio-labeled protein bands were quantified using ImageJ.

Quantification of cell curvature

The cell curvature of *Td* whole cells and *V. cholera* minicells was quantified by analyzing images of cross-sections of the cells where top views of chemotaxis arrays are present. For *Td* cells, the inner membrane curvature and CheA:CheW base-plate curvature was quantified. For *Vc* minicells, the inner membrane curvature was quantified. These images were pre-processed with Fiji by placing points along the desired area with a distance of 10 nm between each point. The curvature of the inner membrane was measured with a pre-built plugin for python, called Sabl_mpl, written by Jewett, A. from the Jensen lab (Pasadena, CA)[23]. The "measure 3-point curvature" function was used to select three adjacent points along the inner membrane of the cell and calculate the radius of these points. The radius of the three selected points allowed for the calculation of the local curvature by dividing 1 with the measured radius (1/R = c). This was repeated for all points with a "sliding window" approach, where the second point of the initial three points would become the first point, until the desired area was covered.

Quantification of array alignment to the cell axis

The angle between the strands of CheA:CheW rings and the *Td* cell axis was quantified using Image J software. 2D images from reconstructions that clearly locate the orientation of the strands and cell axis were chosen for analysis. First, a straight line was drawn from the cell pole down the axis of the cells. Then, a second line was drawn through one of the strands in the array and the angle between the two intersecting lines was quantified. In some cases, the angle was too small (<3°) to be accurately determine so the angle was annotated as 0°.

Residue conservation and molecular modeling of *Td* CheW, CheA P5, and the CheW domain of CheW-CheR$_{like}$

The protein sequences of the two *Td* CheW domains and *Td* P5 were aligned (Clustal Omega), conservation was calculated based on the alignment (JalView[9]), and the highest variable regions were selected based on conservation (10+ adjacent residues with conservation scores lower than 8). Homology models of *Td* CheA P5, CheW and the CheW domain of CheW-CheR$_{like}$ were generated via the Swiss-Model server using complete residue sequences of each protein as a target[24]. The CheW protein from *Thermoanaerobacter tengcongensis* (*Tt*) (PDB ID: 2QDL) had the highest percent identity to the CheW proteins (37% and 31%) and was therefore used as the structural template. The resulting homology models for *Td* CheW and the CheW domain of CheW-CheR$_{like}$ had a QMEAN of -1.3 and -1.2, respectively. The P5 structure from *Escherichia coli* produced the best homology model for *Td* CheA P5, with QMEAN -0.98 (PDB ID: 6S1K). The homology models were then aligned to the CheW protein and P5 domain in a crystal structure containing *Thermotoga maritima* CheW and CheA P4P5 in complex using PyMol (PDB ID: 3UR1).

Crystallization and structural determination of the P3 domain of *T. denticola* CheA

The isolated P3 domain was concentrated to 32 mg/ml and crystallized via hanging drop in 0.1 M Imidazole pH 7.0, 25% PEG 400 using a 1:1 ratio of protein solution to crystallization solution with a final volume of 3 µl. Crystals were apparent within eight hours but increased in size over three days. Crystals were manually picked up in loops, flash cooled and shipped in liquid nitrogen to a beamline (APS, line NE-CAT 24-ID-C, Dectris Pilatus 6M-F Pixel Array detector). The crystals diffracted to ~1.3 Å with a C2 symmetry and data was cut-off at 1.5 Å. The diffraction data was scaled and integrated using XDS[17], and phased by molecular replacement with Phaser MR using *ab initio* search models generated through the QUARK server and then ran through the AMPLE pipeline on the CCP4 web server[18–20]. Model improvement was done by several rounds of manual model improvement in COOT followed by automated refinement using Phenix Refine software[21,22].

1. Ohta, K., Makinen, K. K. & Loesche, W. J. Purification and characterization of an enzyme produced by Treponema denticola capable of hydrolyzing synthetic trypsin substrates. *Infect. Immun.* **53**, 213–20 (1986).
2. Muok, A. R. *et al.* A di-iron protein recruited as an Fe[II] and oxygen sensor for bacterial chemotaxis functions by stabilizing an iron-peroxy species. *Proc. Natl. Acad. Sci.* **116**, 14955–14960 (2019).
3. Goetting-Minesky, M. P. & Fenno, J. C. A simplified erythromycin resistance cassette for Treponema denticola mutagenesis. *J. Microbiol. Methods* **83**, 66–68 (2010).
4. Kurniyati, K. & Li, C. pyrF as a counterselectable marker for unmarked genetic manipulations in Treponema denticola. *Appl. Environ. Microbiol.* **82**, 1346–1352 (2016).
5. Gumerov, V. M., Ortega, D. R., Adebali, O., Ulrich, L. E. & Zhulin, I. B. MiST 3.0: an updated microbial signal transduction database with an emphasis on chemosensory systems. *Nucleic Acids Res.* **48**, D459–D464 (2020).
6. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996 (2018).
7. Ortega, D. R. & Jensen, G. J. Regular Architecture (RegArch): A standard expression language for describing protein architectures. *bioRxiv* 679910 (2019). doi:10.1101/679910
8. Li, W., Jaroszewski, L. & Godzik, A. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**, 282–283 (2004).
9. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
10. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
11. Hodcroft, E. TreeCollapserCL4: Collapse Trees by Bootstrap. Available at: http://emmahodcroft.com/TreeCollapseCL3.html.
12. Crooks, G., Hon, G., Chandonia, J. & Brenner, S. WebLogo: A Sequence Logo Generator. *Genome Res* **14**, 1188–1190 (2004).
13. Mastronarde, D. N. Dual-axis tomography: An approach with alignment methods that preserve resolution. *J. Struct. Biol.* **120**, 343–352 (1997).
14. Xiong, Q., Morphew, M. K., Schwartz, C. L., Hoenger, A. H. & Mastronarde, D. N. CTF determination and correction for low dose tomographic tilt series. *J. Struct. Biol.* **168**, 378–387 (2009).
15. Castaño-Díez, D., Kudryashev, M., Arheit, M. & Stahlberg, H. Dynamo: A flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *J. Struct. Biol.* **178**, 139–151 (2012).
16. Castaño-Díez, D., Kudryashev, M. & Stahlberg, H. Dynamo Catalogue: Geometrical tools and data management for particle picking in subtomogram averaging of cryo-electron tomograms. *J. Struct. Biol.* **197**, 135–144 (2017).
17. Kabsch, W. *et al. XDS*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
18. Bibby, J., Keegan, R. M., Mayans, O., Winn, M. D. & Rigden, D. J. AMPLE: A cluster-and-truncate approach to solve the crystal structures of small proteins using rapidly computed ab initio models. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 1622–1631 (2012).
19. Keegan, R. M. *et al.* Exploring the speed and performance of molecular replacement with AMPLE using QUARK ab initio protein models. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 338–343 (2015).
20. Thomas, J. M. H. *et al.* Routine phasing of coiled-coil protein crystal structures with AMPLE. *IUCrJ* **2**, 198–206 (2015).
21. Adams, P. *et al.* The Phenix Software for Automated Determination of Macromolecular Structures. *Methods* **292**, 342–351 (2009).
22. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
23. Yao, Q. *et al.* Short FtsZ filaments can drive asymmetric cell envelope constriction at the onset of bacterial cytokinesis. *EMBO J.* **36**, 1577–1589 (2017).
24. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).

Supplemental figures

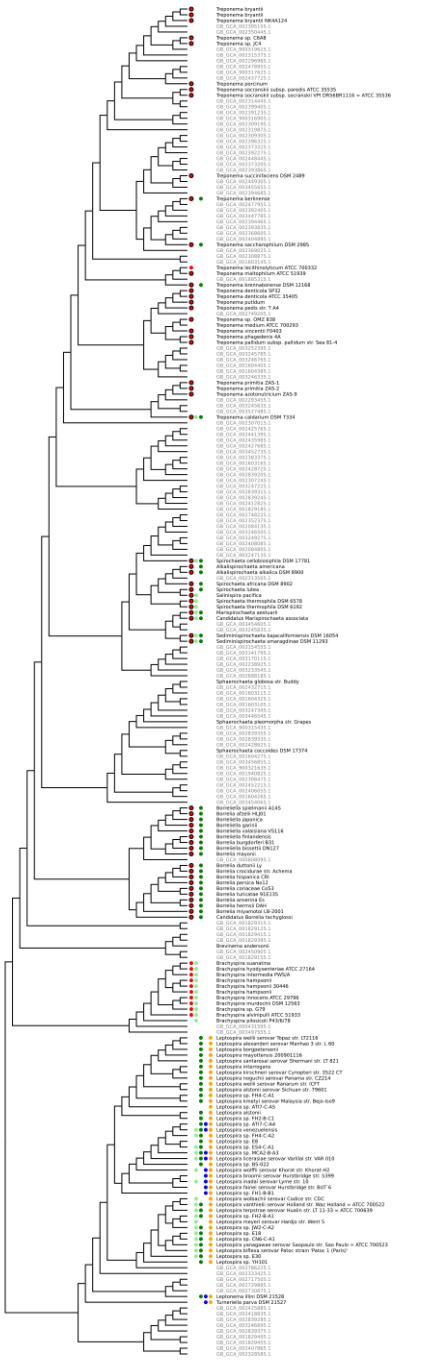


Figure S1 The chemosensory class profile of Spirochaetota. Sequences belonging to the four major classes of Spirochaetota are marked: UBA4802 (dark red), Leptospira (purple), Brachyspirae (cyan), UBA6919 (blue), Brevinematia (yellow), GWE2-31-10 (green) and Spirochaetia (light green). Nodes are marked for presence of classes F1 (orange), F2 (red), F5 (blue), F7 (light green), F8 (green), ACF (light blue), and Tfp (purple). Genomes not present in MiST3 database are marked as grey. F2 systems with CheW-CheR$_{like}$ are marked with a black outline around the red circle.

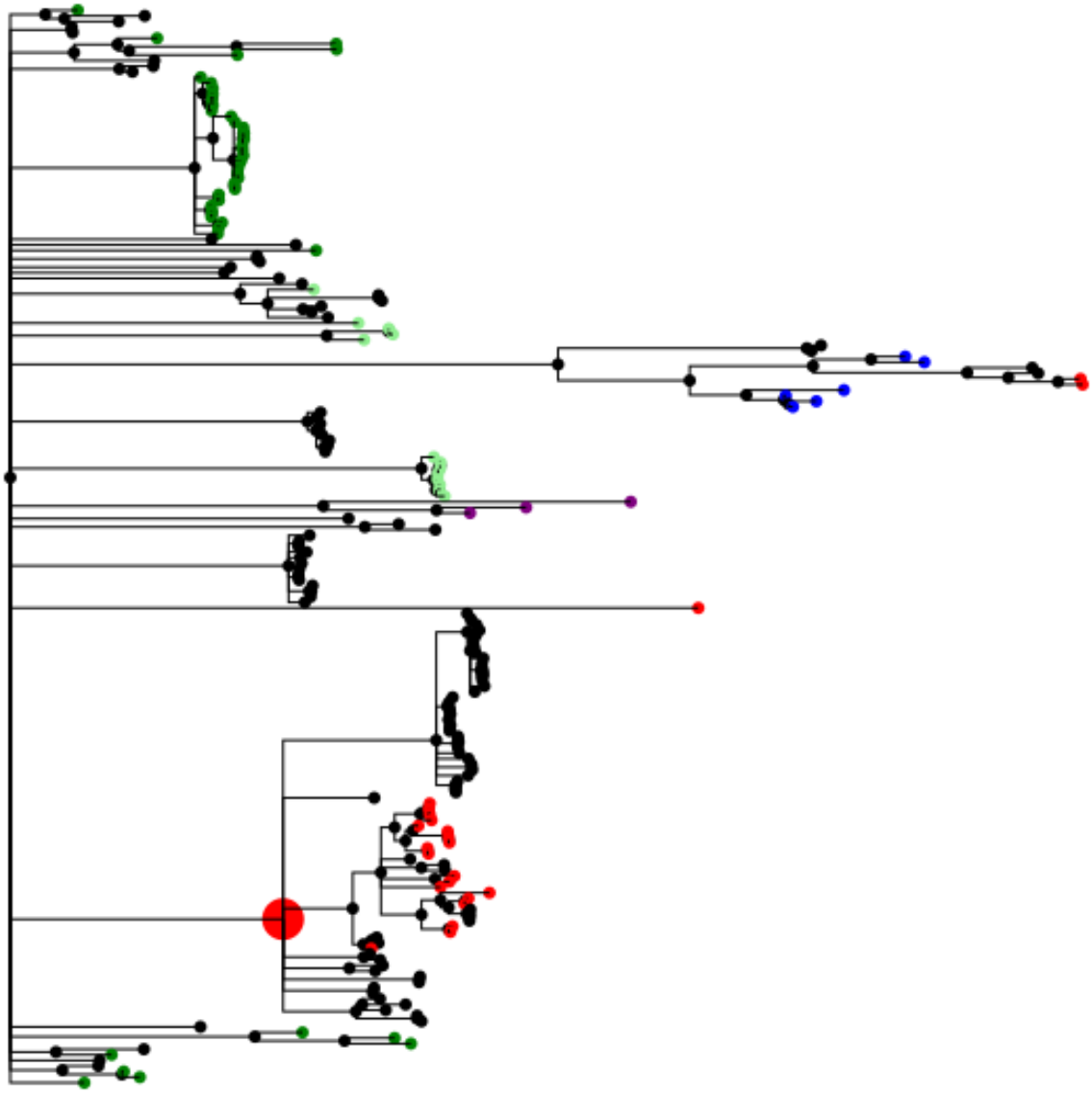


Fig. S2 Phylogenetic tree of CheW sequences in genomes with at least one CheA-F2 suggests a last common ancestor of CheW-F2 sequences. We mapped the sequences of CheW from the classes F2 (red), F5(purple), F7 (light green), F8 (green) and ACF (blue). The larger red internal node marks a candidate of the last common ancestor of CheW-F2.

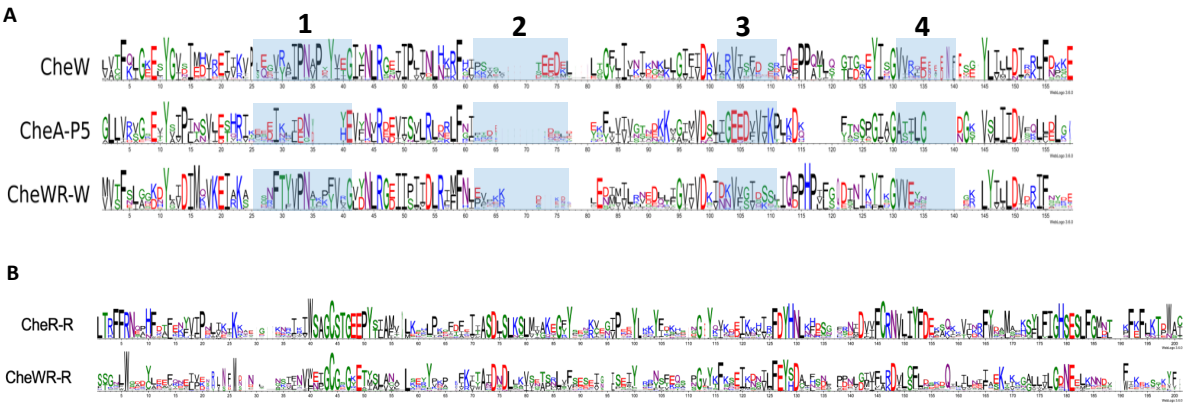
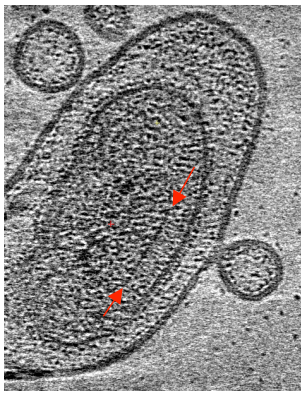


Fig. S3 (A) Sequence logo of representative sequences of the three groups of CheW-containing proteins in the F2 system. Blue boxes denote the location of variable regions identified in the *Td* CheW and P5 homologs. In these locations, all regions possess unique conserved residues with the exception of Region 2, which is not at the CheA:CheW ring interfaces. (B) Sequence logo of representative sequences of the two groups of CheR-containing proteins in the F2 system.

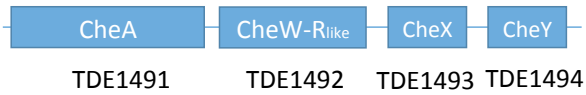
A

Td strain	No. Cells/ Arrays	No. Particles total
WT	11	728
Δ2498	10	894
Δ2498Δ2496	10	554
$\Delta CheR_{like}$	5	194

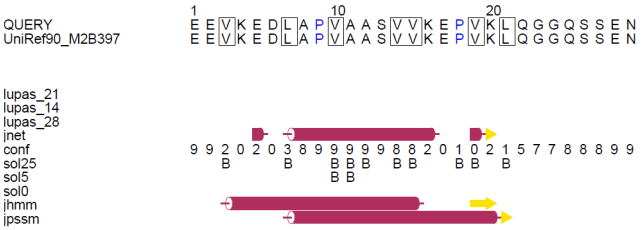
B



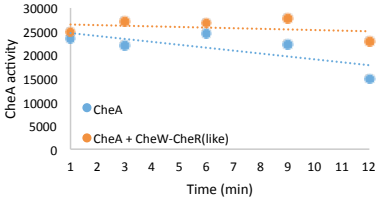
C



D



E



F

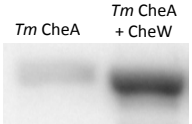


Fig. S4 (A) The number of arrays and particles chosen to generate sub-tomogram averages for each *Td* strain. (B) Lysed *Td* cells still possess chemotaxis arrays. (C) The CheW-$CheR_{like}$ protein is conserved on the same operon as the only CheA, CheX, and CheY proteins in *Td*. (D) The CheW-$CheR_{like}$ linker is predicted to forma single alpha helix flanked by unordered regions (Jpred). (E) Radioisotope assays that monitor CheA autophosphorylation over time in the presence and absence of CheW-$CheR_{like}$ demonstrate a difference in CheA activity. (F) In the *T. maritima* system, the presence of CheW also activates CheA kinase activity. Both samples are at time point 3 minutes.

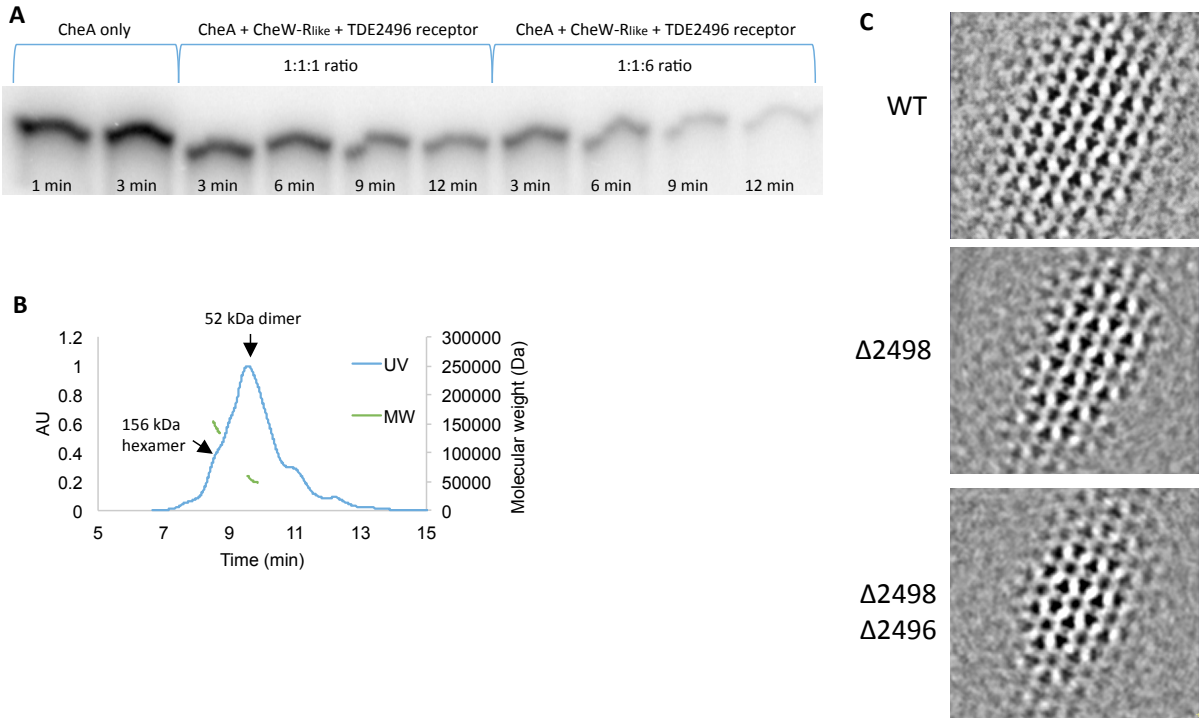


Fig. S5 (A) The soluble receptor TDE2496 deactivates *Td* CheA in a concentration-dependent manner. (B) MALS experiments demonstrate that TDE2496 forms the expected dimer (52 kDa) and hexamer (156 kDa) oligomeric states. (C) Cryo-ET averages of particles from three different *Td* strains.

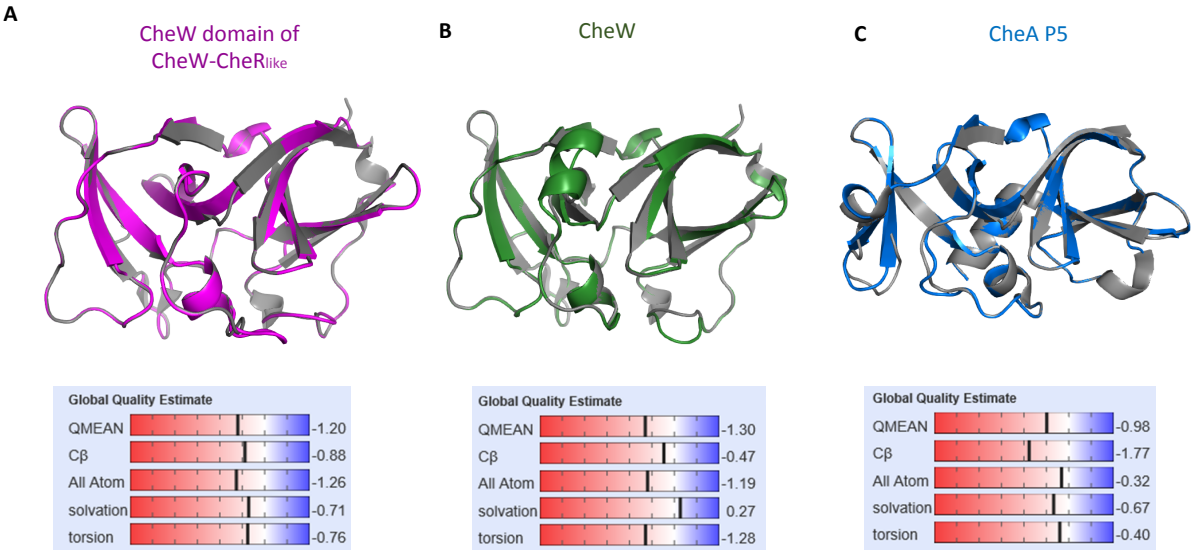
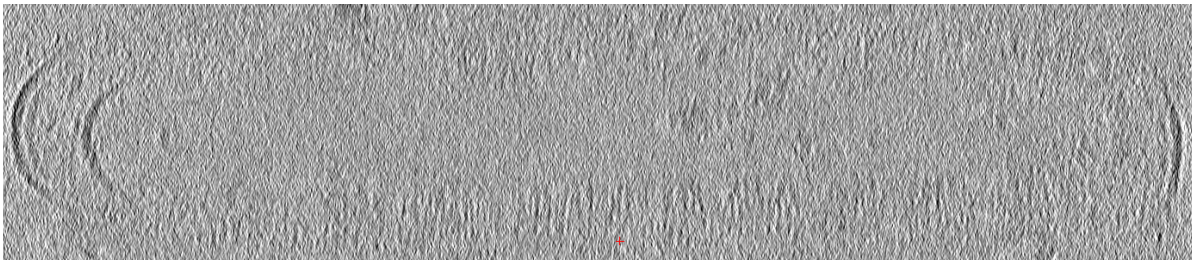


Fig. S6 Homology models of *Td* CheW domains and CheA P5. (A) A homology model of the CheW domain of CheW-CheRlike using *Thermoanaerobacter tengcongensis* CheW (PDB ID: 2QDL) as the template. (B) A homology model of the classical *Td* CheW using *Thermoanaerobacter tengcongensis* CheW (PDB ID: 2QDL) as the template. (C) A homology model of CheA P5 using *E. coli* CheA P5 (PDB ID: 6S1K) as the template.

A



B

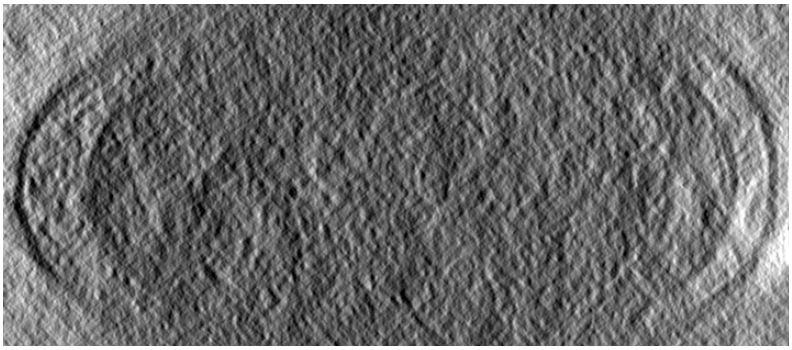


Fig. S7 Curvature of *E. coli* lysed cells and *V. cholerae* minicells used for structural determination of chemotaxis arrays in previous experiments. (A) A 2D image of a lysed *E. coli* cell at a cross-section where array top views are present. The lysed cells are flat. (B) *V. cholerae* minicells have an inner membrane curvature of 9.15 +/- 4.5 /µm (radius 1092 Å).

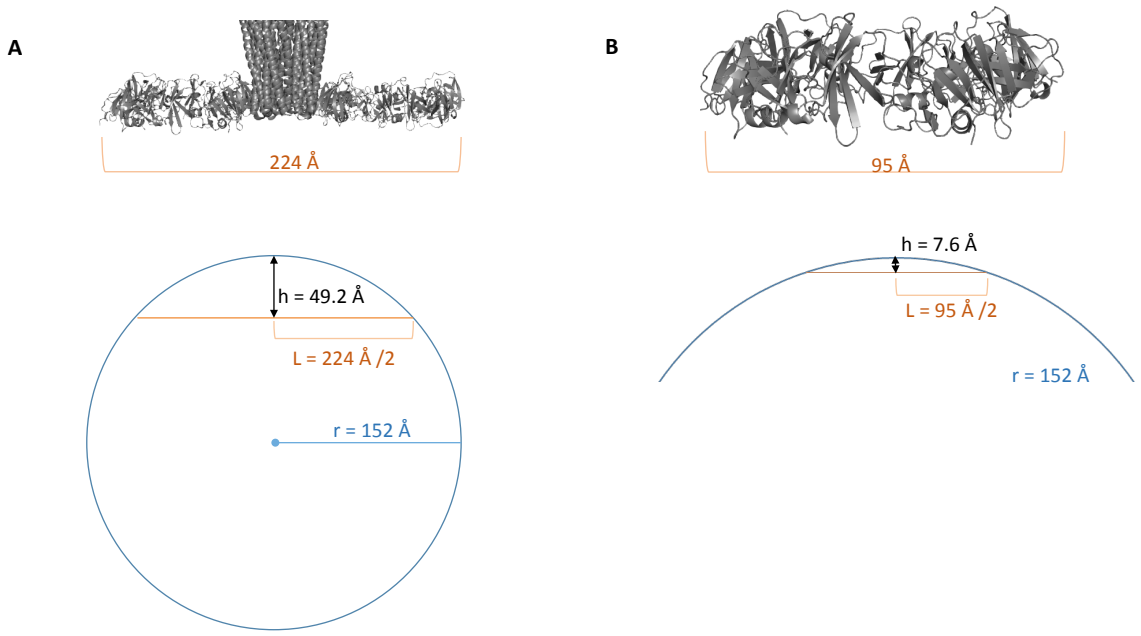


Fig. S8 Modeling of CheA:CheW rings to the curvature of the *Td* baseplate. (A) In order for two linked CheA:CheW rings to run perpendicular to the cell axis, the center of the ring must buckle by 49.2 Å toward the cell membrane. (B) For single CheA:CheW rings to follow the baseplate curvature, it must buckle by an average of 7.6 Å.

A

```
T.denticola      LRVDANRIDYLLNLVSETVITKASLNQSTIEFAELYDKFQNSSTIYKDKTRRLLDKMPEY   60
V.cholerae       LKVDQEKVDLLMDLVGELIVAKNSMQYLAYRAENEFGVR---------------------   39
C.jejuni         IRVEVKRLDHLMNLIGELVLGKNRLLKIYDDVEERYEGE---------------------   39
E.coli           IRVAVEKVDQLINLVGELVITQSMLAQRSSELDPVN------------------------   36
S.enterica       IRVAVEKVDQLINLVGELVITQSMLAQRSNELDPVN------------------------   36
T.maritima       VRVDIEKLGNLMDLMGELVIARSRILETLKK---YN------------------------   33
B.subtilis       IRVNIDRLDSLMNLFEELVIDRGRLEQIAKE---LE------------------------   33
                 ::*  .::. *::*. * :: :  :

T.denticola      LEKIQQGYDINSIKQDVLNEYSSLLEVFGDFDSLMKAAVTKFKSSSQNLGRISGELQEGV   120
V.cholerae       -------------------------------------KLAQDIKAEQSVISRLAEDLQSVV   63
C.jejuni         -------------------------------------KFLEELNQVVSQLSIITTDVQLAV   63
E.coli           ---------------------------------------HGDLITSMGQLQRNARDLQESV   58
S.enterica       ---------------------------------------HGDLITSMGQLQRNARDLQESV   58
T.maritima       ----------------------------------------IKELDESLSHLSRITLDLQNVV   55
B.subtilis       ----------------------------------------HNELTETVERMTRISGDLQSII   55
                                                            .:      :   : ::*  :

T.denticola      MKIRMVP        127
V.cholerae       MQVRMVP        70
C.jejuni         MKTRMQP        70
E.coli           MSIRMMP        65
S.enterica       MSIRMMP        65
T.maritima       MKIRMVP        62
B.subtilis       LNMRMVP        62
                 :. ** *
```

B

```
B.burgdorferi      NLIRIDSKKIDYILNLVSEAVISKSSYNQINSEMITLFYNFNYFYDYQESFQRNFLIDLK 60
B.pilosicoli       SYLRVESDRIDAMMNQVGELVTNKSSYVQYDDDLSSYGKLISNSI-----------NDIK 49
T.denticola        --LRVDANRIDYLLNLVSETVITKASLNQSTIEFAELYDKFQNSS--------------- 43
T.pallidum         SVLRVDAKRIDYLLNLVSETVIIKASLNQSALEFGEVYTLFQNAN--------------- 45
                     :*:::.:** ::* *.* *  *:*  *   ::      :.

B.burgdorferi      IVFKDAGLTLEDEIESH------------INSLMSFKMEKALKDISELRNSFFRLLQNFK 108
B.pilosicoli       KYYRDTIVQILRKFEDY----AQKKEIKDLRANYINSFNSKLNEISKIEDDFKYTLDRYR 105
T.denticola        TIYKDKTRRLLDKMPEYLEKIQQGYDINSIKQDVLNEYSSLLEVFGDFDSLMKAAVTKFK 103
T.pallidum         GAYKERLRKFFDRVPAYLEKVKNGQDADAVRKEMIAEAVGVFDIFSSFENGLKQSVTKFR 105
                     :::    :  ..  :            :.     .    :. :..: . :   : .::

B.burgdorferi      MTSGRLSRIITDLHESVLKTRMLP     132
B.pilosicoli       SSYQLLTRVTNELQETVMKIRMVP     129
T.denticola        SSSQNLGRISGELQEGVMKIRMVP     127
T.pallidum         SSAQNLGRISGELQEGVMKIRMVP     129
                    :   * *:  :*:* *:* **:*
```

Fig. S9 Multiple sequence alignments of the CheA P3 domain from several bacteria demonstrates the presence of additional P3 residues in *Td* and other Spirochetes. (A) CheA P3 alignments of *Td* and P3 from other bacteria with previously characterized chemotaxis proteins. *Td* possesses ~50 residues that are not found in the other homologs and are located in between the traditional dimerization helices. (B) *Td* CheA P3 aligned with other Spirochete P3. The additional residues identified in alignment **A** are highlighted in blue. Figures were made using Clustal Omega.

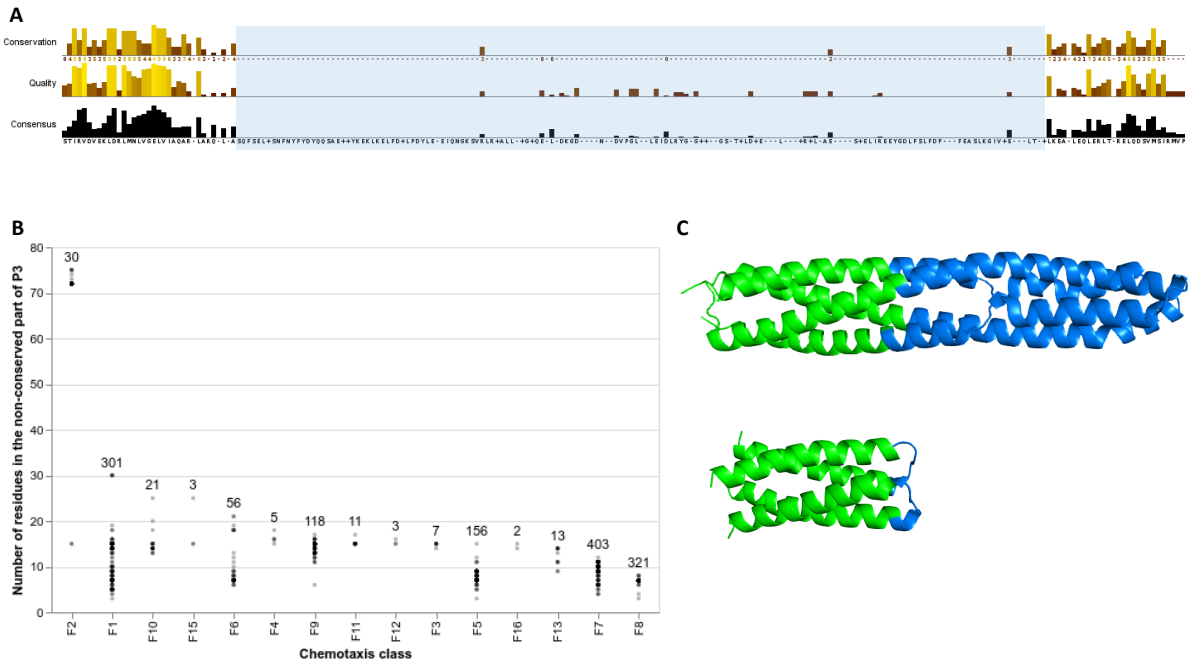


Fig. S10 Analysis of non-redundant CheA P3 domains with a 75% sequence identity cut-off (1450 sequences). (A) Conservation scores of the P3 alignment with the 1450 sequences. The traditional P3 helices are largely conserved, but regions located between the helices are non-conserved. (B) Analyses of CheA from different chemotaxis classes reveals that CheA F2 homologs possess the most residues in the non-conserved region of P3. (C) The non-conserved regions are llustrated with the two known CheA P3 structures. Top: *Td* P3 possesses 71 residues that align to the non-conserved region (PDB ID: 6Y1Y). Bottom: *Tm* P3 possesses seven residues that align to the non-conserved region (PDB ID: 1B3Q).

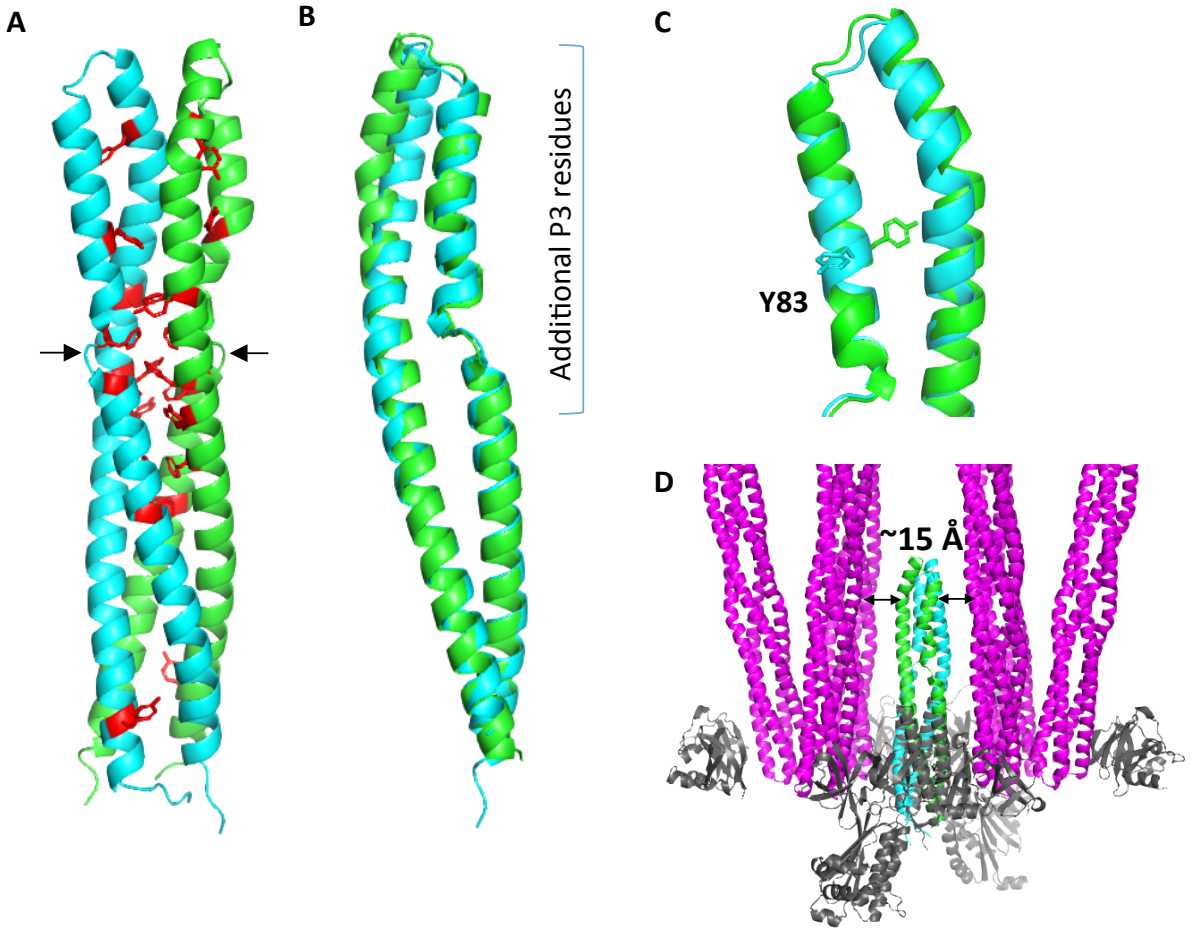


Fig. S11 The P3 domain of *Td* CheA. (A) The CheA P3 dimer contains a cluster of Phe and Tyr residues near the breakages in the helices (black arrows). All Phe and Tyr residues are highlighted in red. (B) The crystal structure of CheA P3 demonstrates asymmetry in the subunits. (C) Repositioning of Y82 in the subunits induces alterations of adjacent residues and may account for subunit asymmetry. (D) Alignment of the P3 structure to a previously determined model of the chemotaxis array in *E. coli* (PDB ID: 3JA6) indicates that the P3 domain lies within ~15 Å of the receptors (when measuring from peptide back-bone). CheA and CheW in this model are shown in grey.

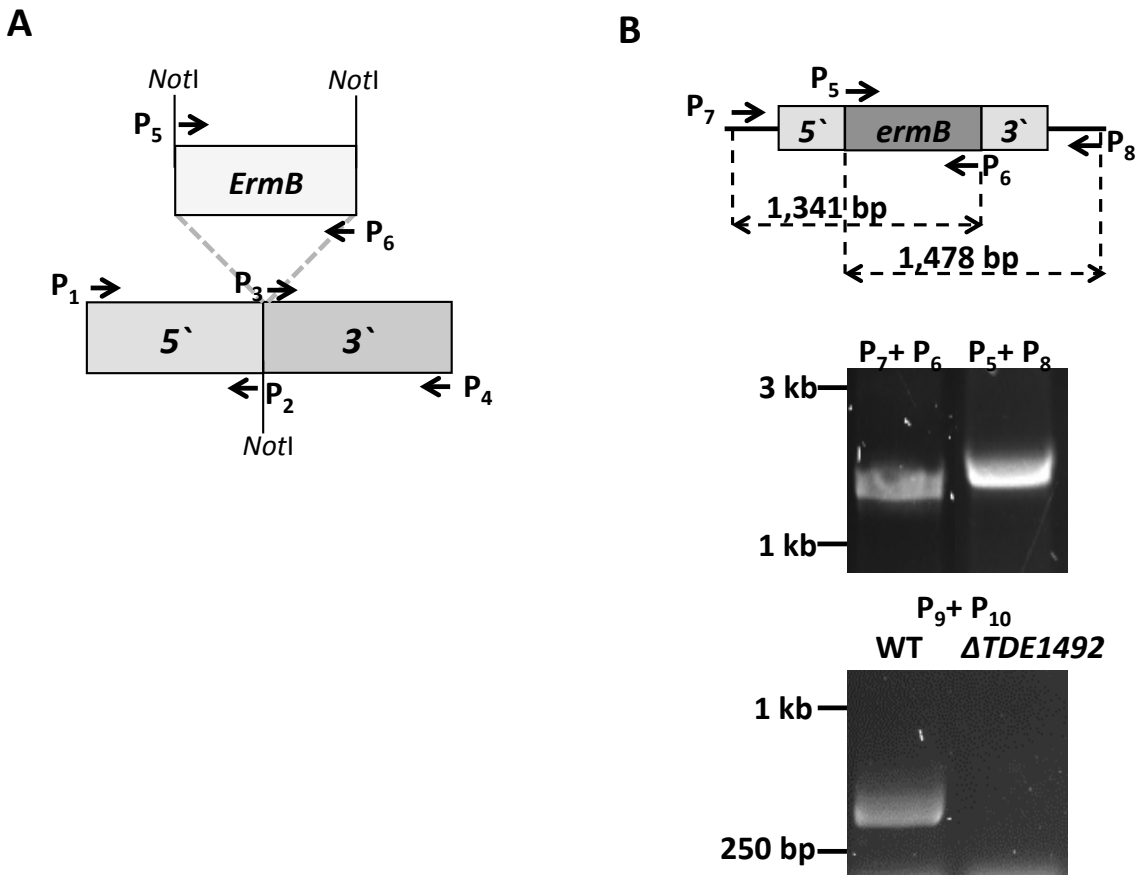


Fig. 12 Diagrams illustrating construction of the *TDE1492::ermB* vector (A) for the targeted mutagenesis of *TDE1492 (781-1,308 nt)* by in-frame replacement of *TDE1492* using *ermB* cassette**.** These constructs were constructed by two-step PCR followed by DNA cloning. Arrows represent the relative positions and orientations of these primers, which are listed in Table S3. *ermB* = erythromycin resistance. (B) Characterization of the *ΔTDE1492* strain by PCR analysis. The top panel illustrates how the PCR analysis is designed; the bottom panel is the PCR results. Arrows represent the relative positions and orientations of these primers; the numbers are predicted sizes of PCR products generated by the corresponding primers. The primer P_7 is located at the 5'-end of *TDE1492*, P_6 at the 3'-end of *ermB*, P_5 at the 5'-end of *ermB*, P_8 at the flanking region of *TDE1492*, P_9 at the middle of *TDE1492*, and P_{10} at the 3'- end of *TDE1492*. The sequences of these primers are listed in Table S3.

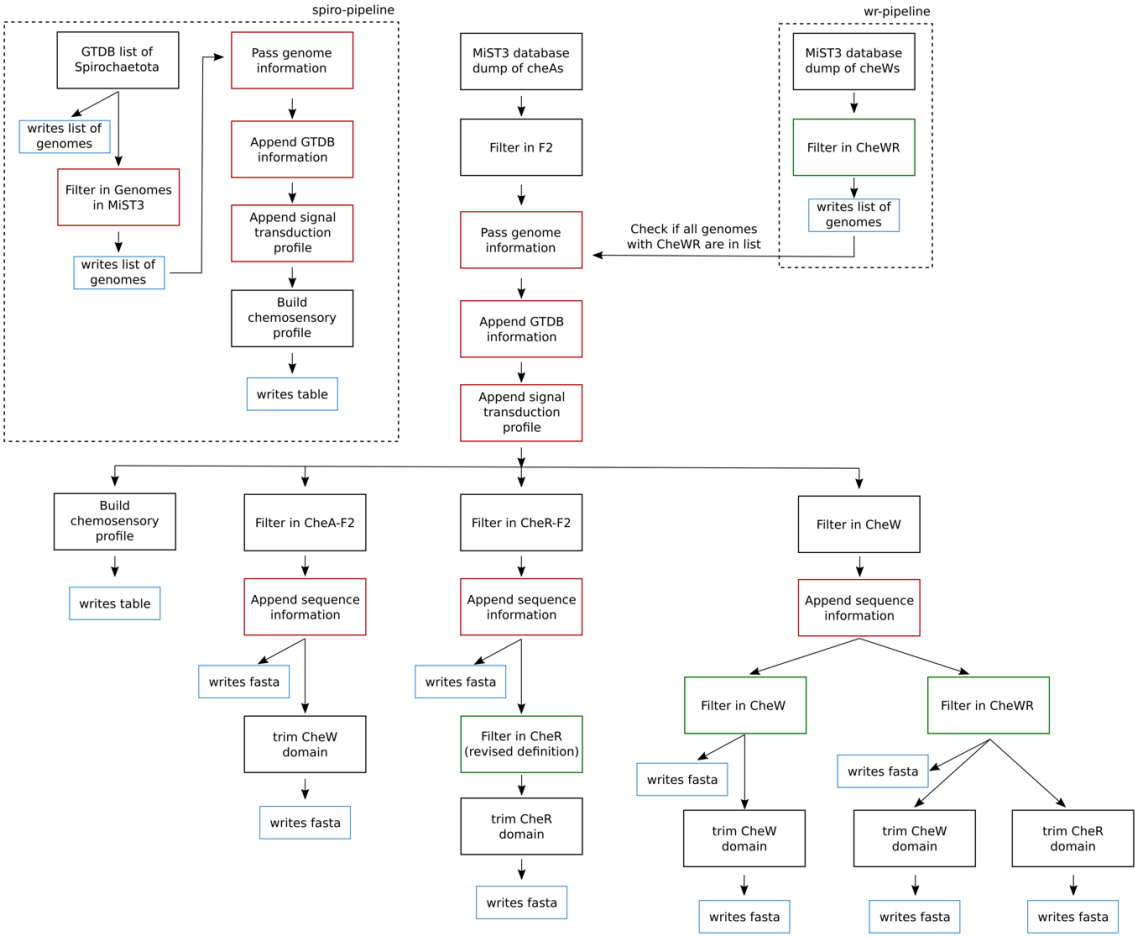


Fig. S13 Flowchart of the three major pipelines used to produce the bioinformatics datasets. Steps marked in red represents fetching information from MiST3 database, in green are steps requiring RegArch as a filter, and in blue indicate writing data to file.

Td cell number	Inner membrane curvature (/µm)	Baseplate curvature (/µm)
1	35.0	49.7
2	25.4	54.0
3	34.2	63.7
4	43.1	96.1
5	30.8	99.7
6	41.1	55.7
7	39.6	74.6
8	33.1	47.7
9	46.2	47.8
10	29.1	67.1
Average:	**35.8 +/- 6.6**	**65.6 +/- 19**

Table S1A. Statistics of inner membrane and baseplate curvature for 10 WT *Td* cells.

Vc cell number	Inner membrane curvature (/µm)
1	6.2
2	15.4
3	14.2
4	7.4
5	4.3
6	7.5
Average:	**9.155 +/- 4.53**

Table S1B. Statistics of inner membrane curvature for 6 *Vc* mini-cells.

Td strain	Average angle (°)
WT (n=9)	9.67 +/- 9.21
Δ2498 (n=9)	9.55 +/- 8.75
Δ2498Δ2496 (n=8)	12.18 +/- 7.62
Combined average:	**10.40 +/- 8.59**

Table S1C. Statistics for the average angle between the *Td* cell axis and the 'strands' of CheA:CheW rings.

Table S1

	***Td* CheA P3 domain** (PDB ID: 6Y1Y)
Wavelength	0.979100
Resolution range	75.36 – 1.50 (1.55 – 1.50)
Space group	C 1 2 1
Unit cell	a, b, c: 101.64, 28.77, 82.90 α, β, γ: 90, 114.75, 90
Total reflections	53542
Unique reflections	35178 (3490)
Multiplicity	6.9 (6.2)
Completeness (%)	96.72 (99.31)
Mean I/sigma(I)	14.7 (5.4)
R-merge	0.063(0.290)
R-meas	0.075(0.350)
R-pim	0.040(0.193)
CC1/2	0.99(0.94)
Reflections for refinement	33864 (3361)
Reflections for R-free	1308 (143)
R-work	0.187 (0.205)
R-free	0.211 (0.244)
No. non-H atoms	2344
macromolecules	2028
ligands	0
solvent	316
Protein residues	253
RMS(bonds)	0.006
RMS(angles)	0.93
Ramachandran favored (%)	100
Ramachandran allowed (%)	0.00
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00
Clashscore	6.87
Average B-factor	19.1
macromolecules	17.17
solvent	28.20

Table S2. Data collection and refinement statistics for *Td* CheA P3 domain

Primers	Sequences (5'-3')	Note [a]
P_1	CGGGCGTAGGCATCGGAGATAC	5` portion for TDE1492 inactivation; [F]
P_2	TTACAATATCACCTGAGCTGGCGGCCGCTTAAGGTGACA AGAAAGATGATA	5` portion for TDE1492 inactivation; [R]
P_3	TATCATCTTTCTTGTCACCTTAAGCGGCCGCCAGCTCAGG TGATATTGTAA	3` portion for TDE1492 inactivation; [F]
P_4	CCCAGAGCACTTATCATAAC	3` portion for TDE1492 inactivation; [R]
P_5	ATGAACAAAAATATAAAATATTCTC	Erythromycin B cassette (ermB); [F]
P_6	TTATTTCCTCCCGTTAAATAATAG	Erythromycin B cassette (ermB); [R]
P_7	ATGGAAGAAATGAAAGAAC	5` flanking region of TDE1492, ΔTDE1492 PCR analysis; [F]
P_8	GATATAGTTCTTGGCTCCAAG	3` flanking region of TDE1492, ΔTDE1492 PCR analysis; [R]
P_9	GAGGCCTATATAAATGCC	TDE1492, ΔTDE1492 PCR analysis; [F]
P_{10}	CTTGGGCATTGCCTCATTATG	TDE1492, ΔTDE1492 PCR analysis; [R]

Table S3. Oligonucleotide primers used in this study

[a] Underlined sequences are engineered restriction cut sites for DNA cloning; [F] forward; [R] reverse.