1   **Atypical chemoreceptor arrays accommodate high membrane curvature**

2

3   **Alise R. Muok[1], Davi R. Ortega[2,3], Kurni Kurniyati[4], Wen Yang[1], Adam Sidi Mabrouk[1],**

4   **Chunhao Li[4], Brian R. Crane[5], Ariane Briegel[1*]**

5   [1]Institute for Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden

6   [2]Department of Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA

7   [3]Laboratory for Research in Complex Systems, 425 California St., Suite 1200, San Francisco, CA 94104, USA

8   [4]Department of Oral and Craniofacial Molecular Biology, Philips Research Institute for Oral Health, Virginia

9   Commonwealth University, Richmond, VA 23298

10   [5]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14850

11

12   **Abstract/Summary**

13   The prokaryotic chemotaxis system is arguably the best-understood signaling pathway in

14   biology, but most insights have been obtained from only a few model organisms and many

15   studies have relied on artificial systems that alter membrane curvature[1–3]. In all previously

16   described species, chemoreceptors organize with the histidine kinase (CheA) and coupling

17   protein (CheW) into a hexagonal (P6 symmetry) extended array that is considered universal

18   among archaea and bacteria[4,5]. Here, for the first time, we report an alternative symmetry

19   (P2) of the chemotaxis apparatus that emerges from a strict linear organization of CheA in

20   *Treponema denticola* cells, which possesses arrays with the highest native curvature

21   investigated thus far. Using cryo-ET, we reveal that the *Td* chemoreceptor arrays assume a

22   truly unusual arrangement of the supra-molecular protein assembly that has likely evolved to

23   accommodate the high membrane curvature. The arrays have several additional atypical

24   features, such as an extended dimerization domain of CheA and a variant CheW-CheR-like

25   fusion protein that is critical for maintaining an ordered chemosensory apparatus in an

26   extremely curved cell. Furthermore, the previously characterized *Td* oxygen sensor ODP

27   influences array integrity and its loss substantially orders CheA. These results suggest a

28   greater diversity of the chemotaxis signaling system than previously thought and

29   demonstrate the importance of examining transmembrane systems *in vivo* to retain native

30   membrane curvature.

31

32   **Introduction**

33   Chemotaxis is a behavior most motile bacteria employ to sense their chemical environment

34   and navigate toward favorable conditions. The main components of the system are

35   transmembrane chemotaxis receptors called methyl-accepting chemotaxis proteins (MCPs),

36   the histidine kinase CheA, and the adaptor protein CheW. The intracellular tips of MCPs

37   bind CheA and CheW and communicate changes from the external chemical environment

38   into the cell by modulating CheA kinase activity (Fig. S1A)[6–8]. Activation of CheA initiates an

39   intracellular phosphorelay that ultimately controls flagellar rotation and cell movement.

40   CheA functions as a dimer and possesses five domains (P1-P5) with distinct roles in

41   autophosphorylation and array integration. The P1 domain is the phosphate substrate

42   domain, P2 interacts with response regulators, P3 is the dimerization domain, P4 binds ATP,

43   and P5 interacts with CheW. In the model species *Escherichia coli* (*Ec*), CheA P5 and CheW

44   are paralogs that interact pseudo-symmetrically to form six-subunit rings. In all bacterial and

45   archaeal species examined thus far, the MCPs are arranged in a trimer-of-dimer oligomeric

46   state and further organize into a hexagonal lattice (Fig. S1B). In *Ec*, the receptors are

47   connected by the highly ordered rings of CheA and CheW bound to the cytoplasmic tips of

48   the receptors[5,9]. These insights have established a widely accepted central model of the

49   chemotaxis array (Fig. S1B)[4,5,10]. However, emerging research has recently revealed

50   divergent components and arrangements of the chemotaxis apparatus in non-canonical

51   organisms. For example, in *Vibrio cholerae* (*Vc*) chemotaxis arrays, CheA and CheW lack

52   an ordered arrangement in the rings[1,11].

53

54   Many of these structural insights have transpired from cryo-electron tomography (cryo-ET)

55   studies that utilize artificial systems for higher resolution data. Specifically, the advent of so-

56   called 'mini-cell' bacterial strains produce extremely small cells that are ideal for cryo-ET[1,3],

57   and lipid-templating methods generate *in vitro* arrays with increased conformational

58   homogeneity[2,12]. However, these methods generate arrays with non-native curvature and it

59   is unclear how this may affect array structure and behavior. Here, we examine the *in vivo*

60   array structure of the pathogenic spirochete *Treponema denticola* (*Td*) by cryo-ET, which

61   possesses the highest native membrane curvature of any bacteria examined for their

62   chemotaxis system so far[4,5]. We demonstrate the presence of a novel array architecture with

63   two-fold (P2) symmetry in *Td*, that likely owes to the high curvature of the cells. Genetics

64   experiments, bioinformatics analyses, structural investigations, and molecular modeling of

65   *Td* chemotaxis proteins reveals adaptations that have likely evolved to accommodate

66   formation of an extended chemotaxis array in a highly curved membrane. We demonstrate

67   that a CheR-like fusion domain in a variant CheW is key for maintaining the structural

68   integrity of the arrays. Furthermore, cryo-ET analysis of *Td* cells lacking the oxygen sensor

69   ODP reveals substantial changes to the ordering or mobility of CheA[13]. Collectively, these

70   data demonstrate a greater diversity of the chemotaxis system than previously assumed,

71   and exemplify the importance of examining biological structures in native *in vivo* conditions.

72

73   **Results**

74   **Conservation of the F2 chemotaxis system**

75   The chemotaxis systems in prokaryotes have been classified into 19 systems based on

76   phylogenomic markers[14]. These classes include 17 'Flagellar' systems (F), one 'Alternative

77   Cellular Function' system (ACF), and one 'Type Four Pilus system' (TFP). The spirochete

78   chemotaxis system belongs to the 'Flagella class 2' (F2) category, which has not been

79   investigated with structural methods[14]. To explore the characteristics of chemotaxis proteins

80   in F2 genomes, we analyzed genomes in the Microbial Signal Transduction Database

81   version 3 (MiST3)[15]. All genomes (306) with at least one CheA of the F2 class (CheA-F2)

82   are in the Spirochaetota phylum, with a few exceptions of lateral gene transfer to other phyla

83   (Dataset 1, see Supplementary Methods). Not all Spirochaetota bacteria possess an F2

84   system however. There are 1096 genomes assigned to the Spirochaetota phylum in GTDB[16]

85   and 804 of these are present in the MiST3 database (Dataset 2)[15]. With these 804 genomes,

86   we mapped the different classes of CheA kinases to the Spirochaetota taxonomy tree

87   (Figure S1). Based on the topology of this tree, it appears that the major chemosensory

88   systems in the genomes from the Spirochaetota phylum are: F1 (Leptospirae), a transitional

89   hybrid F1/F2 system (Brachyspirae), and F2 (Spirochaetia) (Fig. S2). Therefore, we

90   conclude that F2 systems are exclusive to the Spirochaetia class, with a few exceptions of

91   lateral gene transfer.

92

93   The main architectural difference of the F2 system compared to others is the presence of an

94   unusual scaffold protein that consists of an N-terminal CheW domain and a C-terminal

95   CheR-like domain, hereafter referred to as CheW-CheR$_{like}$. Typically, CheR is a

96   methyltransferase that, together with the methylesterase CheB, controls the methylation

97   state of the receptors and thus provides an adaptation system[17]. Our analyses indicate that

98   all Spirochaetia genomes in MiST3 contain CheW-CheR$_{like}$. Furthermore, if we limit our

99   analysis to genomes that are fully sequenced (see Supplementary Methods), CheW-CheR$_{like}$

100   is only found in F2 chemosensory systems.

101

102   To investigate sequence patterns in the CheR protein and the CheR-like domain of CheW-

103   CheR$_{like}$, we produced a sequence dataset with 83 CheR-F2 and 88 CheW-CheR$_{like}$ proteins

104   and summarized them in sequence logos (Fig. S3A). Although the key catalytic residues are

105   conserved in the typical CheR protein, two residues that are essential for CheR to methylate

106   chemoreceptors, R79 and Y218 in *Td* CheR, are modified in the CheW-CheR$_{like}$ protein

107   (R79W and Y218F)[17]. Furthermore, the conserved region at the C-terminus of CheR is not

108   conserved in CheW-CheR$_{like}$. Based on these results we speculate that the CheR$_{like}$ domain

109   does not possess methyltransferase activity. Collectively, our analyses suggest that CheR

110   and CheW-CheR$_{like}$ have different biological functions.

111

112   F2 systems contain three proteins with a CheW domain: the classical scaffold CheW, CheW-

113   CheR$_{like}$, and the P5 domain from the histidine kinase CheA. Phylogenetic mapping of

114   CheW-F2 proteins indicate the location of last common ancestor (Fig. S4, see

115   Supplementary methods). To investigate sequence patterns of the three CheW domains, we

116   analyzed non-redundant sequence datasets of CheW, CheW-CheR$_{like}$, and CheA-F2 from all

117   genomes with at least one CheA-F2. The final alignments for each group contain the CheW

118   domain portion of 74 CheW proteins, 59 CheW-CheR$_{like}$ proteins, and 73 CheA proteins. The

119   sequences of each group are summarized in sequence logos and demonstrate the presence

120   of conserved regions at established interaction interfaces, as well as loop insertions near

121   these interfaces that could confer altered specificity of binding (Fig. S3B).

122

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

124   Cell poles of intact *Td* cells were imaged by cryo-electron tomography (cryo-ET) and used

125   for three-dimensional reconstructions. Top views (cross-sections through the array) and side

126   views (visualizing the long axes of the receptors) of membrane-associated arrays were

127   clearly visible (Fig. 1A, S5A). Sub-volume averaging revealed the conserved receptor trimer-

128   of-dimers in the typical hexagonal arrangement. Remarkably, several novel features of the

129   chemotaxis arrays are apparent (Fig. 1B). Specifically, a density of unknown origin is located

130   in the center of the receptor hexagons and slightly above the plane of the CheA:CheW rings.

131   This density, which will hereafter be referred to as the middle density, extends from two

132   subunits in the rings (Fig. 2A). Additionally, there are small but distinct puncta of density in

133   between some of the trimer-of-dimer modules (Fig. 1B). However, averages of the arrays at

134   the CheA:CheW layer did not reveal discernible CheA density, indicating either a sparse or

135   disordered distribution of CheA or a highly mobile kinase (Fig. 1C). Remarkably, the sub-

136   tomogram averages reveal the axis of the *Td* cells relative to the chemotaxis arrays,

137   demonstrating that the arrays occupy a preferred orientation with respect to the cell axis

138   (Fig. S5B).

139

140   **Arrays in *T. denticola* deletion mutants**

141   The oxygen-binding diiron protein (ODP) functions as an oxygen sensor for chemotaxis in

142   *Td*, however it is uknown if ODP is an integral component of the array[13]. This protein is

143   genetically coupled to an MCP homolog, TDE2496, that lacks both transmembrane and

144   sensing modules. TDE2496 likely integrates into the cytoplasmic regions of the membrane-

145   bound arrays based on the observation that no cytoplasmic arrays were observed in the

146   tomograms. Moreover, the *Td* genome encodes only one CheA homolog, and cytoplasmic

147   receptors often associate with distinct kinases[9,18]. TDE2496 forms trimers-of-dimers *in vitro*

148   and can modulate CheA activity (Fig. S6A,B). To determine if the presence of ODP

149   (TDE2498) and its cognate receptor (TDE2496) impacts array architecture or integrity, we

150   conducted cryo-ET with two *Td* gene knock-out strains, Δ2498 and Δ2498Δ2496[13].

151   Importantly, deletion of ODP does not impact transcription of TDE2496[13]. The sub-volume

152   averages of these strains reveal distinct differences in array densities compared to the wild-

153   type (WT) strain (Fig. 1B,C, Fig. S5A). Namely, the location of CheA at the CheA:CheW

154   layer (Layer 2, Fig 1C) is now clearly visible. Interestingly, CheA arranges in well-ordered

155   linear rows. Placement of CheA necessarily positions the P3 domain in between two of the

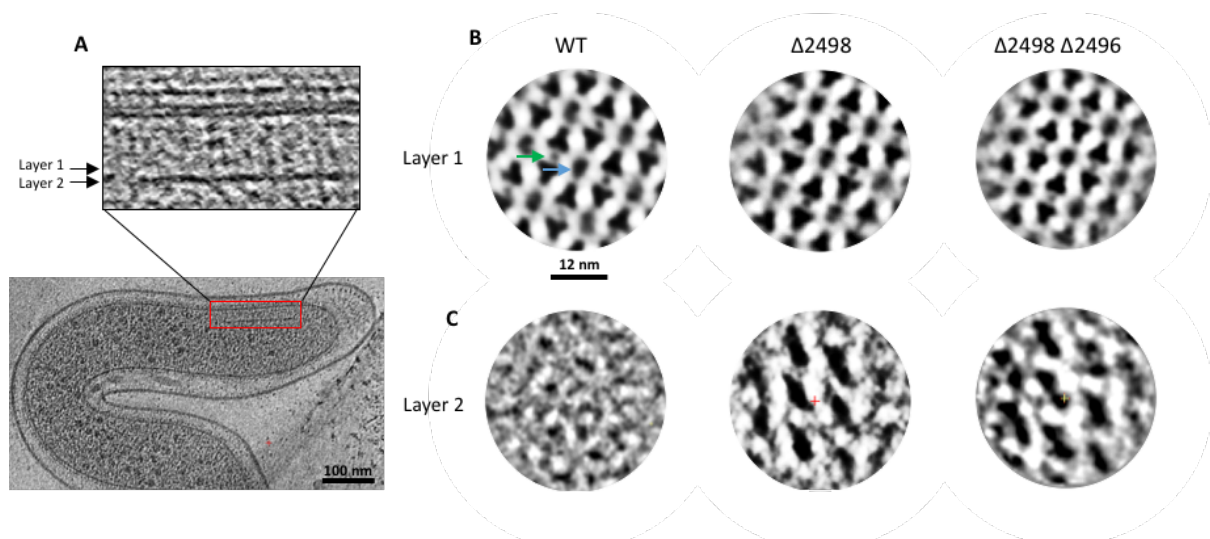156   trimer-of-dimer modules in each hexagon. This position exclusively corresponds to the

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). (C) Sub-volume averages at Layer 2 reveal the organization of CheA. In this arrangement, the density between the trimer-of-dimer modules (green arrow) corresponds to the CheA P3 domain.

157   location of the puncta between receptor trimer-of-dimer modules observed in the WT array,

158   indicating that this density represents the P3 domain (Layer 1, Fig. 1B). Like the WT arrays,

159   the cell axis relative to the chemotaxis arrays is also apparent in these sub-tomogram

160   averages and matches the preferred array orientation in the WT cells (Fig. S5B).

161

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

163   To determine the composition of the middle density in the *Td* arrays, we considered the

164   unique chemotaxis proteins in spirochete

165   genomes. As shown above, CheW-CheR$_{like}$ is

166   a conserved component of the F2 Spirochaetia

167   chemotaxis system (Fig. S2). In *Td,* a 28-

168   residue linker in CheW-CheR$_{like}$ separates the

169   two domains and is predicted to be largely

170   helical (JPred), and the gene is co-transcribed

171   with the only CheA, CheX, and CheY proteins

172   in the genome (Fig. S6C,D)[18]. Furthermore,

173   the purified CheW-CheR$_{like}$ protein dimerizes

174   and interacts with CheA, as evidenced by its

175   ability to alter autophosphorylation activity (Fig.

176   2C,D, Fig. S6E,F). Therefore, we postulated

177   that the middle density may be comprised of

178   two dimerized CheR$_{like}$ domains that extend

179   from the CheW units harbored within the

180   CheA:CheW rings.

181

182   A *Td* strain lacking the CheR$_{like}$ domain

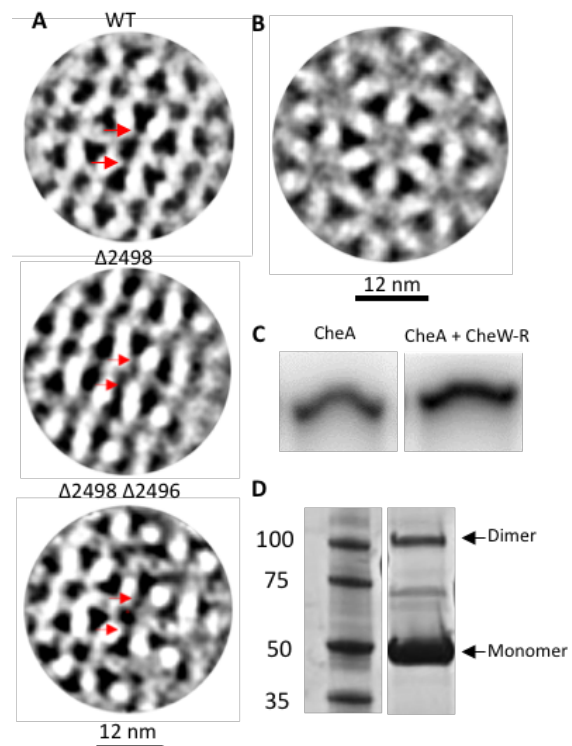183   (ΔCheR$_{like}$) reveals a significant decrease in

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*. 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.

184   the prevalence and size of the arrays (Fig. S5A). Due to the small size of the arrays in

185   ΔCheR$_{like}$, only 194 particles were available for sub-tomogram averaging, but the resulting

186   averages clearly demonstrate that the middle density is no longer present (Fig. 2B, S5A).

187   Like the WT strain, the CheA density below the rings in ΔCheR$_{like}$ is not apparent.

188

189   **Protein interfaces in the CheA:CheW:CheW-CheR$_{like}$ rings**

190   Bioinformatics analyses demonstrate that all functional Spirochaetota F2 chemotaxis

191   systems possess a CheW-CheR$_{like}$ homolog and at least one classical CheW protein. As

192   only two of the CheW subunits in the hexagonal rings extend to the middle density (which

193   probably arises from the CheW-CheR$_{like}$ protein), and two of the ring positions are occupied

194   by CheA P5, it is likely that the other two positions are occupied by the classical CheW

195   protein. (Fig. 2A, S7). Within this arrangement three unique interaction interfaces are

196   possible, interface 1 occurs between CheA P5 and the classical CheW (as seen in canonical

197   systems), interface 2 occurs between CheA P5 and the CheW domain of CheW-CheR$_{like}$,

198   and a third interface (interface 3) occurs between the classical CheW and the CheW domain

199   of CheW-CheR$_{like}$ (Fig. S7). To explore the binding interfaces within the *Td* rings, we

200   analyzed homology models of the classical CheW, the CheW domain of CheW-CheR$_{like}$, and

201   the CheA P5 domain. The CheW models were generated using a crystal structure of

202   *Thermoanaerobacter tengcongensis* (*Tt*) CheW as the template (PDB ID: 2QDL), and the

203   CheA P5 model was generated using a cryo-EM structure of *E. coli* CheA P5 (PDB ID:6S1K)

204   (Fig. S8A-C)[2,19]. Three of the four regions with lowest sequence conservation among the

205   three domains are located at interfaces 1-3 (Fig. S9). Alignment of the *Td* CheW and CheA

206   P5 models to a crystal structure of *Tm* CheW in complex with *Tm* CheA P5 (PDB ID: 3UR1)

207   further illustrates that these regions are located at the CheW:P5 ring interfaces (Fig S9)[20].

208   Mapping the variable regions onto the sequence logos of the F2 CheW domains

209   demonstrates that they evolved different sequence patterns in these regions, with the

210   exception of the variable region that is not located at the interaction interface (region 2, Fig.

211   S9, S3B).

212

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:CheA/W interactions (Fig. 3A,B). The orientation of the strands are apparent in the sub-tomogram averages and run relatively perpendicular to the axis of the cells (Fig. S10A). The array densities at Layer 1 (receptors, P3, and CheR$_{like}$) produce apparent lines in the cryo-ET reconstructions that, like the sub-tomogram averages, run parallel to the cell axis (Fig. 3C,D)[21]. This arrangement allows interactions that hold the strands together to occur with minimal bending (Fig. 3E,F). Indeed, the angle between the cell axis and these lines in the cells is 10.4 +/- 8.6°, (*n* = 26 cells) and no significant difference was found among the three *Td* strains measured (Table S1A). As the CheA:CheW strands are perpendicular to these lines in the sub-tomogram averages, they will be oriented at an angle of 55.4 +/- 8.6° with respect to the cell axis (Fig. S10A).

226

Previous studies in *Ec* and *Vc* demonstrate that the chemotaxis arrays can accommodate vastly different curvatures between lysed and artificially small mini-cell strains[1,3,22]. However, *Td* cells demonstrate a significantly higher curvature of the cell membrane than organisms investigated for their chemotaxis arrangement thus far [1,5,22]. 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. S10B,C, Table S1B,C). Additionally, the measured curvature of the *Td* CheA:CheW baseplate is 65.6 +/- 19 /μm (radius: 152 Å, *n* = 10 cells) (Fig. S10B, Table S1C).

236

The CheA:CheW rings present in a crystal structure (PDB ID: 3UR1) are flat and ~95 Å in diameter (Fig. S11A)[20]. The length across two rings connected by a dimeric CheA is 224 Å (Fig. 11B). To determine the extent of bending that would need to occur in the two

240   CheA:CheW rings if they ran perpendicular to the cell axis, the 224 Å rings were modeled as

241   a chord in a circle with radius 152 Å. Using the equation $h = r - \sqrt{r^2 - L^2}$ (where h is the

242   height of the circular segment, r is the circle radius, and L is half the chord length (224 Å /2)),

243   the height of the circular segment is 49.2 Å (Fig. S11B). Therefore, the center of the two

244   rings (the P3 domain) must bend by an average of 49.2 Å toward the cell membrane to

245   accommodate the cell curvature. Using the same equation above (where L is 95 Å /2), the

246   height of the circular segment is 7.6 Å (Fig. S11A). Therefore, in *Td,* the center of a single

247   ring must bend toward the membrane by an average of 7.6 Å to align to the measured

248   baseplate curvature.
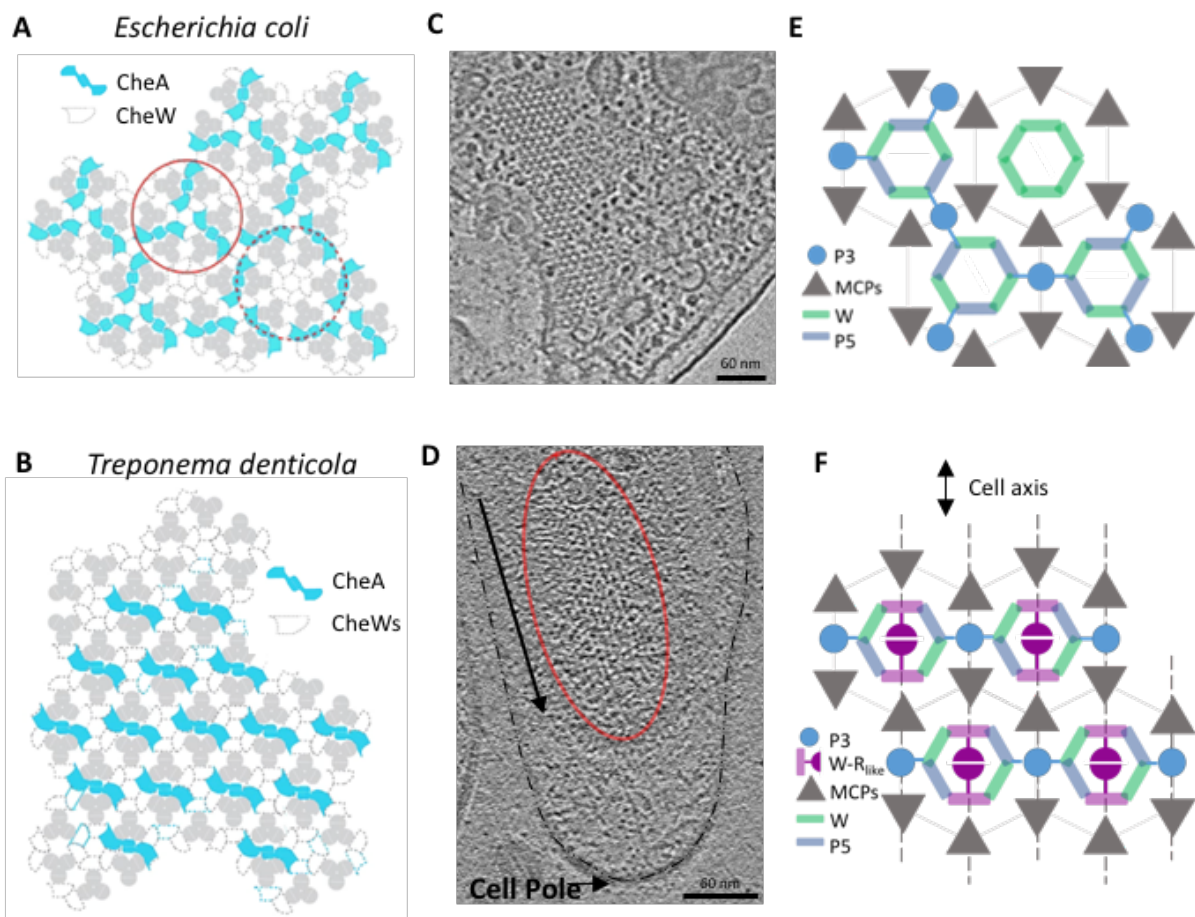
249

Fig. 3 The organization of CheA in *Td* differs from all previously reported arrangements. (A) The hexagonal arrangement of CheA in *E. coli* and other canonical chemotaxis systems. (B) In *Td,* CheA is arranged with a strict linear organization, producing 'strands' of CheA:CheW rings linked by the CheA P3 domain. (C) The receptor densities in *Ec* (Layer 1) are apparent and hexagonal. Figure adapted from Briegel A. et al. 2010.[21] (D) The densities at Layer 1 in *Td* (red oval) produce apparent lines of receptors, P3 and the middle density (CheR-like) that follow the axis of *Td* cells (black arrow). (E) *Ec* arrays possess radial P6 symmetry where all interaction interfaces are equally curved. (F) The *Td* array has a two-fold symmetry arrangement that allows receptor:CheA/W interactions that hold the CheA:CheW strands together to occur with minimal bending. Dashed lines represent the apparent lines in the tomograms.

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 indicate that in *Td* CheA an additional ~50 residues join the canonical dimerization domain (P3) helices (Figure S12A)[4]. CheA homologs from other spirochete genera including Borrelia and Brachyspira also possess additional residues in this region (Fig. S12B)[23]. 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. S13A). Furthermore, CheA-F2 proteins contain the most residues in this non-conserved region (Fig. S13B,C). The x-ray crystal structure of the isolated *Td* P3 domain (PDB ID: 6Y1Y, Fig. 4, Table S2) reveals 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. 4, S14A). Interestingly, aromatic residues (Phe, Tyr) cluster near the helix breakages and unusual core packing of these residues allows for maintenance of a coiled-coil register despite a disruption of helical heptad repetition in the C-terminal helix (Fig. S14A). The net result is a distortion in the alignment of the hairpin tip, the consequence of which is currently unknown. Differing orientations of Tyr83 in the two subunits produces asymmetry in the added tip region (Fig. S14B,C). Fitting the new P3 domain into an all-atom chemotaxis
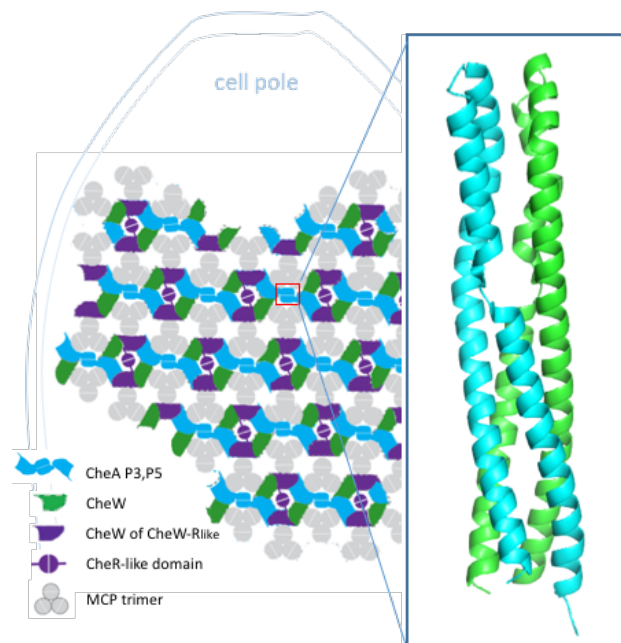

Fig. 4 The arrangement of chemotaxis proteins in *Td*. The linear organization of CheA produces 'strands' of CheA:CheW rings that are perpendicular to the cell axis and are linked by an atypical P3 domain.

277   array that was generated for previous molecular dynamics simulations (PDB ID: 3JA6)

278   shows that these additional helices are within a ~15 Å from receptors (Fig. S14D)[12].

279   Additionally, the handedness of the helix connection in the *Td* P3 domain differs from that of

280   *Thermotoga maritima* CheA and rather matches helix connectivity of sensor kinase DHP

281   domains[24,25].

282

Discussion

284   Here, we reveal the protein arrangement of F2 chemotaxis arrays through cryo-ET of intact

285   *T. denticola* (*Td*) cells. *Td* cells are thin and cyndrillical, producing a cell that is polarized in

286   shape and membrane curvature; the membrane has extremely high curvature perpendicular

287   to the cell axis and lower curvature parallel to the axis. In accordance with this feature, the

288   *Td* arrays are polarized and assume a preferred orientation with respect to the cell axis. In

289   this system, three proteins comprise the rings at the receptor tips: CheA, CheW, and a

290   CheW-CheR$_{like}$ protein. Like the *Ec* system, these proteins in *Td* are integrated into the array

291   with strict organization[1,22]. However, a linear arrangement of CheA is present that generates

292   'strands' of rings interlinked by the CheA dimerization domain (P3). The strands run

293   perpendicular to the axis of the *Td* cells, resulting in substantial bending of inter-linked rings.

294   Presumably, the unique binding interfaces in the CheA:CheW rings generate bent interfaces

295   with the appropriate curvature. Furthermore, this directional distinction allows for some array

296   contacts to follow the cell axis and remain undistorted. The deleterious effect on array

297   integrity with the loss of the CheR$_{like}$ domain strongly suggests that CheR$_{like}$ dimerization

298   plays a key role in array assembly and stabilization. The strict linear arrangement of CheA

299   could be facilitated by the composition of the *Td* rings; three unique protein interfaces are

300   present in the rings and restrict CheA P5 integration (i.e. CheA P5 can only occupy these

301   two positions in the six-member ring). Furthermore, the *Td* CheA P3 domain is clearly

302   discernible in the averaged tomograms, which has not been previously observed *in vivo*[5,9,22].

303   As the CheA:CheW rings have to undergo substantial bending to accommodate the

304   baseplate curvature, the elongated P3 domain may have evolved to stabilize CheA

305 dimerization by increasing the interface area and furthermore may encourage interactions

306 with neighboring receptors[12]. Due to the curvature, the receptor trimer-of-dimer modules are

307 expected to be further splayed, and the elongated P3 may compensate for the increased

308 distance between receptors and P3. As *Td* cells have the smallest average diameter (0.1 –

309 0.4 μm) of all bacteria with determined chemotaxis architectures thus far, and the novel

310 linear arrangement of CheA produces arrays that can better accommodate its unique

311 membrane curvature, we surmise that the array architecture in *Td* is an adaptation that

312 evolved to produce an extended chemoreceptor apparatus in a highly curved and

313 asymmetric membrane[4,5,9]. Importantly, bioinformatics analyses indicate that the unique

314 protein features seen in *Td* are exclusive to Spirochaetia F2 systems.

315

316 Unexpectedly, the placement of CheA in WT *Td* arrays could not be discerned (with the

317 exception of the P3 domain), but was clearly visible in two *Td* mutants (Δ2498,

318 Δ2498Δ2496). As the density corresponding to the P3 domain in the WT strain is clearly

319 discernible, the sparse density corresponding to all other CheA domains (P1, P2, P4, P5) is

320 not attributed to low CheA incorporation in these arrays. These results indicate that the

321 kinase is highly mobile or more disordered in the WT strain, but is more constrained when

322 ODP (TDE2498) is deleted, suggesting that ODP directly affects array structure. However,

323 densities in the three strains do not designate an obvious position for ODP, indicating that

324 ODP may not be an integral component of the array, but rather peripherally interacts with the

325 chemotaxis machinery or influences its architecture through other means, perhaps related to

326 its signaling properties.

327

328 In summary, we illustrate a novel chemotaxis arrangement that has evolved to compliment

329 the spirochetes' high membrane curvature and asymmetry. Therefore, it is likely that the

330 behavior and characteristics of chemoreceptor arrays in general can be influenced by

331 perturbing the shape of the cell membrane. Recent studies with *Ec* ultra-minicells (the

332 smallest mini-cells available to date) produce densities corresponding to portions of

333   chemoreceptors that have not been observed before[3]. However, it is possible that increased

334   membrane curvature limits movement of the transmembrane receptors, resulting in

335   increased receptor localization and resolution. Furthermore, the use of lipid-templating has

336   been extensively used to assemble arrays *in vitro* for cryo-ET and reconstitutes the

337   chemotaxis apparatus in a perfectly flat formation[2,12]. While cryo-ET experiments of *in vivo*

338   *Ec* arrays consistently shows the CheA P3 density to be too sparse to determine its position,

339   cryo-ET of *in vitro* lipid-templated *Ec* arrays clearly defines the P3 position. If the P3 domain

340   does engage receptors, this discrepancy may suggest that lipid-templating abrogates this

341   native behavior. Furthermore, it is unclear how the *Td* chemotaxis arrangements affects

342   cooperative behavior in the arrays as linear CheA:CheW strands are only connected through

343   receptor interactions. Investigations into this system may reveal alternative mechanisms for

344   cooperative behavior than what have been reported. Collectively, our results illustrate the

345   importance of investigating transmembrane systems *in situ*, and that examining systems in

346   non-model organisms can lead to new, unexpected advances for understanding the

347   remarkable signaling system of bacterial chemotaxis.

348

359

360

361 **Author contributions.** A.R.M., D.R.O, K.K., C.L., B.R.C, and A.B. designed research; A.R.M.,

362 D.R.O., and K.K. performed research; A.R.M., D.R.O, W.Y., K.K., and A.S. analyzed data; and

363 A.R.M., D.R.O, and A.B. wrote the paper with input from all authors.

364

365 **Competing interest.** The authors declare no competing interests.

366

367 **Materials and Correspondence.** Correspondence and materials requests should be addressed to A.

368 Briegel at a.briegel@biology.leidenuniv.nl.

369

370

371 1. Yang, W., Alvarado, A., Glatter, T., Ringgaard, S. & Briegel, A. Baseplate variability of

372 *Vibrio cholerae* chemoreceptor arrays. *Proc. Natl. Acad. Sci.* **115**, 201811931 (2018).

373 2. Cassidy, C. K. *et al.* Structure and dynamics of the E. coli chemotaxis core signaling

374 complex by cryo-electron tomography and molecular simulations. *Commun. Biol.* **3**,

375 1–10 (2020).

376 3. Burt, A. *et al.* Complete structure of the core signalling unit of the E. coli

377 chemosensory array in an optimised minicell strain. *Nat. Commun.* 1–9 (2019).

378 doi:10.1038/s41467-020-14350-9

379 4. Briegel, A. *et al.* Universal architecture of bacterial chemoreceptor arrays. *Proc. Natl.*

380 *Acad. Sci. U. S. A.* **106**, 17181–17186 (2009).

381 5. Briegel, A. *et al.* Bacterial chemoreceptor arrays are hexagonally packed trimers of

382 receptor dimers networked by rings of kinase and coupling proteins. *Proc. Natl. Acad.*

383 *Sci.* **109**, 3766–3771 (2012).

384 6. Hazelbauer, G. L., Falke, J. J. & Parkinson, J. S. Bacterial chemoreceptors: high-

385 performance signaling in networked arrays. *Trends Biochem. Sci.* **33**, 9–19 (2008).

386 7. Pedetta, A., Parkinson, J. S. & Studdert, C. A. Signalling-Dependent Interactions

387 Between the Kinase-Coupling Protein CheW and Chemoreceptors in Living Cells. **93**,

388 1144–1155 (2015).

8. 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).

9. Briegel, A. *et al.* Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling. *Elife* **2014**, 1–16 (2014).

10. 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).

11. Zhang, K. *et al.* Two CheW coupling proteins are essential in a chemosensory pathway of Borrelia burgdorferi. *Mol. Microbiol.* **85**, 782–794 (2012).

12. Cassidy, C. K. *et al.* CryoEM and computer simulations reveal a novel kinase conformational switch in bacterial chemotaxis signaling. *Elife* **4**, 1–20 (2015).

13. 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).

14. Wuichet, K. & Zhulin, I. Origins and Diversification of a Complex Signal Transduction System in Prokaryotes. *Sci Signal* **128**, 1–27 (2010).

15. 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).

16. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996 (2018).

17. 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).

18. Greene, S. R. & Stamm, L. V. Molecular characterization of a chemotaxis operon in the oral spirochete, Treponema denticola. *Gene* **232**, 59–68 (1999).

19. Yao, W., Shi, L. & Liang, D. C. Crystal structure of scaffolding protein CheW from

417   thermoanaerobacter tengcongensis. *Biochem. Biophys. Res. Commun.* **361**, 1027–

418   1032 (2007).

419 20. Li, X. *et al.* The 3.2 Å resolution structure of a receptor:CheA:CheW signaling complex

420   defines overlapping binding sites and key residue interactions within bacterial

421   chemosensory arrays. *Biochemistry* **52**, 3852–3865 (2013).

422 21. Briegel, A. *et al.* The mobility of two kinase domains in the Escherichia coli

423   chemoreceptor array varies with signaling state . **14**, 384–399 (2010).

424 22. Yang, W. *et al.* In situ conformational changes of the escherichia coli serine

425   chemoreceptor in different signaling states. *MBio* **10**, 1–14 (2019).

426 23. Sze, C. W., Zhang, K., Kariu, T., Pal, U. & Li, C. Borrelia burgdorferi needs

427   chemotaxis to establish infection in mammals and to accomplish its enzootic cycle.

428   *Infect. Immun.* **80**, 2485–2492 (2012).

429 24. Greenswag, A., Muok, A., Li, X. & Crane, B. Conformational Transitions that Enable

430   Histidine Kinase Autophosphorylation and Receptor Array Integration. *J. Mol. Biol.*

431   **427**, 87–92 (2016).

432 25. Bilwes, A. M., Alex, L. A., Crane, B. R. & Simon, M. I. Structure of CheA, a signal-

433   transducing histidine kinase. *Cell* **96**, 131–141 (1999).

434

435

436 **Supplementary Methods and Materials**

437 **Data availability**

438 The cryo-ET sub-tomogram averages that support these findings are deposited in the Electron

439 Microscopy Data Bank (EMDB) with accession codes EMD-10784, EMD-10785, EMD-10786, EMD-

440 10787, EMD-10788, EMD-10790. The protein x-ray crystallography structure that supports these

441 findings is deposited in the Protein Data Bank (PDB) with accession code 6Y1Y. These depositions

442 will be publicly available with publication.

443

444 **Code availability**

445 All custom-made scripts are available to editors and referees upon request. These scripts will be

446 publicly available with publication.

447

448 **Bacterial strains, culture conditions, and oligonucleotide primers**

449 *Treponema denticola* (*Td*) ATCC 35405 (wild-type) was used in this study. The *Td* deletion

450 mutants, Δ2498 and Δ2498Δ2986, were generated in a previous study[2]. Cells were grown in

451 tryptone-yeast extract-gelatin-volatile fatty acids-serum (TYGVS) medium at 37°C in an

452 anaerobic chamber in presence of 85% nitrogen, 10% carbon dioxide, and 5% hydrogen[1].

453 *Td* mutants were grown with an appropriate antibiotic for selective pressure as needed:

454 erythromycin (50 µg/ml) and gentamycin (20 µg/ml). *Escherichia coli* 5α strain (New England

455 Biolabs, Ipswich, MA) was used for DNA cloning. The *E. coli* strains were cultivated in

456 lysogeny broth (LB) supplemented with appropriate concentrations of antibiotics. The

457 oligonucleotide primers for PCR amplifications used in this study are listed in Table S3.

458 These primers were synthesized by IDT (Integrated DNA Technologies, Coralville, IA).

459

460 **Construction of a CheR truncated mutant (ΔCheR$_{like}$)**

461 *TDE1492::ermB* (Fig. S15) was constructed to replace the CheR-like domain (8781-1,308 nt) in

462 *TDE1492 with* a previously documented erythromycin B resistant cassette (*ermB*)[3]. The

463 *TDE1492::ermB* vector was constructed by two-step PCR and DNA cloning. To construct this vector,

464 the 5` end of *TDE1492* region and the downstream flanking region were PCR amplified with primers

465   P_1/P_2 and P_3/P_4, respectively, and then fused together with primers P_1/P_4, generating Fragment 1. The

466   Fragment 1 was cloned into the pMD19 T-vector (Takara Bio USA, Inc, Mountain View, CA). The

467   *ermB* cassette was PCR amplified with primers P_5/P_6, generating Fragment 2. The Fragment 2 was

468   cloned into the pGEM-T easy vector (Promega, Madison, WI). The Fragment 1 and 2 were digested

469   using *Not*I and ligated, generating the *TDE1492::ermB* plasmid. The primers used here are listed in

470   Table S3. To delete *TDE1492*, the plasmid of *TDE1492::ermB* was transformed into *Td* wild-type

471   competent cells via heat shock, as previously described[4]. Erythromycin-resistance colonies that

472   appeared on the plates were screened by PCR for the presence of *ermB* and absence of *TDE1492*

473   *(781-1,308 nt)* gene**.** The PCR results showed that the *TDE1492 (781-1,308 nt)* gene was replaced by

474   *ermB* cassette as expected (Fig. S15). One positive clone (*ΔTDE1492*) was selected for further study.

475

476   **Bioinformatics software and resources**

477   The datasets used in the bioinformatics analysis were built using data from Microbial Signal

478   Transduction Database v3 (MiST3) accessed February 2020[5] and the Genome Taxonomy Database

479   v89 (GTDB) [6]. We built custom scripts using TypeScript-3.7.5 and NodeJS-12.13. To make these

480   scripts, we also used packages publicly available at the node package manager repository (npm): we

481   used RegArch-1.0.1[7] to separate CheW-CheR$_{like}$ from other CheWs, gtdb-local-0.0.12

482   (https://npmjs.com/package/gtdb-local) to use the GTDB taxonomy, Phylogician-TS-0.10.1-4

483   (https://npmjs.com/package/phylogician-ts) to visualize and manipulate the phylogenetic trees,

484   BioSeq-TS-0.2.4 (https://npmjs.com/package/bioseq-ts) to handle protein sequences and MiST3-TS-

485   0.7.6 (https://npmjs.com/package/mist3-ts) to access MiST3 API. Multiple sequence alignments were

486   produced using L-INS-I algorithm from the MAFFT package. To reduce redundancy in sequence

487   datasets we used CD-HIT v4.6[8] with unaligned sequences and Jalview[9] with aligned sequences.

488   RAxML v8.2.10[10] was used to perform phylogenetic reconstructions, and low support branches in the

489   phylogenetic trees were collapsed with TreeCollapseCL4[11]. Sequence logos were built using Weblogo

490   3.7[12].

491

492   **Bioinformatics scripts and pipelines**

493   We collected all information on the proteins classified as CheA (96,434) and CheW (134,165). To

494   process this dataset we built several scripts and pipelines to produce the tables, figures and datasets

495   used in this analysis (Fig. S16). The scripts are found in Supplementary File 1.

496

497   **Chemosensory profile of Spirochaetotas**

498   The "spiro-pipeline" selects all genomes from Spirochaetota phylum using gtdb-local package to

499   access GTDB v89, then it filters only the genomes that are also present in MiST3 database. It collects

500   the information on MiST for each genome and appends the complete taxonomy information from

501   GTDB and signal transduction profiles. Finally, the pipeline builds the table with the information in

502   markdown (Dataset 2).

503

504   **Chemosensory profile of genomes with at least one CheA-F2**

505   The "chea-pipeline" starts from the raw dataset taken from MiST3 database with information on

506   96,434 CheA genes. Based on MiST3 classification it selects the genomes with at least one CheA-F2

507   sequences and fetches information about these genomes. At this step, it also checks with the list

508   generated by the "wr-pipeline" of the genomes containing CheW-CheR$_{like}$. It proceeds to append

509   chemosensory information for each genome and GTDB taxonomy. At this step the pipeline splits into

510   four pathways, where one parses the information to build the Dataset 1 and the other three build

511   FASTA formatted files with sequences from: CheA-F2, CheR-F2, CheW-CheR$_{like}$ and all CheWs

512   belonging to genomes with at least 1 CheA-F2. We noticed that MiST3 currently misclassifies some

513   CheR proteins, so we used RegArch to filter out false positives. We also used RegArch to separate

514   CheW and CheW-CheR$_{like}$ sequences as MiST3 classifies both as CheW (scaffold). The RegArch

515   definitions can be found in the script 'regArchDefinitions.ts' of the source code in Supplementary File

516   1.

517

518   In MiST3[12], 306 genomes contain at least 1 CheA-F2 (Dataset 1) with three exceptions: two of them

519   belong to Acidobacteriota phylum,and one from Planctomycetota, which suggest that the presence of

520   an F2 system in these genomes (outside of the Spirochaetota phylum) is the consequence of lateral

521   gene transfer. Of the 306 genomes with at least one CheA-F2, only the following lack the CheW-

522   CheR$_{like}$ protein: the three non-Spirochaetota genomes mentioned previously, 40 genomes of the

523   Brachyspirae class, and three genomes from the order Borreliales (Fig. S2). Interestingly, the

524   Borreliales genomes do appear to have the CheW-CheR$_{like}$ gene, except there is no gene product

525   associated with them in the MiST3 database.

526

527   **Classification of CheW**

528   CheW proteins are not classified in MiST3. In order to make comparisons between the sequences of

529   CheW domains in F2 systems, we must first select only CheW-F2. We first assign to the F2 class all

530   the canonical CheW found in genomes with a single CheA-F2. Contrary to the F2 systems where the

531   canonical CheW is not present in the chemosensory gene clusters, other classes do contain their

532   canonical CheW within the rest of the gene cluster. CheW found within 5 genes from a classified

533   CheA were assigned to the same class as CheA. To perform this classification, we selected the 598

534   full-length CheW sequences generated by the chea-pipeline. We used CD-HIT to remove 405

535   redundant sequences (-c 1). Next, we ran the "classify-w" pipeline on the remaining sequences (193).

536   The pipeline reads the identifiers of the sequences and fetches the chemotaxis profile from MiST3 for

537   each genome. It classifies CheWs as F2 classes if there is only 1 CheA of the class F2 in the profile.

538   Next, it fetches the gene neighborhood (5 genes up and downstream) of the remaining CheWs and

539   assigns a matching class if there is a classified CheA within these genes. We also aligned the

540   sequences (193) with the L-INS-I algorithm of the MAFFT package, produced a phylogeny with 1000

541   rapid bootstrap using RAxML (-f a -m PROTGAMMAIAUTO -N 1000) and collapsed the phylogeny

542   using TreeCollapse4 at 50% bootstrap. We mapped the CheW classification to the CheW tree in

543   Figure S2. We expanded the F2 classification to the 74 sequences within the branch with only CheW-

544   F2 sequences.

545

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

547   We put together the sequences from CheA-F2 and CheW-CheR$_{like}$ (both trimmed by the PFAM model

548   for the CheW domain, already annotated in MiST3) and the full sequence of the CheW-F2 selected in

549   the previous step. We then use L-INS-I algorithm to align the sequences and Jalview to manually

550   inspect and eliminate identical redundant sequences. Finally, we trimmed the whole alignment based

551   on the boundaries of CheA-F2 and CheW-CheR$_{like}$ and eliminated one incomplete sequence:

552   GCF_000413015.1-HMPREF1221_RS07250. The final alignment had a total of 206 sequences: 73

553   CheA, 59 CheW-CheR$_{like}$ and 74 CheW. We separated the alignments into individual files and built

554   sequence logos to summarize the amino-acid diversity in each position for each group (Fig. S3B).

555

Comparison of CheR domains

557   First, we added together the trimmed part matching the CheR domain of the CheW-CheR$_{like}$

558   sequences and the 292 sequences of the CheR protein. We then aligned the 550 sequences using L-

559   INS-I. We used Jalview to inspect the alignment and remove identical redundant sequences. The final

560   alignment had the CheR domain of 83 CheR and 83 CheW-CheR$_{like}$ proteins. We separated the

561   alignments into individual files and generated independent sequence logos (Fig. S3A).

562

Analyses of CheA P3 domains

564   The "p3-pipeline" pipeline processes the data for the analysis of the length of P3 domains of all CheA

565   homologs in the MiST3 database. It reads the information for all 96,434 CheAs in the MiST3

566   database, trims the sequence matching the Pfam mode H-kinase_dim and builds FASTA formatted

567   datasets for each chemotaxis class. For each dataset, we used CD-HIT with 75% identity cutoff and

568   aligned them using the L-INS-I algorithm from MAFFT. Using Jalview we manually inspect and edited

569   the alignment to remove divergent sequences that opens major gaps in the alignment. We removed 6

570   F1 sequences, 37 F7 sequences, 20 F8 sequences and 5 F9 sequences. Then we merged the

571   alignment using mafft-profile with each dataset as a seed alignment in a single shot (Fig. S13A). We

572   selected the non-conserved central region of the alignment and measured the number of amino-acids

573   in each sequence (Fig. S13B).

574

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

576   Cells were concentrated by centrifugation, and a 1/10 dilution of protein A- treated 10-nm colloidal

577   gold solution (Cell Microscopy Core, Utrecht University, Utrecht, The Netherlands) was added to the

578   cells and mixed by pipeting. 3 μL aliquots of the cell suspension were applied to glow-discharged

579   R2/2 200 mesh copper Quanti-foil grids (Quantifoil Micro Tools, GmbH), the sample was pre-blotted

580   for 30 seconds, and then blotted for 2 seconds. Grids were pre-blotted and blotted at 20 °C and at

581   95% humidity. The grids were plunge-frozen in liquid ethane using an automated Leica EM GP

582   system (Leica Microsystems).

583   Data collection was achieved on a Titan Krios transmission electron microscope (Thermo Fisher

584   Scientific) operating at 300 kV. Images for three strains (WT, Δ2498, Δ2498Δ2986) were recorded

585   with a Gatan K2 Summit direct electron detector with a GIF Quantum energy filter (Gatan) operating

586   with a slit width of 20 eV. Images were taken at a magnification of 42,000×, which corresponds to a

587   pixel size of 3.513 Å. Tilt series were collected using SerialEM with a modified bidirectional tilt scheme

588   (-20° to 60°, followed by -22° to -60°) with a 2° increment. Images for the ΔCheR$_{like}$ strain were

589   recorded with a Gatan K3 Summit direct electron detector equipped with a GIF Quantum energy filter

590   (Gatan) operating with a slit width of 20 eV. Images were taken at a magnification of 26,000X, which

591   corresponds to a pixel size of 3.27 Å. Tilt series were collected using SerialEM with a bidirectional

592   dose-symmetric tilt scheme (-60° to 60°, starting from 0°) with a 2° increment. For all strains, the

593   defocus was set to − 6μm and the cumulative exposure per cell was 100 e-/A^2.

594   Bead tracking-based tilt series alignment and drift correcting were done using IMOD[13]. CTFplotter

595   was used for contrast transfer function determination and correction[14]. Tomograms were

596   reconstructed using simultaneous iterative reconstruction with iteration number set to 6. Dynamo was

597   used for manual particle picking and sub-tomogram averaging[15,16].

598

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

600   DNA segments encoding the CheA P3 domain, CheW-CheR$_{like}$ protein, and TDE2496 in *T. denticola*

601   were PCR amplified from *Td* genomic DNA using a forward oligonucleotide encoding an NdeI

602   restriction site and a reverse primer encoding an EcoRI restriction site. The CheA protein was

603   amplified using a forward oligonucleotide encoding an NdeI restriction site and a reverse primer

604   encoding an BamHI restriction site. The PCR products were treated with the appropriate restriction

605   enzymes, purified, and ligated into a pet28a plasmid with a poly-Histidine tag and kanamycin

606   resistance marker. The plasmids were transformed into *Escherichia coli* BL21-DE3 cells and 4-8 L of

607   cell culture were grown at 37°C until an O.D. of 0.6 was reached. The flasks were cooled to 21°C and

608   1 mM of IPTG was added to the culture. The cells were harvested after 16 hours of growth. The cells

609   were lysed via sonication in lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM Imidazole) while

610   cooled on ice. The lysate was centrifuged at 20,000 X G for 1 hour at 4°C. The lysate was then run

611   over a gravity-flow purification column containing 3 ml of Nickel-NTA resin. The resin was washed

612   with 10 ml wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 20 mM Imidazole) and the protein was

613   eluted with 10 ml elution buffer (50 mM Tris pH 7.5, 150 mM NaCl, 200 mM Imidazole) and collected

614   in 1 ml fractions. The fractions were assessed for protein concentration via Bradford reagent and the

615   fractions containing protein were run on a size-exclusion s75 and s200 column systems that

616   monitored absorbance at 280 nm and collected 6 ml fractions. Fractions that contain CheA were

617   concentrated to ~20 mg/ml via centrifugation in a protein concentrator containing a regenerated

618   cellulose filter with a 50 kDa molecular-weight cut-off (MWCO). Fractions that contain CheA P3,

619   CheW-CheR$_{like}$ and TDE2496 were concentrated with a 10 kDa MWCO filter to 32 mg/ml, 11 mg/ml

620   and 7 mg/ml, respectively. The protein solutions were aliquoted, flash frozen in liquid nitrogen, and

621   stored at -80°C. For CheW-CheR$_{like}$, CheA, and CheA P3 domain, the purifications were prepared at

622   ambient temperatures. For TDE2496, the purification was prepared at 4°C.

623

Radioisotope assays

625   23 μl samples containing 2 μM CheA alone, or 2 μM CheA and 2 μM CheW-CheR$_{like}$, or 2 μM CheA

626   and 2 μM CheW-CheR$_{like}$ with TDE2496 (2 μM or 12 μM), were incubated in 50 mM MOPS pH 7.5,

627   150 mM KCl, 10 mM MgCl$_2$ for 30 minutes at ambient temperatures. Phosphorylation of CheA was

628   initiated by the addition of 2 μl of a solution containing 1 mM ATP mixed and 2-10 μl radiolabeled γ-

629   P$_{32}$ATP (3000 Ci/mmol, 10 μCi/μL; Perkin Elmer) and quenched with 25 μl of 3X LDS buffer

630   containing 100 mM EDTA pH 8.0 after 1-12 minutes. The samples were run on a native Tris-glycine

631   gel for 2 hours at 120 volts. The gels were dried, placed in a radiocassette for 24 hours, and then

632   imaged with a Typhoon phosphor-imager (GE Healthcare). The intensity of the radio-labeled protein

633   bands were quantified using ImageJ.

634

Quantification of cell curvature

636   The cell curvature of *Td* whole cells and *V. cholera* minicells was quantified by analyzing images of

637   cross-sections of the cells where top views of chemotaxis arrays are present. For *Td* cells, the inner

638   membrane curvature and CheA:CheW base-plate curvature was quantified. For *Vc* minicells, the

639   inner membrane curvature was quantified. These images were pre-processed with Fiji by placing

640   points along the desired area with a distance of 10 nm between each point. The curvature of the inner

641   membrane was measured with a pre-built plugin for python, called Sabl_mpl, written by Jewett, A.

642   from the Jensen lab (Pasadena, CA)[23]. The "measure 3-point curvature" function was used to select

643   three adjacent points along the inner membrane of the cell and calculate the radius of these points.

644   The radius of the three selected points allowed for the calculation of the local curvature by dividing 1

645   with the measured radius ($1/R = c$). This was repeated for all points with a "sliding window" approach,

646   where the second point of the initial three points would become the first point, until the desired area

647   was covered.

648

649   **Quantification of array alignment to the cell axis**

650   The angle between the strands of CheA:CheW rings and the *Td* cell axis was quantified using Image

651   J software. 2D images from reconstructions that clearly locate the orientation of the strands and cell

652   axis were chosen for analysis. First, a straight line was drawn from the cell pole down the axis of the

653   cells. Then, a second line was drawn through one of the strands in the array and the angle between

654   the two intersecting lines was quantified. In some cases, the angle was too small (<3°) to be

655   accurately determine so the angle was annotated as 0°.

656

657   **Residue conservation and molecular modeling of *Td* CheW, CheA P5, and the CheW domain of**

658   **CheW-CheR$_{like}$**

659   The protein sequences of the two *Td* CheW domains and *Td* P5 were aligned (Clustal Omega),

660   conservation was calculated based on the alignment (JalView[9]), and the highest variable regions

661   were selected based on conservation (10+ adjacent residues with conservation scores lower than 8).

662   Homology models of *Td* CheA P5, CheW and the CheW domain of CheW-CheR$_{like}$ were generated

663   via the Swiss-Model server using complete residue sequences of each protein as a target[24]. The

664   CheW protein from *Thermoanaerobacter tengcongensis* (*Tt*) (PDB ID: 2QDL) had the highest percent

665   identity to the CheW proteins (37% and 31%) and was therefore used as the structural template. The

666   resulting homology models for *Td* CheW and the CheW domain of CheW-CheR$_{like}$ had a QMEAN of -

667   1.3 and -1.2, respectively. The P5 structure from *Escherichia coli* produced the best homology model

668   for *Td* CheA P5, with QMEAN -0.98 (PDB ID: 6S1K). The homology models were then aligned to the

669   CheW protein and P5 domain in a crystal structure containing *Thermotoga maritima* CheW and CheA

670   P4P5 in complex using PyMol (PDB ID: 3UR1).

671

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

673   The isolated P3 domain was concentrated to 32 mg/ml and crystallized via hanging drop in 0.1 M

674   Imidazole pH 7.0, 25% PEG 400 using a 1:1 ratio of protein solution to crystallization solution with a

675   final volume of 3 µl. Crystals were apparent within eight hours but increased in size over three days.

676   Crystals were manually picked up in loops, flash cooled and shipped in liquid nitrogen to a beamline

677   (APS, line NE-CAT 24-ID-C, Dectris Pilatus 6M-F Pixel Array detector). The crystals diffracted to ~1.3

678   Å with a C2 symmetry and data was cut-off at 1.5 Å. The diffraction data was scaled and integrated

679   using XDS[17], and phased by molecular replacement with Phaser MR using *ab initio* search models

680   generated through the QUARK server and then ran through the AMPLE pipeline on the CCP4 web

681   server[18–20]. Model improvement was done by several rounds of manual model improvement in COOT

682   followed by automated refinement using Phenix Refine software[21,22].

683

684   1.   Ohta, K., Makinen, K. K. & Loesche, W. J. Purification and characterization of an enzyme

685   produced by Treponema denticola capable of hydrolyzing synthetic trypsin substrates. *Infect.*

686   *Immun.* **53**, 213–20 (1986).

687   2.   Muok, A. R. *et al.* A di-iron protein recruited as an Fe[II] and oxygen sensor for bacterial

688   chemotaxis functions by stabilizing an iron-peroxy species. *Proc. Natl. Acad. Sci.* **116**, 14955–

689   14960 (2019).

690   3.   Goetting-Minesky, M. P. & Fenno, J. C. A simplified erythromycin resistance cassette for

691   Treponema denticola mutagenesis. *J. Microbiol. Methods* **83**, 66–68 (2010).

692   4.   Kurniyati, K. & Li, C. pyrF as a counterselectable marker for unmarked genetic manipulations

693   in Treponema denticola. *Appl. Environ. Microbiol.* **82**, 1346–1352 (2016).

694   5.   Gumerov, V. M., Ortega, D. R., Adebali, O., Ulrich, L. E. & Zhulin, I. B. MiST 3.0: an updated

695   microbial signal transduction database with an emphasis on chemosensory systems. *Nucleic*

696   *Acids Res.* **48**, D459–D464 (2020).

697   6.   Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny

698   substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996 (2018).

699   7.   Ortega, D. R. & Jensen, G. J. Regular Architecture (RegArch): A standard expression

700   language for describing protein architectures. *bioRxiv* 679910 (2019). doi:10.1101/679910

701   8.   Li, W., Jaroszewski, L. & Godzik, A. Clustering of highly homologous sequences to reduce the

702   size of large protein databases. *Bioinformatics* **17**, 282–283 (2004).

703   9.   Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version

704   2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–

705   1191 (2009).

706   10.   Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large

707   phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

708   11.   Hodcroft, E. TreeCollapserCL4: Collapse Trees by Bootstrap. Available at:

709   http://emmahodcroft.com/TreeCollapseCL3.html.

710   12.   Crooks, G., Hon, G., Chandonia, J. & Brenner, S. WebLogo: A Sequence Logo Generator.

711   *Genome Res* **14**, 1188–1190 (2004).

712   13.   Mastronarde, D. N. Dual-axis tomography: An approach with alignment methods that preserve

713   resolution. *J. Struct. Biol.* **120**, 343–352 (1997).

714   14.   Xiong, Q., Morphew, M. K., Schwartz, C. L., Hoenger, A. H. & Mastronarde, D. N. CTF

715   determination and correction for low dose tomographic tilt series. *J. Struct. Biol.* **168**, 378–387

716   (2009).

717   15.   Castaño-Díez, D., Kudryashev, M., Arheit, M. & Stahlberg, H. Dynamo: A flexible, user-friendly

718   development tool for subtomogram averaging of cryo-EM data in high-performance computing

719   environments. *J. Struct. Biol.* **178**, 139–151 (2012).

720   16.   Castaño-Díez, D., Kudryashev, M. & Stahlberg, H. Dynamo Catalogue: Geometrical tools and

721   data management for particle picking in subtomogram averaging of cryo-electron tomograms.

722   *J. Struct. Biol.* **197**, 135–144 (2017).

723   17.   Kabsch, W. *et al.* XDS. Acta Crystallogr. Sect. D Biol. Crystallogr. **66**, 125–132 (2010).

724   18.   Bibby, J., Keegan, R. M., Mayans, O., Winn, M. D. & Rigden, D. J. AMPLE: A cluster-and-

725   truncate approach to solve the crystal structures of small proteins using rapidly computed ab

726   initio models. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 1622–1631 (2012).

727   19.   Keegan, R. M. *et al.* Exploring the speed and performance of molecular replacement with

728   AMPLE using QUARK ab initio protein models. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**,

729   338–343 (2015).

730   20.   Thomas, J. M. H. *et al.* Routine phasing of coiled-coil protein crystal structures with AMPLE.

731   *IUCrJ* **2**, 198–206 (2015).

732   21.   Adams, P. *et al.* The Phenix Software for Automated Determination of Macromolecular

733   Structures. *Methods* **292**, 342–351 (2009).

734   22.   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta*

735   *Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).

736   23.   Yao, Q. *et al.* Short FtsZ filaments can drive asymmetric cell envelope constriction at the onset

737   of bacterial cytokinesis. *EMBO J.* **36**, 1577–1589 (2017).

738   24.   Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: An automated protein

739   homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).

740

741
742
743
744
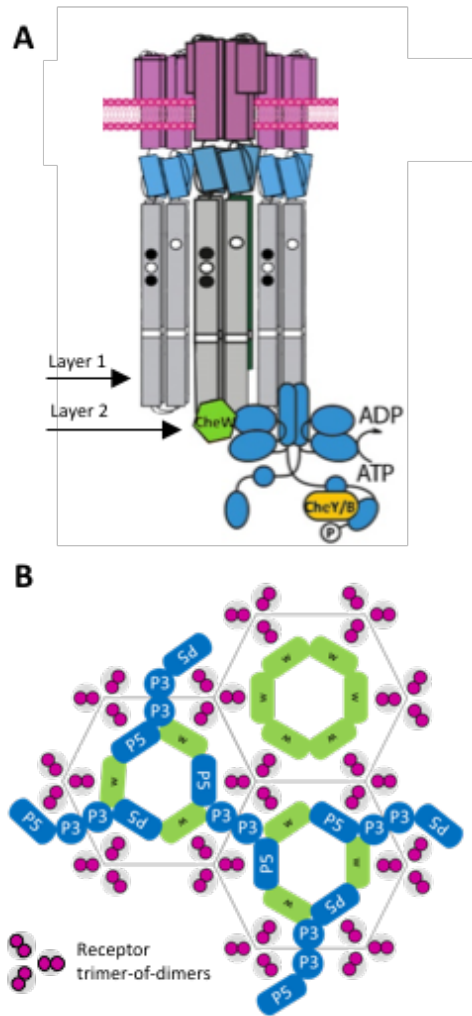
Fig. S1 The arrangement of the canonical chemoreceptor array, exemplified in *Escherichia coli* (*Ec*). (A) Transmembrane chemoreceptors form dimers that function as a trimer-of-dimers module to modulate the activity of the histidine kinase CheA (blue) via a coupling protein CheW (green). In cryo-ET experiments, cross-sections at Layer 1 reveals the arrangement of chemoreceptors and Layer 2 reveals the position of the kinase. (B) In all organisms investigated previously, CheA is integrated into the array with hexagonal (pseudo-P6) symmetry.

745
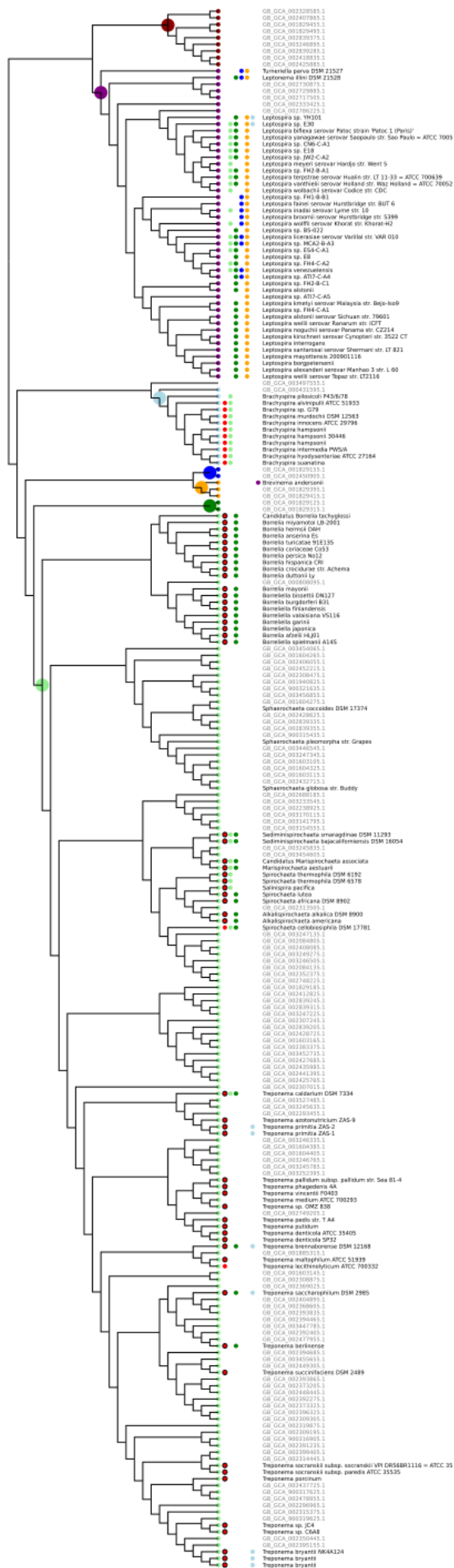
Fig. S2 The chemosensory class profile of Spirochaetota. Sequences belonging to the four major classes of Spirochaetota are marked: UBA4802 (dark red), Leptospira (purple), Brachyspirae (cyan), UBA6919 (blue), Brevinematia (yellow), GWE2-31-10 (green) and Spirochaetia (light green). Nodes are marked 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.
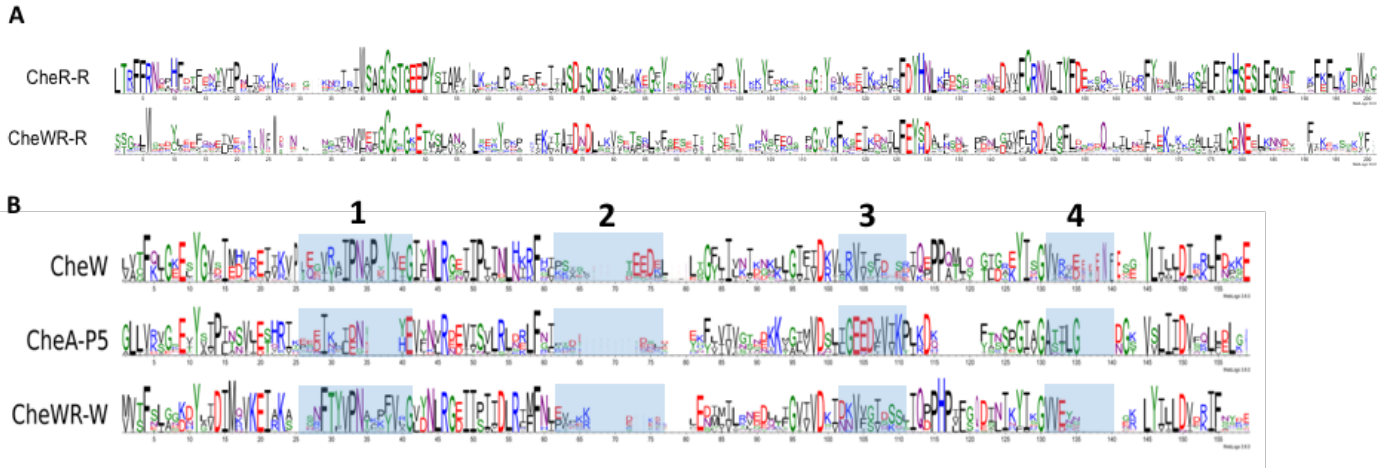
Fig. S3 (A) Sequence logo of representative sequences of the two groups of CheR-containing proteins in the F2 system. (B) Sequence logo of representative sequences of the three groups of CheW-containing proteins in the F2 system. Blue boxes denote the location of variable regions identified in the *Td* 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.
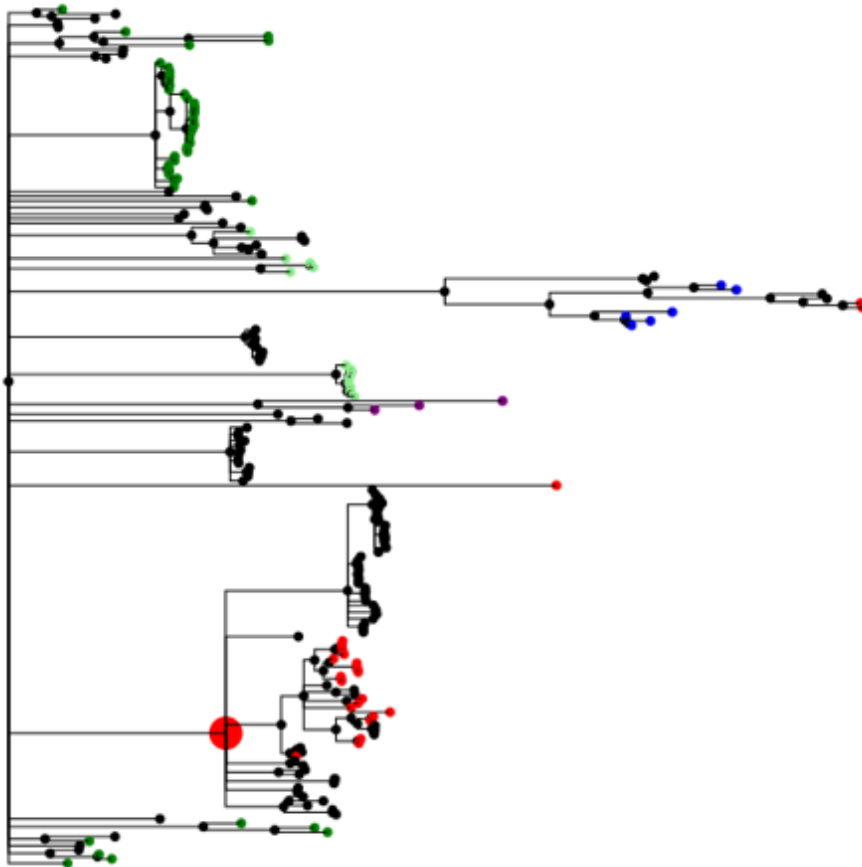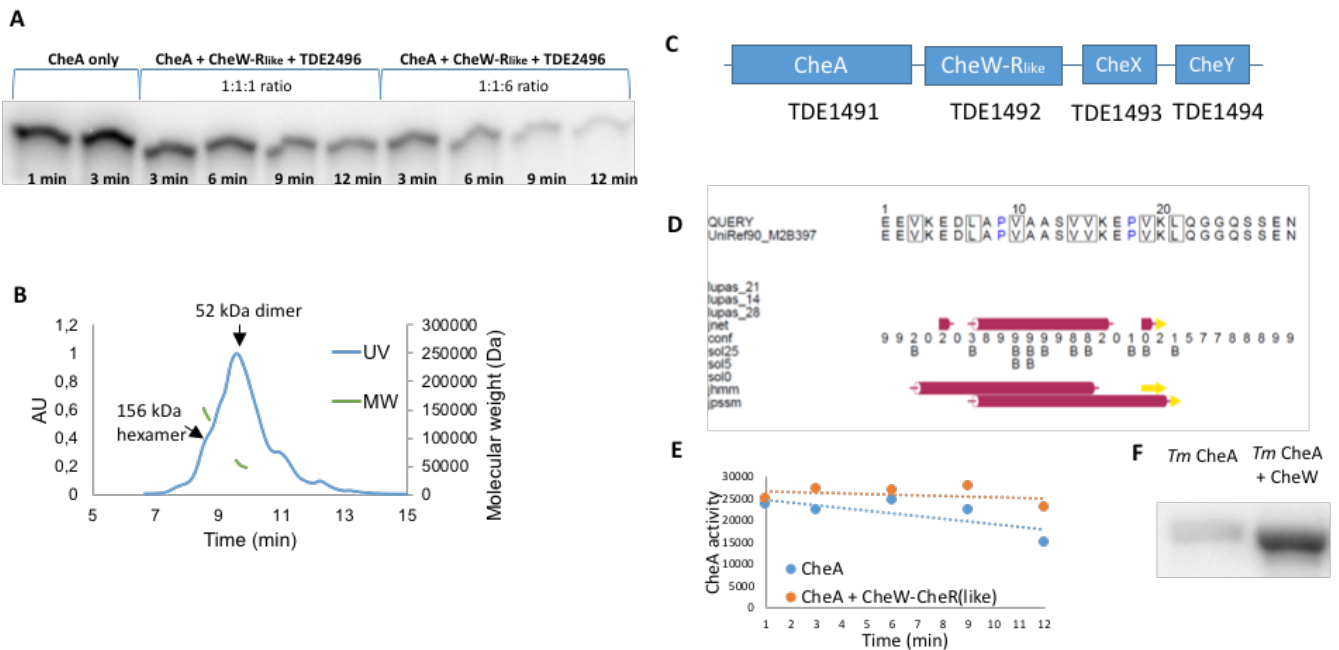

Fig. S4 Phylogenetic tree of CheW sequences in genomes with at least one CheA-F2 suggests a last common ancestor of CheW-F2 sequences. We mapped the sequences of CheW from the classes F2 (red), F5(purple), F7 (light green), F8 (green) and ACF (blue). The larger red internal node marks a candidate of the last common ancestor of CheW-F2.

A

Td strain	No. Cells/Arrays	No. Particles
WT	11	728
Δ2498	10	894
Δ2498Δ2496	10	554
ΔCheR$_{like}$	5	194

B

Fig. S5 (A) The number of arrays and particles selected to generate sub-tomogram averages for each *Td* strain. (B) Sub-tomogram averages reveal the orientation of the cell axis with respect to the orientation of the chemotaxis arrays. Intriguingly, the arrays have a preferred orientation in the cells and this orientation is conserved in WT and deletion mutants.

763
764

765
766 **Fig. S6** (A) The soluble receptor TDE2496 deactivates *Td* CheA in a concentration-dependent manner. (B) MALS
767 experiments demonstrate that TDE2496 forms the expected dimer (52 kDa) and hexamer (156 kDa) oligomeric states.
768 (C) The CheW-CheR$_{like}$ protein is conserved on the same operon as the only CheA, CheX, and CheY proteins in *Td*. (D)
769 The CheW-CheR$_{like}$ linker is predicted to forma single alpha helix flanked by unordered regions (Jpred). (E) Radioisotope
770 assays that monitor CheA autophosphorylation over time in the presence and absence of CheW-CheR$_{like}$ demonstrate a
771 difference in CheA activity. (F) In the *T. maritima* system, the presence of CheW also activates CheA kinase activity.
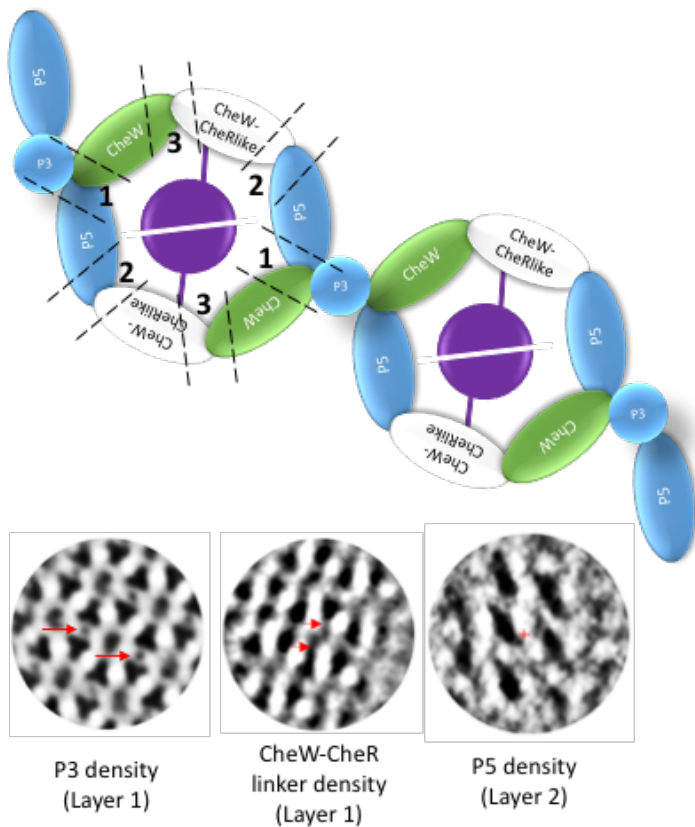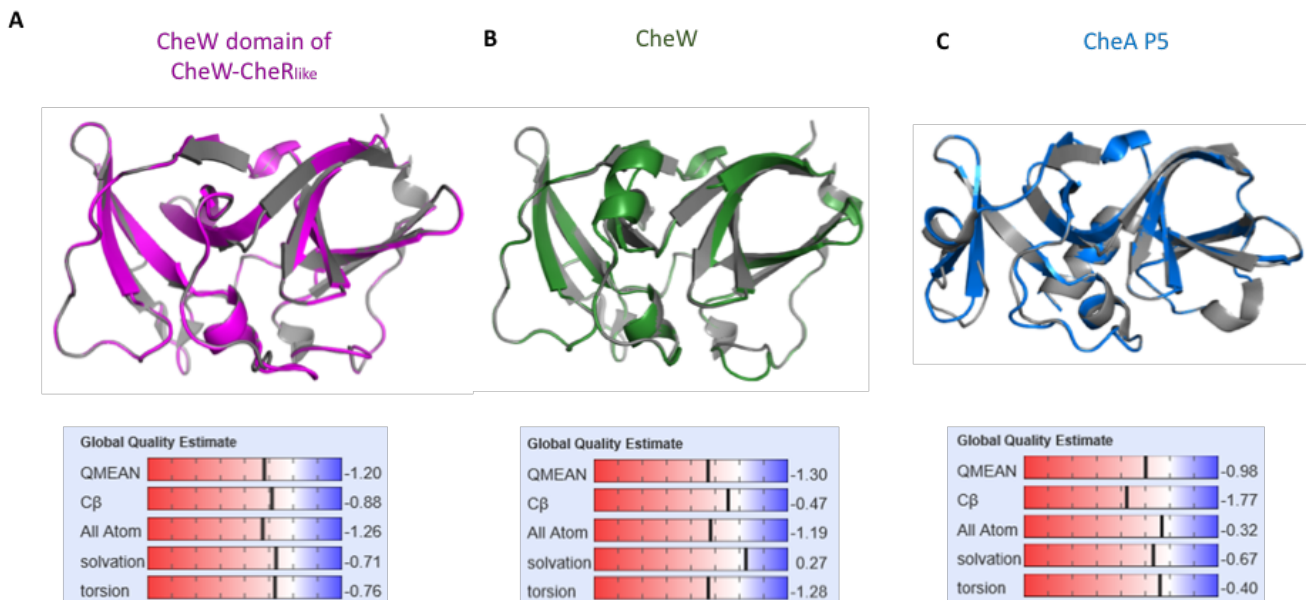Both samples are at time point 3 minutes.

772
773
774
775
776

Fig. S7 The arrangement of chemotaxis CheA:CheW rings in *Td.* 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.

P3 density (Layer 1)

CheW-CheR linker density (Layer 1)

P5 density (Layer 2)

777
778
779
780

Fig. S8 Homology models of *Td* CheW domains and CheA P5. (A) A homology model of the CheW domain of CheW-CheRlike using *Thermoanaerobacter tengcongensis* (*Tt*) CheW (PDB ID: 2QDL) as the template. (B) A homology model of the classical *Td* CheW using *Tt* 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.

781
782
783
784
785
786
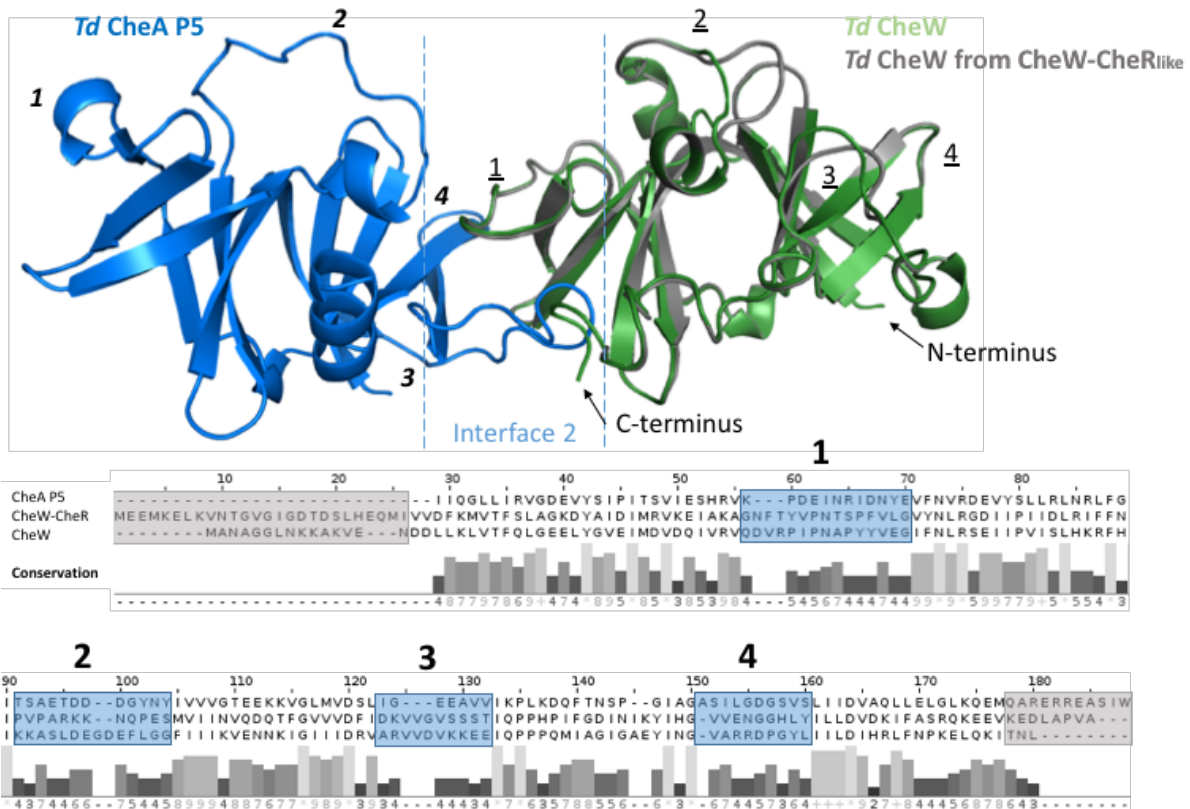
Fig. S9 Homology models of *Td* CheA P5, CheW, and the CheW domain of CheW-CheR_like mapped onto a previously determined crystal structure of *Thermotoga maritima* CheA P5 and CheW (PDB ID: 3UR1). Variable regions between the three domains (blue boxes) were determined by sequence alignments of the three 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.
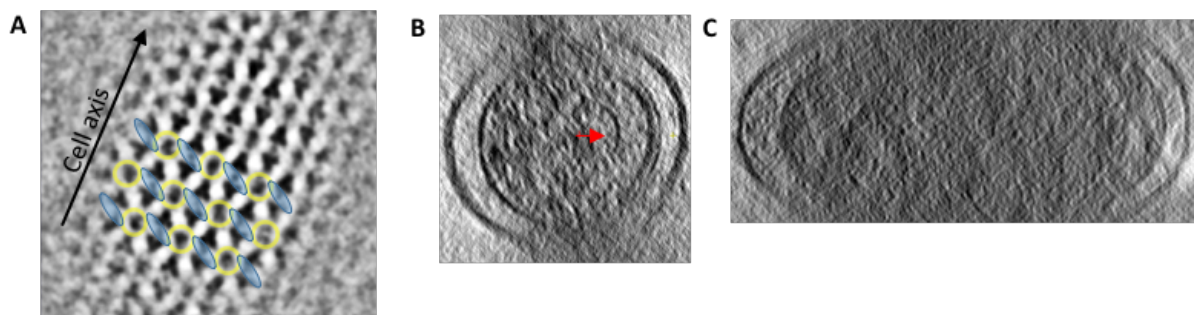
Fig. S10 (A) Sub-tomogram averaging of *Td* WT and mutant strains reveal that linked CheA:CheW rings (yellow) run perpendicular to the cell axis via a strict linear orientation of CheA (blue). Averages from WT *Td* are illustrated here but apply to all strains. (B) The curvature of the inner membrane 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). (C) *V. cholerae* minicells have an inner membrane curvature of 9.15 +/- 4.5 /μm (radius 1092 Å).
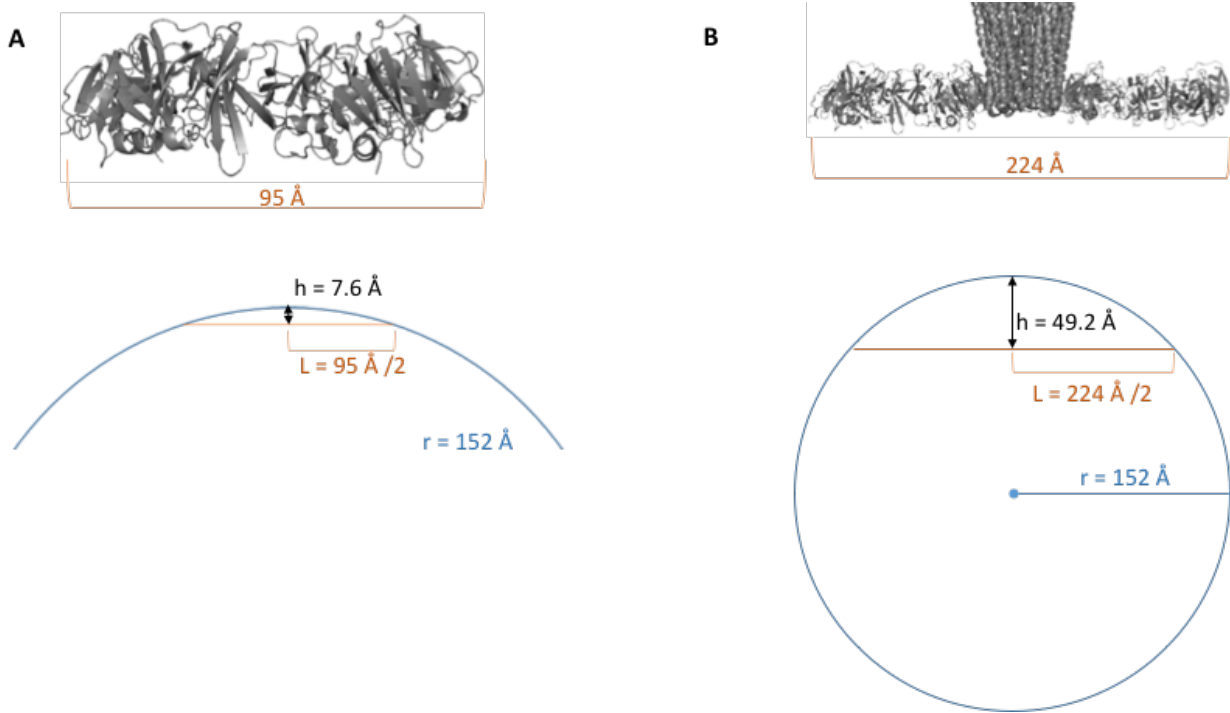
810
811 **Fig. S11** Modeling of CheA:CheW rings to the curvature of the *Td* baseplate. (A) For single CheA:CheW rings to
812 follow the baseplate curvature, it must buckle by an average of 7.6 Å toward the membrane. (B) In order for two
813 linked CheA:CheW rings to run perpendicular to the cell axis, the center of the rings (P3) must buckle by 49.2 Å
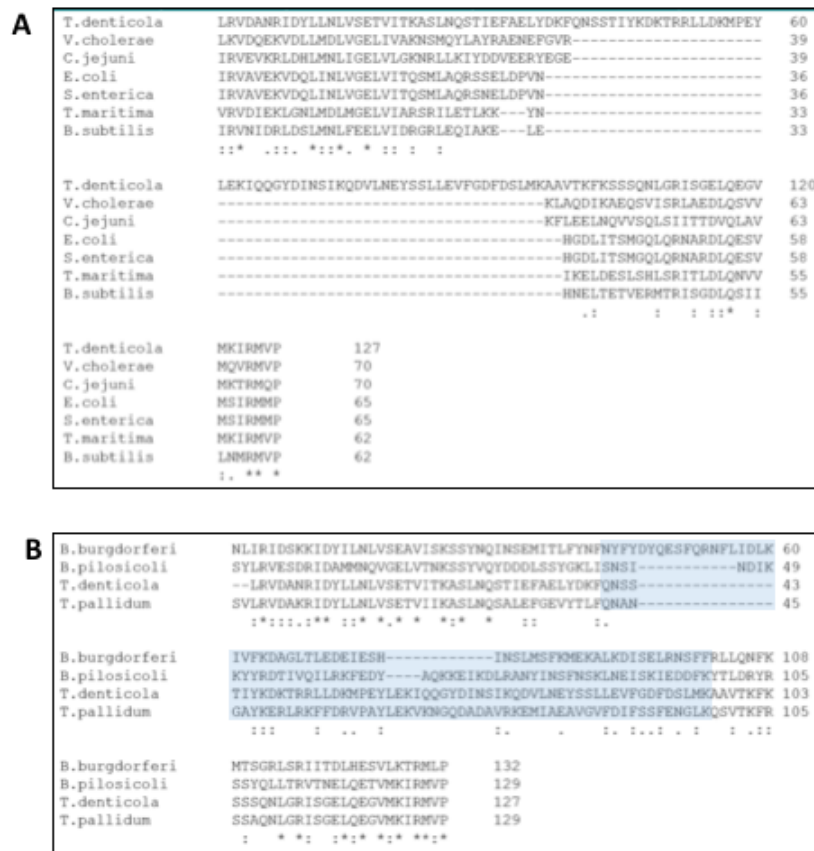814 toward the cell membrane.
815
816

Fig. S12 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.

817
818
819
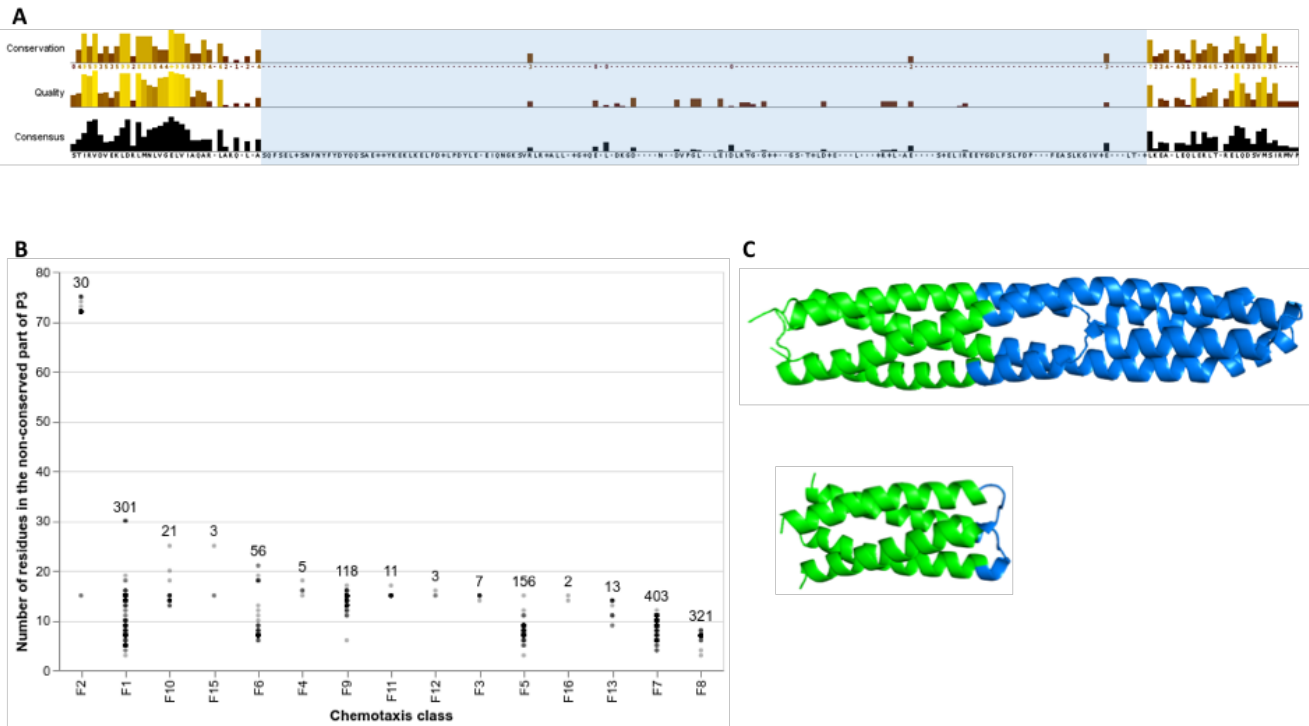
820

821 **Fig. S13** Analysis of non-redundant CheA P3 domains with a 75% sequence identity cut-off (1450 sequences). (A)
822 Conservation scores of the P3 alignment with the 1450 sequences. The traditional P3 helices are largely conserved, but
823 regions located between the helices are non-conserved. (B) Analyses of CheA from different chemotaxis classes
824 reveals that CheA F2 homologs possess the most residues in the non-conserved region of P3. (C) The non-conserved
825 regions are llustrated with the two known CheA P3 structures. Top: *Td* P3 possesses 71 residues that align to the non-
826 conserved region (PDB ID: 6Y1Y). Bottom: *Tm* P3 possesses seven residues that align to the non-conserved region
(PDB ID: 1B3Q).

827

828

829

830

831
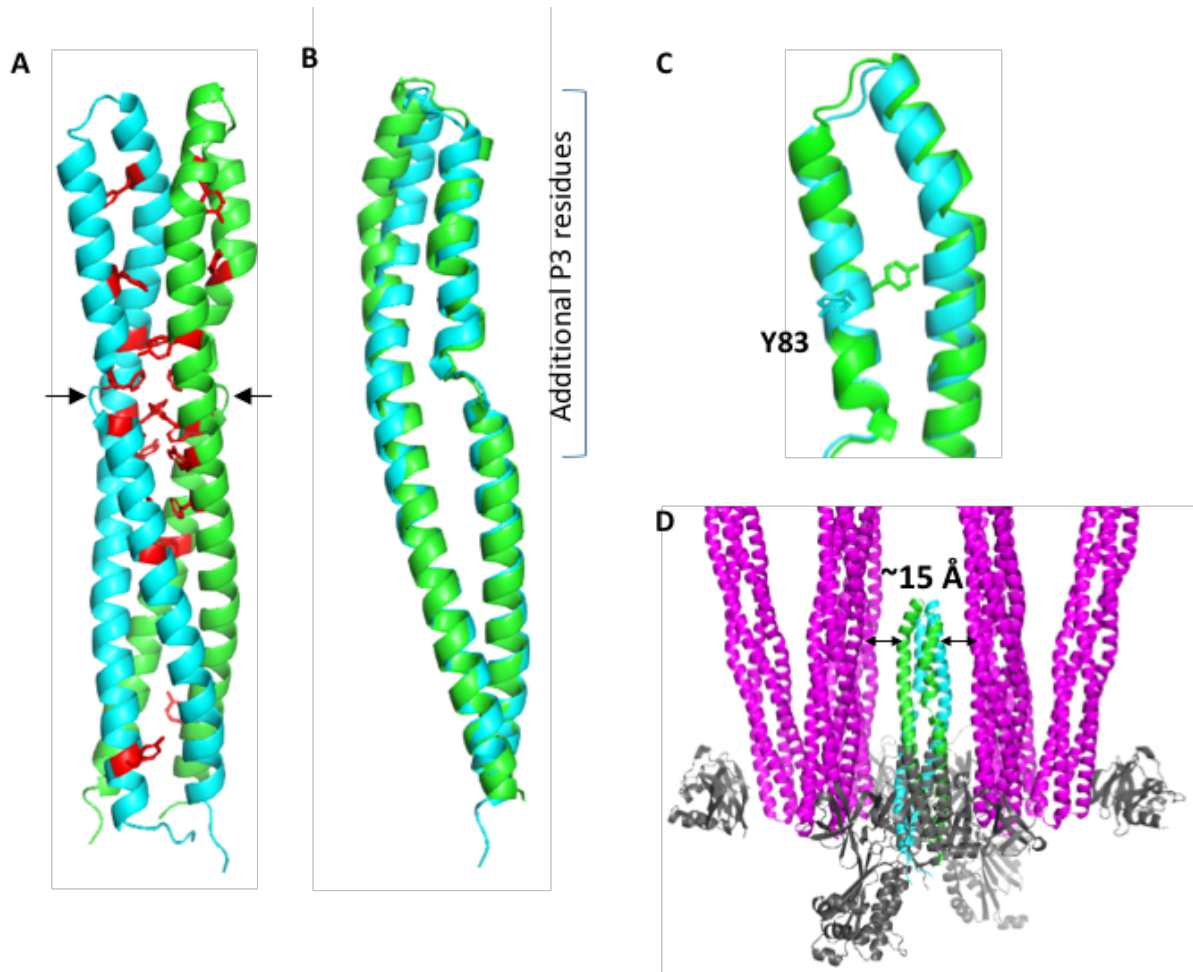
Fig. S14 The P3 domain of *Td* CheA. (A) The CheA P3 dimer contains a cluster of Phe and Tyr residues near the breakages in the helices (black arrows). All Phe and Tyr residues are highlighted in red. (B) The crystal structure of CheA P3 demonstrates asymmetry in the subunits. (C) Repositioning of Y82 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.
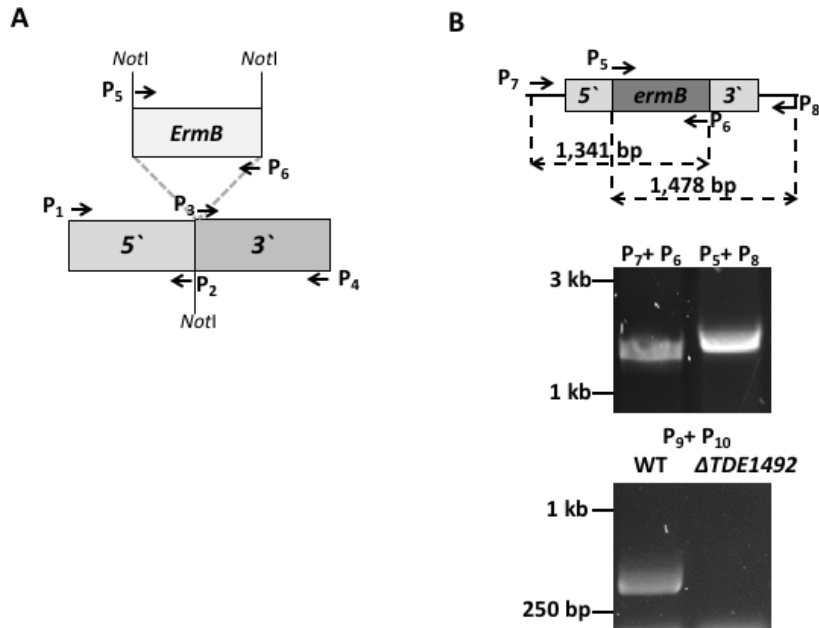
Fig. S15 Diagrams illustrating construction of the *TDE1492::ermB* vector (A) for the targeted mutagenesis of *TDE1492 (781-1,308 nt)* by in-frame replacement of *TDE1492* using *ermB* cassette**.** These constructs were constructed by two-step PCR followed by DNA cloning. Arrows represent the relative 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.

867
868
869
870
871

Fig. S16 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 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 S1A. Statistics for the average angle between the *Td* cell axis and the 'strands' of CheA:CheW rings.

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 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**

872 **Table S1C.** Statistics of inner membrane and baseplate curvature for 10 WT *Td*
873 cells.
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904

	Td P3 domain (PDB: 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	17.2
solvent	28.2

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	TTACAATATCACCTGAGCTGGCGGCCGCTTAAGGTGACAAGAAAGATGATA	5` portion for TDE1492 inactivation; [R]
P_3	TATCATCTTTCTTGTCACCTTAAGCGGCCGCCAGCTCAGGTGATATTGTAA	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.

911
912
913
914
915