AMERICAN SOCIETY FOR MICROBIOLOGY | mBio

Check for updates

| Structural Biology | Research Article

Structure of the native chemotaxis core signaling unit from phage E-protein lysed *E. coli* cells

C. Keith Cassidy,[1,2] Zhuan Qin,[3] Thomas Frosio,[1] Khoosheh Gosink,[4] Zhengyi Yang,[1] Mark S. P. Sansom,[5] Phillip J. Stansfeld,[6] John S. Parkinson,[4] Peijun Zhang[1,3,7]

AUTHOR AFFILIATIONS See affiliation list on p. 14.

ABSTRACT Motile bacteria employ conserved chemotaxis networks to detect chemical gradients in their surroundings and effectively regulate their locomotion, enabling the location of essential nutrients and other important biological niches. The sensory apparatus of the chemotaxis pathway is an array of core-signaling units (CSUs) composed of transmembrane chemoreceptors, the histidine kinase CheA and an adaptor protein, CheW. Although chemotaxis pathways represent the best understood signaling systems, a detailed mechanistic understanding of signal transduction has been hindered by the lack of a complete structural picture of the CSU and extended array. In this study, we present the structure of the complete CSU from phage φX174 E protein lysed *Escherichia coli* cells, determined using cryo-electron tomography and sub-tomogram averaging to 12-Å resolution. Using AlphaFold2, we further predict the atomic structures of the CSU's constituent proteins as well as key protein-protein interfaces, enabling the assembly an all-atom CSU model, which we conformationally refine using our cryo-electron tomography map. Molecular dynamics simulations of the resulting model provide new insight into the periplasmic organization of the complex, including novel interactions between neighboring receptor ligand-binding domains. Our results further elucidate previously unresolved interactions between individual CheA domains, including an anti-parallel P1 dimer and non-productive binding mode between P1 and P4, enhancing our understanding of the structural mechanisms underlying CheA signaling and regulation.

IMPORTANCE Bacterial chemotaxis is a ubiquitous behavior that enables cell movement toward or away from specific chemicals. It serves as an important model for understanding cell sensory signal transduction and motility. Characterization of the molecular mechanisms underlying chemotaxis is of fundamental interest and requires a high-resolution structural picture of the sensing machinery, the chemosensory array. In this study, we combine cryo-electron tomography and molecular simulation to present the complete structure of the core signaling unit, the basic building block of chemosensory arrays, from *Escherichia coli*. Our results provide new insight into previously poorly-resolved regions of the complex and offer a structural basis for designing new experiments to test mechanistic hypotheses.

KEYWORDS chemotaxis, cryoET, sub-tomogram averaging, *in situ*, cryoEM, chemosensory, phage lysis, MD simulation

Motile bacteria detect ambient chemical gradients and control locomotion via conserved chemotaxis signaling networks, which enable them to seek out nutrients, potential hosts, and other important biological niches (1). The chemotaxis network of *Escherichia coli* has served for decades as a model system for the study of sensory signal transduction and motility behaviors (2). In this system, thousands

Editor Matthew Parsek, University of Washington, Seattle, Washington, USA

Address correspondence to Peijun Zhang, peijun.zhang@strubi.ox.ac.uk.

C. Keith Cassidy and Zhuan Qin contributed equally to this article. Author order was determined both alphabetically and in order of increasing seniority.

The authors declare no conflict of interest.

See the funding table on p. 14.

Received 28 March 2023
Accepted 9 August 2023
Published 29 September 2023

of copies of transmembrane (TM) chemoreceptors, the histidine autokinase CheA and adaptor protein CheW, form highly ordered structures known as chemosensory arrays (3), which cooperatively integrate the sensory inputs of multiple receptors to regulate the autophosphorylation activity of CheA. CheA, in turn, donates its phosphoryl groups to the diffusible intracellular response regulator CheY to modulate the cell's flagellar motors.

Core signaling units (CSU) of the chemosensory array contain six receptor dimers, organized as two trimers of dimers (TODs), a single CheA dimer and two CheW adaptor proteins. The CSU is the minimal molecular unit needed to couple CheA autophosphorylation to receptor control (4). The constituent proteins of CSUs have complex, modular structures. *E. coli* has four canonical chemoreceptors (Tsr, Tar, Trg, and Tap), known as methyl-accepting chemotaxis proteins (MCPs), that form homodimers of largely helical protomers. MCP molecules contain a periplasmic ligand-binding domain (LBD), TM four-helix bundle, and cytoplasmic kinase control domain (2). The cytoplasmic domain itself has a membrane-proximal HAMP (*h*istidine kinases, *a*denylyl cyclases, *M*CP and some *p*hosphatases) domain coupled to a long coiled-coil four-helix bundle, the membrane-distal end of which contains the contact surfaces for TOD formation and CSU assembly.

CheA functions as a homodimer in the CSU; each protomer comprises five domains (P1–P5) joined by flexible linkers (5). The P3, P4, and P5 domains form a compact core through the principal dimerization determinant P3. That dimeric core is integrated into the larger CSU via P5 interactions with the receptor tips and CheW. The P4 domain binds ATP and catalyzes autophosphorylation at a histidine residue in the P1 domain. The P1 and P2 domains are attached to the P3-P4-P5 core via long, disordered linkers and mediate phosphoryl group transfer to CheY. The monomeric CheW protein has a fold homologous to the CheA.P5 domain and interacts with P5 to couple CheA to the receptor trimers. We note that two additional CheW molecules can associate with the periphery of the CSU and contribute to the formation of hexameric CheW rings in the larger array (6–8). Those CheW molecules are not essential for kinase regulation in the CSU.

High-resolution CSU structures are central to understanding their molecular signaling mechanisms, particularly receptor-mediated kinase control. Toward this end, cryo-electron tomography (cryoET) has played an essential role in revealing that chemoreceptors in diverse microbial species form TODs organized into a hexagonal lattice with identical spacing (9, 10). In *E. coli*, this hexagonal packing was shown to extend to the CheA/CheW baseplate region at the receptor tip, where rings containing CheA.P5 and CheW were seen to inter-lock adjacent CSUs (6, 7, 11). CryoET studies employing sub-tomogram averaging (STA) have produced increasingly detailed structures of the CSU in thinner array samples, including those of lysed cells (12), minicells (13), and *in vitro* reconstituted arrays (7, 14). Such cryoET density maps have been combined with integrative modeling and molecular simulation techniques to achieve residue-level characterization of CSU structure and dynamics (7, 13, 14). Nevertheless, shortcomings remain: reconstructions from lysed cells have so far achieved only modest resolutions (>20 Å), and minicells can still be relatively thick (~400 nm), which limits the maximum resolution possible. Moreover, although reconstructions of monolayer arrays reached 8-Å resolution, they were formed using soluble receptor molecules that lacked the periplasmic and transmembrane regions.

Here we utilized controlled lysis triggered by induction of a phage E gene to create thin *E. coli* ghost cells containing wild-type chemosensory arrays. We further used STA to obtain a density map of the complete CSU and AlphaFold2 to produce full-length atomic models of the CSU proteins that could be flexibly fitted to the cryoET density map. All-atom molecular dynamics (MD) simulations of the resulting complete CSU model provided new insights into the periplasmic organization of the CSU as well as previously unresolved interactions between individual CheA domains, suggesting a new generation of structure-function experiments to elucidate the workings of chemosensory arrays.

RESULTS

Imaging native chemosensory arrays in wild-type *E. coli* ghosts

The thickness of native *E. coli* cells (~1.5 µm) prohibits acquisition of quality cryo-tomograms for high resolution structural analysis of chemosensory arrays. We previously developed a method to produce *E. coli* 'ghost' cells just before plunge-freezing (15). Upon induction of lysis gene E from the φX174 bacteriophage, spot lesions form in the cell membranes that release the cellular contents while maintaining the overall integrity of the cell envelope. Vitrification of such samples produces thin, flattened cell ghosts (<200 nm) (Fig. 1A). In light microscopy, the *E. coli* ghosts were easily distinguished from unlysed cells due to minimal cytoplasmic green fluorescent protein (GFP)-labeled CheY and had distinct chemosensory arrays as visualized via associated CheY (16) (Fig. 1B, arrows). The *E. coli* ghosts were also easy to identify in cryoEM projection images as they appeared semi-transparent. The resulting reconstructed tomograms display well-ordered hexagonal polar arrays about 200–300 nm across with a 12-nm lattice spacing (Fig. 1C and D; Movie S1).

CryoET STA structure of the complete CSU in *E. coli* ghosts

CryoET data were collected from the E protein lysed *E. coli* ghosts. The data collection and processing statistics are summarized in Table S1, and the cryoET STA process is illustrated in a workflow schematic (Fig. S1A). A low-resolution structure of the CSU lattice was used for template matching in emClarity (17). The resulting convolution

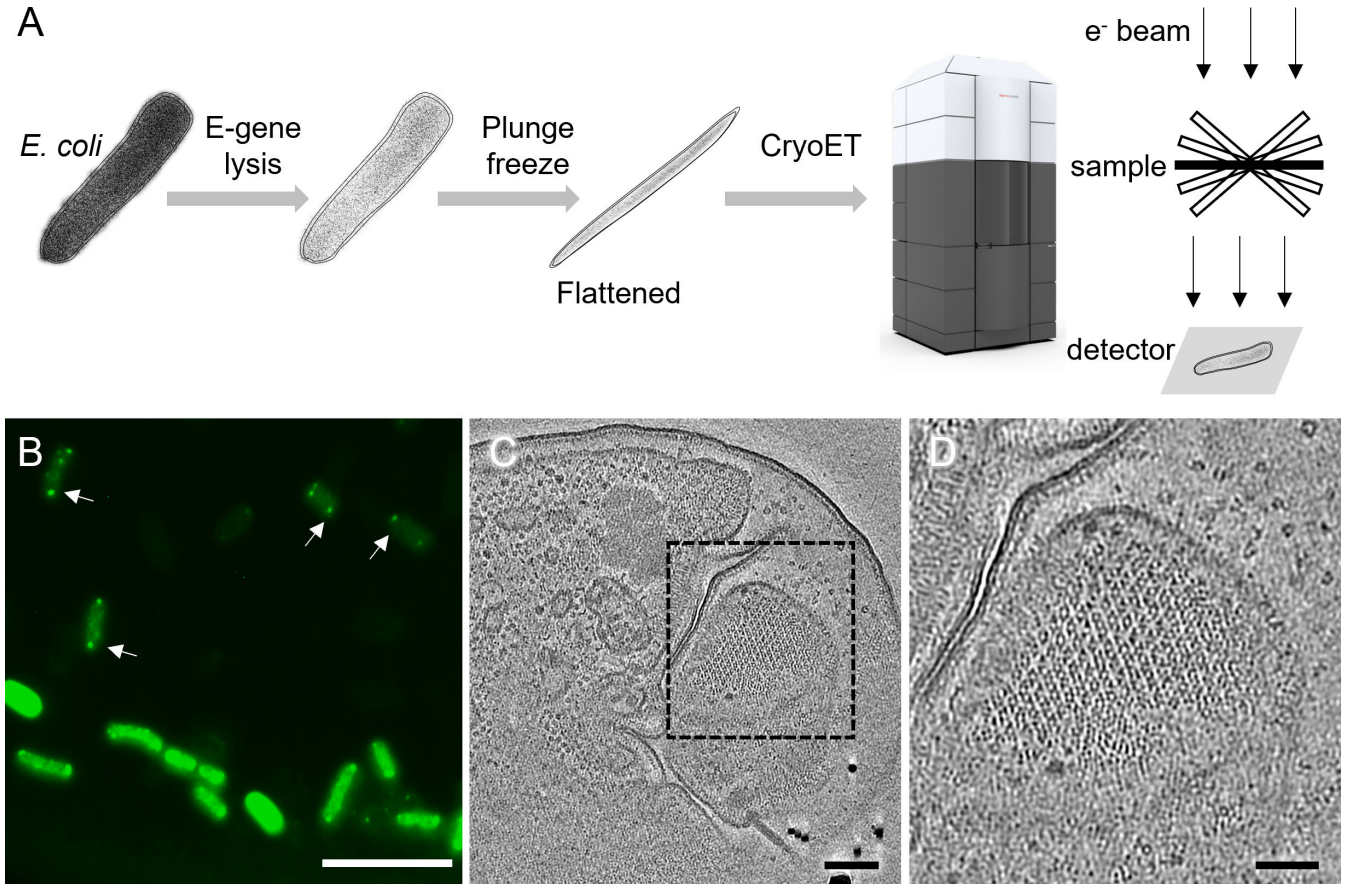


FIG 1 CryoET of native chemosensory arrays in wild-type *E. coli* ghost cells. (A) A schematic diagram of the cryoET workflow used in this study. (B) A fluorescent light microscopy image of GFP-labeled CheY highlights chemosensory arrays (arrows) near the poles of both E-lysed cells (transparent green) and unlysed cells (bright green). (C) A tomographic slice of an E-lysed *E. coli* cell highlighting a patch of the chemosensory array lattice (dashed box). (D) A zoom-in view of the area marked by the dashed box in panel C. Scale bars, 10 µm in panel B, 100 nm in panel C, and 50 nm in panel D.

maps (Fig. S1B and S1C; Movie S2) display cross-correlation peaks for the selection of sub-tomograms (Fig. S1D and S1E). Although the template-matched subtomograms all contained six receptor trimers in a hexagonal lattice (Fig. S1Aiii and Fig. S1E), they were heterogeneous at the CheA/CheW baseplate level and resulted in four structural classes (Fig. S1Aiv). Sub-tomograms centered on a CSU trimer (Fig. S1Aiv, classes with green checks) were selected (Fig. S1F) and subjected to further refinement (Fig. S1A-v), from which sub-tomograms containing a single CSU were iteratively aligned and averaged to give rise to the final CSU map (Fig. S1A-vi).

The map of the full CSU, determined at an overall resolution of 12 Å (Fig. 2A; Fig. S2A), clearly resolves six full-length chemoreceptor dimers (Fig. 2A, red), a central CheA dimer containing distinct regions for P3–P5 (Fig. 2A, blue), four CheW monomers (Fig. 2A, orange and yellow), and the lipid bilayer headgroups (Fig. 2A, gray). The cytoplasmic organization of the complex is consistent with the recently reported CSU structure at 16-Å resolution from *E. coli* minicells (13) (Fig. S3A through S3C). Both maps show similar splaying of the receptors within each TOD as well as comparable baseplate densities, particularly regarding the shape and orientation of a CheA "keel" density that contains the P4 domain (Fig. S3C). The improved resolution of the present map, however, provides better localization of individual protein domains throughout the structure but particularly within the membrane-proximal region, especially the MCP LBDs (Fig. S3A and S3B). We note that while our biological samples contained wild-type receptors of all types,

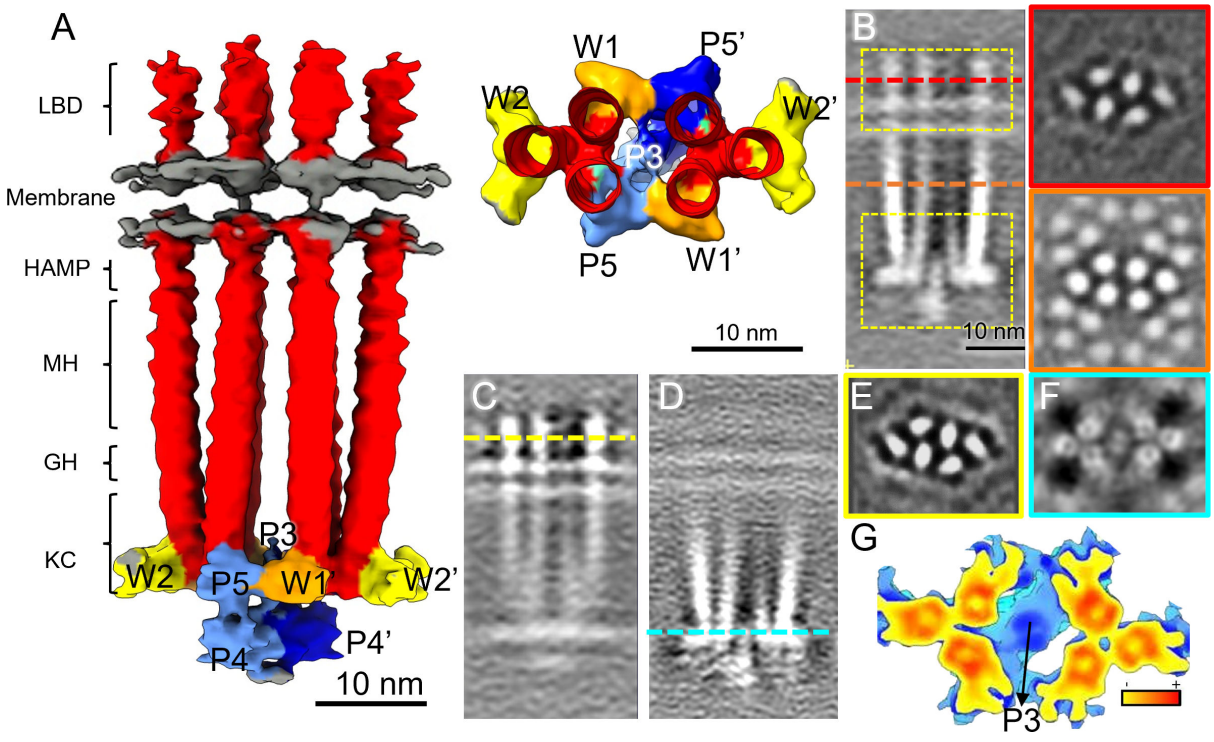


FIG 2 CryoET sub-tomogram averaging of the core signaling unit (CSU) of native chemosensory arrays. (A) Surface rendering of the CSU density map. Densities corresponding to the receptors are colored in red, CheA dimer in light/dark blue, CheW in yellow/orange, and lipid headgroups in gray. Specific regions of the receptors are labeled, including the ligand-binding domain (LBD), HAMP, methylation helix bundle (MH), glycine hinge (GH), and the kinase control region (KC). Individual CheA domains and CheW molecules are labeled accordingly. On the right side, a top view of the CSU from the glycine hinge downwards is displayed. (B) The CSU density map is shown in sectional side view (left) and top views (right). The top views are presented at the LBDs (red line) (top right) and at the cytoplasmic domains (orange line) (bottom right). (C and D) Focused alignment and averaging centered on the receptor LBDs (top dashed yellow box in panel B) (C) and on the baseplate region containing CheA, CheW, and the receptor signaling domain (bottom dashed yellow box in panel B) (D). (E and F) Sectional views of the averaged maps from C and D at the receptor LDBs (yellow line in panel C) (E) and baseplate (blue line in panel D) (F). Note the red and yellow lines in panels B and C, respectively, correspond to the same LBD position. (G) An enlarged view of the refined CSU baseplate, colored from blue to red, according to the density value.

albeit carrying potentially variable methylation states, we are not able to distinguish different receptor types at the present resolution.

In the baseplate region, we further observed additional keel density apparently not attributable to CheA.P4. Comparison of the present CSU structure with the one derived from *in vitro* monolayer arrays at 8 Å (14) shows that the latter structure does not contain corresponding density in this region (Fig. S3D through S3G). Notably, the monolayer system utilized truncated CheA molecules not possessing P1 and P2 domains, suggesting that the excess density likely corresponds to these domains. Docking of the molecular model derived from the *in vitro* system (PDB 6S1K) (14) allowed further localization of the excess density, positioning it directly between and below the two P4 domains (Fig. S3H).

Focused density refinement of the CSU periplasmic and baseplate regions

Local resolution analysis of the full CSU suggested that the complex was quite flexible in the periplasmic and CheA.P4 regions (Fig. S2B). We, therefore, carried out a focused refinement of the periplasmic and baseplate regions separately (Fig. 2B, yellow boxes). Indeed, as the density of each localized region improved during focused refinement, the remainder of the complex became poorly resolved (Fig. 2C and D), suggesting considerable structural flexibility between these two regions. The focused refinements also resolved additional structural details in both regions. In the periplasmic-focused map, the ellipsoidal shape of the individual LBDs became prominent and clearly indicated the orientation of each LBD (Fig. 2E), a feature not discernible in previous studies. These ellipsoidal densities showed a pseudo-threefold symmetry within each TOD, although no such symmetry was applied during refinement. Similarly, in the baseplate-focused map, the individual α-helices in the receptor signaling domain and CheA P3/P3′ dimerization domains became better resolved (Fig. 2F and G). Notably, the flexibility increases continuously from Gly hinge to methylation helices, and to the HAMP domain (Fig. 2D; Fig. S4A and S4B). Moreover, further refinement focused on the HAMP domain suggests flexible hinges on either side of the HAMP domain (Fig. S4C through S4E).

All-atom model of the complete *E. coli* CSU

Existing high-resolution information on the structures of full-length chemotaxis proteins is sparse. Previous *E. coli* structural studies have largely relied on the use of analogous structures from *Thermotoga maritima*, a distantly related organism, either directly or as templates for homology modeling, to assess residue-level details. Recently, AlphaFold2 (18) has been shown to produce high-fidelity atomistic structures of proteins and complexes using machine learning, providing a powerful alternative means of obtaining high-resolution structural information, especially in experimentally challenging cases. Accordingly, we used AlphaFold to construct models of the *E. coli* CSU proteins and sub-complexes, including full-length Tsr and CheA as well as the $(CheA.P5/CheW)_3$ and $(CheW)_6$ hexameric rings comprising the array baseplate (Fig. 3A through C). In all cases, wild-type sequences were used. Overall, the predicted structures are of high quality as assessed by pLDDT, a per-residue confidence score output by AlphaFold, and agree well with existing high-resolution structures and previous modeling results as described below.

The predicted full-length Tsr structure (Fig. 3A) is consistent with X-ray crystal structures of the Tsr cytoplasmic domain (PDB 3Z X6) (19) and APO Tsr periplasmic domain (PDB 2D4U) (20). Additionally, the Tsr TM, control cable, and HAMP regions, which were previously unresolved, closely resemble those seen in a crystal structure of the symmetric APO state of the NarQ sensor kinase (PDB 5JEQ) (21) and agree well with previous modeling results based on TM cross-linking data from *E. coli* Tar (13, 22). In particular, the predicted TM bundle displays a central TM1/TM1′ dimer with flanking TM2 helices, as previously proposed for *E. coli* chemoreceptors generally (23), and the critical control-cable segment connecting TM2 to the AS1 helix of HAMP is observed to be helical and kinked as previously suggested based on mutagenesis data (24).

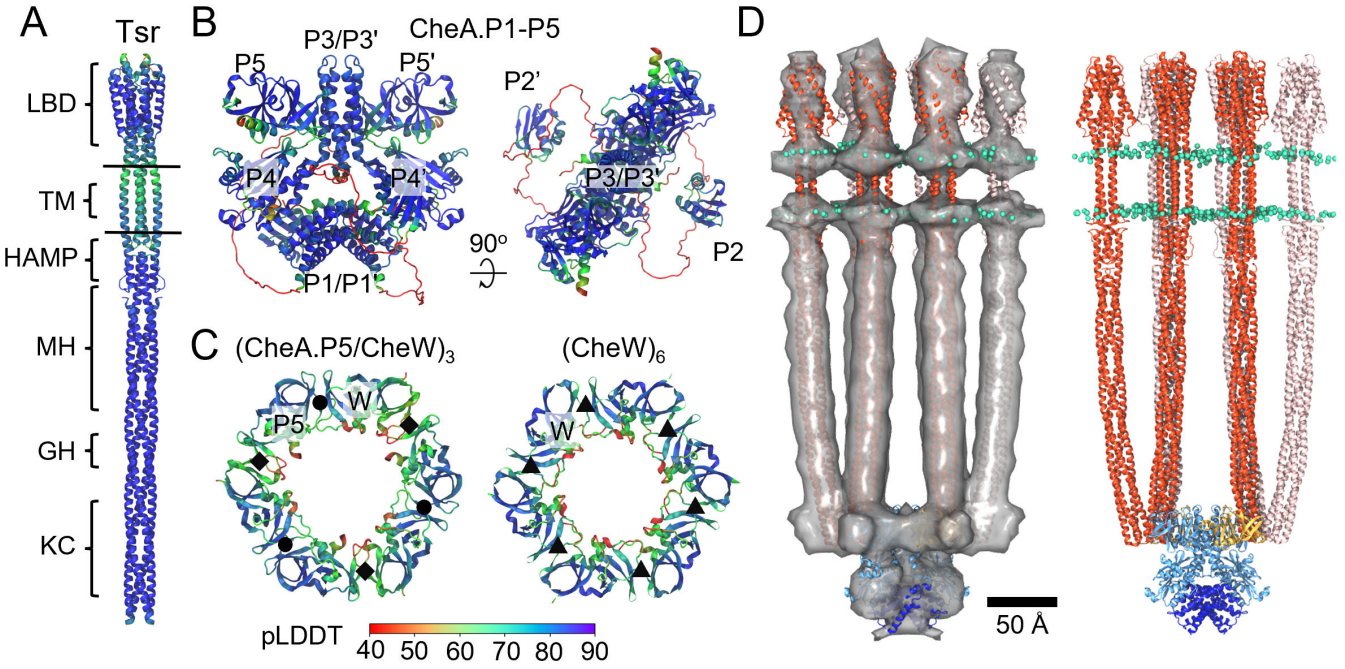


FIG 3 All-atom model of the full-length *E. coli* core signaling unit (CSU). (A–C) AlphaFold2 models of *E. coli* Tsr (A), CheA (B), and intact $(CheA.P5/CheW)_3$ and $(CheW)_6$ rings (C). Individual domains are labeled as in Fig. 2. Models are colored by pLDDT, a residue-based pLDDT confidence score. In panel C, interfaces I, II, and III are denoted by black circles, diamonds, and triangles, respectively. (D) Molecular dynamics flexible fitting-refined CSU model with (left) and without (right) overlayed density map. Tsr dimers are shown in red/pink; CheA.P3-5 is shown in light blue, CheA.P1 in dark blue, CheW in gold, and lipid headgroups in cyan.

The predicted folds of the individual *E. coli* CheA domains and their inter-domain arrangements (Fig. 3B) agree well with existing crystal structures from *T. maritima* (25, 26). The P3 and P4 domains adopt a planar arrangement similar to that seen in a crystal structure of a soluble *T. maritima* CheA.P3-P4 construct (PDB 4XIV) (26) and closely resemble the 'undipped' CheA conformation deduced from cryoET maps of *in vitro* monolayer arrays (PDB 3JA6 and 6S1K) (7, 14). The P5 regulatory domains are rotated slightly compared to the orientation required to form interface I with CheW. Intriguingly, the P1 domains were observed to form a dimer wedged between the two P4 domains (Fig. 3B) where the aforementioned excess density in our cryoET map is localized (Fig. 3D). The locations of the P2 domains, on the other hand, vary considerably between AlphaFold predictions (Fig. S5). The functional implications of these observations are discussed in more detail below. Additionally, in line with the disordered nature of the long, flexible linkers connecting P1 and P2 to one another and P3 (27), AlphaFold predicts these regions to be unstructured with low pLDDT score (Fig. S6).

In the case of the $(CheA.P5/CheW)_3$ ring (Fig. 3C), which contains interfaces I and II between CheA.P5 and CheW, the AlphaFold-predicted structure agrees well with the analogous *T. maritima* crystal structure (PDB 4JPB) (28), as well as *in vivo* cross-linking studies in *E. coli* (29). Additionally, the predicted $(CheW)_6$ ring structure is in good agreement with molecular models derived via analogy with the $(CheA.P5/CheW)_3$ ring and validated with *in vivo* cross-linking data (8, 30).

To aid in the computational assembly of the CSU, we also utilized AlphaFold to construct models of the other key protein-protein interfaces, including the trimeric interface between the Tsr protein-interaction regions as well as the Tsr/CheA.P5 and Tsr/CheW interfaces. The constituent models were assembled via rigid docking using the predicted interfaces and our cryoET map to produce a preliminary model of the CSU, which was subsequently refined using molecular dynamics flexible fitting (MDFF) (31)

(Fig. S7). Finally, the refined CSU model was embedded in an atomistic lipid bilayer (Fig. 3D) with a composition chosen to mimic that of the *E. coli* inner membrane (32). Full details regarding the modeling procedures are provided in the Materials and Methods.

Neighboring receptors form transient interactions within the periplasmic space

The improved characterization of MCP LBD orientations in our cryoET STA map permitted positioning these domains within the full CSU structure with added confidence (Fig. 4A). To test the robustness of the obtained fits, we conducted a series of MDFF simulations in which the receptor LBDs were rotated by 30°, 60°, or 90° from their initial positions. Indeed, in each simulation, the LBDs returned to their original fitted orientations. We additionally noted potential interactions between neighboring LBDs as they rotated past one another. We therefore decided to carry out an all-atom simulation of the CSU to explore the possibility of specific inter-receptor interactions within the periplasmic space. The resulting trajectory (Movie S3) identified interactions between two co-planar clusters of charged residues on the Tsr LBDs, involving K99, E102, and K103 on helix 2 (cluster 1) as well as E124, K126, R127, and D130 on helix 3 (cluster 2) (Fig. 4B). In addition to interactions between receptors within the same TOD, we also observed interactions between the two TODs (Fig. 4C). Given the symmetry of the extended array architecture, especially in the periplasmic space (30), such interactions presumably also occur between receptors from neighboring CSUs. Notably, however, the observed interactions were transient (<10 ns) and did not form a single, long-range pattern across the CSU on the timescale of our simulation. Rather, these interactions appear to form

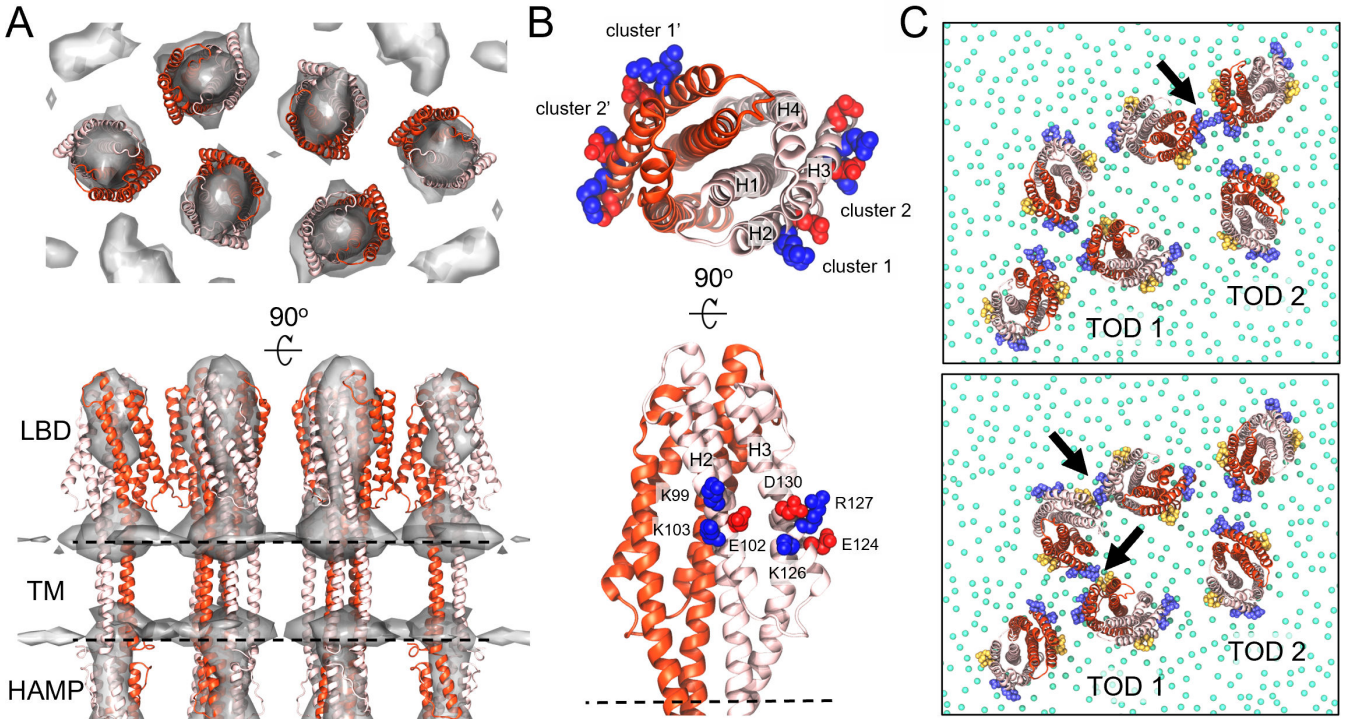


FIG 4 Periplasmic inter-receptor interactions within CSU. (A) Top and side views of the periplasmic region of the refined CSU model, involving the receptor ligand-binding domain (LBD), transmembrane (TM) bundle, and HAMP domain, overlayed with the cryoET STA map (shown in Fig. 2C and E). Receptor monomers are differentiated using light and dark red. The location of the membrane, not shown for clarity, is illustrated with dashed lines. (B) Top and side views of a single Tsr LBD, highlighting the two clusters of residues observed in an MD simulation to form inter-receptor interactions. Individual residues within each cluster are labeled and shown as Van der Waals (VDW) spheres in red (acidic) or blue (basic). LBD helices are labeled H1–H4. (C) Snapshots from an MD trajectory illustrating interactions (arrows) between receptors within the same trimer of dimers (TOD) (top) and between TODs (bottom). For clarity, residues forming clusters 1 and 2 have been colored yellow and purple, respectively, and only the headgroup of each lipid is shown (green spheres).

a pseudo-lattice by which the Tsr LBDs roughly maintain their relative orientations and inter-domain distances.

CheA.P1 forms an anti-parallel dimer that interacts non-productively with CheA.P4

The central P1 dimer predicted as part of the full-length CheA structure by AlphaFold describes well the excess density between the P4 domains observed in our map (Fig. 5A), which was previously noted within a cryoET map from *E. coli* minicells but could not be assigned atomic structure (13). Moreover, the size and shape of the P1 dimer suggest it alone accounts for this excess, in line with the variable placement of the P2 domains by AlphaFold. The apparent lack of P2 density therefore suggests that it may not form stable, functional interactions with the rest of the complex, an observation in line with both solution NMR measurements (33) as well as studies involving P2-deletion CheA mutants demonstrating that P2 is not strictly required for effective kinase regulation and chemotaxis (34). The P1 domains themselves are predicted to interact in an anti-parallel fashion mediated by P1 helices A and B, such that the substrate histidines face away from the P4 catalytic site and toward the P1/P1′ interface (Fig. 5B). The interaction between P1 and P4 is primarily mediated by a string of oppositely charged residues residing on helix D and the C-terminus of helix A in P1 and the α3 helix in P4 (Fig. 5C). Additionally, there exist interactions between the N-terminus of P1 helix A and the N-terminus of the P3 dimerization bundle, particularly between the previously noted P1 residues D14 and E18 (35) and P3 residue R265, which has been shown to play a vital role in CheA function (7, 36). Such interactions may also account for the observed increase in *E. coli* CheA dimer affinity attributed to the presence of P1 (37).

In line with this predicted P1/P4 binding surface, a previous NMR analysis of *T. maritima* CheA, designed to monitor the binding of free P1 to a soluble P3/P4 construct, showed large chemical shifts on the P1 D helix as well as on the P4 α4 helix, a region distal to the catalytic site, leading to the proposal of a non-productive P1/P4 binding mode (38). In addition, sizable changes in chemical shifts were also reported on the P1 A and B helices, which form interactions with both P4 and the neighboring P1 monomer in our predicted structure, while the P1 C and E helices, which showed little change in chemical shifts, do not form any inter-domain contacts. In contrast, the predicted CheA structure primarily shows the P4 α3 helix mediating non-productive P1 interactions, whereas the NMR analysis suggested the P4 α4 helix fulfills this role. It is possible, however, that this discrepancy may arise due to slight alterations in the relative orientations of the P3 and P4 domains due to the soluble and truncated nature of the CheA constructs used in the NMR analysis.

Recently, Muok and co-workers also observed close interactions between the P1 domains within *T. maritima* 'foldon' complexes, especially in the inhibited state, and further reported a crystal structure of a *T. maritima* P1 dimer (39). Like the AlphaFold-predicted *E. coli* P1 dimer in full-length CheA, this structure showed the substrate histidines facing toward the dimer interface in a non-productive fashion; however, the P1 domains adopted a parallel configuration as opposed to an anti-parallel one. Indeed, for two isolated *E. coli* P1 monomers, AlphaFold predicts a parallel arrangement similar to that reported in *T. maritima* (Fig. S8), although the relatively low Predicted Aligned Error (PAE) scores (Fig. S8) between P1 monomers suggests that AlphaFold is not confident in this predicted arrangement. We note, however, that this does not necessarily suggest the interface is incorrect but rather that it should not be interpreted on its own.

In light of the above observations, we wondered how well a parallel P1 dimer configuration would agree with the density map. We therefore rigidly docked the parallel P1 dimer predicted by AlphaFold into the excess density between the P4 domains (Fig. S9). Despite the excess density being relatively featureless, the parallel P1 dimer fit is considerably poorer than the anti-parallel configuration in terms of overlap with the density. Additionally, there do not appear to be any potential strong contacts to be formed between P1 and P4. Nevertheless, we cannot rule out that changes in the

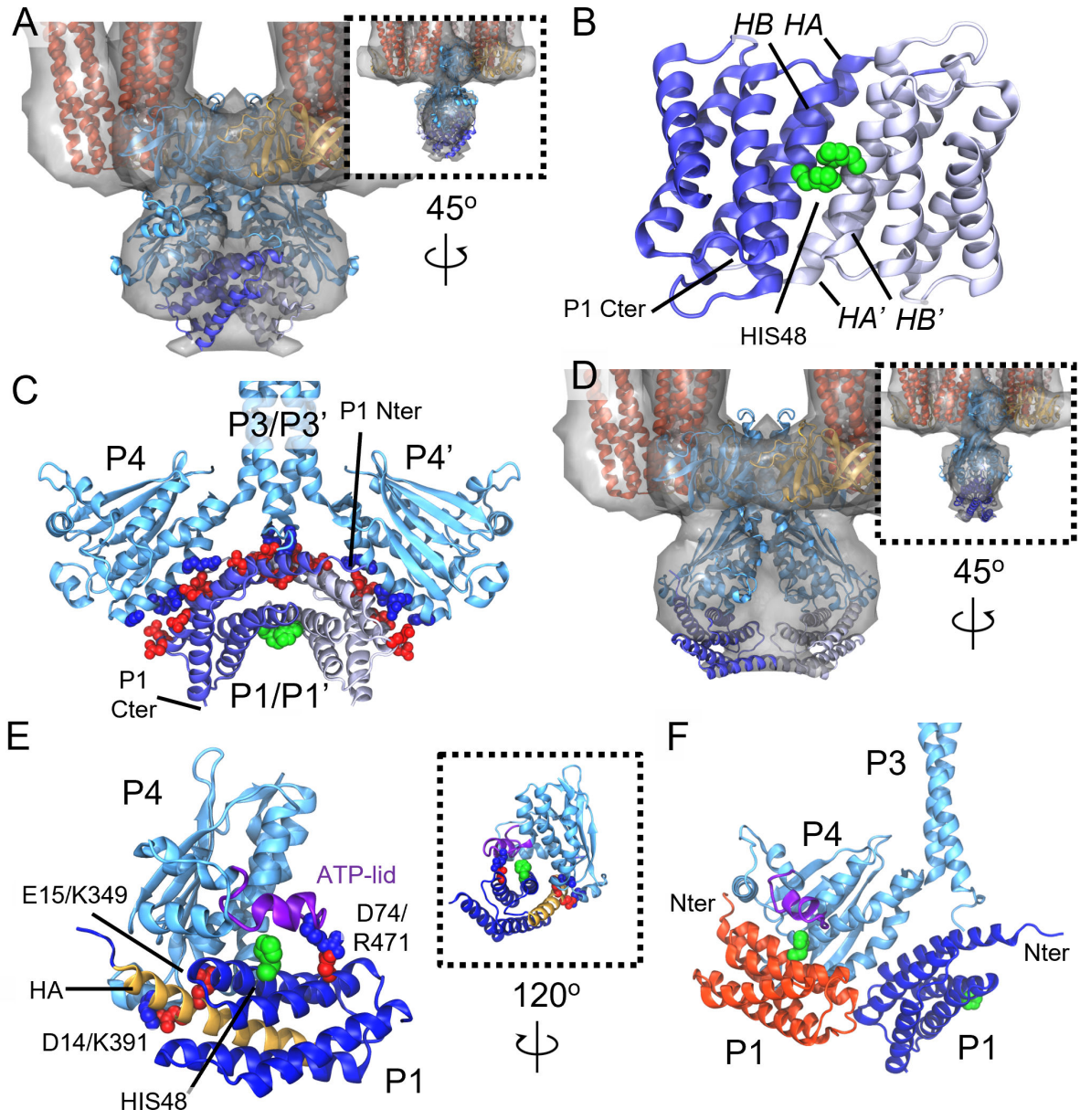


FIG 5 CheA.P1/P1 and CheA.P1/P4 interactions within the CSU. (A) Baseplate region of the MDFF-refined CSU model with density map overlayed, illustrating the fit of the AF2-predicted P1 dimer bound non-productively to P4/P4′. (B) Anti-parallel P1 dimer structure predicted by AF2 for full-length CheA shown in isolation. The individual P1 monomers are depicted in light and dark blue color and the substrate histidines are colored green and shown in VDW representation. The A and B helices mediating the dimer interaction are labeled for each monomer. (C) Predicted interaction interface between P1 and P3-P4. Acidic and basic residues forming potential inter-domain salt bridges are shown in red and blue, respectively. Note: the surrounding regions of the CSU model have been removed for clarity. (D) Baseplate region of the CSU in which a model of the dipped CheA conformation bound productively to P1 has been docked into the overlayed density map. (E) Predicted interaction between CheA.P1 and CheA.P4 for the productive transfer of phosphate from bound ATP (not shown) to HIS48 (green VDW spheres). Residues stabilizing the interaction are labeled and shown as VDW spheres in red (acidic) or blue (basic). CheA.P1 and CheA.P4 are shown in dark and light blue, respectively. P1 helix A and the ATP-lid are shown in gold and purple, respectively. (F) Overlay of non-productive (blue) and productive (red) P1/P4 binding modes. The N-terminus of each is labeled.

conformations of P4 and/or at the P1 dimer interface might produce alternate dimer configurations that describe well the density. Along these lines, analysis of the PAE plot for full-length CheA (Fig. S6) suggests that AlphaFold is relatively confident in its prediction for the inter-domain arrangement between P1 and P4 but less confident in the arrangement between the two P1 domains. We suggest therefore that important, non-productive interactions between P4 and P1 drive the dimerization of P1 in the

spatially constrained environment of the kinase core. Whether P1 also dimerizes away from the kinase core and the particular configuration such a dimer might acquire will require future studies to elucidate.

Keel density can accommodate dipped CheA.P4 with productively bound P1

Previously, we predicted the existence of a "dipped" P4 conformation using molecular simulations of the *T. maritima* CSU (7), which we also subsequently identified in a sub-nanometer cryoET map of the *E. coli* CSU from *in vitro* monolayer arrays (14). Recently, multiple experimental studies have lent strong evidence to the existence of such a conformation and demonstrated its importance for the CheA catalytic cycle (5, 36, 40). Although the present resolution prevents explicitly isolating distinct P4 conformations within our cryoET map, a rigid docking of our dipped CheA model is also consistent with the density (Fig. 5D). However, the dipped P4 domains, which are rotated downward and toward one another, considerably reduce the volume attributable to the P1 dimer and prevent the formation of predicted stabilizing interactions with P4, suggesting that the postulated non-productive binding of a P1 dimer is not compatible with a dipped P4 conformation. We therefore explored what implications this conformational change might have on the productive binding of P1 to P4.

Previous molecular docking (41) and mutagenesis (35) studies have led to models for the productive interaction of P1 with P4, which involve primarily P1 helices A and B oriented in such a way as to permit the highly constrained spatial requirements for chemical transfer of the gamma-phosphoryl group from bound ATP to the substrate histidine on P1 helix B (42). Additionally, a previous crystal structure of *T. maritima* CheA.P4 bound to an ATP analog showed that the so-called ATP lid, a typically unstructured span of ~20 residues near the nucleotide binding site, adopted an ordered conformation, thereby altering the probable interface for productive P1 binding (43). This observation led to the proposal that folding of the ATP lid is a prerequisite for productive P1 binding, a notion which has been further supported by recent kinetic analyses of CheA activation, suggesting an ordered sequential mechanism (44).

To assess the implications of the above observations, we constructed a model of the productive P1/P4 binding mode for *E. coli* CheA. As AlphaFold is currently not able to predict coordinates for ligands and co-factors when folding protein structures, it does not capture the folded ATP lid conformation induced by ATP binding to CheA.P4. We therefore opted to construct a homology model of *E. coli* P4 with a folded ATP lid based on PDB 1I58 and use HADDOCK (45) to carry out an extensive exploration of potential P1/P4 interactions. During the docking procedure, the distance between the substrate histidine and gamma-phosphoryl group was restrained to between 2 and 5 Å to encourage productive binding orientations. The top-ranked prediction (Fig. 5E) visually agrees with the productive P1/P4 binding mode proposed for *T. maritima* CheA (41) and shows interactions mediated primarily by P1 helices A and B, including salt bridges between residues D14/K391 and E15/K349, as well as a potential salt bridge between D74 on P1 helix C and R471 on the P4 ATP lid. Thus, P1 helix A appears to play an important role in both productive and non-productive P1 binding (Fig. 5F), although the N-terminus faces inward toward the P3 bundle in the non-productive binding mode and is outward facing in the productive binding mode. Notably, the proximity of the productive and non-productive P1 binding modes (Fig. 5F) further suggests there may be some degree of steric hindrance to simultaneous binding, an observation that may be relevant to studies involving the use of liberated P1 domains to assess CheA signaling mechanism (44, 46, 47).

Finally, we used our model of the productive P1/P4 binding mode to map P1 onto the dipped CheA structure, producing an overall CheA configuration that is fully accommodated by the keel density (Fig. 5D). Unlike the non-productive binding of P1 to undipped P4, however, which poises the P1 domains to bury a sizable surface area through dimerization, the productively bound P1 domains do not. Rather they face one another through a small cross-sectional area formed by the turns connecting helices A with B

and C with D, suggesting that P1 dimerization likely does not play an important role in the productive P1/P4 interaction. However, as we have only modeled a single dipped P4 conformation here and have assumed an overall symmetric CheA configuration for simplicity, future work will be required to characterize potential P1/P1 interactions when productively bound to P4.

DISCUSSION

In this study, we have reported the structure of the chemotaxis CSU from E protein lysed *E. coli* cells at an overall resolution of 12 Å by cryoET STA. Using AlphaFold2, we have further predicted the full-length structures of the CSU's constituent proteins as well as key protein-protein interfaces, enabling the assembly an atomistic CSU model, which we conformationally refine using our cryoET data. The improved resolution in the periplasmic region allowed for a more precise localization of the positions and relative orientations of the receptor LBDs. An MD simulation of the CSU model in an atomistic lipid bilayer permitted us to directly observe the local diffusion of receptors within the membrane for the first time and identified clusters of charged residues forming interactions between neighboring receptor LBDs. In line with these observations, previous *in vivo* work using a Tar-only *E. coli* strain observed that the efficiency of disulfide cross-linking between residues in periplasmic helices 2 and 3 varied as a function of methylation state and exposure to the chemoattractant aspartate (48). However, analysis of the cross-linking results was based on a crude picture of CSU structure and array architecture, potentially limiting functional interpretations. Our present model and observations therefore provide a high-resolution structural basis for designing experiments to test for concerted, signal-dependent changes in the periplasmic space.

Additionally, integration of AlphaFold2 predictions of full-length CheA into our cryoET density map allowed us to incorporate the P1 domains into the CSU model as an anti-parallel dimer, non-productively bound between the P4 domains. As previously noted, this prediction is in line with findings in *T. maritima*, which suggest both the existence of a non-productive P1-P4 binding mode and P1 dimerization (38, 39). Moreover, previous *in vivo* cryoET images of *E. coli* CSUs in distinct signaling states imposed by mutagenesis showed that the overall volume of CheA density tended to increase in the kinase-OFF state, suggesting that regulation of P1 and P2 mobility might be a key feature of kinase control (49). Along these lines, the non-productive P1/P4 contact interfaces presented here require both P4 domains to adopt an undipped conformation, which we previously showed to be stabilized by direct P4-P5 contacts (14). The P1 dimer may therefore act as a wedge to further reduce P4 mobility and CheA dynamics overall in the kinase-OFF state. Our results suggest, however, that P2 is not integral to this ordered state, and thus, its mobility is unlikely to be directly regulated.

Studies involving the kinetic analysis of CheA autophosphorylation using liberated P1 domains, on the other hand, have discounted the role of P1 sequestration in CheA regulation, proposing that receptors primarily mediate kinase regulation by altering the apparent rate constant of autophosphorylation through control of ATP binding (44, 47, 50). Nevertheless, our present observations can be reconciled with such a picture if, for instance, ATP binds preferentially to either the undipped or dipped P4 conformation, thereby allowing receptors to indirectly mediate ATP binding through the regulation of P4 dynamics, which could be in turn affected by P1 dimerization and sequestration. It has been previously noted that the undipped P4 conformation situates the unfolded ATP-lid near the P5 regulatory domain where it can form stabilizing interactions (5, 14). One possibility, therefore, is that such interactions discourage the folding ATP-lid, preventing stabilization of ATP binding and the formation of a suitable interface for productive P1 binding (41, 43). Another possibility is that differences in the P3-P4 and P4-P5 linker conformations, which have previously been shown to be critical for kinase function and coupled to P4 dipping (40, 51–53), might induce subtle changes in the ATP-binding pocket that increase ATP binding affinity for the dipped P4 state. Given that P4 dipping

would also give rise to an overall CheA configuration that destabilizes the binding of a non-productive P1 dimer, such a conformational change should ultimately lead to disassociation of the P1 dimer, freeing it for productive binding to P4.

Moving forward, the use of strategic functional mutations combined with the improved resolutions enabled by cryoET analysis of E-lysis and minicell constructs will provide a promising tool to reveal critical details of signaling mechanisms such as kinase regulation. The overall similarity between present CSU map and that obtained from *E. coli* minicells (13) suggests that insights from both contexts may be transferrable. Additionally, our atomistic CSU model will greatly enhance the potential for the investigation of signaling mechanisms using molecular simulation, providing a unique way to obtain high-resolution mechanistic insight.

MATERIALS AND METHODS

Bacterial strains, plasmids, cell culture, and cryoEM sample preparation

The pRY100 plasmid, carrying the phage φX174 lysis E gene under a tacP promoter and a lacIQ repressor control (54), was transformed into the wild-type *E. coli* K-12 strain RP437 cells using a standard protocol (55). For fluorescence imaging, cells were also transduced with a CheY-GFP construct. Cells were grown in the TB broth (1% tryptone and 0.4 NaCl, pH 7.0) supplemented with 100-µg/mL ampicillin. E gene expression was induced by addition of 0.5-mM isopropyl β-D-thiogalactopyranoside at an OD_{600} of 0.6 as previously described (15). Ten minutes after induction, an aliquot of 4-µL culture, mixed with 1-µL fiducial gold beads (10-nm size), was applied to glow-discharged Quantifoil *R*2/2 grids and vitrified with a GP2 plunge-freezing device (Leica).

Cryo-ET data acquisition

The cryoET titled series data were collected using a Thermo Fisher Titan Krios G3 instrument, with a K2 summit detector at a pixel size of 2.8 Å and a K3 summit detector at a pixel size of 2.1 Å (see Table S1). Data were collected using the SerialEM software (56, 57).

Cryo-ET reconstruction and sub-tomogram averaging

Motion correction was performed using MotionCor2 (58). The data collected by the K3 detector were rescaled to 2.8 Å/pixel by Fourier cropping in MotionCor2. The tilt series were aligned using fiducial gold beads in Etomo (59). The aligned tilt series were then imported into emClarity (17) for contrast transfer function (CTF) correction and weighted backprojection (WBP) reconstruction. About 15% of the analyzed tomograms contained arrays of CSUs. A total of 33 array-containing tomograms were selected for cryoET sub-tomogram averaging.

Template matching as implemented in emClarity was used to for sub-tomogram particle picking. The position and orientation of picked particles were plotted for visual inspection, and bad particles were manually removed. The resulting sub-tomograms were then aligned and classified according to the workflow described in Fig. S1. A total of 950 CSU trimers from the K2 data set processed using emClarity initially yielded an average structure at 8.8-Å resolution but with a strong preferred orientation. A second data set was collected with a K3 detector, and additional sub-tomograms containing side views were added. In addition, single CSU sub-volumes were extracted from the CSU trimers through symmetry expansion and were subjected to further alignment and refinement using i3 (60).

Molecular modeling and simulations

Atomic models of the CSU's constituent proteins and interfaces were predicted using AlphaFold-Multimer v.3 (18, 61) via the ColabFold (62) *AlphaFold2_mmseqs2* notebook

at https://github.com/sokrypton/ColabFold. Structural templates were not explicitly used for any of our AlphaFold predictions. A preliminary CSU model was assembled via rigid docking of the component models using UCSF ChimeraX (63). Breifly, this process was carried out by first constructing a model of the Tsr TOD using AlphaFold models of a full-length Tsr homodimer and trimer of Tsr protein-interaction regions. As the receptor C-termini (residues 517–551) were predicted to be unstructured by AlphaFold and accordingly did not have corresponding density within the cryoET map, we excluded them from further modeling. Two copies of the Tsr TOD model were then rigidly docked into our cryoET density map followed by two copies each of the $(CheA.P5/CheW)_3$ and $(CheW)_6$ hexameric ring models. Finally, a model of the CheA homodimer, excluding the CheA.P2 domain due to its variable placement within the AlphaFold models, was docked. Manual rearragements were required only for the CheA.P5 domain, whose positioning in the homodimer model was not consistent with interface 1. We therefore removed the P5 domains from the homodimer model and reattached the P4-P5 linker to the P5 domains present within the $(CheA.P5/CheW)_3$ ring models. The resulting Tsr/CheA.P5 and Tsr/CheW interactions were checked using AlphaFold models of the respective interfaces. Finally, the entire structure (Fig. S5) was conformationally refined to our cryoET data through a series of symmetry-restrained MDFF simulations (31). All MDFF simulations were carried out in NAMD v.2.14 (64, 65) and performed in the NVT ensemble at 300 K with a coupling factor of 0.3 applied to the protein backbone. Additional harmonic restraints were applied during fitting to prevent the loss of secondary structure as well as the formation of cis-peptide bonds and chirality errors. The stereochemistry of the entire CSU model was then refined using ISOLDE (66) in ChimeraX and validated using MolProbity (67).

To prepare the refined CSU model for MD simulation, an atomistic lipid bilayer consisting of 70% PVPE (1-palmitoyl 2-cis-vaccenic phosphatidylethanolamine), 20% PVPG (1-palmitoyl 2-cis-vaccenic phosphatidylglycerol), and 10% cardiolipin (1-palmitoyl 2-cis-vaccenic 3-palmitoyl 4-cis-vaccenic diphosphatidylglycerol) was assembled around the transmembrane region using CHARMM-GUI (68). The resulting structure was then solvated with TIP3P (69) water molecules and 150-mM KCl and subjected to conjugant gradient energy minimization followed by a series of equilibration simulations conducted as follows. First, solvent and lipids were permitted to relax, while the full protein was harmonically restrained following the multi-step equilibration protocol recommended by CHARMM-GUI. Next, restraints on the protein structure were slowly removed over a span of 10 ns by lowering the associated spring constants in increments of 2 ns. Finally, a 120-ns production simulation without restraints was carried out with analyses being conducted on the last 100 ns. All MD simulations were performed using NAMD v.2.14 (64, 65) and the CHARMM36 force field (70). Production simulations were conducted in the NPT ensemble with conditions maintained at 1 atm and 310 K using the Nosé-Hoover Langevin piston and Langevin thermostat, respectively. The r-RESPA integrator scheme was employed with an integration time step of 2 fs, and SHAKE constraints were applied to all hydrogen atoms. Short-range, non-bonded interactions were calculated every 2 fs with a cut-off of 12 Å; long-range electrostatics were evaluated every 6 fs using the particle-mesh Ewald method.

ACKNOWLEDGMENTS

We are grateful to Drs. Tao Ni, Jiwei Liu, Luiza Mendonca, Loic Carrique, and James Gilchrist for support.

We acknowledge Diamond for access and support of the CryoEM facilities at the UK National Electron Bio-Imaging Centre (proposal NT21004, CM22941, and BI22910), funded by the Wellcome Trust, MRC, and BBSRC. Computational aspects of this research were supported by the Wellcome Trust Core Award (grant bumber 203141/Z/16/Z) and the NIHR Oxford BRC. This project also made use of computational resources on HPC granted via the UK High-End Computing Consortium for Biomolecular Simulation, supported by EPSRC (grant no. EP/R029407/1). This work was further supported by

the UK Wellcome Trust Investigator Award (206422/Z/17/Z), the UK Biotechnology and Biological Sciences Research Council grant (BB/S003339/1), and the ERC AdG grant (101021133).

P.Z. conceived the research and designed the experiments. Z.Y. prepared samples. Z.Q., T.F., and Z.Y. collected cryoET data. Z.Q. and T.F. performed tomography reconstruction, sub-tomogram averaging, and analysis. C.K.C. performed and analyzed the modeling and MD simulations, supervised by M.S.P.S., P.J.S., and P.Z. C.K.C., Z.Q. and P.Z. analyzed the data. C.K.C., Z.Q., J.S.P., and P.Z. wrote the paper with support from all authors.

The authors declare that they have no competing interests.

AUTHOR AFFILIATIONS

[1]Diamond Light Source, Didcot, United Kingdom
[2]Department of Physics and Astronomy, University of Missouri-Columbia, Columbia, Missouri, USA
[3]Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom
[4]School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA
[5]Department of Biochemistry, University of Oxford, Oxford, United Kingdom
[6]School of Life Sciences, University of Warwick, Coventry, United Kingdom
[7]Chinese Academy of Medical Sciences Oxford Institute, University of Oxford, Oxford, United Kingdom

PRESENT ADDRESS

Zhuan Qin, Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, Connecticut, USA

AUTHOR ORCIDs

C. Keith Cassidy http://orcid.org/0000-0001-8106-1187
Zhuan Qin http://orcid.org/0000-0001-9112-7645
Peijun Zhang http://orcid.org/0000-0003-1803-691X

FUNDING

Funder	Grant(s)	Author(s)
HECBioSim	EP/R029407/1	Peijun Zhang
Wellcome Trust (WT)	206422/Z/17/Z	Peijun Zhang
UKRI | Biotechnology and Biological Sciences Research Council (BBSRC)	BB/S003339/1	Peijun Zhang
EC | H2020 | PRIORITY 'Excellent science' | H2020 European Research Council (ERC)	101021133	Peijun Zhang

AUTHOR CONTRIBUTIONS

C. Keith Cassidy, Data curation, Formal analysis, Investigation, Software, Validation, Visualization, Writing – original draft | Zhuan Qin, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft, Validation, Writing – review and editing, Software | Thomas Frosio, Data curation, Formal analysis, Methodology, Software | Khoosheh Gosink, Data curation | Zhengyi Yang, Data curation | Mark S. P. Sansom, Resources, Supervision | Phillip J. Stansfeld, Resources, Software, Supervision | John S. Parkinson, Writing – review and editing | Peijun Zhang, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review and editing

DATA AVAILABILITY

All data needed to evaluate the conclusions in the paper are present in the paper and/or the supplemental materials. The cryoET STA map of the full CSU has been deposited in the EMDB under accession code EMD-15641. The cryoET STA maps of the CSU periplasmic and baseplate regions after focused refinement and classification have been deposited in the EMDB under accession code EMD-15643 and EMD-15642, respectively. Atomic coordinates for the associated CSU model are deposited in the PDB under accession code 8C5V.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental material (mBio00793-23-s0001.pdf). Supplemental figures, table, text, and movie captions.
Movie S1 (mBio00793-23-s0002.mp4). Aligned tilt series of an E protein lysed native *E. coli* cell displaying a patch of the chemosensory array lattice.
Movie S2 (mBio00793-23-s0003.mp4). A convolution map resulted from template-matching in emClarity, overlapped with picked subtomograms (green circles).
Movie S3 (mBio00793-23-s0004.mp4). Extract from MD simulation (50-ns) of CSU model in an atomistic lipid bilayer, illustrating interactions between neighbouring receptor ligand binding domains (LBD). Clusters of residues mediating LBD interactions are shown as VDW spheres and colored in blue (residues E124, K126, R127, D130) or yellow (residues K99, E102, and K103).

REFERENCES

1. Keegstra JM, Carrara F, Stocker R. 2022. The ecological roles of bacterial chemotaxis. Nat Rev Microbiol 20:491–504. https://doi.org/10.1038/s41579-022-00709-w
2. Parkinson JS, Hazelbauer GL, Falke JJ. 2015. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. Trends Microbiol 23:257–266. https://doi.org/10.1016/j.tim.2015.03.003
3. Falke JJ, Piasta KN. 2014. Architecture and signal transduction mechanism of the bacterial chemosensory array: progress, controversies, and challenges. Curr Opin Struct Biol 29:85–94. https://doi.org/10.1016/j.sbi.2014.10.001
4. Li M, Hazelbauer GL. 2011. Core unit of chemotaxis signaling complexes. Proc Natl Acad Sci U S A 108:9390–9395. https://doi.org/10.1073/pnas.1104824108
5. Muok AR, Briegel A, Crane BR. 2020. Regulation of the chemotaxis histidine kinase CheA: a structural perspective. Biochim Biophys Acta Biomembr 1862:183030. https://doi.org/10.1016/j.bbamem.2019.183030
6. Liu J, Hu B, Morado DR, Jani S, Manson MD, Margolin W. 2012. Molecular architecture of chemoreceptor arrays revealed by cryoelectron tomography of *Escherichia coli* minicells. Proc Natl Acad Sci U S A 109:E1481–8. https://doi.org/10.1073/pnas.1200781109
7. Cassidy CK, Himes BA, Alvarez FJ, Ma J, Zhao G, Perilla JR, Schulten K, Zhang P. 2015. CryoEM and computer simulations reveal a novel kinase conformational switch in bacterial chemotaxis signaling. Elife 4:e08419. https://doi.org/10.7554/eLife.08419
8. Piñas GE, DeSantis MD, Cassidy CK, Parkinson JS. 2022. Hexameric rings of the scaffolding protein CheW enhance response sensitivity and cooperativity in *Escherichia coli* chemoreceptor arrays. Sci Signal 15:718. https://doi.org/10.1126/scisignal.abj1737
9. Briegel A, Ortega DR, Tocheva EI, Wuichet K, Li Z, Chen S, Müller A, Iancu CV, Murphy GE, Dobro MJ, Zhulin IB, Jensen GJ. 2009. Universal architecture of bacterial chemoreceptor arrays. Proc Natl Acad Sci U S A 106:17181–17186. https://doi.org/10.1073/pnas.0905181106
10. Hadjidemetriou K, Kaur S, Cassidy CK, Zhang P. 2022. Mechanisms of *E. coli* chemotaxis signaling pathways visualized using cryoET and computational approaches. Biochem Soc Trans 50:1595–1605. https://doi.org/10.1042/BST20220191
11. Briegel A, Li X, Bilwes AM, Hughes KT, Jensen GJ, Crane BR. 2012. Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. Proc Natl Acad Sci U S A 109:3766–3771. https://doi.org/10.1073/pnas.1115719109
12. Yang W, Cassidy CK, Ames P, Diebolder CA, Schulten K, Luthey-Schulten Z, Parkinson JS, Briegel A. 2019. *In situ* conformational changes of the *Escherichia coli* serine chemoreceptor in different signaling states. mBio 10:e00973-19. https://doi.org/10.1128/mBio.00973-19
13. Burt A, Cassidy CK, Ames P, Bacia-Verloop M, Baulard M, Huard K, Luthey-Schulten Z, Desfosses A, Stansfeld PJ, Margolin W, Parkinson JS, Gutsche I. 2020. Complete structure of the chemosensory array core signalling unit in an *E. coli* minicell strain. Nat Commun 11:743. https://doi.org/10.1038/s41467-020-14350-9
14. Cassidy CK, Himes BA, Sun D, Ma J, Zhao G, Parkinson JS, Stansfeld PJ, Luthey-Schulten Z, Zhang P. 2020. Structure and dynamics of the *E. coli* chemotaxis core signaling complex by cryo-electron tomography and molecular simulations. Commun Biol 3:24. https://doi.org/10.1038/s42003-019-0748-0
15. Fu X, Himes BA, Ke D, Rice WJ, Ning J, Zhang P. 2014. Controlled bacterial lysis for electron tomography of native cell membranes. Structure 22:1875–1882. https://doi.org/10.1016/j.str.2014.09.017
16. Phillips GJ. 2001. Green fluorescent protein--a bright idea for the study of bacterial protein localization. FEMS Microbiol Lett 204:9–18. https://doi.org/10.1111/j.1574-6968.2001.tb10854.x
17. Himes BA, Zhang P. 2018. emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. Nat Methods 15:955–961. https://doi.org/10.1038/s41592-018-0167-z
18. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior

AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2
19. Ferris HU, Zeth K, Hulko M, Dunin-Horkawicz S, Lupas AN. 2014. Axial helix rotation as a mechanism for signal regulation inferred from the crystallographic analysis of the *E. coli* serine chemoreceptor. J Struct Biol 186:349–356. https://doi.org/10.1016/j.jsb.2014.03.015
20. Tajima H, Imada K, Sakuma M, Hattori F, Nara T, Kamo N, Homma M, Kawagishi I. 2011. Ligand specificity determined by differentially arranged common ligand-binding residues in bacterial amino acid chemoreceptors Tsr and Tar. J Biol Chem 286:42200–42210. https://doi.org/10.1074/jbc.M111.221887
21. Gushchin I, Melnikov I, Polovinkin V, Ishchenko A, Yuzhakova A, Buslaev P, Bourenkov G, Grudinin S, Round E, Balandin T, Borshchevskiy V, Willbold D, Leonard G, Büldt G, Popov A, Gordeliy V. 2017. Mechanism of transmembrane signaling by sensor histidine kinases. Science 356:eaah6345. https://doi.org/10.1126/science.aah6345
22. Pakula AA, Simon MI. 1992. Determination of transmembrane protein structure by disulfide cross- linking: the *Escherichia coli* Tar receptor. Proc Natl Acad Sci U S A 89:4144–4148. https://doi.org/10.1073/pnas.89.9.4144
23. Falke JJ, Hazelbauer GL. 2001. Transmembrane signaling in bacterial chemoreceptors. Trends Biochem Sci 26:257–265. https://doi.org/10.1016/s0968-0004(00)01770-9
24. Kitanovic S, Ames P, Parkinson JS. 2011. Mutational analysis of the control cable that mediates transmembrane signaling in the *Escherichia coli* serine chemoreceptor. J Bacteriol 193:5062–5072. https://doi.org/10.1128/JB.05683-11
25. Bilwes AM, Alex LA, Crane BR, Simon MI. 1999. Structure of CheA, a signal-transducing histidine kinase. Cell 96:131–141. https://doi.org/10.1016/s0092-8674(00)80966-6
26. Greenswag AR, Muok A, Li X, Crane BR. 2015. Conformational transitions that enable histidine kinase autophosphorylation and receptor array integration. J Mol Biol 427:3890–3907. https://doi.org/10.1016/j.jmb.2015.10.015
27. Gloor SL, Falke JJ. 2009. Thermal domain motions of CheA kinase in solution: disulfide trapping reveals the motional constraints leading to trans-autophosphorylation. Biochemistry 48:3631–3644. https://doi.org/10.1021/bi900033r
28. Li X, Fleetwood AD, Bayas C, Bilwes AM, Ortega DR, Falke JJ, Zhulin IB, Crane BR. 2013. The 3.2 Å resolution structure of a receptor:CheA:CheW signaling complex defines overlapping binding sites and key residue interactions within bacterial chemosensory arrays. Biochemistry 52:3852–3865. https://doi.org/10.1021/bi400383e
29. Piñas GE, Frank V, Vaknin A, Parkinson JS. 2016. The source of high signal cooperativity in bacterial chemosensory arrays. Proc Natl Acad Sci U S A 113:3335–3340. https://doi.org/10.1073/pnas.1600216113
30. Burt A, Cassidy CK, Stansfeld PJ, Gutsche I. 2021. Alternative architecture of the *E. coli* chemosensory array. Biomolecules 11:495. https://doi.org/10.3390/biom11040495
31. McGreevy R, Teo I, Singharoy A, Schulten K. 2016. Advances in the molecular dynamics flexible fitting method for cryo-EM modeling. Methods 100:50–60. https://doi.org/10.1016/j.ymeth.2016.01.009
32. Amin DN, Hazelbauer GL. 2012. Influence of membrane lipid composition on a transmembrane bacterial chemoreceptor. J Biol Chem 287:41697–41705. https://doi.org/10.1074/jbc.M112.415588
33. McEvoy MM, de la Cruz AF, Dahlquist FW. 1997. Large modular proteins by NMR. Nat Struct Biol 4:9. https://doi.org/10.1038/nsb0197-9
34. Jahreis K, Morrison TB, Garzón A, Parkinson JS. 2004. Chemotactic signaling by an *Escherichia coli* CheA mutant that lacks the binding domain for phosphoacceptor partners. J Bacteriol 186:2664–2672. https://doi.org/10.1128/JB.186.9.2664-2672.2004
35. Nishiyama S, Garzón A, Parkinson JS. 2014. Mutational analysis of the P1 phosphorylation domain in *Escherichia coli* CheA, the signaling kinase for chemotaxis. J Bacteriol 196:257–264. https://doi.org/10.1128/JB.01167-13
36. Piñas GE, Parkinson JS. 2019. Identification of a kinase-active CheA conformation in *Escherichia coli* chemoreceptor signaling complexes. J Bacteriol 201:1–14. https://doi.org/10.1128/JB.00543-19
37. Kott L, Braswell EH, Shrout AL, Weis RM. 2004. Distributed subunit interactions in CheA contribute to dimer stability: a sedimentation equilibrium study. Biochim Biophys Acta 1696:131–140. https://doi.org/10.1016/j.bbapap.2003.10.001
38. Hamel DJ, Zhou H, Starich MR, Byrd RA, Dahlquist FW. 2006. Chemical-shift-perturbation mapping of the phosphotransfer and catalytic domain interaction in the histidine autokinase CheA from *Thermotoga maritima*. Biochemistry 45:9509–9517. https://doi.org/10.1021/bi060798k
39. Muok AR, Chua TK, Srivastava M, Yang W, Maschmann Z, Borbat PP, Chong J, Zhang S, Freed JH, Briegel A, Crane BR. 2020. Engineered chemotaxis core signaling units indicate a constrained kinase-off state. Sci Signal 13:eabc1328. https://doi.org/10.1126/scisignal.abc1328
40. Maschmann Z, Chandrasekaran S, Chua TK, Crane BR. 2022. Interdomain linkers regulate histidine kinase activity by controlling subunit interactions. Biochemistry 61:2672–2686. https://doi.org/10.1021/acs.biochem.2c00326
41. Zhang J, Xu Y, Shen J, Luo X, Chen J, Chen K, Zhu W, Jiang H. 2005. Dynamic mechanism for the autophosphorylation of CheA histidine kinase: molecular dynamics simulations. J Am Chem Soc 127:11709–11719. https://doi.org/10.1021/ja051199o
42. Quezada CM, Hamel DJ, Gradinaru C, Bilwes AM, Dahlquist FW, Crane BR, Simon MI. 2005. Structural and chemical requirements for histidine phosphorylation by the chemotaxis kinase CheA. J Biol Chem 280:30581–30585. https://doi.org/10.1074/jbc.M505316200
43. Bilwes AM, Quezada CM, Croal LR, Crane BR, Simon MI. 2001. Nucleotide binding by the histidine kinase CheA. Nat Struct Biol 8:353–360. https://doi.org/10.1038/86243
44. Jun SY, Pan W, Hazelbauer GL. 2020. ATP binding as a key target for control of the chemotaxis kinase. J Bacteriol 202:1–13. https://doi.org/10.1128/JB.00095-20
45. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ. 2016. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. J Mol Biol 428:720–725. https://doi.org/10.1016/j.jmb.2015.09.014
46. Garzón A, Parkinson JS. 1996. Chemotactic signaling by the P1 phosphorylation domain liberated from the CheA histidine kinase of *Escherichia coli*. J Bacteriol 178:6752–6758. https://doi.org/10.1128/jb.178.23.6752-6758.1996
47. Pan W, Dahlquist FW, Hazelbauer GL. 2017. Signaling complexes control the chemotaxis kinase by altering its apparent rate constant of autophosphorylation. Protein Sci 26:1535–1546. https://doi.org/10.1002/pro.3179
48. Irieda H, Homma M, Homma M, Kawagishi I. 2006. Control of chemotactic signal gain via modulation of a pre-formed receptor array. J Biol Chem 281:23880–23886. https://doi.org/10.1074/jbc.M600018200
49. Briegel A, Ames P, Gumbart JC, Oikonomou CM, Parkinson JS, Jensen GJ. 2013. The mobility of two kinase domains in the *Escherichia coli* chemoreceptor array varies with signalling state. Mol Microbiol 89:831–841. https://doi.org/10.1111/mmi.12309
50. Mello BA, Pan W, Hazelbauer GL, Tu Y. 2018. A dual regulation mechanism of histidine kinase CheA identified by combining network-dynamics modeling and system-level input-output data. PLoS Comput Biol 14:e1006305. https://doi.org/10.1371/journal.pcbi.1006305
51. Wang X, Wu C, Vu A, Shea JE, Dahlquist FW. 2012. Computational and experimental analyses reveal the essential roles of interdomain linkers in the biological function of chemotaxis histidine kinase CheA. J Am Chem Soc 134:16107–16110. https://doi.org/10.1021/ja3056694
52. Wang Xiqing, Vallurupalli P, Vu A, Lee K, Sun S, Bai W-J, Wu C, Zhou H, Shea J-E, Kay LE, Dahlquist FW. 2014. The linker between the dimerization and catalytic domains of the CheA histidine kinase propagates changes in structure and dynamics that are important for enzymatic activity. Biochemistry 53:855–861. https://doi.org/10.1021/bi4012379
53. Ding X, He Q, Shen F, Dahlquist FW, Wang X. 2018. Regulatory role of an interdomain linker in the bacterial chemotaxis histidine kinase CheA. J Bacteriol 200:10. https://doi.org/10.1128/JB.00052-18
54. Roof WD, Fang HQ, Young KD, Sun J, Young R. 1997. Mutational analysis of *slyD*, an *Escherichia coli* gene encoding a protein of the FKBP immunophilin family. Mol Microbiol 25:1031–1046. https://doi.org/10.1046/j.1365-2958.1997.5201884.x

55. Hanahan D, Jessee J, Bloom FR. 1991. Plasmid transformation of *Escherichia coli* and other bacteria. Methods Enzymol 204:63–113. https://doi.org/10.1016/0076-6879(91)04006-a
56. Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol 152:36–51. https://doi.org/10.1016/j.jsb.2005.07.007
57. Hagen WJH, Wan W, Briggs JAG. 2017. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. J Struct Biol 197:191–198. https://doi.org/10.1016/j.jsb.2016.06.007
58. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods 14:331–332. https://doi.org/10.1038/nmeth.4193
59. Mastronarde DN, Held SR. 2017. Automated tilt series alignment and tomographic reconstruction in IMOD. J Struct Biol 197:102–113. https://doi.org/10.1016/j.jsb.2016.07.011
60. Winkler H. 2007. 3D reconstruction and processing of volumetric data in cryo-electron tomography. J Struct Biol 157:126–137. https://doi.org/10.1016/j.jsb.2006.07.014
61. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, Ronneberger O, Bodenstein S, Zielinski M, Bridgland A, Potapenko A, Cowie A, Tunyasuvunakool K, Jain R, Clancy E, Kohli P, Jumper J, Hassabis D. 2022. Protein complex prediction with AlphaFold-multimer. Bioinformatics. https://doi.org/10.1101/2021.10.04.463034
62. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. Nat Methods 19:679–682. https://doi.org/10.1038/s41592-022-01488-1
63. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci 27:14–25. https://doi.org/10.1002/pro.3235
64. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. 2005. Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802. https://doi.org/10.1002/jcc.20289
65. Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Hénin J, Jiang W, McGreevy R, Melo MCR, Radak BK, Skeel RD, Singharoy A, Wang Y, Roux B, Aksimentiev A, Luthey-Schulten Z, Kalé LV, Schulten K, Chipot C, Tajkhorshid E. 2020. Scalable molecular dynamics on CPU and GPU architectures with NAMD. J Chem Phys 153:044130. https://doi.org/10.1063/5.0014475
66. Croll TI. 2018. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps . Acta Crystallogr D Struct Biol 74:519–530. https://doi.org/10.1107/S2059798318002425
67. Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB 3rd, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci 27:293–315. https://doi.org/10.1002/pro.3330
68. Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, Wei S, Buckner J, Jeong JC, Qi Y, Jo S, Pande VS, Case DA, Brooks CL 3rd, MacKerell AD Jr, Klauda JB, Im W. 2016. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. J Chem Theory Comput 12:405–413. https://doi.org/10.1021/acs.jctc.5b00935
69. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935. https://doi.org/10.1063/1.445869
70. Huang J, MacKerell AD Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem 34:2135–2145. https://doi.org/10.1002/jcc.23354

Supplementary materials

for

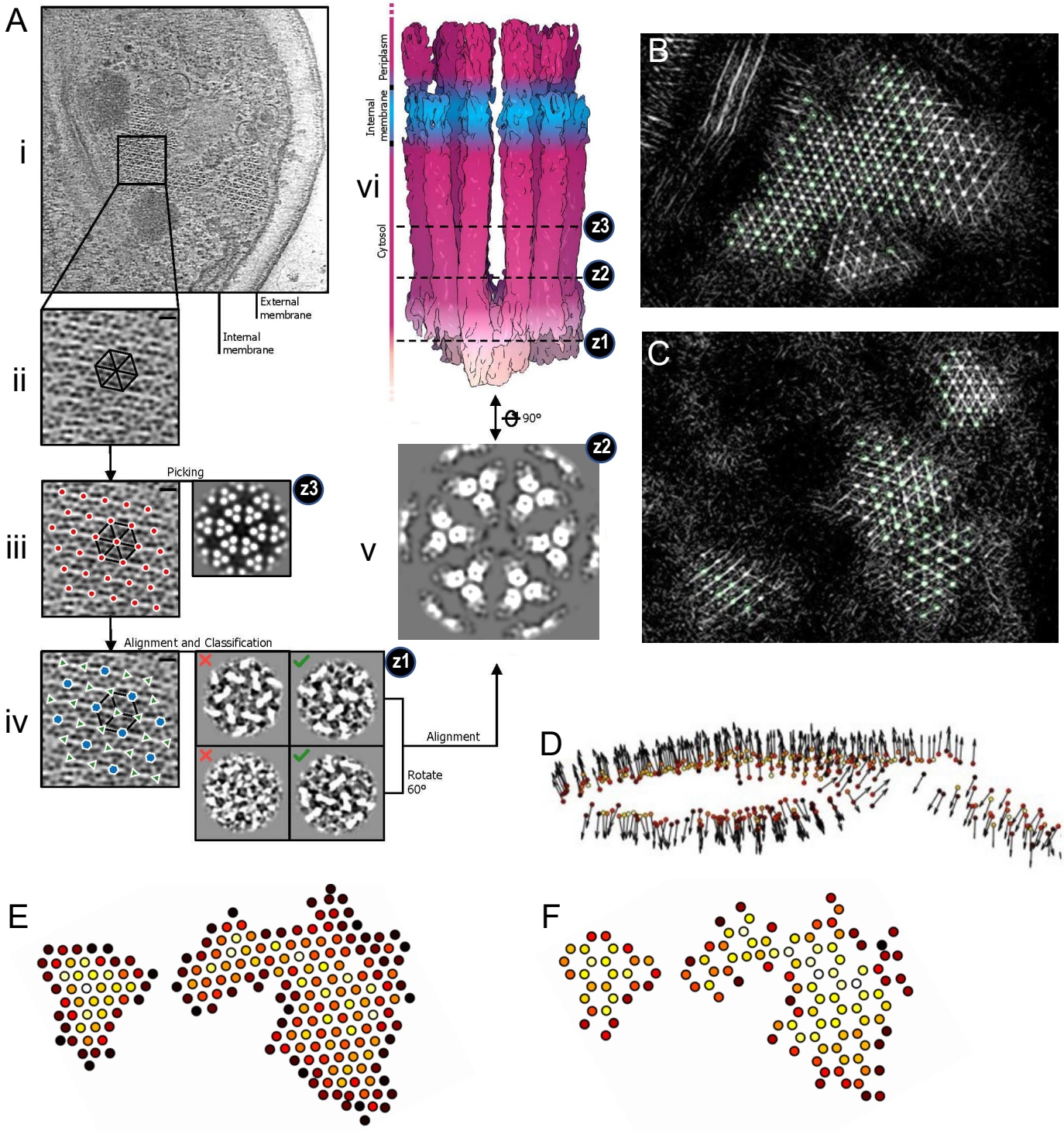
Structure of the Native Chemotaxis Core Signalling Unit from E-gene lysed *E. coli* cells

C. Keith Cassidy, Zhuan Qin, Thomas Frosio, Khoosheh Gosink, Zhengyi Yang, Mark Sansom, Phillip J. Stansfeld, John S Parkinson, Peijun Zhang

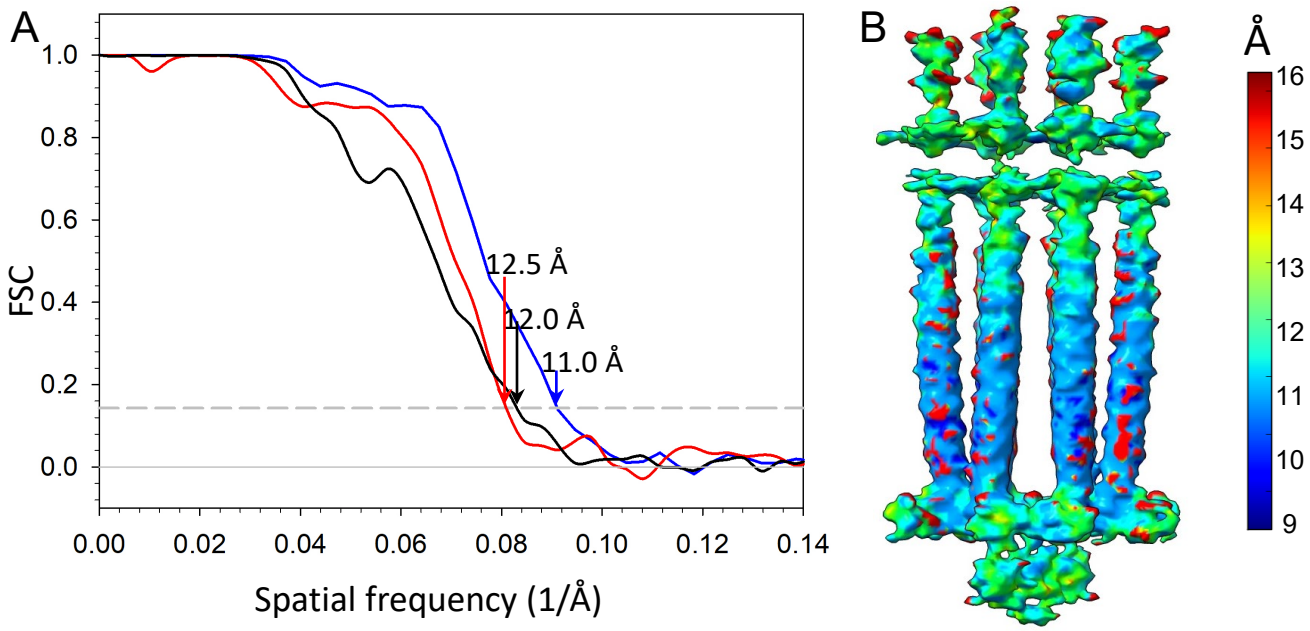
List of Supplementary materials

Supplementary Table 1 | CryoET data collection and structure determination of CSU from native *E. coli* cells.

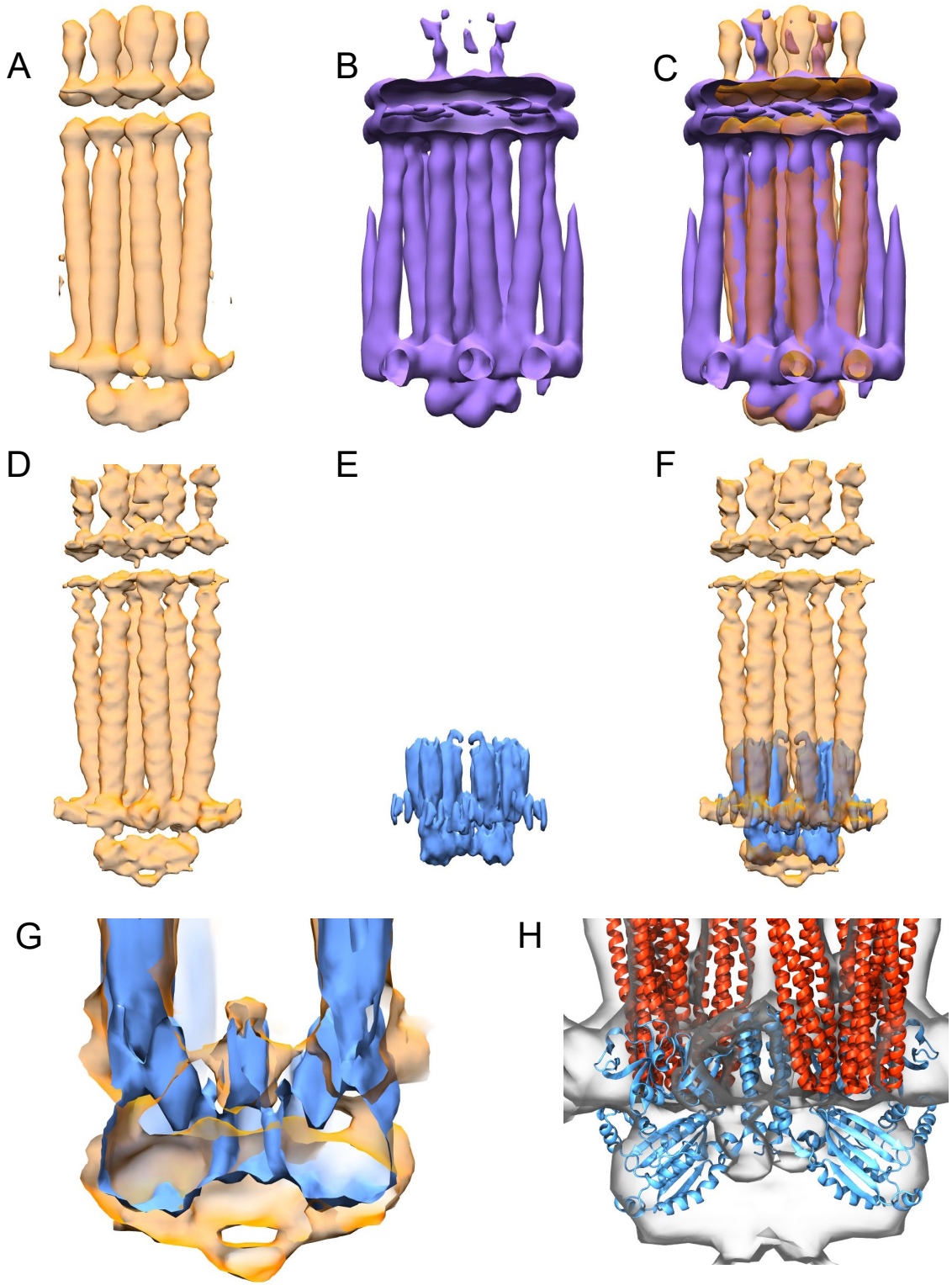
Data acquisition			
Dataset	Titan Krio (K2)	Titan Krio (K3)	
Voltage (keV)	300	300	
Detector	Gatan Quantum K2 Direct Electron Detector	Gatan Quantum K3 Direct Electron Detector	
Energy-filter	Yes	yes	
Slit width (eV)	20	20	
Super-resolution mode	no	yes	
Pixel size (Å/pixel)	2.8	2.1	
Defocus range (μm)	-2 to -3	-0.5 to -3	
Acquisition scheme	Dose-Symmetric, -60/60, 3^{o} step, group 3	Dose-Symmetric, -60/60, 3^{o} step, group 3	
Total Dose (electrons/Å^{2})	120	120	
Number of Frames	8	10	
Number of Tomograms	21	12	
Structure determination			
No. of subotomograms	3596	1509	
Resolution (Å)	Full CSU	periplasmic	baseplate
	12.0	12.5	11.0
Data deposited	EMD-15641 PDB-8C5V	EMD-15643	EMD-15642



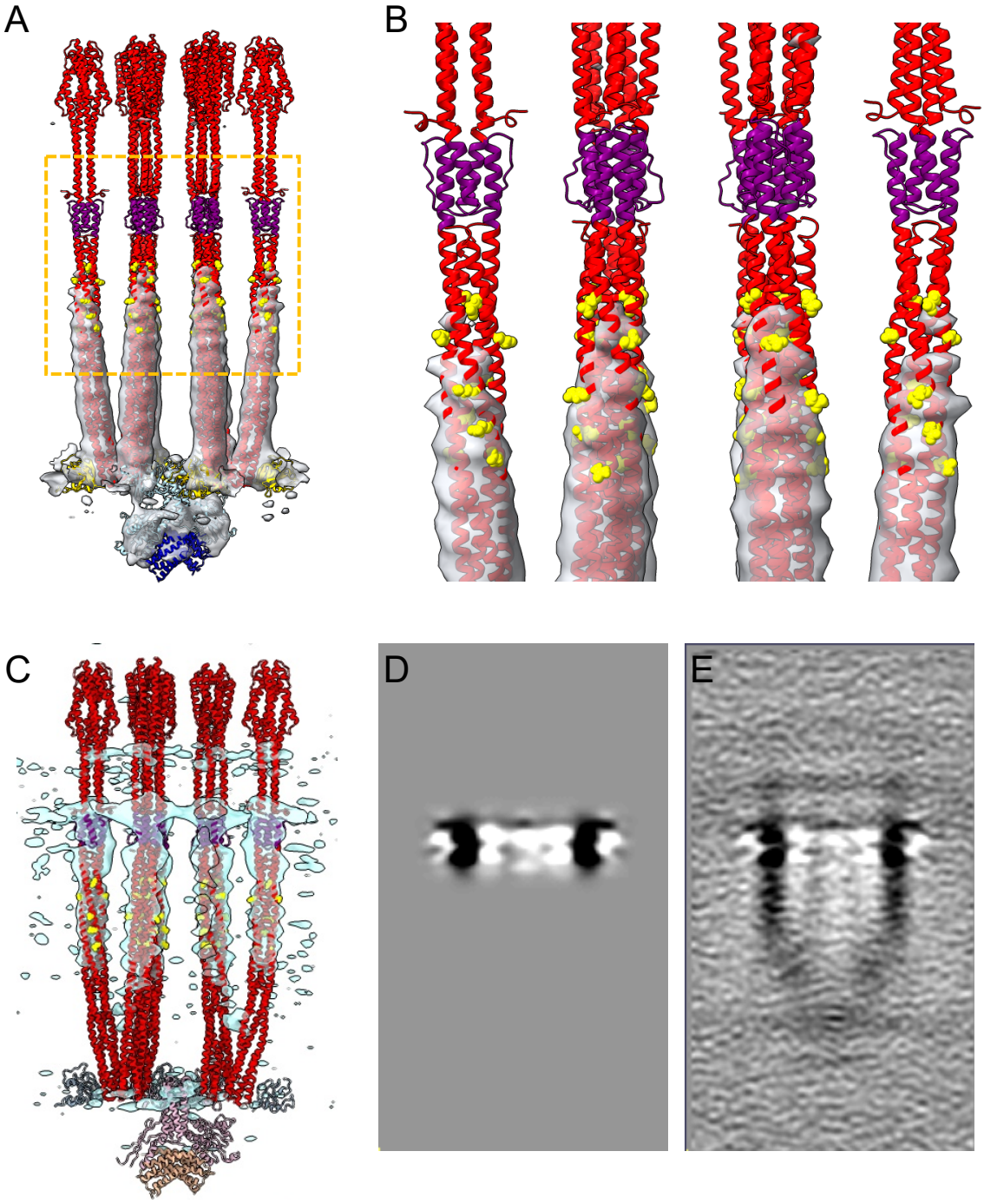
Supplementary Figure 1 | CryoET STA workflow for the CSU trimer. (A) Overview of the cryoET STA workflow. Z1 and Z2 indicate the positions of cross-sections. Steps are indicated as i) raw tomogram, ii) template matching, iii) subvolume picking (the average is shown on the right), iv) alignment and classification (four classes are shown on the right), v) average of CSU trimer, and vi) symmetry expansion and final average of a single CSU. (B-C) Two representative sections of the same conversion map overlapped with picked subvolumes (green circles). (D) Orientations of the picked subvolumes plotted with arrows. (E) Positions of subvolumes after picking in A (iii), where receptor signal dictates. (F) Positions of subvolumes after classification in A (iv), where CSU trimer classes are selected (green checks). The colors indicate cross correlation values, from lowest (black) to highest (white).



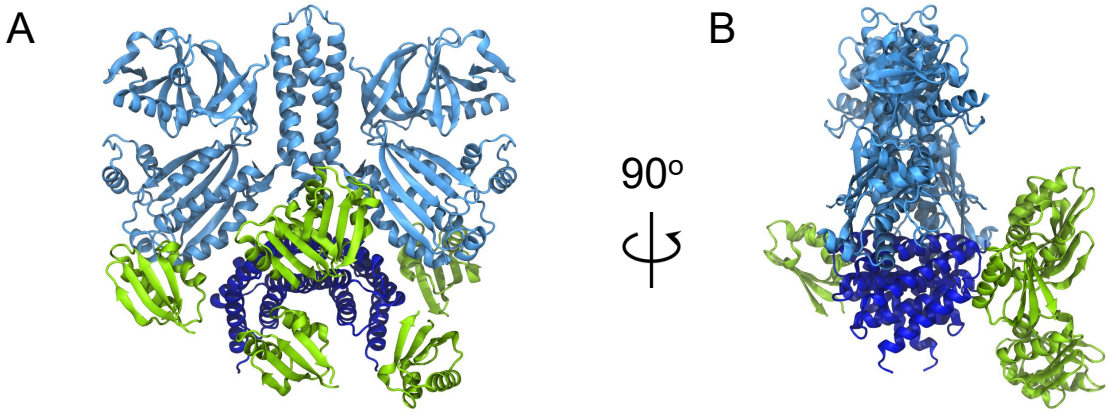
Supplementary Figure 2 | Fourier Shell Correlation (FSC) of the CSU subtomogram average. (A) FSC plots of the overall CSU (black), chemoreceptor ligand binding domains (red), and baseplate region (blue). (B) Local resolution map of the overall CSU, colored from blue to red as indicated in the key.



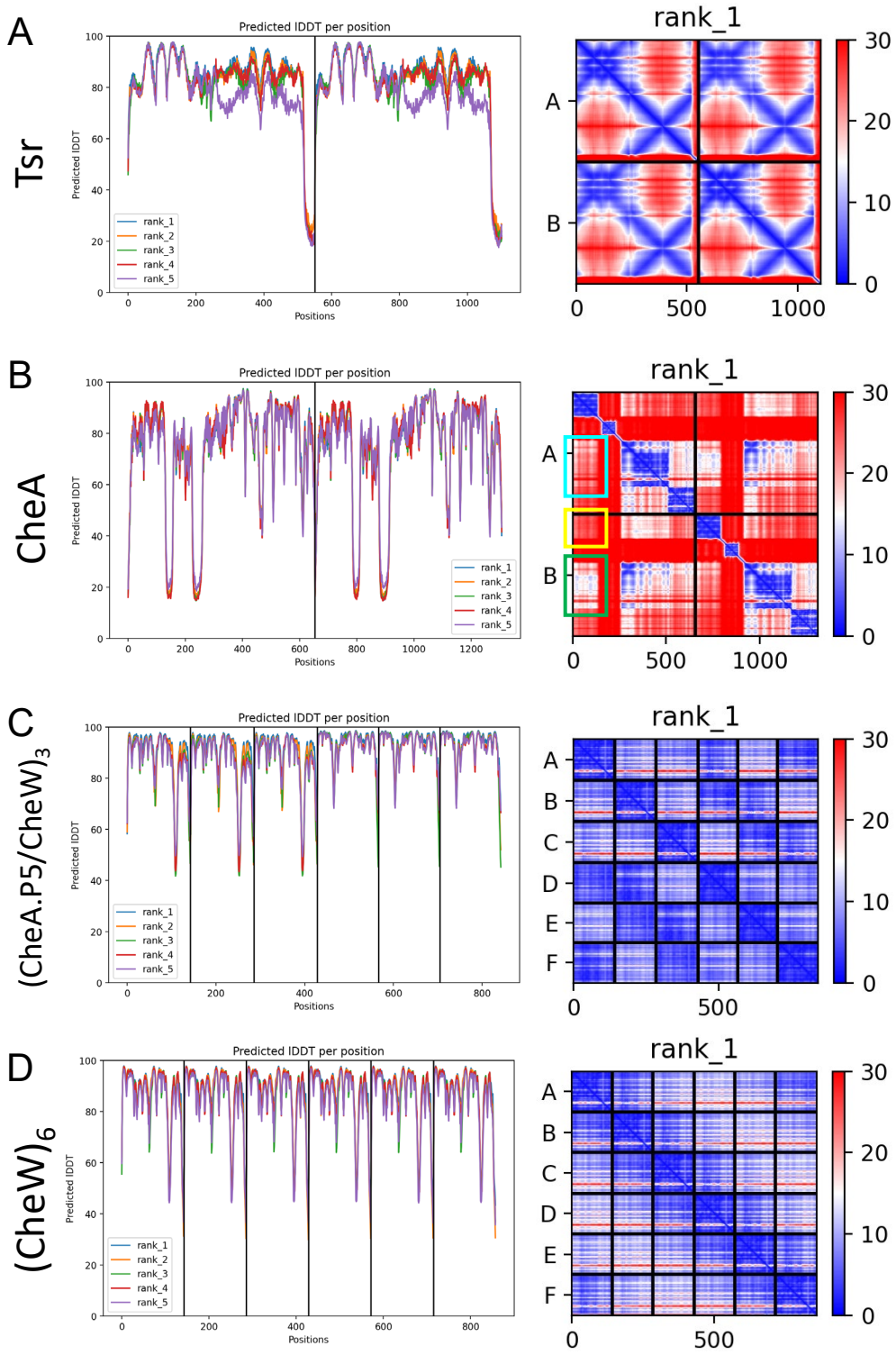
Supplementary Figure 3 | Comparison of CSU maps. (A) CSU map from *E. coli* ghost cells, low-pass filtered to 16 Å resolution. (B) CSU map at 16 Å resolution from *E. coli* minicells (EMD-10160). (C) Overlay of A and B. (D) CSU map from *E. coli* ghost cells, median filtered at 12 Å resolution. (E) CSU map from *E. coli* 4Q monolayer arrays (EMD-10050), low-pass filtered to 12 Å resolution. Note the P1 and P2 domains of CheA are not present in the monolayer sample. (F) Overlay of the D and E. (G) A slice through a zoomed view of (F), highlighting excess density between CheA P4 domains. (H) Overlay of CSU map from *E. coli* ghost cells with model constructed from 4Q monolayer arrays (PDB 6S1K). For clarity, CheW is not shown.



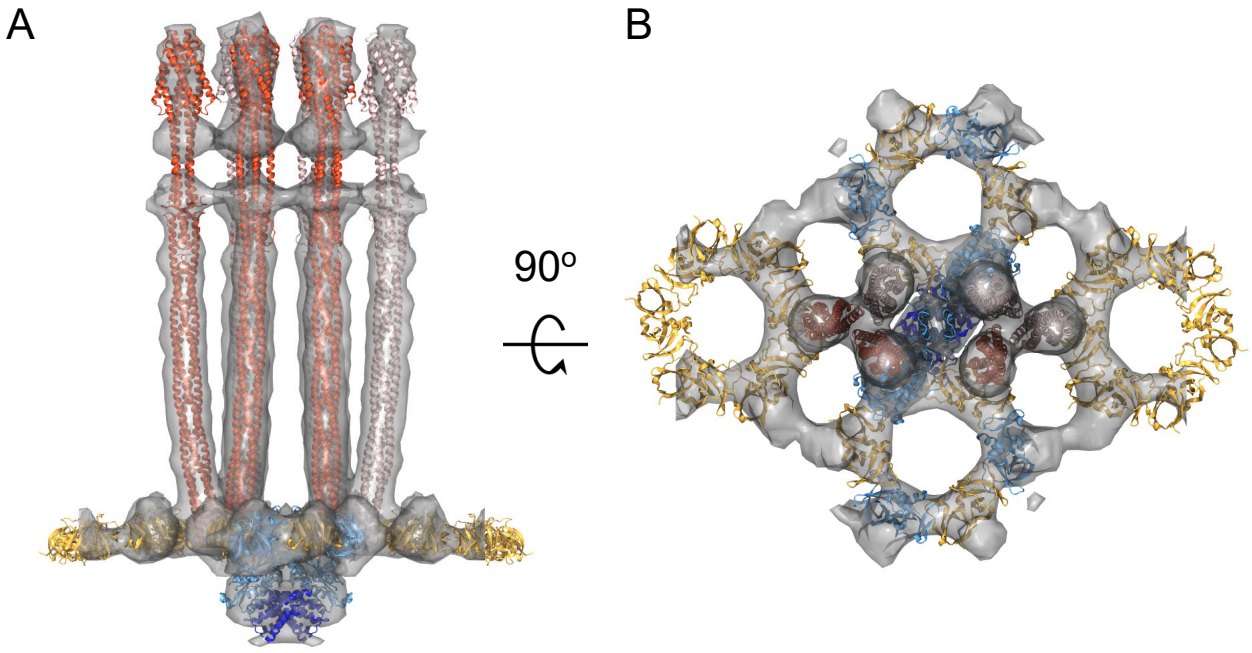
Supplementary Figure 4 | Baseplate- and HAMP-focused refinements. (A) Overlay between the CSU model and the baseplate-focused refined map. (B) Zoom onto region outlined by the dashed orange box in (A). In the model, the receptors are colored in red, the HAMP domains in purple, and the methylation sites (Q297, E304, Q311, E493, E502) are shown as VDW spheres and highlighted in yellow. (C) Overlay between the CSU model and the HAMP-focused refined map. (D-E) The HAMP focused refinement mask (D) and the resulting map (E).



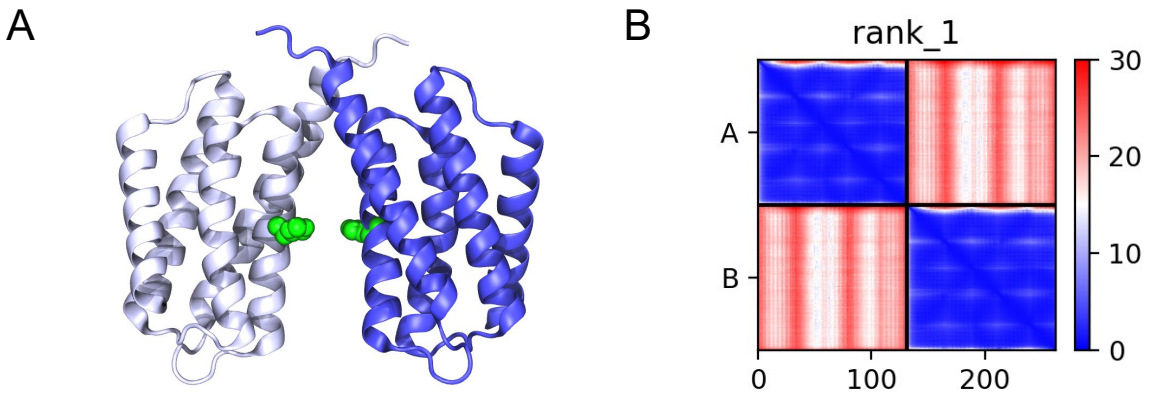
Supplementary Figure 5 | Variable placement of CheA.P2 by AlphaFold2. (A-B) Overlay of five full-length CheA dimer models obtained from Alphafold2, shown from the side (A) and rotated by 90 degrees (B). The position of the P2 domain (green) from one monomer in each model is shown. Note, P2' (not shown) is placed symmetrically with respect to the opposing side of the CheA dimer. For clarity, CheA.P3-P5 (light blue) and CheA.P1 (dark blue), whose positions are unaltered between the models, are shown here for a single model and the disordered linkers connecting P1 to P2 and P2 to P3 are not shown.



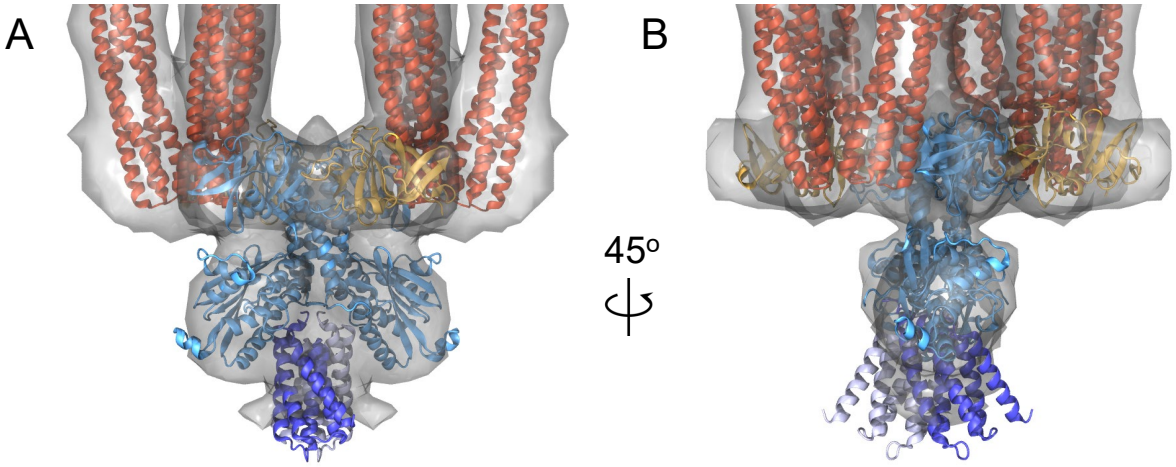
Supplementary Figure 6 | pLDDT and PAE plots for AlphaFold predictions. (A-D) Plots of the Predicted Local Distance Difference Test (pLDDT) and Predicted Aligned Error (PAE) scores output by AlphaFold for full-length Tsr (A), full-length CheA (B), the $(CheA.P4/CheW)_3$ hexameric ring (C), and the $(CheW)_6$ hexameric ring (D). The PAE plot is provided for the top model only. In (B), for one of the P1 domains, PAE regions corresponding to P1/P1, intra-chain P1/P4, and inter-chain P1/P4 interactions are highlighted by yellow, cyan, and green boxes respectively.



Supplementary Figure 7 | Molecular Dynamics Flexible Fitting (MDFF). (A-B) Overlay between the CSU density map and MDFF-refined CSU model, shown from the side (A) and top (B). The fitted structure included six Tsr homodimers (red/pink), a central CheA.P3-5 dimer (light blue) with associated P1 domains (dark blue), two $(CheA.P5/CheW)_3$ rings (blue and gold), and two $(CheW)_6$ rings (gold).



Supplementary Figure 8 | Parallel P1 dimer predicted by AlphaFold for isolated P1 monomers. (A) Coordinates of parallel P1 dimer predicted for isolated *E. coli* P1 domains. P1 monomers are colored in light and dark blue; the substrate histidine is colored green and shown in VDW representation. (B) PAE plot for top model from prediction.



Supplementary Figure 9 | CheA keel density with docked parallel P1 dimer. (A-B) Two views of a parallel P1 dimer model rigidly docked into the excess keel density. Molecules are colored as in other panels.

Movie legends:

Movie 1. Aligned tilt series of an E-gene lyszed native *E. coli* cell displaying a patch of the chemosensory array lattice.

Movie 2. A convolution map resulted from template-matching in emClarity, overlapped with picked subtomograms (green circles).

Movie 3. Extract from MD simulation (50-ns) of CSU model in an atomistic lipid bilayer, illustrating interactions between neighbouring receptor ligand binding domains (LBD). Clusters of residues mediating LBD interactions are shown as VDW spheres and coloured in blue (residues E124, K126, R127, D130) or yellow (residues K99, E102, and K103).