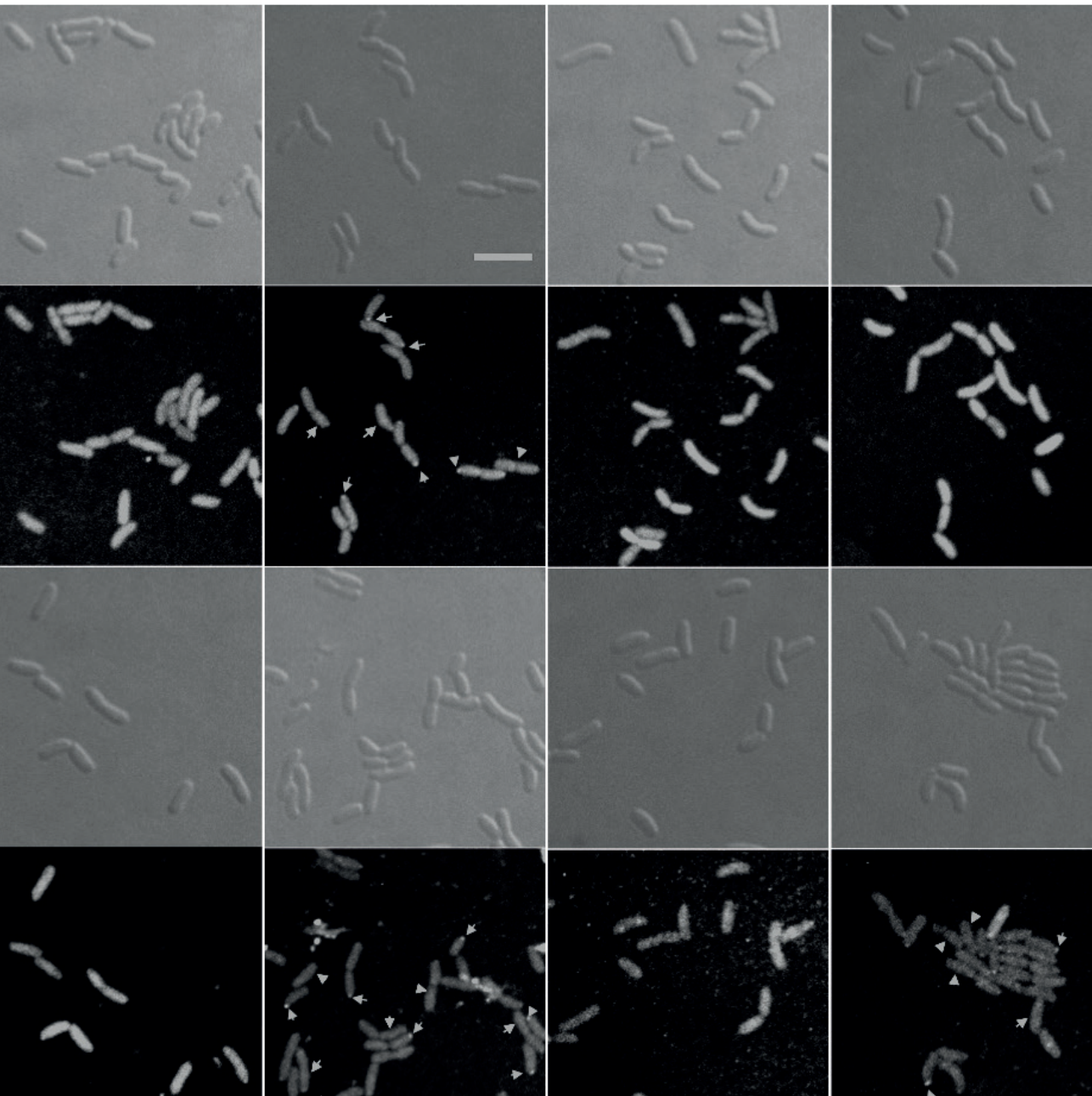
Cryo electron tomography studies of bacterial chemosensory arrays

Yang W.

CHAPTER 3



Baseplate variability of chemosensory arrays in *Vibrio cholerae*

This chapter is published as:
Yang W*[1], Alvarado A*[2], Glatter T[3], Ringgaard S[2], Briegel A[1] (2018) Proceedings of the National Academy of Sciences 115(52):13365-13370.

1 Institute of Biology, Leiden University, Leiden, The Netherlands
2 Department of Ecophysiology, Max Planck Insitute for Terrestrial Microbiology, Marburg, Germany
3 Core Facility of Mass Spectrometry and Proteomics, Max Plankc Insitute for Terrestrial Microbiology, Marburg, Germany
* These authors contribute equally to this work

Abstract

The chemoreceptor array, a remarkably ordered supramolecular complex, is composed of hexagonally packed trimers-of-receptor-dimers networked by a histidine kinase and one or more coupling proteins. Even though the receptor packing is universal among chemotactic bacteria and archaea, the array architecture has only been extensively studied in selected model organisms. Here, we show that even in the complete absence of the kinase, the cluster II arrays in *Vibrio cholerae* retain their native spatial localization and the iconic hexagonal packing of the receptors with 12 nm spacing. Our results demonstrate that the chemotaxis array is versatile in composition, a property that allows auxiliary chemotaxis proteins, such as ParP and CheV, to integrate directly into the assembly. Along with its compositional variability, cluster II arrays exhibit a low degree of structural stability compared to the ultrastable arrays in *E. coli*. We propose that the variability in chemoreceptor arrays is an important mechanism that enables the incorporation of chemotaxis proteins based on their availability.

Introduction

Chemotactic bacteria constantly assess their environment and compare their current situation with that of the recent past in order to bias their movements towards favorable surroundings. Changes in their environment are typically detected by transmembrane receptors known as methyl-accepting chemotaxis proteins (MCPs). Nutrients, toxins and biological signaling molecules bind to the periplasmic domains of the receptors either directly or via periplasmic binding proteins (PBPs) (68, 138, 139). MCPs communicate their binding state through the membrane to a HAMP (histidine kinases, adenylyl cyclases, MCPs, and some phosphatases) domain, which in turn communicates with their cytoplasmic tips to regulate the autophosphorylation of a histidine kinase, CheA. The kinase transfers phosphoryl groups to a response regulator, CheY, which then diffuses throughout the cell, binds to flagellar motors, and regulates their direction of rotation (140). In *E. coli*, either an increase of attractants or a decrease of repellents will lead to a reduced CheY-P production. In the presence of low CheY-P levels, the flagellar motors rotate counterclockwise, which results in a smooth-swimming motion ("run"). Increase of CheY-P promotes a change of spinning direction to clockwise, resulting in a random re-orientation of the cell ("tumbling"). Adjustments in the length of runs cause cells to follow gradients toward attractants and away from repellents.

The arrangement of chemoreceptors in hexagonal arrays has been shown to be universal in both Bacteria and Archaea (51, 55). In *E. coli*, the chemoreceptors are networked by rings of CheA and CheW into hexagonally packed arrays with a 12 nm spacing (43, 44, 99, 100, 141). Each set of three CheA dimers that link one receptor hexagon together lie at the vertices of a larger hexagonal lattice with a spacing of 21 nm (58). This ordered arrangement results in ultrastable receptor arrays that retain their highly ordered architecture even after cell lysis (43, 54, 58, 95, 142).

V. cholerae contains 3 chemotaxis operons (clusters 1-3), each of which forms its own structurally distinct arrays. Of these, clusters I and III are only expressed under certain growth conditions (low oxygen and as a general stress response, respectively) (80, 143, 144). The proteins from cluster I form cytoplasmic arrays, in which two baseplate layers sandwich cytosolic chemoreceptors and are stabilized by a specialized receptor with two signaling domains (104, 105). Cluster III is structurally not well understood at present. Neither of these systems has been implicated in canonical chemotactic behavior, and their cellular function remains elusive. In contrast, cluster II has been shown to be responsible for chemotactic behavior under all tested conditions so far (82). The arrays formed by cluster II proteins are expressed under standard growth conditions as well as under conditions in which clusters I and III are also expressed

(51, 104). Electron cryotomography (ECT) has previously reveled a typical hexagonal packing with a 12 nm spacing in the cluster II arrays (51). The same hexagonal packing order was also found in both receptor layers in the cytoplasmic cluster I arrays as well (104). *V. cholerae* possesses 43 MCPs genes that are scattered throughout the genome. 38 of these MCPs belong to the 40-heptad class (40H) (91), and this class of receptors integrates into the cluster II array (70). The cluster II arrays are strictly localized to the bacterial cell poles by the ParC/ParP system. ParC acts as a polar determinant that directs localization of arrays to the cell pole via its cognate partner protein ParP. ParP integrates into the array via its C-terminal AIF-domain through interactions with MCPs and CheA. The integration of ParP stimulates array formation and prevents the release of chemotaxis proteins from already-formed arrays (112, 145, 146).

The fact that ParP is not part of the chemotaxis system in *E. coli* reveals a structural variability of the chemotaxis arrays among species. Here, we show that that the CheA/CheW-only array, as well as the ultrastability of receptor arrays, is characteristic for the chemotaxis system in *E. coli* rather than a universal feature shared among all chemotactic bacteria and archaea. We further show that, although the hexagonal packing of the chemoreceptors is universal among species, the baseplate that binds the receptor tips differs in its composition. Our studies reveal that the baseplate, as a variable structure, allows for the incorporation of alternative coupling proteins. This feature likely facilitates the exchange of different chemotaxis proteins within existing arrays in response to environmental cues. Our results on *V. cholerae* emphasize the importance of understanding the structure and function of chemoreceptor arrays in non-model organisms, especially in organisms in which highly versatile chemotaxis might be important for pathogenicity.

Results

CheA is not necessary for chemoreceptor array structure and formation

Because the baseplate of the chemoreceptor array of *V. cholerae* is different from the model organism *E. coli* (145), we set out to study the effect of baseplate composition on array formation, structure and stability. Because *V. cholerae* contains 3 chemotaxis operons, with one CheA encoded by each system, we constructed a triple *cheA* deletion mutant (Δvc1397 Δvc2063 Δvca1095). We considered it was important to delete the *cheA* gene from each chemotaxis cluster in order to avoid the possibility that CheA proteins of clusters I and III might substitute for CheAII in formation of cluster II arrays. Like the *cheA2* single-deletion mutant, this CheA-free strain was no longer capable of chemotaxis in soft agar assay (Fig.S3.1).

In order to investigate whether the receptors were still able to form ordered hexagonal arrays in the complete absence of all three CheA proteins, we performed electron cryotomography (ECT) analysis of this CheA-free mutant. The data revealed that arrays were still present, but with a 40% reduction in the number of cells with observable arrays compared to the wild-type. This result supports the importance of CheA in normal array formation. Notably, subtomogram averaging showed that, even in the complete absence of CheA, the chemoreceptors were still packed in a hexagonal order with the typical 12 nm spacing and were indistinguishable in structure from the wild-type Cluster II arrays (Fig. 1 A-B).

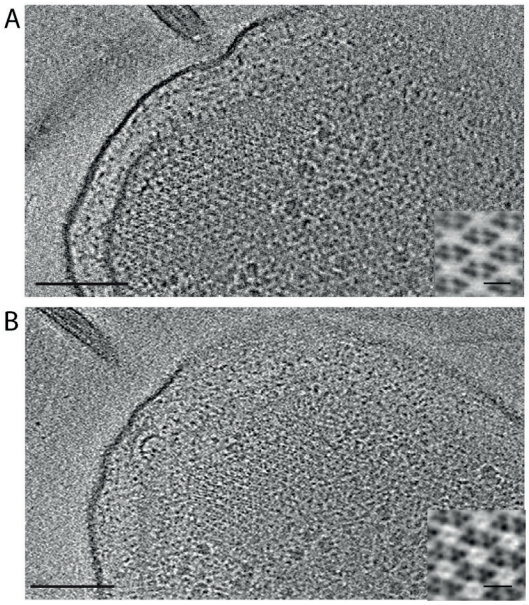


Figure 1. Cluster II chemoreceptor array observed in wild-type and CheA-free *V. cholerae*. (A) Hexagonally packed chemoreceptor array with 12 nm spacing near the flagellar pole in wild type *V. cholerae*. (B) Hexagonal receptor array with 12 nm spacing in a CheA-free cell. Subtomogram averaging map is shown in the inserts for both stains. The scale bar is 200 nm for both tomoslices and 10 nm for both inserts.

Previous reports, based on fluorescence microscopy (FLM) studies in *E. coli*, suggest that CheW alone might be sufficient for receptor clustering, so that chemoreceptor arrays might still form in the absence of CheA (147). However, to our knowledge, our study is the first substantial proof that the receptors not only cluster together, but that they indeed form the typical ordered arrays in the complete absence of CheA. Thus, even though CheA is indispensable for the chemotactic function of the array, it is not indispensable for array formation or the hexagonal packing of chemoreceptors in *V. cholerae*.

CheA is not necessary for array localization

To study the role of CheA protein in the localization of cluster II arrays, we analyzed array localization in both the wild-type and the triple *cheA* deletion strain using a functional YFP-CheW1 fusion (148) as a marker. Ectopic expression and visualization of YFP-CheW1 revealed, as previously described (112, 148), that in wild-type cells YFP-CheW1 localized as distinct clusters at the bacterial cell poles in a uni- and bi-polar manner (Fig. 2A). Additionally, demographic analysis showed that short cells contained a single uni-polar YFP-CheW1 cluster, whereas longer cells contained a bi-polar localization of YFP-CheW1 clusters (Fig. 2B). Particularly, YFP-CheW1 was uni-polarly localized in 53 % of cells and bi-polarly localized in 44 % of cells (Fig. 2C). YFP-CheW1 showed an identical cell length-dependent polar localization pattern in the triple cheA deletion background (Fig. 2A-B) and a similar ratio of cells with uni- (52 %) and bi-polar (42%) localization (Fig. 2C) when compared to wild-type cells. In direct support of the FLM results, all the CheA-free arrays observed by ECT were localized very near the flagellar basal body (Fig. S3.2). Thus, the correct localization of the array lattice does not depend on the presence of CheA. Given that either CheA2 or ParP alone was shown to be sufficient for CheW1 clustering, and thus for array formation (112), we believe ParP is still present in the CheA-free array and effectively assumes the role of CheA2 in facilitating array formation.

CheA does not influence recruitment of new chemotaxis proteins

To further understand the role of CheA in array formation, we performed fluorescence-recovery-after-photo bleaching (FRAP) experiments with YFP-CheW1 in wild-type and the triple *cheA* deletion background. Bleached clusters of YFP-CheW1 recovered their fluorescence intensity in the triple *cheA* deletion strain in the same manner as in the wild-type, demonstrating a continuous recruitment of new CheW1 to the chemotaxis arrays in both strains (Fig. 2D-E). This observation clearly shows that CheW1 does not depend on the presence of CheA to extend the pre-formed array lattice, which indicates that CheA does not play an indispensable role in recruiting new chemotaxis proteins into the existing arrays. Given that the formation of arrays is detected much less frequently in the total absence of CheA proteins, it confirms that CheA is more likely to play a crucial role in the initiation of array formation rather than regulating the incorporation of new baseplate protein in pre-formed arrays.

V. cholerae Cluster II arrays are highly unstable

In order to achieve an improved resolution (2-3.5 nm) of *E. coli* chemoreceptor arrays by ECT, several laboratories have applied cell lysis either by antibiotic treatment or by inducing a phage lysis gene to flatten the cells (54, 55, 58, 142). Both methods resulted in lysed cells with chemoreceptor arrays that clearly retained their architecture (Fig. 3A). In fact, it was reported the architecture of the arrays in lysed cells was indistinguishable

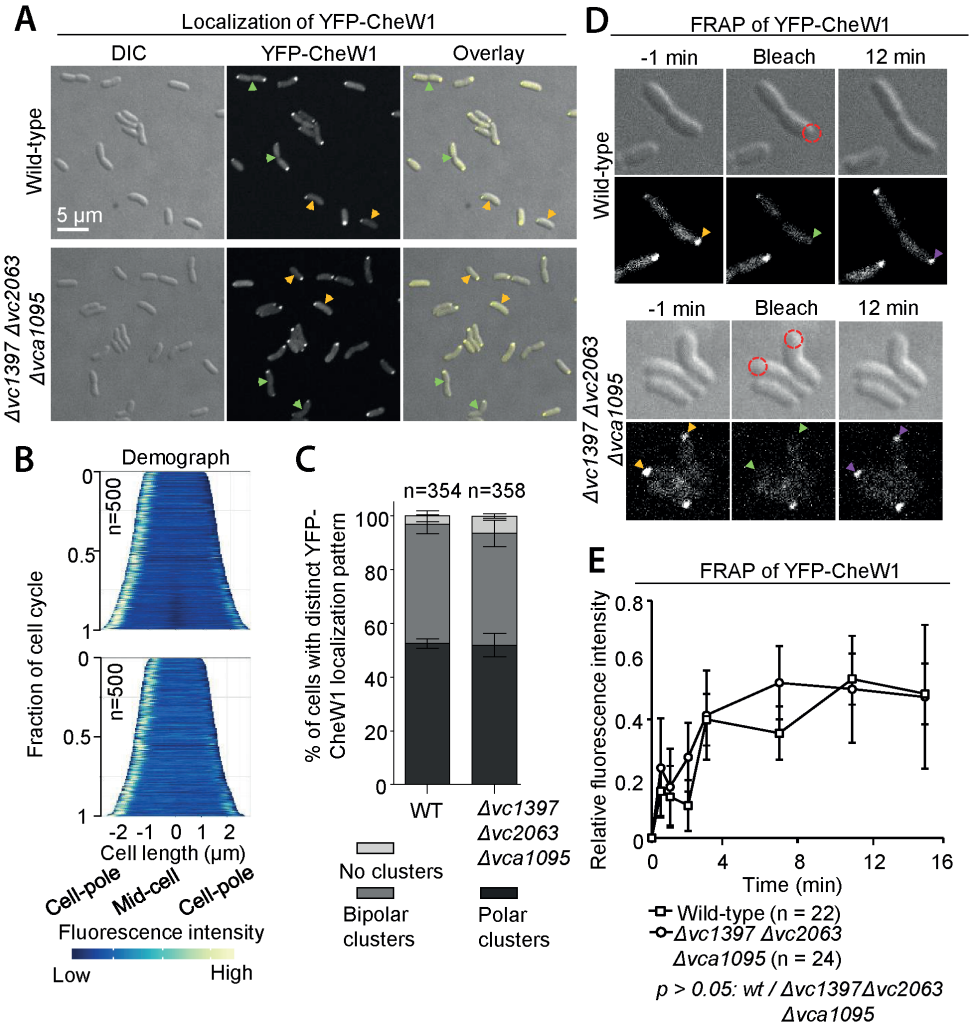


Figure 2. Formation of YFP-CheW1 clusters in the absence of CheA proteins. (A) Fluorescence microscopy images show the intracellular localization of YFP-CheW1 in wild type and the indicated *V. cholerae* mutants. Orange arrows indicate cells with uni-polar localization of YFP-CheW1, and green arrows indicate cells with bi-polar localization of YFP-CheW1. The scale bar represents 5µm. (B) Demographs showing the fluorescence intensity of YFP-CheW1 along the cell length in a population of *V. cholerae* cells relative to cell length. The n-value indicates the number of cells analyzed. (C) Bar graph showing the percentage of cells with distinct YFP-CheW1 localization patterns in the indicated *V. cholerae* strains. The n-value indicates the total number of cells analyzed from three independent experiments. (D) Fluorescence- recovery-after-photobleaching (FRAP) experiment of YFP-CheW1 clusters. Clusters recover post-bleaching in wild type and a strain deleted for all CheA proteins. Numbers indicate minutes pre- and post-bleaching. The dashed circle outlines the bleached region. Yellow arrows indicate the pre-bleaching clusters, green arrows indicate the bleached clusters and the purple arrows indicate clusters with a recovered YFP signal. (E) Graph depicting the

fluorescence intensity of YFP-CheW1. The YFP-CheW1 post-bleach signal intensity the cell pole is plotted relative to the initial pre-bleach at the pole intensity normalized to 1. The average fraction recovery is shown. In panels C and E, error bars indicate standard of the error mean (149).

from that of arrays analyzed in intact minicells (43, 44). However, instead of one continuous superlattice, the arrays are more often found in patches in lysed *E. coli*. In contrast, the same antibiotic-induced gentle cell lysis resulted in a quick loss of hexagonal packing in the *V. cholerae* chemoreceptor arrays (Fig. 3B). Even though the receptors clearly remain localized at the cell pole close to the flagellar motor, their ordered packing is almost completely disrupted. Similar to the array side views in lysed *E. coli* (Fig. 3C), as well as side views of cluster II arrays in intact *V. cholerae* (Fig. 3D), the receptors appear to be intact and clustered in lysed *V. cholerae* (Fig. 3E). The baseplate is still visible in the side views, suggesting that the occupancy of receptors' membrane distal end was not completely lost. However, instead of a continuous baseplate layer in intact cells, the density of the baseplate is discontinuous in the lysed *V. cholerae*. Tomographic results from lysed *V. cholerae* cells also revealed micelle-like zipper structures, where the receptors bend the inner membrane through the association of their membrane distal ends (Fig. 3F). Similar structures, which represent a different mode of receptor clustering, were frequently observed in *E. coli* when the chemoreceptors were disproportionally overexpressed relative to CheA and CheW (58, 150). Overall, these observations demonstrate that the cluster II array in *V. cholerae* does not exhibit an ultrastable structural integrity like the arrays in *E. coli*. It is noteworthy that this disruption of receptor packing order is only visible with ECT imaging and cannot be detected by fluorescence light microcopy. As a further support for the baseplate disruption, quantitative analysis of chemotaxis proteins in the membrane fraction of lysed *V. cholerae* cells clearly showed that baseplate components were not co-enriched with the membrane bound MCPs (Fig. S3.3).

The composition of the *V. cholerae* Cluster II array is variable and exhibits a distinct stoichiometry of chemotaxis proteins

To further understand the higher degree of instability of the *V. cholerae* cluster II arrays, we set out to determine the stoichiometry between the baseplate chemotaxis proteins CheW, CheA and ParP using targeted LC-MC proteomics on wild-type *V. cholerae* cells. Initial proteomic analysis was used to determine the synthetic heavy peptides used as standards for quantification of CheW1, CheA2 and ParP ratios. Peptide samples were spiked with identical amounts of the heavy peptides in order to calculate the relative ratio between the identical light peptides from CheW, CheA and ParP. The analysis revealed a stoichiometry between CheW1:CheA2:ParP of 35:5.3:1 (Fig. 4A), showing that CheW1 is highly abundant compared to CheA2 and especially ParP, indicating that

the cluster II baseplate is primarily composed of CheW1, to a lesser extend of CheA2, and with an even lower level of ParP.

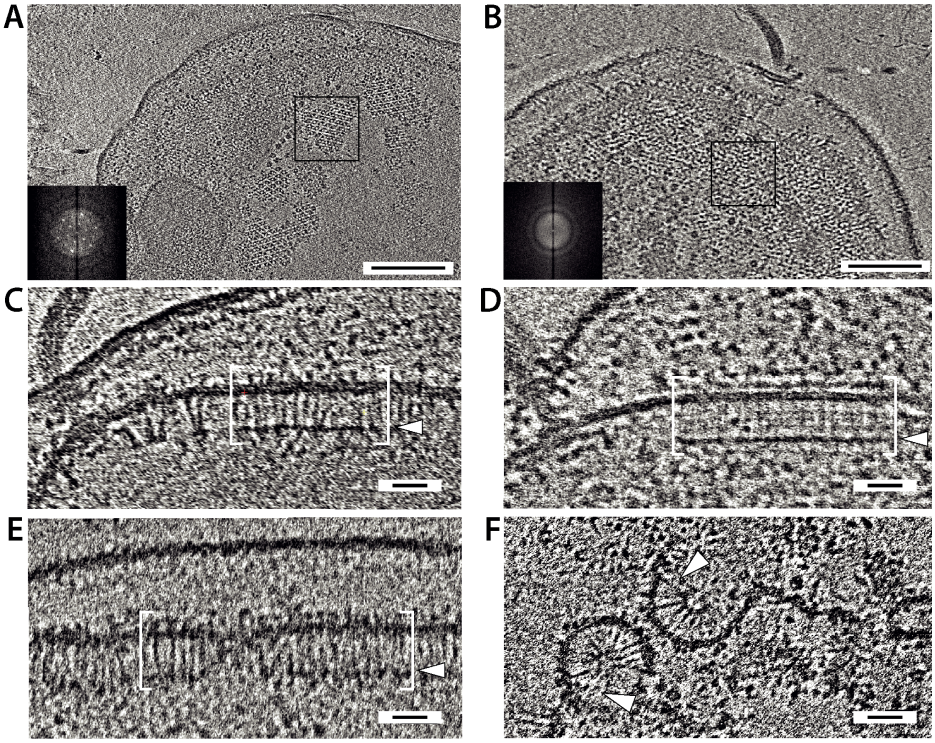


Figure 3. Chemoreceptor arrays in lysed *V. cholerae* and *E. coli*. (A) The hexagonal packing order can be clearly identified in a lysed *E. coli* cell. The insert is the power spectrum that displays a strong diffraction pattern in the boxed region of the receptor array. (B) Lysis disrupts the chemoreceptor array packing order in *V. cholerae*. Strong density representing the receptors is still clustered near the flagellar pole, yet no hexagonal order can be detected in either the tomoslice or the power spectrum (insert). Scale bar for A and B is 100nm. (C) Side view of chemoreceptor array in a lysed *E. coli* cell. (D) Side view of cluster II array in an intact *V. cholera* cell. (E) Side view of the cluster II array in lysed *V. cholerae* showing a discontinuous occupancy at the baseplate compared to the continuous density shown in panel C and D. White brackets and arrows highlight the receptor array and baseplate, respectively in panels C-E. (F) Tomoslice of a lysed *V. cholerae* cell where receptors associated through their membrane distal ends forming micelle- like structures. White arrows point at the receptors. Scale bar is 20 nm for panels C-F.

V. cholerae also encodes four predicted CheV proteins (VC1602, CheV1; VC2006, CheV2; VC2202, CheV3; and VCA0954, CheV4) in its genome. CheV is a fusion protein between CheW and CheY and integrates into the baseplate similar to CheW (151). Sequence alignment showed that all four CheVs possess the hydrophobic residues that mediate interaction between CheW, CheA-P5 and ParP-AIF, respectively, with the MCP

interaction tip (Fig. S3.4), suggesting that all four CheVs have the potential to integrate into the cluster II array baseplate. In a global proteomic analysis, we were able to detect the presence of all four predicted CheV proteins (Table S3.4), showing that they are all expressed under the conditions assayed. Furthermore, we detected all proteins from cluster II but none from Cluster I and III. This result suggests that CheV proteins are continuously expressed similar to cluster II proteins and thus have the potential to contribute to the structure of the baseplate of the cluster II array and consequently to array formation and/or stability.

We investigated the intracellular localization of all four CheVs by ectopically expressing CFP-tagged versions of each protein individually. Fluorescence microscopy showed that, in wild-type cells, only CheV2-CFP localized in clusters at the cell poles whereas CheV1-CFP, CheV3-CFP and CheV4-CFP were diffusely localized in the cytoplasm (Fig. 4B). Thus, CheV2 is the primary CheV integrated into cluster II arrays in wild-type cells. Targeted LC-MS analysis revealed that the stoichiometry of all the baseplate proteins, CheW1:CheA2:ParP:V1:V2:V3:V4, is 50:7.5:1.4:3.8:1:4.3:4.9 (Fig. 4D). CheW is clearly the most abundant of the baseplate proteins, followed by CheA2. Furthermore, CheV2 was the least abundant of all of the baseplate proteins, even though it appears to be the primary CheV protein in the cluster II arrays. In the absence of all CheA proteins, we observed that not only CheV2-CFP formed polar clusters in ~56% of cells (a level identical to that observed for wild-type cells), but also CheV4-CFP and CheV1-CFP were polarly localized in approximately ~57% and ~25% of the population, respectively (Fig. 4C). Thus, it appears that both CheV1 and CheV4 are able to integrate into the arrays in the absence of CheA proteins. This recruitment of different CheV proteins under certain conditions suggests that the baseplate is a highly variable structure and is capable of adjusting its composition in order to accommodate changes in the dynamic accessibility of different chemotaxis proteins. Indeed, CheVs are predicted to coordinate with certain receptors to integrate into the array and to modulate receptor function (110).

Discussion

The architecture of bacterial chemoreceptor arrays has been predominantly studied in the model organism *E. coli*, in which the structural core of the array is composed of rings formed by alternating P5 domains of CheA and CheW. These rings network the trimers of receptor dimers in the typical hexagonal lattice. The architecture of these arrays, in which six rings of CheA/CheW surround a ring lacking CheA predicts a stoichiometry of 1:1:6 (CheA:CheW:MCPs) in these array, or 1:2:6 if the CheA-less hexagons are not empty but instead filled with CheW only (43). This architecture

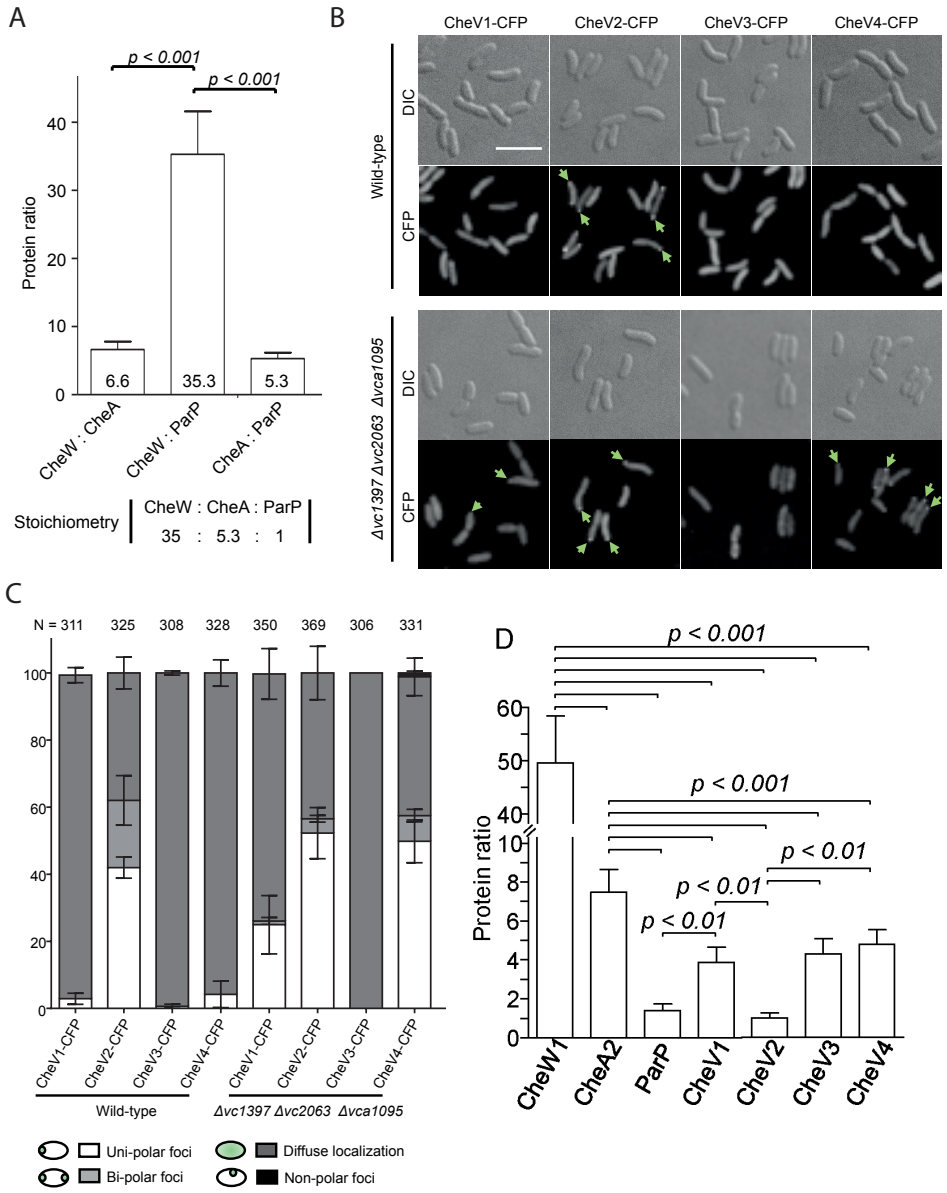


Figure 4. Stoichiometry and variability in the *V. cholerae* baseplate. (A) Bar graph showing the ratio between the chemotaxis proteins CheW1, CheA2 and ParP from cluster II determined by targeted LC-MS. Error bars indicate SEM. (B) Fluorescence microscopy of cells expressing the four CheVs in wild-type *V. cholerae* and a strain deleted for all three cheA genes. Scale bar indicates 5 µm. Green arrows indicate polar clusters of CheV proteins. (C) Bar graph depicting the percentage of cells with specific localization patterns of the four CheV proteins in wild-type *V. cholerae* and a strain deleted for all three cheA genes. (D) Bar graph showing the ratio between the chemotaxis protein CheW1, CheA2, ParP and CheV1-4 from cluster II array determined by targeted LC-MS. Error bars indicate standard deviation of the mean in panels A, E and F. P values were calculate by Student's t-test.

agrees with the experimentally determined protein ratios that have been published (94, 95, 152). It is possible that because of the flexible stoichiometry among the ternary components, the direct visualization of the CheW-only rings *in vivo* has not yet been reported. Meanwhile, in the recombinant array assembled *in vitro*, the array formation was commonly promoted with CheW in molar excess of CheA, typically at the ratio of 1:2 (103, 141). In those studies, such a high concentration of CheW may strongly favor CheW only rings in the CheA-less hexagons. As a result, ECT studies and subsequent subtomogram averaging indeed revealed CheW-only rings *in vitro* (141). However, it is still unclear whether all rings lacking CheA are filled up with 6 CheW monomers to create a complete ring *in vivo*. Our results with CheA-free *V. cholerae* provided an extreme case in which the majority of the receptor trimers-of-dimers bind to CheW at the baseplate. The baseplate composed predominantly of CheW is clearly sufficient to arrange the receptor in a hexagonal packing (Fig. 5A). This result certainly favors the possibility that CheW-only rings are sufficient to network the receptor trimers into the native hexagonal packing to form a super-lattice.

The observation of CheA-free arrays that we report here does not diminish the structural importance of CheA protein in the array. On the contrary, it highlights the significance of CheA for maintaining the overall structural integrity of chemoreceptor arrays. The CheA homodimer integrates into two neighboring receptor hexagons in the lattice, linked by its P3 dimerization domain (47). Thus, CheA dimerization is crucial for interlinking neighboring CheA/CheW rings and also for establishing the whole allosteric network of the complex (92). In *E. coli*, the ratio of CheA dimer to the total amount of coupling protein is 1:4, assuming that all CheA-free rings are filled with CheW. The high CheA occupancy ensures that each trimer of receptor dimers directly binds to a P5 domain of CheA. Consequently, the hexagonal packing of the receptors is guaranteed regardless of the presence of the CheW-only rings (Fig. 5C). Our stoichiometry data show that the ratio of CheA dimer to other coupling proteins in *V. cholarae* is 1:14, which is dramatically lower than in *E. coli*. Therefore, there are much fewer CheA dimers in the baseplate that can function as "structural staples" to interlink the rings, which are predominantly formed by the coupling protein CheW (Fig. 5B). It is worth mentioning that the low abundance of CheA does not eliminate the existence of complete CheA/CheW rings in cluster II arrays. Yet, this structure is likely to be interspersed in the lattice given the low CheA/CheW ratio, which would not provide an equivalent stability to the array architecture as extensively networked CheA/CheW rings do. Ultrastability of *in vitro* array complexes has been reported previously, which was proposed to arise from the multiple linkages between the individual core complexes (95). ECT studies later provided experimental evidence for the array stability based on the universal appearance of the chemoreceptor array lattice after cell lysis, which has been shown for example in *E. coli*, *B. subtilis* and *T. maritima* (43, 51, 142). The stoichiometry of CheW:

CheA: ParP : CheV in *V. cholerae* provides a plausible explanation for why cluster II arrays are not as stable as those of *E. coli*. We further expect a variance of the chemoreceptor array stability exist among other species in which the stoichiometry of the baseplate components deviates from the stoichiometry established in *E. coli*.

ParP from *V. parahaemolyticus* forms dimers (146) similar to those formed by ParP from *V. cholerae*, as suggested by a bacterial-two-hybrid assay (Fig. S3.5). ParP may also form dimers through the flexible linker between its C-terminal AIF domain and the N-terminal ParC interaction domain. Thus, ParP may substitute for the CheA homodimer in the baseplate instead of competing with CheW monomers. If so, then ParP is potentially capable of contributing to array stability in the same way as CheA dimers despite its comparatively low abundance.

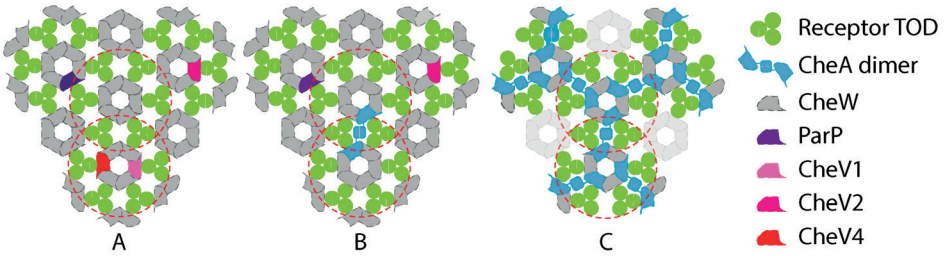


Figure 5. CheA dimer occupancy in different arrays. (A) CheA-free array in *V. cholerae*. Baseplate is predicted to be composed of CheW-only rings with sporadic insertion of ParP, CheV1 and CheV4. (B) Cluster II array in wild-type *V. cholerae*. CheA2 dimers serve as a structural "staple" that interlinks the ring networks in the baseplate. Due to the low occupancy of CheA2 dimer (1:14) compared to the other coupling proteins, only a few neighboring receptor hexagons are structurally interlinked. (C) *E. coli* array. The1:2 ratio of CheA dimer to CheW result a highly structured packing of receptors and base plate components. The red dashed circles outline individual receptor hexagons in the receptor array super-lattice.

The stoichiometry of the *V. cholerae* cluster II baseplate components reported here suggests a high level of dynamics in baseplate architecture during array formation. Due to a more diverse composition of the baseplate, it is conceivable that the array lattice develops at the cell pole through a dynamic recruitment of receptors that bind CheA, ParP, CheV and predominantly, CheW. This is in direct contrast to *E. coli*, in which the core units are thought to assemble first, and subsequently associate to form the extended receptor arrays with a strict stoichiometry and ordered arrangement of both the receptors and the baseplate protein CheA and CheW (58). This assembly process may reflect the relative simplicity of the *E. coli* chemotaxis system. The compositional diversity of baseplates can also be found in other organisms for which the ratios of chemotaxis proteins ratios have been determined. For example, in *Bacillus subtilis*, the baseplate also contains CheV, and the array stoichiometry was determined to be

1:1:3:23 (CheA:CheW:CheV:MCP) (115). These differences in protein ratios indicate that the protein arrangement in the baseplate varies significantly depending on the organism and perhaps the growth conditions, even though the receptors are packed in the typical hexagonal lattice (43, 51, 55). It seems that the architecture of the baseplate tolerates different levels of baseplate proteins. This conclusion is supported by the loss of ordered CheA distribution in arrays where the array proteins are overexpressed from different plasmids (58).

Taken together, our results emphasize that there are significant differences in the composition and architecture of the chemoreceptor array among species. In the case of *V. cholerae*, the less-stable arrays may be the result of the variable composition of the baseplate proteins. This finding is consistent with the high number of different MCPs encoded by *V. cholerae*. Their integration into the array super-lattice is believed to depend on the presence of specific coupling proteins (110). Additionally, the array variability may enable a swift adaptation by exchange of new chemoreceptors within the existing array lattice during in response to altered environmental conditions. In contrast, the *E. coli* system might have evolved to form ultrastable arrays, which provide a robust generic chemotaxis response without the need for adding or exchanging array components once assembled. Our findings highlight the need to study chemoreceptor array structure and composition in different organisms to properly understand the diversity and biological significance of chemotaxis signaling.

Materials and Methods

Construction of strain AA26

The wild-type strain of *V. cholerae* used was the El Tor clinical isolate C6706 and all mutants are derivatives or this strain. Strain AA26 was constructed by sequential in-frame deletions of vc1397, vc2063 and finally vc1095. To perform deletions, *E. coli* strain SM10λpir was used to transfer DNA into *V. cholerae* cells. Derivatives of suicide vector pCVD442 were employed to construct AA26 by standard allele exchange techniques. Stains and the plasmids and primers used for construction are listed as Table S3.1 - S3.3 in the supplementary information.

Construction of plasmids

Plasmid pAA100. Genomic DNA of *V. cholerae* was used to amplify *vc1602* gene with primers 3039 and 3024. Similarly, cfp was amplified from plasmid pMF391 using primers 3025 and 565. Both PCR products were used as template for a third PCR using primers using primers 3039 and 565. The resulting product was digested with enzymes XbaI and SphI and inserted in the corresponding sites of vector pBAD33, resulting in

plasmid pAA100.

Plasmid pAA101. The gene vc2006 was amplified using genomic DNA as template with primers 3027 and 3028. Using vector pMF391, *cfp* was amplified using primers 3025 and 565. Both PCR products were used as template for a third PCR using primers using primers 3027 and 565. The resulting product was digested with enzymes XbaI and SphI and inserted in the corresponding sites of vector pBAD33, resulting in plasmid pAA101.

Plasmid pAA102. To create this plasmid vc2202 was amplified from genomic DNA using primers 3029 and 3037. Vector pMF391 was used for template of *cfp* amplified using primers 3025 and 565. Both PCR products were used as template for a third PCR using primers using primers 3029and 565. The resulting product was digested with enzymes XbaI and SphI and inserted in the corresponding sites of vector pBAD33.

Plasmid pAA102. Similarly, *vca0954* was amplified from genomic DNA with primers 3031 and 3038. Vector pMF391 was used for template of *cfp* amplified using primers 3025 and 565. Both PCR products were used as template for a third PCR using primers using primers 3031 and 565. The resulting product was digested with enzymes XbaI and SphI and inserted in the corresponding sites of vector pBAD33 giving pAA102.

Growth conditions and media

V. cholerae and *E. coli* were grown in LB media or on LB agar plates at 30°C or 37°C containing antibiotics in the following concentrations: streptomycin 200 µg/ml; ampicillin 100 µg/ml; chloramphenicol 20 µg/ml for *E. coli* and 5 µg/ml for *V. cholerae*.

Cell Lysis

For lysis, *V. cholerae* and *E. coli* were first cultured overnight with 200 rpm shaking at 30 °C in LB and TB media, respectively. Overnight cultures were diluted with fresh media at 1:500, then incubated at 30 °C with shaking for another 3 hours. 2000 UI/ml penicillin was added to the culture when OD_{600} was reached. After an additional hour of incubation, cells were harvested by centrifugation. Lysis was monitored under the light microscope.

Fluorescence microscopy

Fluorescence microscopy was performed essentially as described earlier (80). *V. cholerae* cells were grown at 37 °C with shaking in LB medium including chloramphenicol for selection of plasmid pSR1033. Ectopic expression of YFP-CheW1 was induced by addition of L-arabinose to a final concentration of 0.2% v/v when the cell culture reached OD_{600} 0.8-1.0. Subsequently, cell cultures were incubated an additional hour. To collect images, cells were mounted on a 1% agarose pad, which included 10% LB

and 20% PBS. Images were collected using a Nikon eclipse Ti inverted Andor spinning-disc confocal microscope equipped with a 100x lens and an Andor Zyla sCMOS cooled camera. Acquisition settings (exposure time and laser intensity), were kept identical throughout all experiments. Images were analyzed using ImageJ imaging software (http://rsbweb.nih.gov/ij) and Metamorph Offline (v 7.7.5.0, Molecular Devices, Union City, CA). When counting the percentage of cells with distinct localization patterns of YFP-CheW1, three biological experiments were performed. Cells were enumerated by hand and for each experiment >100 cells were counted. The mean of the three experiments was then plotted with error bars indicating the standard-error-mean (79). A t-test was performed to calculate the p value. Demographic analyses were performed as previously described (79). Software R studio version 3.0.1 (http://www.rstudio.com/) along with ggplot2 version 0.9.3.1 (Hadley Wickham, Department of Statistics, Rice University) were employed. For demographic analysis, the data from three independent biological experiments were pooled. For each experiment >100 cells were analyzed. The total number of cells included in the demograph (n) is mentioned in the figure.

Determination of isotopically labeled reference peptides for CheA, CheW and ParP

Vibrio cholerae C6706 (lacZ-) was grown at 37 °C in liquid LB medium until OD_{600} reached ~0.5-0.7. Cells were pelleted by centrifugation and washed twice with PBS buffer (20%). Sample preparation, LC-MS and data analysis were carried out as described previously (153, 154) with the exception of the proteolytic digest, which was carried out using a tandem LysC/trypsin digest. Following protein solubilization using 2% Sodiumlauroylsarcosinate (SLS) the detergent concentration was diluted to 0.5% with 100 mM Ammoniumbicarbonate. To 50 µg total protein extract, 500 ng LysC (Wako) was added and incubated for 4 hours at 37 °C. Then digest was continued by adding 1 µg trypsin (Promega) for overnight at 30 °C. Detergent depletion and preparation of peptides for LC-MS analysis was carried out as described previously (153).

The peptides were then analyzed using liquid-chromatography-mass spectrometry carried out on a Q-Exactive Plus instrument connected to an Ultimate 3000 RSLC nano and a nanospray flex ion source (all Thermo Scientific). Peptide separation was performed on a reverse phase HPLC column (75 µm x 42 cm) packed in-house with C18 resin (2.4 µm; Dr. Maisch). The following separating gradient was used: 98 % solvent A (0.15 % formic acid) and 2 % solvent B (99.85 % acetonitrile, 0.15 % formic acid) to 32 % solvent B over 90 minutes at a flow rate of 300 nl/min.In order to identify the most suitable peptides for peptide standard synthesis, digests from *Vibrio cholerae* lysates were analyzed using data-dependent acquisition (153). In brief, the MS parameters were set as follows: MS1 resolution of 60 000 (m/z 200), scan range from 375 to 1500 m/z, MS/MS scans of the 10 most intense ions with 17 500 (m/z 200). The ion accumulation

time was set to 50 ms (both MS and MS/MS). The automatic gain control (AGC) was set to 3 × 106 for MS survey scans and 1 × 105 for MS/MS scans. MS raw files were searched using MASCOT (v 2.5, Matrix Science) with the following criteria: semitryptic tryptic specificity; two missed cleavages were allowed; carbamidomethylation (C) was set as a fixed modification; and oxidation (M) and deamidation (N,Q) were set as a variable modification. The mass tolerance was set to 10 ppm for precursor ions and 0.02. Search results were loaded into Scaffold 4 (Proteome software) and peptides were chosen as candidates for peptides synthesis when only fully cleaved and no miscleaved version of this peptide was observed.

Targeted Liquid chromatography-mass spectrometry (LC-MS)

Before solid-phase extraction, isotopically labeled reference peptides (TQL, JPT Peptide Technologies) for CheA, CheW and ParP were prepared according to the manufactures' instruction and added to the digested lysate with a concentration of 20 fmol/µl. For targeted-MS analysis the mass spectrometer first acquired a full MS-SIM scan with an MS1 resolution of 70.000, ACG target setting of 1e6 and 100 ms max injection time. Then PRM scans were carried out with a MS2 resolution of 35.000, AGC target settings of 2e5, 100 ms and an isolation window of 2 m/z. Normalized collusion energy was set to 27%. The analysis was performed unscheduled. For data analysis the results were imported into Skyline (v. 4.1.0.111714) (155).

Electron Cryotomography

Penicillin treated cells were enriched by centrifugation and protein A-treated 10 nm colloidal gold solution (Cell Microscopy Core, Utrecht University, Utrecht, The Netherlands) was added. After vortexing, aliquots of 3 µl was applied to freshly plasma-cleaned R2/2, 200mesh copper Quantifoil grids (Quantifoil Micro Tools GmbH, Germany). Plunge freezing was carried out in liquid ethane using a Leica EMGP (Leica microsystems, Wetzlar, Germany). Grids were blotted at 20°C and at 95% humidity. Data acquisition was performed on a Titan Krios transmission electron microscope (Thermo Fisher Scientific, OR, USA) operating at 300 kV. Images were recorded with a Gatan K2 Summit direct electron detector (Gatan, Pleasanton, CA) equipped with a GIF-quantum energy filter (Gatan) operating with a slit width of 20eV. Images were taken at a nominal magnification of 42,000 x, which corresponded to a pixel size of 3.513 Å. Tilt series were collected using UCSFtomo with a bidirectional tilt scheme (0° to –60°, followed by 0° to 60°) with a 2° increment. Defocus was set to -8 µm. The cumulative dosage used was 120 e-/A^2 and 80 e-/A^2 for unlysed and lysed cells , respectively.

Tomogram Reconstruction and Subtomogram Averaging

Drift correction, bead-tracking based tilt series alignment were done using IMOD (156). CTFplotter was employed for CTF determination and correction (157). Tomograms

were reconstructed using both weighted back-projection and simultaneous iterative reconstruction (SIRT) with iteration number set to 9. Dynamo was used for particle picking and subtomogram averaging (158, 159).

Sample size and analysis

For fluorescence microscopy experiments sample size and demographic analysis was performed as previously described (112). The total number of cells included (n) is mentioned for each experiment and demograph in the respective figures. For global transcriptomics and targeted LC-MC analysis the results are based on a minimum of four independent experiments. The mean was then plotted with error bars representing the standard deviation. The p-value was calculated performing a Student's t-test.

Supplementary Information

Table S3.1. Strain list

Strain ID	Genotype	Reference
V. cholerae C6706	*lacZ-*	
AA26	Δvc1397 Δvc2063 Δvca1095 lacZ-	This Work

Table S3-2. Plasmid list

Plasmid ID	Description	Reference
pCVD442	Suicide vector for gene deletion	
pAK13	Plasmid to delete *vc2063*	(112)
pCVD442-VCA1095-flank	Plasmid to delete *vca1095*	This work
pPM041	Plasmid to delete *vc1397*	(276)
pMF391	*P*BAD::*cfp*	This work
pBAD33	Arabinose induced expression vector	(277)
pSR1033	*P*BAD::*yfp-vc2059*	(148)
pAA100	*P*BAD:: *vc1602-cfp*	This work
pAA101	*P*BAD::*vc2006-cfp*	This work
pAA102	*P*BAD::*vc2202-cfp*	This work
pAA103	*P*BAD::*vca0954-cfp*	This work

Table S3.3. Primer list

Primer ID	Sequence (5'-)
565	CCCCCGCATGCTCACTTGTACAGCTCGTCCATGCC
3024	TCCTCGCCCTTGCTCACCAT GAGCTCGAGGATGTC CATATTCAGGATATATTTATCCAT
3025	GACATCCTCGAGCTC ATGGTGAGCAAGGGCGAGGA
3027	CCCCC TCTAGA ATGTCAGGTGTTTTGAACACGGTA
3028	TCCTCGCCCTTGCTCACCAT GAGCTCGAGGATGTC GGAAAGCACTTCTCGCAATCTT
3029	CCCCC TCTAGA ATGACGGGTATTCTTGATT
3031	CCCCC TCTAGA ATGGCTAAAGTCGTCAGTAA
3037	TCCTCGCCCTTGCTCACCAT GAGCTCGAGGATGTC TTTAACTAGCGCAGCTTTCAC
3038	TCCTCGCCCTTGCTCACCAT GAGCTCGAGGATGTC CGATTTTTGCAAACGCTGCGTCT
3039	CCCCC TCTAGA ATG AGCAACCCAAGCAGTACTA

Table S3.4. Abundance of chemotaxis proteins detected by global LC-MS proteomic analysis.

Presence of chemotaxis proteins in V. cholerae

			Detection				Detection
Cluster I	VC1394	MCP	No	**Cluster III**	VCA1088	MCP	No
	VC1395	CheY1	No		VCA1089	CheB3	No
	VC1396	HP	No		VCA1090	CheD	No
	VC1397	CheA1	No		VCA1091	CheR3	No
	VC1398	CheY2	No		VCA1092	MCP	No
	VC1399	CheR1	No		VCA1093	CheW2	No
	VC1400	HP	No		VCA1094	CheW3	No
	VC1401	CheB1	No		VCA1095	CheA3	No
	VC1402	CheW	No		VCA1096	CheY4	No
	VC1403	MCP	No				
	VC1404	HP	No	**CheVs**	VC1602	CheV1	Yes
	VC1405	MCP	No		VC2006	CheV2	Yes
	VC1406	MCP	No		VC2202	CheV3	Yes
					VCA0954	CheV4	Yes
Cluster II	VC2059	CheW1	Yes	**Others**	VC2201	CheR2	No
	VC2060	ParP	Yes				
	VC2061	ParC	Yes				
	VC2062	CheB2	Yes				
	VC2063	CheA2	Yes				
	VC2064	CheZ	Yes				
	VC2065	CheY3	Yes				

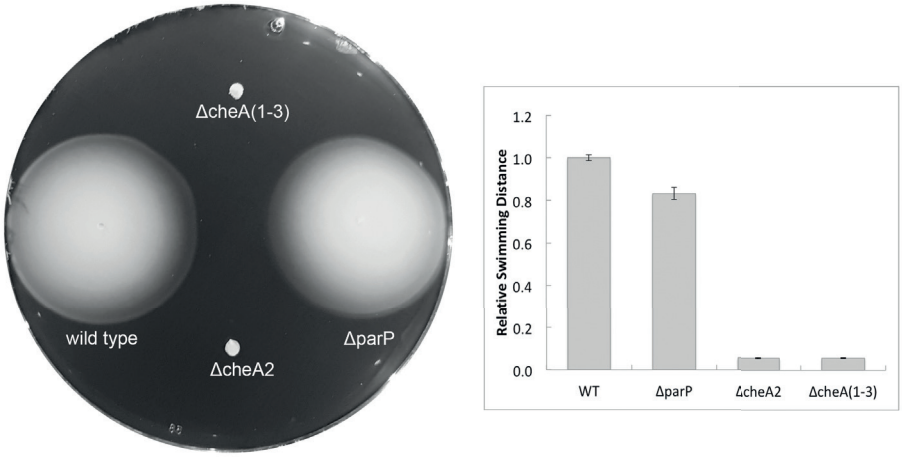


Figure S3.1. Soft agar assay shows chemotatic behavior in *V. cholera* wild-type and stains with selective chemotaxis protein deletion. ParP deletion had little impact on the chemotaxis ability (N=25). No chemotactic ring was observed for *cheA2* deletion strain or *cheA(1-3)* deletion strain , which suggested the chemotactic behavior in those two strains was largely impaired.

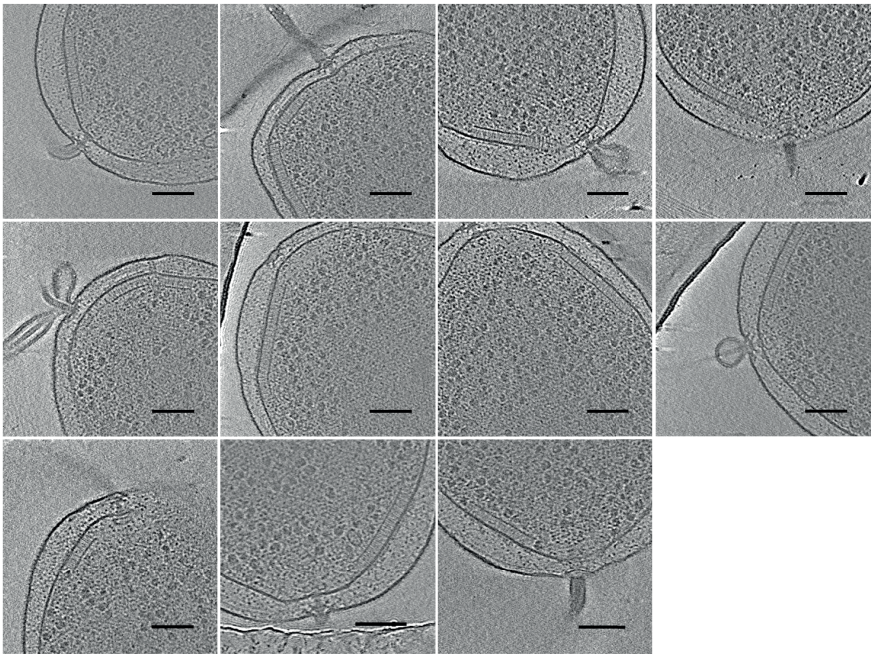


Figure S3.2. Localization of cheA-free cluster II chemoreceptor arrays in *V. cholerae*. Among 27 cell poles examined, all 11 chemoreceptor arrays observed were closely localized to the polar flagellar. Scale bar is 100 nm.

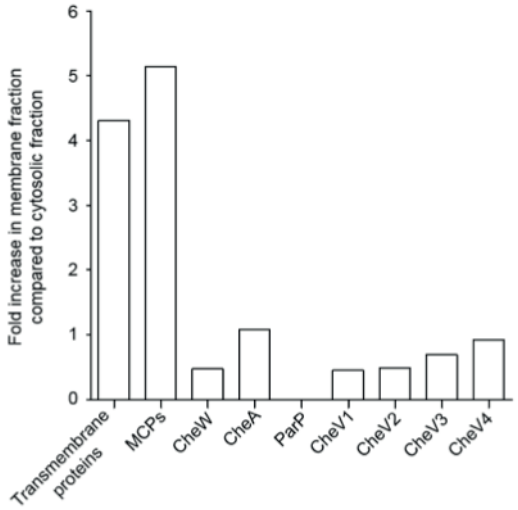


Figure S3.3. Baseplate proteins do not co-fractionate with MCP proteins in membrane fractions. The bar graph show the average (average of four experiments) fold increase in the membrane fraction compared to the cytosolic fraction of total annotated transmembrane proteins, total detected MCP proteins and the baseplate proteins including CheW, CheA, ParP, CheV1-4. Notably, ParP protein was not detected at all in the membrane fraction.

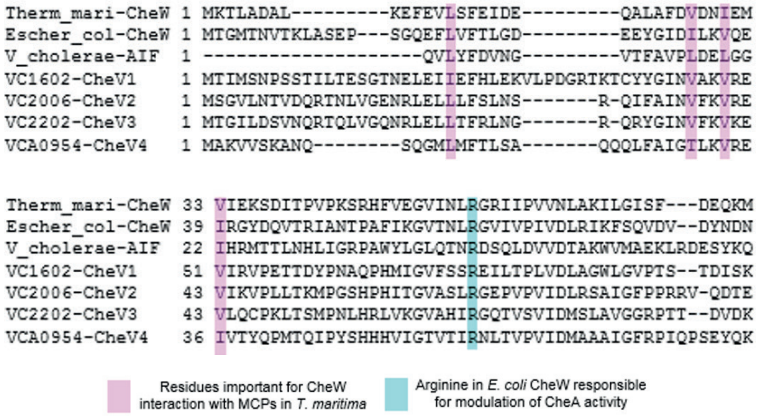


Figure S3.4. Alignment showing the conservation of residues in *V. cholerae*. CheVs, CheW, and ParP-AIF that are required for interaction with MCP proteins. The alignment suggests that all four CheVs should be able to integrate into the chemoreceptor array.

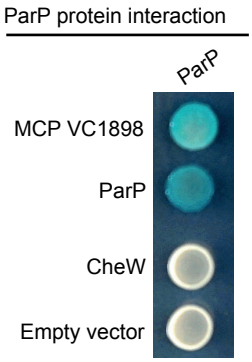


Figure S3.5. Bacterial-two-hybrid assay for screening protein interaction between the *V. cholerae* ParP with MCP VC1898, ParP and CheW. Blue coloration of bacterial colonies indicates an interaction.