

**CryoEM and Computer Simulations Reveal a Novel Kinase Conformational Switch in
Bacterial Chemotaxis Signaling**

C. Keith Cassidy^{1*}, Benjamin A. Himes^{2*}, Frances J. Alvarez^{2*}, Jun Ma², Gongpu Zhao², Juan R.
Perilla¹, Klaus Schulten^{1†} and Peijun Zhang^{2†}

¹Department of Physics and Beckman Institute, University of Illinois at Urbana-Champaign,
Urbana, IL 61801, USA

²Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA
15260, USA

*These authors contributed equally to this work

[†]To whom correspondence may be addressed. Email: pez7@pitt.edu; kschulte@ks.uiuc.edu

Abstract

Chemotactic responses in bacteria require large, highly ordered arrays of sensory proteins to mediate the signal transduction that ultimately controls cell motility. A mechanistic understanding of the molecular events underlying signaling, however, has been hampered by the lack of a high-resolution structural description of the extended array. Here, we report a novel reconstitution of the array, involving the receptor signaling domain, histidine kinase CheA, and adaptor protein CheW, as well as a density map of the core-signaling unit at 11.3 Å resolution, obtained by cryo-electron tomography and sub-tomogram averaging. Extracting key structural constraints from our density map, we computationally construct and refine an atomic model of the core array structure, exposing novel interfaces between the component proteins. Using all-atom molecular dynamics simulations, we further reveal a distinctive conformational change in CheA. Mutagenesis and chemical cross-linking experiments confirm the importance of the conformational dynamics of CheA for chemotactic function.

Introduction

Bacterial chemotaxis is a ubiquitous, two-component signal transduction system that allows cells to extract information from environmental chemical gradients and place themselves within the nutrient-optimal portion of their habitat (Wadhams and Armitage 2004, Capra and Laub 2012, Eisenbach and Lengeler 2004). Though the topology and complexity of the protein networks employed in bacterial chemotaxis vary by species, each uses the histidine kinase CheA (component 1) and response regulator CheY (component 2) to set up an intracellular phosphorylation cascade that regulates the motile behavior of the cell (Szurmant and Ordal 2004). CheA, in particular, is a multi-domain protein, consisting of five separate and functionally distinct domains (P1-P5): P1-phosphoryl transfer domain, P2-substrate binding domain, P3-dimerization domain, P4-kinase domain and P5-regulatory domain. In addition to CheA and CheY, an expanded set of molecules assist in the mechanics of signal reception, transmission, and regulation. Specifically, bacteria utilize dedicated chemoreceptors (also known as methyl-accepting chemotaxis proteins, MCPs) to recognize ambient chemicals and transmit mechanical signals across the cell membrane to affect CheA kinase activity (Ortega, Yang, et al. 2013, Parkinson, Hazelbauer, and Falke 2015). The adaptor protein CheW universally participates in the coupling of conformational changes within receptors to kinase regulation (Szurmant and Ordal 2004, Liu and Parkinson 1989). Bacteria, moreover, have evolved the ability to tune or adapt their chemotactic sensitivity to stimulus intensity, giving rise to short-term molecular memory and allowing an appropriate system response over wide ranges of chemical concentrations (Hazelbauer and Lai 2010, Parkinson, Hazelbauer, and Falke 2015). In the case of the model organism, *Escherichia coli*, the adaptation mechanism involves the use of two

enzymes, CheR and CheB, which reversibly modify specific residues in the receptor molecules (Hazelbauer, Falke, and Parkinson 2008, Hazelbauer and Lai 2010, Goy, Springer, and Adler 1977, Ortega, Yang, et al. 2013).

The tunable control of chemotactic activity requires the assembly of collaborative core-signaling units, involving the chemoreceptor trimer of dimers (TOD) (Amin and Hazelbauer 2010, Li et al. 2011), CheA dimer and CheW monomer (Li and Hazelbauer 2011, Falke and Piasta 2014). Through the formation of large, highly organized clusters known as chemosensory arrays, thousands of core-signaling units establish a network of cooperative interactions that dramatically affect signal transmission and regulation and endow the basic two-component chemotaxis infrastructure with heightened information processing and control capabilities (Hazelbauer and Lai 2010, Falke and Piasta 2014, Sourjik and Armitage 2010, Bray, Levin, and Morton-Firth 1998, Tu 2013). Important progress has been made in the characterization of localized portions of array structure using a battery of genetic, biochemical, and biophysical techniques. This progress includes the derivation of atomic structures of the individual core signaling components (Kim, Yokota, and Kim 1999, Bilwes et al. 1999, Park et al. 2006, Li et al. 2007, Griswold et al. 2002) and several of their sub-complexes (Park et al. 2006, Li et al. 2013, Briegel et al. 2012) as well as the elucidation of key interactions between the core signaling components in soluble multi-protein complexes (Bhatnagar et al. 2010, Vu et al. 2012, Wang et al. 2012) and in reconstituted, attractant-regulated core complexes (Li and Hazelbauer 2011, Piasta et al. 2013, Natale et al. 2013, Piasta and Falke 2014, Li and Hazelbauer 2014).

Recently, a global view of the extended structural organization of chemosensory arrays has

emerged from cryo-electron tomography (cryoET) studies of native bacterial cells (Briegel et al. 2009, Briegel et al. 2012, Liu et al. 2012, Zhang et al. 2007). Specifically, chemoreceptor TODs were observed to form hexagonal arrays with a 12 nm lattice spacing conserved across several, distantly related bacterial species including *E. coli* and *T. maritima* (Briegel et al. 2009, Briegel, Wong, et al. 2014, Briegel et al. 2012, Liu et al. 2012, Zhang et al. 2007). The conservation of this hexagonal organization has also been demonstrated in non-membrane spanning cytoplasmic chemosensory arrays (Briegel, Wong, et al. 2014, Briegel, Ladinsky, et al. 2014). Additionally, studies using cryoET with sub-tomogram averaging, in tandem with crystallographic structures of portions of the core complex, have reported the extended structure of the array to consist of receptor TODs packed in a two-facing-two fashion about kinase-filled and kinase-empty rings (Briegel, Wong, et al. 2014, Briegel et al. 2012, Liu et al. 2012). However, due to the thickness of the cells as well as cellular crowding and heterogeneity, past cellular tomography studies have been limited to discerning only the overall arrangement of the core signaling components.

The lack of a high-resolution description of the intact and extended chemosensory array structure has hindered the development of a detailed understanding of molecular events occurring within the array during signaling. To address this problem, we have taken a joint experimental-computational approach. In particular, we have developed a novel reconstitution method yielding ultra-thin monolayer samples of core-signaling complex arrays, from which we derived a three-dimensional density map of the reconstituted core-signaling complex at 11.3 Å resolution using cryoET and sub-tomogram classification and averaging. Through the computational synthesis of existing X-ray crystallography data and our new cryoET data, we have constructed an atomic model of the extended chemosensory array. Our model highlights

novel interaction interfaces between the receptor, CheA, and CheW and permits the use of large-scale, all-atom molecular dynamics (MD) simulations (Perilla et al. 2015) to further illuminate the molecular details of a key kinase-signaling event.

Results:

Reconstitution of bacterial chemotaxis core-signaling complex arrays

To overcome the limitations imposed by cellular tomography of native chemosensory arrays (Briegel et al. 2009, Briegel et al. 2012, Liu et al. 2012, Zhang et al. 2007), we elected to establish an *in vitro* reconstituted system for high-resolution structural analysis of the signaling complex. Inspired by the template-directed method to assemble functional signaling complexes on lipid vesicles (Montefusco et al. 2007, Shrout, Montefusco, and Weis 2003), we designed a Ni^{2+} -NTA lipid containing monolayer system (Taylor et al. 2007, Taylor and Taylor 1999) to reconstitute the two-dimensional (2D) arrays of signaling complexes for structural analysis. To this end, we expressed and purified to high homogeneity *E. coli* chemotaxis proteins: CheA, CheW, and a His-tagged cytoplasmic signaling domain of the wild-type (wt) Tar receptor (TarCF). His-tagged TarCF can be readily incorporated into the Ni^{2+} -NTA lipid monolayers, seen as homogeneous particles in the EM micrographs of negatively stained specimen (Figure 1A). Only in the presence of all three components (TarCF, CheA, and CheW) were ordered arrays evident (Figure 1B), and even then, these microcrystalline 2D arrays were only formed under strictly constrained input ratios of the three components, a finding that is consistent with previous results indicating that the chemotactic function of the complex is diminished when one of the components is reduced or over-produced (Zhang et al. 2007). The optimal condition for array formation was established to be a mixture of TarCF, CheA and CheW with a molar ratio of

9:18:18 μM for TarCF:CheA:CheW in a lipid monolayer containing 2:1 DOPC:DOGS-NTA-
 Ni^{2+} lipids (33% Ni^{2+} -NTA lipid). Notably, the input molar ratio of the reconstitution mixture
does not reflect the actual ratio of components incorporated into the monolayer, as illustrated in
Figure 1C. The resulting arrays are organized in hexagonal lattices with 12 nm spacing (Figure
1B inset, white arrow), resembling the arrays formed in native cells (Briegel et al. 2012, Liu et
al. 2012).

CryoET of the chemotaxis core-signaling complex arrays

Compared to previous cellular tomography studies (Briegel, Wong, et al. 2014, Briegel et al.
2012, Liu et al. 2012, Zhang et al. 2007), the reconstituted monolayer system is ideal for high
resolution structural analysis of chemosensory arrays by cryoET for several reasons: 1) the *in*
vitro reconstituted monolayer array is thin (25 nm) and pseudo-crystalline, compared to cells
with thicknesses ranging from 500 nm to 1 μm ; 2) the monolayer arrays are reconstituted with
purified components, hence the system is well-defined, in contrast to native arrays in the
crowded cellular environment; 3) the reconstituted system allows for control over which array
components are present as well as manipulation of their signaling state; 4) the *in vitro* system
provides large numbers of sub-tomogram volumes (~ 3000 core-signaling units/tomogram),
thereby improving the noise statistics of the sub-tomogram averaging process central to
achieving a high resolution structure. Using cryoET, we collected and reconstructed, correcting
for the contrast transfer function (CTF) of the microscope (Fernandez, Li, and Crowther 2006),
20 tomograms of monolayers containing reconstituted core-signaling complex arrays. Figure 2A
(Video 1) shows a typical raw tomographic slice (without CTF correction) of a reconstituted
monolayer, illustrating patches of 2D lattices with information extending beyond 22 Å (inset,

arrow). By extracting and classifying CTF-corrected sub-tomograms, centered on each hexagon of receptor TODs (Figure 2-figure supplement 1, yellow circle), we obtained two major classes of the receptor hexagons: one containing a trimer of core-signaling units (CheA₂-trimer, Figure 2B and Figure 2-figure supplement 1, cyan boxes) and one containing a hexamer of core-signaling units (CheA₂-hexamer, Figure 2C and Figure 2-figure supplement 1, orange box). By mapping the individual sub-tomograms from the above two classes onto the original contributing tomograms, we were able to extract the extended lattice organization of the subunits in the monolayer (Figure 2D), revealing an interlocking of the CheA₂-trimer and CheA₂-hexamer classes (Figure 2E) consistent with that seen in cellular tomograms (Briegel, Wong, et al. 2014, Briegel et al. 2012, Liu et al. 2012).

To directly compare the lattice organization in our reconstituted monolayer system with that in native *E. coli* cells, we obtained three CTF-corrected tomograms from wt *E. coli* cells that were partially lysed, using a phage-gene-induced instant lysis method that we developed recently (Fu et al. 2014), to reduce the sample thickness. Extracting and classifying sub-tomograms containing receptor hexagons from the native *E. coli* cells revealed the same two classes that were observed in the monolayer system. As with the *in vitro* monolayer system, the two receptor hexagon classes observed in native *E. coli* cells also formed an interlocking lattice (Figure 2-figure supplement 2A). Extracting sub-tomograms with an extended unit that contained both the CheA₂-trimer and CheA₂-hexamer, we obtained an average density map for the *in situ* native chemosensory arrays that overlapped very well with the map from the monolayer system (Figure 2-figure supplement 2B). Therefore, the *in vitro* reconstituted monolayer system with purified *E. coli* proteins faithfully reproduces the lattice organization found in native *E. coli* cell membranes.

3D density maps of CheA₂-trimer and CheA₂-hexamer

The 3D classification process further improved the resolution of the class-averaged sub-tomograms of CheA₂-trimers (Figure 2B) to 11.3 Å and CheA₂-hexamers (Figure 2C) to 17.5 Å resolution, as measured by gold-standard Fourier shell correlation (FSC) (Figure 3-figure supplement 1A). A uniform distribution of in-plane orientations of the sub-tomograms and a relatively well sampled, out-of-plane angle enhanced the quality of the averaged density maps (Figure 3-figure supplement 1B&C). Nevertheless, some resolution anisotropy exists, with 11 Å in X and Y directions and 15.8 Å in Z direction (Figure 3-figure supplement 1A). To take the effect of the anisotropic resolution into account, we loss-pass filtered the density map according to the FSC of the Fourier conical shells along various directions (Diebolder et al. 2015). The resulting maps of the CheA₂-trimer and CheA₂-hexamer clearly delineate the density regions corresponding to the receptor, the CheA-P5/CheW ring at the receptor tip, and CheA kinase domain (Figure 3B&C, Figure 3-figure supplement 2)), and further display a number of new features. In particular, the individual receptor dimers are thoroughly resolved, allowing the kinase and CheW/receptor interactions to be isolated to a specific receptor dimer (Figure 3A-C). Moreover, the position of the previously unobserved four-helix bundle of the CheA-P3 dimerization domain is clearly discerned to run parallel to the receptor and is positioned close to CheW-interacting receptor dimers (Figure 3A-C). In addition, our maps dramatically refine the area of density projecting below the CheA-P5 domain, suggesting that the CheA-P4 kinase domain alone occupies this density region (Figure 2B, 3A, and Video 2). The CheA-P1 and CheA-P2 domains, on the other hand, are not resolved, likely due to their conformational flexibility.

189
190 Regarding the CheA-P5/CheW ring, our density map clearly shows a pseudo three-fold
191 symmetry (Figure 2B) in which the density at the CheA-P5/CheW interface between core-
192 signaling units (interface 2) is considerably weaker than the density at the CheA-P5/CheW
193 interface within core-signaling units (interface 1) (Figure 3A&C) (Briegel et al. 2012). This
194 finding is in contrast to the previously described pseudo six-fold symmetry of the CheA-
195 P5/CheW ring (Li et al. 2013). Most importantly, the previously described “empty hexagon” that
196 is surrounded by six CheA-occupied hexagons (Briegel et al. 2012) is not empty, but rather
197 contains a well-ordered continuous ring of densities (Figure 2C) that we were able to
198 unambiguously assign to individual CheW monomers (Figure 3D&E). This ring of CheW, as
199 previously speculated (Liu et al. 2012), provides additional interactions that couple neighboring
200 receptor TODs and strengthens the interlocking baseplate. Hence, our maps confirm the
201 existence of the CheW ring and establish its participation in the structural foundation responsible
202 for the ultra-stability of the chemosensory array (Liu et al. 2012, Erbse and Falke 2009) and for
203 the high cooperativity and extraordinary sensitivity measured in chemotaxis responses
204 (Goldman, Levin, and Bray 2009).

205 206 **All-atom model of the *T. maritima* chemosensory array.**

207 The resolution of our cryoET data permitted the unambiguous assignment of distinct regions
208 of density to specific protein components, enabling the construction of all-atom models of the
209 chemosensory array substructures and extended lattice (Figure 4-figure supplement 1). A
210 schematic overview of the modeling procedures carried out in this study is provided in Figure 4-
211 figure supplement 2 with a more detailed discussion of these procedures located in the Methods

212 section. Briefly, we first constructed models of the receptor TOD, CheA-P3P4 dimer, CheA-
213 P5/CheW ring, and CheW-only ring, taking advantage of existing high-resolution X-ray
214 structures from the thermophilic bacterium *Thermotoga maritima* (Kim, Yokota, and Kim 1999,
215 Bilwes et al. 1999, Park et al. 2006, Li et al. 2013). We then heuristically-arranged, using a 12
216 nm lattice constant, the resulting component models to produce models of the CheA₂-trimer and
217 CheA₂-hexamer subunits identified by sub-tomogram classification (Figure 4-figure supplement
218 1D&E). To further refine the key protein-protein interfaces within our atomic models, we
219 adopted a dual MD-based strategy, utilizing both unbiased MD and electron-density-biased
220 molecular dynamics flexible fitting (MDFF) simulations (Trabuco et al. 2008). For the subject of
221 our unbiased refinement simulations, we extracted from the CheA₂-hexamer model a portion
222 corresponding to the array unit cell, including six receptor TODs, three CheA dimers, and 12
223 CheW monomers all together arranged as three coupled core-signaling units (Figure 4-figure
224 supplement 1C, herein the “unit-cell model”). For our density-biased refinement simulations, we
225 focused our efforts on the CheA₂-trimer model, owing to the higher-resolution of its associated
226 density map, and hence, better resolved MDFF biasing forces. Because the CheA-P4 density is
227 not as well defined as the other parts of the complex, likely due to its conformational flexibility,
228 we carried out a rigid-body docking of the CheA-P4 domain, starting from 10,000 random
229 angular orientations and up to 20Å shifts from the center of the mass. This fitting exercise
230 resulted in 23 classes separated by 3° and 3Å (Figure 3-figure supplement 3A), generating a
231 metric for the goodness of fit of the P4 domain positioning. In addition to the class of “best fit”
232 (Figure 3-figure supplement 3B, panel 1), one other class, in which P4 is flipped relative to the
233 best fit, was seemingly structurally possible (Figure 3-figure supplement 3B, panel 5). However,
234 compared to the best-fit class, this alternative class had a lower cross-correlation value, lower

occupancy with only a third the number of contributing fits, and the positions of P4-N and C termini are reversed (flipped), making it hard to connect the P3 and P5 termini with short linkers. Thus, we have focused our efforts and resources on the highest ranking class of P4 position. It should be noted, though, that use of the alternative P4 positioning might produce considerably different MD trajectories. Solvation and ionization of the unit cell and CheA₂-trimer models produced systems of size 1.25 million and 1.75 million atoms, respectively, which were subsequently energy minimized and equilibrated for 10 ns, as described in the Methods section. The unit-cell model was then subjected to an 80 ns unconstrained production simulation (Video 3), while a 70 ns symmetry-constrained MDFF simulation was used to computationally bias the tertiary structure of the protein components within the CheA₂-trimer model according to our 11.3 Å CheA₂-trimer density map (Figure 3, Video 2).

The resulting unit-cell and CheA₂-trimer models agreed well with previous structural studies, in particular with respect to the residues participating in the CheA-P5/receptor and CheW/receptor interaction interfaces, as defined by NMR (Vu et al. 2012, Wang et al. 2012, Ortega, Mo, et al. 2013), crystallography (Li et al. 2013), and disulfide mapping studies (Piasta et al. 2013, Natale et al. 2013). For succinctness, specific residues participating in the various protein-protein interfaces within the array have been listed in Supplementary File 1. In addition, the equilibrated model of the *T. maritima* TOD maintained the conserved trimer-forming contacts observed in the X-ray structure of the *E. coli* serine receptor (Tsr) (Kim, Yokota, and Kim 1999) and revealed two additional trimer-stabilizing salt bridges, namely E387/R389 (conserved as E402/R404 in *E. coli* Tsr) and E351/R403 (structurally homologous to D363/R415 in *E. coli* Tsr) (Figure 4-figure supplement 3A). Moreover, in both models, the CheA-P4 kinase

domain was seen to stably occupy the region of density directly below the plane defined by the CheW and CheA-P5/CheW rings. Finally, in tandem with the direct visualization of the CheA-P3 dimerization domain in our cryoET density maps, the all-atom model further revealed previously uncharacterized specific interactions between the P3 bundle and adjacent receptors, involving D333/K390 and D345/R379 contact pairs (I304/N405 and D316/R394 in *E. coli* respectively) (Figure 4B, Supplementary File 1).

A conformational change of the CheA kinase domain

The construction of atomic models of the array unit cell and subunits permitted the use of equilibrium all-atom MD simulations to further investigate the molecular details of dynamic events potentially relevant to signaling. An overview of the key MD simulations conducted in this study is given in Figure 4-figure supplement 4A. The unit cell system contains the minimal set of components needed to represent the full receptor signaling array, which was made possible through the use of periodic boundary conditions to mimic the bulk symmetry of the chemosensory array, preventing the need to interpret potentially problematic effects due to unconstrained boundaries (Figure 4-figure supplement 2C, Video 3). We conducted a series of nine simulations of 450 ns each, using the equilibrated unit-cell model; additionally, we ran ten, 120 ns simulations of the equilibrated CheA₂-trimer system for comparison with the unit-cell simulations. Intriguingly, our simulations of both models revealed an ensemble of distinct core-signaling unit conformations (Figure 4 B&C), including structures in which the associated CheA dimer displayed either an undipped conformation (Figure 4B, top) or dipped conformation (Figure 4B, bottom). In the latter case, the P4 domain of one CheA monomer adopted a “dipped” state through rotations about the P3-P4 and P4-P5 flexible linkers, significantly affecting its

contacts with neighboring receptor dimers and the P5 domain (Video 4). As many biochemical, biophysical, and mutational studies have implicated dynamic structural changes within these regions of the core-signaling unit during the propagation of signals (Piasta et al. 2013, Natale et al. 2013, Wang et al. 2014, Briegel et al. 2013), we systematically identified the distinct structural classes of core-signaling unit conformations present in our MD simulations and isolated them for comparative analysis. Specifically, we used the UPGMC hierarchical clustering method (Müllner 2013) to assign the conformations of the 27 core-signaling units sampled in our unit cell simulations (3 units/unit cell) to groups of similar structure based on their pairwise root-mean-square deviation (RMSD). Cross-examination of structures within the resulting core-signaling unit clusters revealed the formation of two new salt bridges stabilizing the “dipped” state, namely R297/E397 (R265/E368 in *E. coli*) between the P3 and P4 domains and E390/R379 (E361/R394 in *E. coli*) between the P4 domain and nearby receptor tip (Figure 4B, bottom). Moreover, to accommodate the reorientation of the P4 domain, the P3 dimerization bundle was observed to break the receptor contacts (D333/K390 and D345/R379) observed in the “undipped” state (Figure 4B, Video 4), suggesting that the mobility of the P3 bundle plays a key role in the conformational dynamics of the CheA dimer.

We next sought to examine the temporal evolution of the dipping motion in each of the CheA dimers present in our simulations. For this purpose, we used Principal Component Analysis (PCA) to systematically derive, from the trajectory of a single dipping CheA dimer, a pseudo reaction coordinate by which to easily monitor the progression towards the “dipped” conformation. A total of four “dipping” events were observed in our unit cell simulations, as illustrated by projection of the conformations of the 27 CheA dimer time series onto the first

principal component (Figure 4C, top). Importantly, an additional two dipping events were observed in the 30 CheA dimers of the relatively shorter simulations of CheA₂-trimer model (Figure 4-figure supplement 5), demonstrating that the ability of the conformational change to occur is not an artifact of the particular choice of CheA P4 positioning during modeling. Interestingly, the three *extended* “dipping” events observed in the unit-cell simulations (Figure 4C; red, blue, and green traces) as well as the two events observed in the CheA₂-trimer simulations were accompanied by the formation of the R297/E397 contact. Notably, this contact was not formed in the one *short* dipping event, which returned to the “undipped” bulk state (Figure 4C; gold trace), suggesting that the R297/E397 contact may play a role in stabilizing the “dipped” state. To further investigate the significance of the R297/E397 contact for the conformational dynamics of CheA, we launched nine additional unit cell simulations with an R297A mutation to prevent the potential formation of the R297/E397 salt bridge. Indeed, while two CheA dimers exhibited the dipping motion in these simulations, including one dimer that underwent two dips, the mutants quickly return to the bulk (Figure 4C, bottom).

Biochemical validation of CheA conformational change in *E. coli* cells

To determine if the CheA-P4 dipping motion observed in the MD simulations of the *T. maritima* chemosensory array is sampled in the native chemotactic response of *E. coli*, we carried out cysteine disulfide cross-linking experiments. In particular, we tested the interaction interface for contacts existing in the undipped state (I304/N405 and D316/R394) or only in the dipped state (E361/R394) (Figure 5B). Notably, in the simulations, R394 of Tsr switches its contact with D316 of CheA-P3 to E361 of CheA-P4 during the transition of the CheA dimer from “undipped” to “dipped” (Video 4).

327
328 Using soft-agar assays, it was seen that the chemotactic ability of the I304C/N405C double
329 cysteine mutant is appreciably compromised compared to that of the control (cysless CheA/wt
330 Tsr), any of the single mutants (I304C/wt Tsr, cysless CheA/N405C, cysless CheA/N405S), and
331 when one half of the pair has been mutated to serine (I304C/N405S) (Figure 5A), suggesting that
332 dynamic interaction between CheA-P3 and the receptor is important for chemotactic function.
333 Moreover, *in vivo* cross-linking and western blot analysis showed a high molecular weight band
334 present only in the double cysteine mutant, suggesting the presence of species formed by cross-
335 linking between CheA-P3 and Tsr (Figure 5B). We also examined cross-linking residue pairs
336 that involve Tsr-R394 interactions with CheA, one in the “undipped” state (CheA-E361C/Tsr-
337 R394) and the other in the “dipped” state (CheA-E361C/Tsr-R394). When Tsr-R394 is replaced
338 by a cysteine or serine, either as a single mutant (cysless CheA/R394C, cysless CheA/R394S) or
339 in the context of a double mutant (D316C/R394C, D316C/R394S, E361C/R394C,
340 E361C/R394S), the chemotaxis function of *E. coli* is partially inhibited. On the other hand, the
341 chemotactic ability of CheA-E361C as a single mutant (E361C/wt Tsr) is also partially inhibited,
342 while CheA-D316C (E316C/wt Tsr) mutation bears no effect on the function (Figure 5A).
343 Furthermore, the cross-linking pattern of both Tsr-R394 mutant pairs showed two high molecular
344 weight bands corresponding to distinct cross-linked species, one with a disulfide formed between
345 Tsr and CheA (Figure 5B, upper band, blue arrows) and the other with a disulfide between two
346 Tsr molecules with the R394C mutation (lower band). Interestingly, the cross-linking of CheA-
347 P4/Tsr (E361C/R394C) in the predicted “dipped” conformation is much weaker than the cross-
348 linking of CheA-P3/Tsr (D316C/R394C) in the “undipped” state, though both involve the same
349 R394 residue of Tsr (Video 4). The lower cross-linking efficiency could be due to the relatively

infrequent occurrence of the CheA “dipped” conformation, and/or because the residues are further apart in a dominant conformation, suggesting that the CheA-P4 “dipped” conformation observed *in silico* may have been sampled within the native chemosensory complex of *E. coli*.

Our MD simulations of the *T. maritima* unit cell further indicated that R297 on the CheA-P3 domain is potentially involved in the stabilization of the conformational transition of the CheA-P4 (Figure 4C, Video 4). Indeed, substitution of the corresponding residue in *E. coli* (R265) with several amino acids of different properties (R265C/S/A/E) were all detrimental to the chemotactic function of *E. coli* as measured by the soft-agar assay, without affecting the cluster formation (Figure 5C, Figure 5-figure supplement 1). Since this residue is located at the N-terminus of the four-helix P3 dimerization motif, R265 could direct the P2-P3 linker away from the cis subunit and toward the trans subunit, thus anchoring CheA-P1P2 to CheA-P4’ for trans-interaction and phosphorylation (Bilwes et al. 1999). A more complete model of the core-signaling complex for *E. coli* may be necessary to fully interpret the drastic impact of this single CheA residue on the entire chemotactic machinery.

Discussion:

While much effort has been expended in the derivation of models to describe the transduction of ligand-binding events within the receptor proteins, including an established piston mechanism and a hypothesized alternating static–dynamic “yin-yang” on-off switch model (Falke and Piasta 2014), how the structure and dynamics of the CheA kinase are affected during signaling remains poorly understood. In this study, we identified, using MD simulations, a dipping motion of the CheA P4 domain, which was functionally characterized using swim

373 assay and cross-linking experiments. While the role of the predicted conformational change in
374 CheA is not immediately clarified in the preliminary biochemical experiments carried out here,
375 our model highlights the importance of CheA dynamics for signaling and suggests that the
376 dynamics of the P4 kinase domain, in particular, warrants special investigation. More
377 importantly, the atomic model presented here, in general, provides improved knowledge of the
378 positioning of the P3 and P4 domains, incorporates the presence of the CheW only ring, and
379 identifies probable novel side-chain contacts within the extended chemosensory architecture.
380 Further improving the resolution of our cryoET data to better than 8Å using the novel lipid-
381 monolayer system described above would allow generation of an atomic homology model of the
382 *E. coli* chemosensory array, greatly facilitating the use of the wealth of existing biochemical and
383 biophysical data and providing directly transferrable structural and dynamical predictions. We
384 hope that the findings presented here will inspire further experimental and computational studies
385 towards the elucidation of a complete mechanistic description of signal transduction and
386 amplification within this truly impressive biological sensory apparatus.

387

Materials and Methods:

Materials. Plasmids and cell strains used in this study were gifts from Dr. Parkinson, University of Utah, except for plasmid pHTCF (kind gift from Dr. Weis, University of Massachusetts, Amherst). Plasmid pHTCF is an IPTG-inducible expression vector for the N-terminal His₆-tagged cytoplasmic fragment of wt aspartate receptor (TarCF) that contains residues 257-553. Plasmids pKJ9 and PPA770 are IPTG-inducible for the expression of CheA and CheW, respectively. Plasmid pRR53 is an IPTG-inducible expression vector (amp^R) for the wt serine receptor (Tsr). Plasmid pGP26 is a sodium salicylate (Na-S)-inducible expression vector (cam^R) for cysteine-less CheA and wt CheW. Plasmids pRR53 and pGP26 were used to generate mutations in Tsr and CheA, respectively.

Protein expression and purification. *E. coli* strain RP3098, which lacks all Che proteins and chemoreceptors, was transformed with plasmid pKJ9 or PPA770 for CheA or CheW expression, respectively. CheA expression was induced at an OD₆₀₀ of 0.6-0.8, with 1 mM IPTG, overnight at 15°C. CheA was purified using an Affi-gel Blue column (Bio Rad, Hercules, CA) followed by gel filtration on a Superdex 200 column. Further purification with a Mono Q ion exchange column resulted in >99% homogeneity with an overall yield of 50 mg/L of cells. CheW expression was induced by the addition of IPTG (0.5 mM), at an OD₆₀₀ of 0.4-0.6, at 37°C. CheW was purified through 20%-40% ammonium sulfate precipitation, a DEAE column followed by a MonoQ ion exchange column and a Superdex 75 size exclusion column. This procedure resulted in highly purified CheW with a yield of 6 mg/L of cells. His₆-tagged wt TarCF (His₆-TarCF_{QEQE}) was expressed in DH5alpha cells with plasmid pHTCF. TarCF was induced by the addition of IPTG (0.5 mM) at an OD₆₀₀ of 0.4-0.6 at 37°C and purified with a

411 Ni²⁺-NTA affinity column followed with a mono Q column for quick removal of imidazole,
412 without dialyzing overnight. The yield for TarCF was excellent (120 mg/L of cells).
413

414 ***Monolayer reconstitution.*** A Ni²⁺ lipid containing monolayer system was used to reconstitute
415 the chemotaxis core-signaling complex arrays. A mixture of 9:18:18 μ M of TarCF:CheA:CheW
416 in a buffer containing 75 mM Tris-HCl, pH 7.4, 100 mM KCl, 5 mM MgCl₂ was applied to a
417 Teflon well, over which we immediately lay a lipid monolayer containing 2:1 DOPC:DOGS-
418 NTA-Ni²⁺ lipid mixture, at 2 mg/ml concentration. The monolayer set up was left undisturbed in
419 a humidity chamber overnight. The monolayer specimen was picked up with holey carbon grids,
420 stained with 1% uranyl acetate, and examined with an FEI T12 microscope operated at 120KV.
421

422 ***Cryo-electron tomography.*** Reconstituted monolayers using the best conditions identified by
423 negative staining (Figure 1B), were picked up with perforated R2/2 Quantifoil grids (Quantifoil
424 Micro Tools, Jena, Germany) pre-coated with 10 nm fiducial gold beads on the backside of the
425 grid and plunge-frozen using a manual gravity plunger. This method prevents disruption of the
426 monolayer by supporting single-side blotting which eliminates the contact between the blotting
427 filter paper and the delicate monolayer. The frozen-hydrated EM grids were loaded into FEI
428 Polara cartridges and imaged under low-dose conditions using a Tecnai Polara microscope (FEI
429 Corp., OR.) operating at 200kv. A series of low dose projection images were recorded with tilt
430 angles ranging from 70° to -70° with a Gatan 4K \times 4K CCD camera (Gatan, Inc., PA), at a
431 nominal magnification of 39,000 \times , with a defocus value of 5-8 μ m and an accumulated dose of
432 ~ 60 e⁻/Å². A total of 32 tomographic tilt series were collected using an FEI automated
433 tomography software.

3D reconstruction, sub-tomogram classification and averaging. Of the 32 tilt series collected, 20 tilt series with negligible mechanical or physical artifacts were selected for image processing and tomographic volume reconstruction. The monolayer produces an ideal EM specimen: it is thin (25 nm) and also provides strong signals in power spectra, due to near-crystalline packing of the protein components (Figure 2A inset), allowing accurate determination of the Contrast Transfer Function (CTF) using strip-based periodogram averaging in TomoCTF (Fernandez, Li, and Crowther 2006). The tilted projection series were roughly aligned using IMOD (Kremer, Mastronarde, and McIntosh 1996), and the alignment parameters were further refined using fiducial-free Area Matching with Geometry Refinement as implemented in Protomo (Winkler 2007). Using the refined geometry parameters, the raw projections were centered and rotated so the tilt azimuth was coincident with the Y-axis using the IMOD “newstack” function. These rotated stacks were corrected for the CTF with phase flipping, and volume reconstructions were made using SIRT as implemented in IMOD. These were calculated using a GPU, thereby removing an additional interpolation in the reconstruction step, by avoiding the use of cosine stretching of the input projections. Reconstructed volumes calculated from 20 SIRT iterations, providing higher contrast, were used for the initial cycles of sub-tomogram extraction and alignment, while those from 60 SIRT iterations were used for the final cycles.

To extract sub-tomograms, initial positions of the receptor complexes, respective to a Cartesian grid defined by each tomogram, were approximated by using a template matching algorithm implemented in Matlab with a reference that emphasized the receptor dimers with little influence from CheA. Both the template and tomograms were low-pass filtered to 4 nm and binned by 3. This resolution, as well as a coarse angular search, were chosen to eliminate any

statistical correlation of high resolution information between half data sets in later image processing steps. Following template matching sub-volume extraction, the data were randomly segregated into two groups, which were processed independently for all subsequent steps.

Sub-tomogram alignment and classification were carried out using Protomo's i3 image processing utilities (Winkler 2007). Using Multivariate Statistical Analysis and Hierarchical Ascendant Classification, eight class averages were produced from each half data set by focusing the analysis on the CheA portion of the complex. Initial references for each half set were generated by choosing averages from eight classes. These references were then used to align class averages chosen to each have ~50 contributing sub-volumes. In the following cycle, the raw sub-tomograms were subject to multi-reference alignment, but only a small in-plane and translational adjustment was allowed. This alignment by classification was repeated five times, while allowing the automatic exclusion of high variance outliers after the second cycle. In addition to the CheA₂-trimer and CheA₂-hexamer classes (Figure 2B. C), divergent organizations of CheA/receptor complex were also included as references (Figure 2-figure supplement 1).

After the final cycle, class averages containing either CheA₂-trimer or CheA₂-hexamer were manually selected and averaged together for each half data set, and the corresponding gold-standard FSC was calculated to evaluate the reliability of the data. Soft cylindrical masks were used, rather than spherical masks, given the extended slab like nature of the specimen. The final averages of CheA₂-trimer or CheA₂-hexamer from two half data sets of 3,000 sub-volumes or 300 sub-volumes, respectively, were combined and an empirical correction for the CTF envelope was applied for sharpening, which helped to further clarify the receptor dimers, as well as the P3 dimerization domain.

To access the degree of resolution anisotropy, conical Fourier shell correlations from the

two independent half data sets of CheA₂-trimer, along each of the principal axes, as well as the 10 axes bisecting them, were calculated (Diebolder et al. 2015). The averaged density map of CheA₂-trimer was then low-pass filtered according the conical FSCs along three principle axes by using cones with a 42° half-angle, adjusted for any overlapping regions in reciprocal space.

Computational modeling of T. maritima core components: Receptor trimer-of-dimers (TOD):

A model of the cytoplasmic portion of the *T. maritima* receptor dimer was taken from the X-ray crystal structure of the TM1143 chemoreceptor (PDB 2CH7) (Park et al. 2006). Using the *E. coli* receptor TOD (PDB 1QU7) (Kim, Yokota, and Kim 1999) as a reference, a *T. maritima* receptor TOD model (Figure 4-figure supplement 1A, Figure 4-figure supplement 2A) was obtained by arranging individual receptor dimer models from the previous step so that homologous trimer-forming contacts were preserved. *CheA-P34*: An atomic model of the soluble *T. maritima* CheA dimer, including the dimerization (P3) and kinase (P4) domains, was based on atomic coordinates from the X-ray crystal structure PDB 1B3Q (Bilwes et al. 1999). *CheA-P5/CheW and CheW rings*: Atomic models for both the CheA-P5/CheW and CheW rings were based on the X-ray crystal structure of the Receptor/CheA-P5/CheW ternary complex, PDB 4JPB (Li et al. 2013). In the case of the CheW ring model, the P5 domains of the CheA-P5/CheW ring model were exchanged with CheW monomers, using the dual-SH3-like fold shared between by CheA-P5 and CheW, to obtain an appropriate placement and orientation with respect to the neighboring monomers. Figure 4-figure supplement 3 schematically summarizes the modeling procedures described above. All missing loops were added using MODELLER (Sali and Blundell 1993). The TOD, CheA-P5/CheW, and CheW ring core component models were subjected to 150 ns of equilibration to ensure their structural integrity.

Construction of *T. maritima* array subunit models: The CheA₂-trimer and CheA₂-hexamer subunits models (Figure 4-figure supplement 1D&E) were constructed heuristically; using as a visual reference the extended organization of kinase-filled and kinase-empty rings evident in our density maps to arrange the components, also assuming an approximate 12 nm lattice constant (Figure 4-figure supplement 3B). Next, we made use of the CheA-P5/receptor interface from the ternary complex structure PDB 4JPB (Li et al. 2013) to model the CheW/receptor interface, assuming a receptor-binding mode homologous to that of CheA-P5. Using the CheA-P5 and CheW monomer/receptor models from the previous step, positional constraints on the receptor TODs were set relative to the height and orientation of the protein rings. Finally, CheA-P3,4 core component models were placed between adjacent TODs in accordance with the patterns observed in our density maps and joined to nearby ring-bound regulatory domains (P5) at the P4-P5 flexible linker. From the CheA₂-hexamer model we then extracted a portion corresponding to the array unit cell (Figure 4-figure supplement 2C, Figure 4-figure supplement 3C) for further study with all-atom MD simulations. In addition, symmetry-constrained molecular dynamics flexible fitting (MDFF) simulations (Trabuco et al. 2008) were used to refine the overlap between our experimental densities and heuristically constructed CheA₂-trimer and CheA₂-hexamer subunit models (Figure 4-figure supplement 3D, Video 2). The Situs modeling package (Wriggers 2010), was used to rigidly dock the subunit models into their respective cryoET maps to provide the initial overlap for our MDFF simulations.

Molecular dynamics simulations. The array unit cell model was hydrated with TIP3P water molecules using VMD's solvate plugin (Humphrey, Dalke, and Schulten 1996), producing a

simulation box defined by hexagonal lattice parameters $a=208\text{\AA}$, $b=208\text{\AA}$, $c=334\text{\AA}$, $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=120^\circ$. Using VMD's autoionize plugin, the hydrated system was then neutralized and subsequently ionized with sodium and chloride ions to the physiological concentration of 150 mM, resulting in a model containing 1,153,756 atoms. The unit cell model was then subjected to a series of conjugant gradient energy minimizations (300,000 steps in total) and restrained NPT equilibration simulations (10 ns in total). In the same fashion, the CheA₂-trimer and CheA₂-hexamer subunit models were hydrated and ionized to produce systems of size 1,751,375 atoms (245x245x310 Å) and 4,588,588 atoms (385x405x310 Å) respectively. Each subunit model was then subjected to the same minimization (300,000 steps) and restrained NPT equilibration (10 ns) scheme as the unit cell model. An outline of subsequent equilibration and production simulations is given in Figure 4-figure supplement 4A. Production simulations of the unit cell and MDFF-refined CheA₂-trimer models were conducted with weak (spring constant = 0.1 kcal/mol*nm²) harmonic restraints placed on the alpha carbons of the first five membrane-proximal receptor residues to maintain TOD splay in the absence of membrane and crowding agents. In the case of the post-MDFF production simulations of the CheA₂-trimer, additional weak harmonic constraints were placed on the outermost CheW and CheA-P5 domains to enforce the bulk array boundary conditions, as the trimer organization does not permit the use of periodic boundary conditions to represent the necessary symmetry.

All molecular dynamics simulations were performed using the parallel molecular dynamics code, NAMD 2.9 (Phillips et al. 2005) and CHARMM22 force field (MacKerell et al. 1998) with CMAP corrections (Mackerell, Feig, and Brooks 2004). Equilibrium simulations were conducted in the NPT ensemble with isobaric and isothermal conditions maintained at 1 atm and 330 K for equilibration, or 350 K for production using the Nosé-Hoover Langevin piston, with a period 200

femtoseconds (fs) and relaxation time of 50 fs, and the Langevin thermostat with a temperature coupling of 5 ps⁻¹. The r-RESPA integrator scheme with an integration time step of 2 fs was used (Phillips et al. 2005). SHAKE constraints were applied to all hydrogen atoms. Short-range, non-bonded interactions were calculated every 2 fs with a cutoff of 12 Å and long-range electrostatics were evaluated every 6 fs using the particle-mesh-Ewald (PME) method with a grid size of 1 Å. Periodic boundary conditions with fixed cross-sectional area (x-y plane) were used. MDFF simulations were performed in the NVT ensemble at 330 K using the settings described above with additional restraints applied to prevent loss of secondary structure, chirality errors, and the formation of cis-peptide bonds.

Simulation Analysis. Visualization and extraction of raw trajectory data for analysis were performed using VMD. Principal Components Analysis (PCA) was carried out using custom scripts (source code file *PCA.py*) involving the Numpy, Scipy, and MDAnalysis python packages (Michaud-Agrawal et al. 2011). For the PCA analysis, a single dip-exhibiting CheA dimer was isolated from one of our wild-type unit cell simulations, and each frame (23,331 frames in total) was aligned to the initial CheA dimer model using the P5 domains (residues 543-671). Principal components were computed using the alpha carbons of the P4 domains (residues 352 to 542). The fractional variances accounted for by the top three modes were 41.8%, 31.1%, and 8.1% respectively. Subsequently, the three CheA dimers from each replica of the wild-type unit cell model (27 dimers total), R297A unit cell model (27 dimers total), and CheA₂-trimer model (30 dimers total) simulations were extracted, aligned to the P5 domains, and projected on to the top principal component of the wild-type dip-exhibiting CheA dimer. These projections were grouped according to model type to create Figure 4C (top and bottom) and Figure 4-figure

supplement 5. Illustrations of the PCA results were produced using the python-plotting package, Matplotlib. Clustering analysis was performed using custom scripts (source code file *clustering.py*) involving the python packages noted above as well as the implementation of the UPGMC hierarchical, agglomerative clustering algorithm from the fastcluster package (Müllner 2013). For the clustering analysis, we first extracted the three core-signaling units from each of the nine wild-type unit cell replica simulations, using 1500 frames/core-signaling unit for a total of 40,500 frames. The RMSD distance matrix was then computed using the ‘rms_fit_trj’ function from the MDAnalysis package. Our analysis identified four major clusters of structures within the above distance matrix with relative populations of 80%, 10%, 10% and 2%, representing the undipped and dipped CheA dimer states as well as two intermediate states respectively.

Mutagenesis. Specific mutations on CheA and Tsr were generated by site-directed mutagenesis on the background of cysteine-less CheA (pGP26) (Miller et al. 2006, Zhao and Parkinson 2006) and wt Tsr (pRR53), respectively. Each mutation was introduced using a pair of primers (IDT), complementary to the template except for the site of mutation, and PfuUltra II Fusion HS DNA polymerase (Agilent), following the manufacturer’s thermocycling parameters. The presence of the mutations was confirmed by DNA sequencing.

Cross-linking and western blot analysis. Starter cultures were grown in LB broth (10% tryptone/5% yeast extract/10% NaCl), supplemented with appropriate antibiotics, overnight at 37°C with 250-rpm shaking. Subsequently, the overnight cultures were diluted 1:50 into a 5-ml LB broth, supplemented with the appropriate antibiotics and allowed to grow at 37°C with 250-rpm shaking. When the optical density at 600 nm of the cultures reached ~0.8, cells were

595 induced with 100 μ M IPTG and 0.6 μ M Na-S, in the presence of 100 μ M serine, for 1 hr at 37°C.
596 After induction, cells were collected by centrifugation (3000 x g, 4°C for 10 mins) and then re-
597 suspended in cold PBS, in the presence of 100 μ M serine. Cross-linking was initiated by addition
598 of 60 μ M copper (II) sulfate and 200 μ M phenanthroline (1 hr, RT) and stopped by addition of
599 20 mM iodoacetamide and 3.7 μ M neocuproin. Cells were immediately mixed with 4 \times NuPAGE
600 lithium dodecyl sulfate/PAGE sample buffer (Invitrogen), with or without reducing agent (DTT),
601 and then boiled for 5 min before electrophoresis. Samples were analyzed on 4–12% SDS-PAGE
602 gels in MES running buffer (Invitrogen). Gels were transferred to nitrocellulose membranes,
603 blocked, and immunoblotted by using antiserum against Tsr (1:2500) and CheA (1:1250) (gifts
604 from Dr. Subramaniam, NIH), followed by an alkaline phosphatase conjugated anti-rabbit
605 antibody (1:50,000, Sigma). Bands were detected on the membrane using an NBT/BCIP kit from
606 Promega following the manufacturer's instructions.

607

608 ***Soft agar Assays.*** The UU2682 strain does not express any of the chemoreceptors, CheA or
609 CheW, rendering it non-chemotactic despite the presence of an intact flagellar system. Presence
610 of both pRR53 (wt-Tsr) and pGP26 (cysless CheA and wt-cheW) is required to rescue the
611 chemotaxis of UU2682, observed as formation of attractant rings on a soft-agar media. To assess
612 the effect of mutations on Tsr and/or CheA on the chemotactic ability of *E. coli*, the mutant
613 plasmids were introduced into the UU2682 strain and assayed for formation of attractant rings.
614 The soft agar assay protocol used here is adapted from the Parkinson laboratory (University of
615 Utah). Fresh colonies were plated on LB-agar media (10% tryptone/5% yeast extract/10%
616 NaCl/10% agar), supplemented with carbenicillin (100 μ g/ml) and chloramphenicol (34 μ g/ml),
617 and grown overnight at 37°C. Next day, using a fine-tip toothpick, colonies were picked from the

618 fresh LB-agar plate and stabbed into a soft-agar media (10% Tryptone/5% yeast
619 extract/5%NaCl/0.27% agar) containing antibiotics (carbenicillin 50 µg/ml, chloramphenicol 17
620 µg/ml), inducers (100 µM IPTG and 0.6 µM Na-S) and 100 µM serine. Plates were then
621 incubated at 32°C for ~8 hrs and the diameter of attractant rings immediately measured after
622 incubation.

Acknowledgments:

We thank Drs. J.S. Parkinson for bacterial strains and plasmids and S. Subramaniam for antisera against chemotaxis components, D. Bevan and P. Greer for computer technical support and A. Makhov for cryoEM instrumental support. In addition, we thank Drs. Wei Han and Davi Ortega for insightful discussions regarding the computational aspects of this study. We also thank Dr. T. Brosenitsch for reading the manuscript. This work was supported by the National Institutes of Health NIGMS Grant R01GM085043 (P.Z.), P50GM082251-7518 (P.Z.), 9P41GM104601 (K.S.), and 5R01GM098243 (K.S.) as well as the National Science Foundation PHY-1430124 (K.S.). Large-scale molecular dynamics simulations were performed on the Blue Waters supercomputer, financed by the National Science Foundation (awards OCI-0725070 and ACI-1238993).

Author contributions

This study was conceived by P.Z and K.S. *In vitro* reconstitution and cryoEM experiments were designed by P.Z.; J.M., G.Z. and B.H. performed experiments and data analysis; J.M. purified proteins and performed reconstitution; J.M., G.Z. and P.Z. collected cryoEM tomography data, B.H. processed, analyzed and interpreted the tomography data. K.S. developed computational methodology. Computer simulations were designed by C.K.C., J.R.P. and K.S.; C.K.C. constructed structural models and performed molecular dynamics simulations. C.K.C., J.R.P., K.S., B.H. and P.Z. analyzed atomic models and interpreted simulation results. The biochemical and functional assays were designed by P.Z. and F.J.A.; F.J.A. performed the experiments and P.Z, F.J.A., and C.KC. analyzed the results. C.K.C., B.H., F.J.A., K.S. and P.Z. wrote the paper with support from all the authors.

646

647 **Additional information**

648 Accession codes:

649 Cryo-EM structural data have been deposited in the EM Data Bank under accession codes EMD-
650 3234 for a monolayer tomogram, EMD-6319 for the CheA₂-trimer and EMD-6320 for the
651 CheA₂-hexamer. The atomic coordinates of the core-signaling complex model have been
652 deposited in the RCSB Protein Data Bank under accession code 3JA6.

653

654 **Competing Financial Interests:**

655 The authors declare no competing financial interests.

References:

- Amin, D. N., and G. L. Hazelbauer. 2010. "The chemoreceptor dimer is the unit of conformational coupling and transmembrane signaling." *J Bacteriol* no. 192 (5):1193-200. doi: 10.1128/JB.01391-09.
- Bhatnagar, J., P. P. Borbat, A. M. Pollard, A. M. Bilwes, J. H. Freed, and B. R. Crane. 2010. "Structure of the ternary complex formed by a chemotaxis receptor signaling domain, the CheA histidine kinase, and the coupling protein CheW as determined by pulsed dipolar ESR spectroscopy." *Biochemistry* no. 49 (18):3824-41. doi: 10.1021/bi100055m.
- Bilwes, A. M., L. A. Alex, B. R. Crane, and M. I. Simon. 1999. "Structure of CheA, a signal-transducing histidine kinase." *Cell* no. 96 (1):131-41.
- Bray, D., M. D. Levin, and C. J. Morton-Firth. 1998. "Receptor clustering as a cellular mechanism to control sensitivity." *Nature* no. 393 (6680):85-8. doi: 10.1038/30018.
- Briegel, A., P. Ames, J. C. Gumbart, C. M. Oikonomou, J. S. Parkinson, and G. J. Jensen. 2013. "The mobility of two kinase domains in the Escherichia coli chemoreceptor array varies with signalling state." *Mol Microbiol* no. 89 (5):831-41. doi: 10.1111/mmi.12309.
- Briegel, A., M. S. Ladinsky, C. Oikonomou, C. W. Jones, M. J. Harris, D. J. Fowler, Y. W. Chang, L. K. Thompson, J. P. Armitage, and G. J. Jensen. 2014. "Structure of bacterial cytoplasmic chemoreceptor arrays and implications for chemotactic signaling." *Elife* no. 3:e02151. doi: 10.7554/eLife.02151.
- Briegel, A., X. Li, A. M. Bilwes, K. T. Hughes, G. J. Jensen, and B. R. Crane. 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* no. 109 (10):3766-71. doi: 10.1073/pnas.1115719109.
- Briegel, A., D. R. Ortega, E. I. Tocheva, K. Wuichet, Z. Li, S. Chen, A. Muller, C. V. Iancu, G. E. Murphy, M. J. Dobro, I. B. Zhulin, and G. J. Jensen. 2009. "Universal architecture of bacterial chemoreceptor arrays." *Proc Natl Acad Sci U S A* no. 106 (40):17181-6. doi: 10.1073/pnas.0905181106.
- Briegel, A., M. L. Wong, H. L. Hodges, C. M. Oikonomou, K. N. Piasta, M. J. Harris, D. J. Fowler, L. K. Thompson, J. J. Falke, L. L. Kiessling, and G. J. Jensen. 2014. "New insights into bacterial chemoreceptor array structure and assembly from electron cryotomography." *Biochemistry* no. 53 (10):1575-85. doi: 10.1021/bi5000614.
- Capra, E. J., and M. T. Laub. 2012. "Evolution of two-component signal transduction systems." *Annu Rev Microbiol* no. 66:325-47. doi: 10.1146/annurev-micro-092611-150039.
- Diebolder, C. A., F. G. Faas, A. J. Koster, and R. I. Koning. 2015. "Conical Fourier shell correlation applied to electron tomograms." *J Struct Biol* no. 190 (2):215-23. doi: 10.1016/j.jsb.2015.03.010.
- Eisenbach, Michael, and Joseph W. Lengeler. 2004. *Chemotaxis*. London, : Imperial College Press.
- Erbse, A. H., and J. J. Falke. 2009. "The core signaling proteins of bacterial chemotaxis assemble to form an ultrastable complex." *Biochemistry* no. 48 (29):6975-87. doi: 10.1021/bi900641c.
- Falke, J. J., and K. N. Piasta. 2014. "Architecture and signal transduction mechanism of the bacterial chemosensory array: progress, controversies, and challenges." *Curr Opin Struct Biol* no. 29:85-94. doi: 10.1016/j.sbi.2014.10.001.
- Fernandez, J. J., S. Li, and R. A. Crowther. 2006. "CTF determination and correction in electron

cryotomography." *Ultramicroscopy* no. 106 (7):587-96. doi: 10.1016/j.ultramic.2006.02.004.

Fu, X., B. A. Himes, D. Ke, W. J. Rice, J. Ning, and P. Zhang. 2014. "Controlled bacterial lysis for electron tomography of native cell membranes." *Structure* no. 22 (12):1875-82. doi: 10.1016/j.str.2014.09.017.

Goldman, J. P., M. D. Levin, and D. Bray. 2009. "Signal amplification in a lattice of coupled protein kinases." *Mol Biosyst* no. 5 (12):1853-9. doi: 10.1039/B903397a.

Goy, M. F., M. S. Springer, and J. Adler. 1977. "Sensory transduction in *Escherichia coli*: role of a protein methylation reaction in sensory adaptation." *Proc Natl Acad Sci U S A* no. 74 (11):4964-8.

Griswold, I. J., H. Zhou, M. Matison, R. V. Swanson, L. P. McIntosh, M. I. Simon, and F. W. Dahlquist. 2002. "The solution structure and interactions of CheW from *Thermotoga maritima*." *Nat Struct Biol* no. 9 (2):121-5. doi: 10.1038/nsb753.

Hazelbauer, G. L., J. J. Falke, and J. S. Parkinson. 2008. "Bacterial chemoreceptors: high-performance signaling in networked arrays." *Trends Biochem Sci* no. 33 (1):9-19. doi: 10.1016/j.tibs.2007.09.014.

Hazelbauer, G. L., and W. C. Lai. 2010. "Bacterial chemoreceptors: providing enhanced features to two-component signaling." *Curr Opin Microbiol* no. 13 (2):124-32. doi: 10.1016/j.mib.2009.12.014.

Humphrey, W., A. Dalke, and K. Schulten. 1996. "VMD: visual molecular dynamics." *J Mol Graph* no. 14 (1):33-8, 27-8.

Kim, K. K., H. Yokota, and S. H. Kim. 1999. "Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor." *Nature* no. 400 (6746):787-92. doi: 10.1038/23512.

Kremer, J. R., D. N. Mastronarde, and J. R. McIntosh. 1996. "Computer visualization of three-dimensional image data using IMOD." *J Struct Biol* no. 116 (1):71-6. doi: 10.1006/jsbi.1996.0013.

Li, M., and G. L. Hazelbauer. 2011. "Core unit of chemotaxis signaling complexes." *Proc Natl Acad Sci U S A* no. 108 (23):9390-5. doi: 10.1073/pnas.1104824108.

Li, M., and G. L. Hazelbauer. 2014. "Selective allosteric coupling in core chemotaxis signaling complexes." *Proc Natl Acad Sci U S A* no. 111 (45):15940-5. doi: 10.1073/pnas.1415184111.

Li, M., C. M. Khursigara, S. Subramaniam, and G. L. Hazelbauer. 2011. "Chemotaxis kinase CheA is activated by three neighbouring chemoreceptor dimers as effectively as by receptor clusters." *Mol Microbiol* no. 79 (3):677-85. doi: 10.1111/j.1365-2958.2010.07478.x.

Li, X., A. D. Fleetwood, C. Bayas, A. M. Bilwes, D. R. Ortega, J. J. Falke, I. B. Zhulin, and B. R. Crane. 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* no. 52 (22):3852-65. doi: 10.1021/bi400383e.

Li, Y., Y. Hu, W. Fu, B. Xia, and C. Jin. 2007. "Solution structure of the bacterial chemotaxis adaptor protein CheW from *Escherichia coli*." *Biochem Biophys Res Commun* no. 360 (4):863-7. doi: 10.1016/j.bbrc.2007.06.146.

Liu, J. D., and J. S. Parkinson. 1989. "Role of CheW protein in coupling membrane receptors to the intracellular signaling system of bacterial chemotaxis." *Proc Natl Acad Sci U S A* no. 86 (22):8703-7.

- Liu, J., B. Hu, D. R. Morado, S. Jani, M. D. Manson, and W. Margolin. 2012. "Molecular architecture of chemoreceptor arrays revealed by cryoelectron tomography of *Escherichia coli* minicells." *Proc Natl Acad Sci U S A* no. 109 (23):E1481-8. doi: 10.1073/pnas.1200781109.
- MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. "All-atom empirical potential for molecular modeling and dynamics studies of proteins." *J Phys Chem B* no. 102 (18):3586-616. doi: 10.1021/jp973084f.
- Mackerell, A. D., Jr., M. Feig, and C. L. Brooks, 3rd. 2004. "Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations." *J Comput Chem* no. 25 (11):1400-15. doi: 10.1002/jcc.20065.
- Michaud-Agrawal, N., E. J. Denning, T. B. Woolf, and O. Beckstein. 2011. "MDAnalysis: a toolkit for the analysis of molecular dynamics simulations." *J Comput Chem* no. 32 (10):2319-27. doi: 10.1002/jcc.21787.
- Miller, A. S., S. C. Kohout, K. A. Gilman, and J. J. Falke. 2006. "CheA Kinase of bacterial chemotaxis: chemical mapping of four essential docking sites." *Biochemistry* no. 45 (29):8699-711. doi: 10.1021/bi060580y.
- Montefusco, D. J., A. L. Shrout, T. Y. Besschetnova, and R. M. Weis. 2007. "Formation and activity of template-assembled receptor signaling complexes." *Langmuir* no. 23 (6):3280-9. doi: 10.1021/la062717r.
- Müllner, Daniel. 2013. "fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python." *Journal of Statistical Software* no. 53 (9):1-18.
- Natale, A. M., J. L. Duplantis, K. N. Piasta, and J. J. Falke. 2013. "Structure, function, and on-off switching of a core unit contact between CheA kinase and CheW adaptor protein in the bacterial chemosensory array: A disulfide mapping and mutagenesis study." *Biochemistry* no. 52 (44):7753-65. doi: 10.1021/bi401159k.
- Ortega, D. R., G. Mo, K. Lee, H. Zhou, J. Baudry, F. W. Dahlquist, and I. B. Zhulin. 2013. "Conformational coupling between receptor and kinase binding sites through a conserved salt bridge in a signaling complex scaffold protein." *PLoS Comput Biol* no. 9 (11):e1003337. doi: 10.1371/journal.pcbi.1003337.
- Ortega, D. R., C. Yang, P. Ames, J. Baudry, J. S. Parkinson, and I. B. Zhulin. 2013. "A phenylalanine rotameric switch for signal-state control in bacterial chemoreceptors." *Nat Commun* no. 4:2881. doi: 10.1038/ncomms3881.
- Park, S. Y., P. P. Borbat, G. Gonzalez-Bonet, J. Bhatnagar, A. M. Pollard, J. H. Freed, A. M. Bilwes, and B. R. Crane. 2006. "Reconstruction of the chemotaxis receptor-kinase assembly." *Nat Struct Mol Biol* no. 13 (5):400-7. doi: 10.1038/nsmb1085.
- Parkinson, J. S., G. L. Hazelbauer, and J. J. Falke. 2015. "Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update." *Trends Microbiol.* doi: 10.1016/j.tim.2015.03.003.
- Perilla, J. R., B. C. Goh, C. K. Cassidy, B. Liu, R. C. Bernardi, T. Rudack, H. Yu, Z. Wu, and K. Schulten. 2015. "Molecular dynamics simulations of large macromolecular complexes." *Curr Opin Struct Biol* no. 31:64-74. doi: 10.1016/j.sbi.2015.03.007.

- Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. "Scalable molecular dynamics with NAMD." *J Comput Chem* no. 26 (16):1781-802. doi: 10.1002/jcc.20289.
- Piasta, K. N., and J. J. Falke. 2014. "Increasing and decreasing the ultrastability of bacterial chemotaxis core signaling complexes by modifying protein-protein contacts." *Biochemistry* no. 53 (35):5592-600. doi: 10.1021/bi500849p.
- Piasta, K. N., C. J. Ulliman, P. F. Slivka, B. R. Crane, and J. J. Falke. 2013. "Defining a key receptor-CheA kinase contact and elucidating its function in the membrane-bound bacterial chemosensory array: a disulfide mapping and TAM-IDS Study." *Biochemistry* no. 52 (22):3866-80. doi: 10.1021/bi400385c.
- Sali, A., and T. L. Blundell. 1993. "Comparative protein modelling by satisfaction of spatial restraints." *J Mol Biol* no. 234 (3):779-815. doi: 10.1006/jmbi.1993.1626.
- Shrout, A. L., D. J. Montefusco, and R. M. Weis. 2003. "Template-directed assembly of receptor signaling complexes." *Biochemistry* no. 42 (46):13379-85. doi: 10.1021/bi0352769.
- Sourjik, V., and J. P. Armitage. 2010. "Spatial organization in bacterial chemotaxis." *EMBO J* no. 29 (16):2724-33. doi: 10.1038/emboj.2010.178.
- Szurmant, H., and G. W. Ordal. 2004. "Diversity in chemotaxis mechanisms among the bacteria and archaea." *Microbiol Mol Biol Rev* no. 68 (2):301-19. doi: 10.1128/MMBR.68.2.301-319.2004.
- Taylor, D. W., D. F. Kelly, A. Cheng, and K. A. Taylor. 2007. "On the freezing and identification of lipid monolayer 2-D arrays for cryoelectron microscopy." *J Struct Biol* no. 160 (3):305-12. doi: 10.1016/j.jsb.2007.04.011.
- Taylor, K. A., and D. W. Taylor. 1999. "Structural studies of cytoskeletal protein arrays formed on lipid monolayers." *J Struct Biol* no. 128 (1):75-81. doi: 10.1006/jsbi.1999.4167.
- Trabuco, L. G., E. Villa, K. Mitra, J. Frank, and K. Schulten. 2008. "Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics." *Structure* no. 16 (5):673-83. doi: 10.1016/j.str.2008.03.005.
- Tu, Y. 2013. "Quantitative modeling of bacterial chemotaxis: signal amplification and accurate adaptation." *Annu Rev Biophys* no. 42:337-59. doi: 10.1146/annurev-biophys-083012-130358.
- Vu, A., X. Wang, H. Zhou, and F. W. Dahlquist. 2012. "The receptor-CheW binding interface in bacterial chemotaxis." *J Mol Biol* no. 415 (4):759-67. doi: 10.1016/j.jmb.2011.11.043.
- Wadhams, G. H., and J. P. Armitage. 2004. "Making sense of it all: bacterial chemotaxis." *Nat Rev Mol Cell Biol* no. 5 (12):1024-37. doi: 10.1038/nrm1524.
- Wang, X., P. Vallurupalli, A. Vu, K. Lee, S. Sun, W. J. Bai, C. Wu, H. Zhou, J. E. Shea, L. E. Kay, and F. W. Dahlquist. 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* no. 53 (5):855-61. doi: 10.1021/bi4012379.
- Wang, X., A. Vu, K. Lee, and F. W. Dahlquist. 2012. "CheA-receptor interaction sites in bacterial chemotaxis." *J Mol Biol* no. 422 (2):282-90. doi: 10.1016/j.jmb.2012.05.023.
- Winkler, H. 2007. "3D reconstruction and processing of volumetric data in cryo-electron tomography." *J Struct Biol* no. 157 (1):126-37. doi: 10.1016/j.jsb.2006.07.014.
- Wriggers, W. 2010. "Using Situs for the integration of multi-resolution structures." *Biophys Rev* no. 2 (1):21-27. doi: 10.1007/s12551-009-0026-3.
- Zhang, P., C. M. Khursigara, L. M. Hartnell, and S. Subramaniam. 2007. "Direct visualization of

Escherichia coli chemotaxis receptor arrays using cryo-electron microscopy." *Proc Natl Acad Sci U S A* no. 104 (10):3777-81. doi: 10.1073/pnas.0610106104.
Zhao, J., and J. S. Parkinson. 2006. "Cysteine-scanning analysis of the chemoreceptor-coupling domain of the Escherichia coli chemotaxis signaling kinase CheA." *J Bacteriol* no. 188 (12):4321-30. doi: 10.1128/JB.00274-06.

Figure legends:

Figure 1. Reconstitution of 2D arrays of the receptor signaling complex on lipid monolayers.

(A&B) Negatively stained electron micrographs of reconstituted lipid monolayers with TarCF only (A) or with TarCF/CheA/CheW (B). Inset, Fourier transform of a region from the monolayer array, indicating a hexagonal lattice with a 12 nm repeat (white arrow). Scale bars, 100nm. (C) SDS-PAGE gel analysis of the reconstituted monolayer sample (lane 2) and the protein solution (input mixture) used to generate the monolayer arrays (lane 3). Molecular weight markers are indicated (lane 1), and the input mixture contained TarCF:CheA:CheW in a ratio of 9:18:18 μ M.

Figure 2. CryoET of monolayer arrays of TarCF/CheA/CheW ternary signaling complex. (A) A tomographic slice (1.2 nm thick) through the reconstituted monolayer arrays of TarCF/CheA/CheW, without CTF correction. Inset, The Fourier transform of a selected region, displaying Thon rings with information extended to at least 22 Å resolution (arrow). (B&C) Averaged density maps of two sub-volume classes containing receptor hexagons (6 TODs) (red), one with a trimer of CheA dimers (CheA₂-trimer) (B) and the other with a hexamer of CheA dimers (CheA₂-hexamer) (C). Maps were generated following sub-tomogram volume classification and class-averaging, are contoured at 1.5 σ , and are colored according to the height, from the receptor at the top (red) to CheA (blue) below. (D) Spatial arrangement of the CheA₂-

trimer (cyan) and CheA₂-hexamer (orange) in the monolayer lattice array, after mapping the classified sub-volumes back onto the tomogram. The array is formed by interlocking CheA₂-trimer and CheA₂-hexamer subunits. (E) A schematic lattice model for the chemosensory arrays. Small circles represent receptor dimers; arrows represent CheA dimers (CheA₂). Dashed cyan and orange circles highlight a CheA₂-trimer and CheA₂-hexamer respectively. The lattice unit cell is outlined in black. Related to figure 2-figure supplement 1 and 2.

Figure 3. CheA₂-trimer and CheA₂-hexamer density maps with molecular dynamics flexible fitting (MDFF) of computationally constructed *T. maritima* subunit models. (A) Overall fitting of the CheA₂-trimer density map contoured at 1.5 σ . The three core signaling complexes are colored in pink, blue and green. (B) A sectional view of the boxed region in A, rotated 90°. The protein components are labeled at the indicated height of the complex (gray boxes). (C) Sectional views of the gray-boxed regions in B at the receptor level (top), the CheA-P3 and P5/CheW ring region (middle), and CheA-P4 region (bottom). (D) Overall fitting of the CheA₂-hexamer density map contoured at 1.5 σ . (E) A sectional view at the CheA-P3 and CheW-ring region of the CheA₂-hexamer density map. In (A-E), CheA-P3, P4, P5, CheW and receptor are labeled as P3, P4, P5, W and R, respectively, and the CheA-P5/CheW interfaces 1 and 2 are indicated. Related to figure 3-figure supplement 1.

Figure 4. CheA dimer conformational switch. (A) Top and side views of the core-signaling unit, consisting of two receptor TODs (red), a single CheA dimer (blue), and four CheW monomers (green). (B) Two distinct classes, undipped (top) and dipped (bottom), of core-signaling unit structures are present in our MD simulations. Classes differ especially in the orientation of

890 CheA-P4 domain with respect to the rest of the CheA dimer and core signaling unit. Specific
891 contacts that stabilize either conformation are indicated for *T. maritima* and in parentheses for
892 the corresponding residues in *E. coli*. CheA-P5 and CheW have been removed for clarity. (C)
893 Time series of CheA dimer conformations extracted from unit cell simulations.. Traces track the
894 projection of the conformations of 27 CheA dimers from the wt (top) and R297A mutant
895 (bottom) unit cell simulations onto the first principal component of the “dipping” motion.
896 Colored traces track CheA dimers that undergo an extended (>10 ns) “dipping” motion.
897 Horizontal dashed lines visually demarcate the undipped and dipped CheA dimer classes.
898 Vertical dashed lines separate the initial 80 ns equilibration simulation from nine 450 ns
899 production simulations. Related to figure 4-figure supplement 1, 2, 3, 4 and 5, and
900 Supplementary File 1.

901
902 **Figure 5.** Biochemical validation of alternative CheA conformations in *E. coli*. (A) Swimming
903 ability of *E. coli* cells with mutations in the CheA-P3 and Tsr interface (I304/N405 and
904 D316/R394) and in the “dipped” CheA-P4 and Tsr interface (E361/R394). Swimming activities
905 are normalized to the cysless CheA and wt Tsr, \pm standard deviation ($n=6$). Inset, representative
906 images of soft agar plates for swimming ability, with specific constructs labeled in red. (B)
907 Disulphide cross-linking of the CheA-P3 and Tsr interface (I304C/N405C and D316C/R394C) in
908 the undipped CheA dimer conformation (top panel of Figure 5B) and the CheA-P4 and Tsr
909 interface (E361C/R394C) occurring in the dipped CheA-P4 “dipped” conformation (bottom
910 panel of Figure 5B). Non-reducing (top) and reducing (bottom) SDS-PAGE gels were analyzed
911 by immunoblotting for Tsr and CheA. Cross-linked species were indicated with blue arrows. (C)
912 Swimming ability of *E. coli* cells with mutations at R265 of CheA-P3 domain, normalized to the

wt, \pm standard deviation ($n=8$). Related to figure 5-figure supplement 1.

Supplement figure legends:

Figure 2-figure supplement 1. Classification of the sub-tomogram volumes. The sub-tomograms containing the receptor hexagon (6 TODs, yellow circles) were subjected to alignment and classification. Sections of the sub-tomogram classes are shown at the receptor region (top row in each set) and the CheA region (bottom row in each set). Two major configurations emerged, one with a trimer of CheA dimers (CheA₂-trimer, cyan boxes) and the other with a hexamer of CheA dimers (CheA₂-hexamer, orange box). All the CheA₂-trimer classes were combined for final refinement of the density map.

Figure 2-figure supplement 2. Comparison of the chemotaxis arrays from native cells and from *in vitro* reconstituted monolayers. (A) A tomographic slice (5.2 nm thick) of the native chemotaxis arrays from wild-type *E. coli* cells, overlaid with the positions of each sub-tomogram classified as either the CheA₂-trimer unit (cyan) or the CheA₂-hexamer unit (orange). The lattice organization with interlocking CheA₂-trimer and CheA₂-hexamer is the same as the reconstituted arrays (see Figure 2A). (B) Sub-tomogram averaged density maps of the extended lattice (CheA₂-trimer surrounded by six interlocking CheA₂-hexamers) from native clusters (light blue surface rendering, contoured at 2σ), overlaid with the extended lattice from reconstituted monolayer arrays (mesh surface, contoured at 1.5σ). The two lattices match very well. (C) Surface rendering of the sub-tomogram averaged extended lattice map from reconstituted monolayer arrays at 18 Å resolution, showing the interlocking CheA₂-trimer (cyan dashed circles) and CheA₂-hexamer (orange dashed circles). The map is contoured at 1.5σ and colored

according to the height, from receptor (red) to CheA (blue).

Figure 3-figure supplement 1. Gold-standard Fourier shell correlation (FSC) of the CheA₂-trimer (left) and CheA₂-hexamer (right) density maps. At FSC=0.143, the overall resolution of the CheA₂-trimer map is 11.3 Å and that of the CheA₂-hexamer is 17.5 Å. The FSC curves for the conical Fourier shells along the X, Y, and Z directions are in solid green, blue and dark-red, respectively, and along the 10 other directions are in dotted lines. The CheA₂-trimer map was calculated from two independent data sets of 3000 sub-volumes each, and the CheA₂-hexamer was from two independent data sets of 300 sub-volumes each. (B) Angular distributions of contributing sub-tomograms for the in-plane rotation angle (left) and the maximum tilt angle relative to electron beam direction (right).

Figure 3-figure supplement 2. X-Z sectional views of the CheA₂-trimer density map with MDFF model. The positions of the sections are indicated in the last panel with an orthogonal view (X-Y plane).

Figure 3-figure supplement 3. A metric for the goodness of fit for the docking of the CheA-P4 domain. (A) Distribution of 23 classes of fits for the P4 domain starting from random orientations. (B) The models from the top 9 highest cross-correlation classes are shown in panels 1-9, with the cross-correlation values and number of contributing fits (%) indicated below. These 9 classes constitute 90.1% of total fits. Panel 10 is an overlay of the #1 fit (blue) with the MDFF model (orange). The red and blue spheres indicate the N and C termini of P4, respectively. The corresponding connecting termini from P3 and P5 are in light blue and pink,

respectively. P3 and P5 domains are in gray.

Figure 4-figure supplement 1. Nomenclatures. (A) TOD: receptor trimer of dimers. (B) Core-signaling unit/complex: 2 TOD, 1 CheA dimer, 4 CheW. (C) Unit cell: 3 coupled core-signaling units. (D) CheA₂-trimer: 3 core-signaling units, pseudo three-fold symmetry. (E) CheA₂-hexamer: 6 core-signaling units, pseudo six-fold symmetry. (F) Extended lattice: interlocked CheA₂-trimer and CheA₂-hexamer.

Figure 4-figure supplement 2. Overview of molecular modeling and simulation strategy taken in this study. (A) High resolution X-ray structures from *T. maritima* were taken as inputs for the generation of models corresponding to the array's core components, namely the receptor trimer-of-dimers, coupled CheA/CheW rings, and CheW-only ring. (B) The resulting core-component models were arranged heuristically, assuming a 12 nm lattice constant, to produce models of the CheA₂-hexamer and CheA₂-trimer array substructures. For simplicity, only the CheA₂-hexamer organization is shown. (C) A portion of the heuristically-constructed CheA₂-hexamer model corresponding to the array unit cell was extracted for further study with all-atom MD simulations. (D) MDFF simulations were conducted to refine the CheA₂-hexamer and CheA₂-trimer models utilizing their respective density maps. A core-signaling unit was taken from the MDFF-refined CheA₂-trimer model and deposited in the PDB data bank under accession code 3JA6. The full MDFF-refined CheA₂-trimer model was subjected to further investigation using all-atom MD simulations.

Figure 4-figure supplement 3. Computational modeling of the extended chemosensory array

structure. (A) Molecular dynamics (MD) simulations show that *T. maritima* receptors form a stable trimer-of-dimers (TOD). Side view (left) and top view (right) of the highly conserved protein interaction tip, highlighting the inter-receptor salt bridge network formed by E351/R403 and E387/R389. Symmetry-related monomers within individual receptor dimers are distinguished by red and grey coloring. (B) MDFF-refined, all-atom model of the array subunits combining CheA₂-trimer (cyan circle) and CheA₂-hexamer (orange circle) maps. (C) All-atom model of the *T. maritima* lattice containing 3 x 3 unit cells. A portion of the lattice corresponding to a single unit cell is outlined in black. Top (top left) and side (bottom left) views of the array unit cell model arranged as three coupled core-signaling units. Receptor TODs are shown in red, CheA dimers in blue, and CheW monomers in green.

Figure 4-figure supplement 4. Overview of key all-atom molecular dynamics simulations conducted in this study. (A) Table summarizing simulations for each molecular system with atom number reported roughly in millions (M) of atoms (including protein, solvent, and ions) and simulation duration in nanoseconds (ns). (B) Schematic detailing the organization of unit cell simulations. An initial equilibration simulation of 80 ns provided a base structure from which nine wild type (purple) and nine R297A mutant (red) unit cell simulations were subsequently launched in parallel. Each of the 18 production simulations were 450 ns in length, totaling over 8 microseconds of sampling on the unit cell system.

Figure 4-figure supplement 5. Time series of CheA dimer conformations extracted from CheA₂-trimer simulations. Traces track the projection of the conformations of 30 CheA dimers onto the first principal component of the “dipping” motion. Colored traces track CheA dimers

that undergo an extended (>10 ns) “dipping” motion. Horizontal dashed line visually demarcates the undipped and dipped CheA dimer classes. Vertical dashed line separates initial 70 ns MDFF simulation from ten, 120 ns production simulations.

Figure 5-figure supplement 1. CryoEM images of plunge-frozen *E. coli* cells expressing WT Tsr and WT CheA (A&B), R265A CheA (C), R265C CheA (D), R265S CheA (E), and R265E CheA (F). The arrays are marked with white curved arrows. Scale bars, 100nm.

Supplementary File 1. Summary of protein-protein interactions at key interfaces of equilibrated *T. maritima* unit cell model. Residues participating in a given interface but not associated with particular partners are listed separately for each domain. Residues that interact significantly (>50% of frames) are listed as a pair in a separate row. Interactions unique to this study are listed in green. Where ambiguous, residue pairs involving a receptor bound to CheA-P5, CheW from a CheA-P4/CheW ring or CheW from a CheW-only ring are denoted with a (1), (2) or (3) respectively. ** Signifies interfaces taken directly from experimental structures. Recent references pertaining to each protein-protein interface are given.

Video legends:

Video 1. Tomographic slices of monolayer arrays. Related to Figure 2.

Video 2. MDFF model fitting of the CheA2-trimer density map. Related to Figure 3.

1029

1030 **Video 3.** Molecular dynamics simulation of array unit cell. Shown here is a 75 ns clip of a wild
1031 type unit cell trajectory, illustrating the dynamics of the 1.2 million atom model, including 6
1032 receptor TODs (red), 3 CheA dimers (blue), and 12 CheW monomers (green). Periodic images,
1033 shown here with reduced opacity, enforce the boundary conditions of the extended array
1034 architecture but are not simulated explicitly. Solvent and ions have been removed for clarity.
1035 Related to Figure 4.

1036

1037 **Video 4.** Molecular dynamics simulations reveal conformational switch in CheA P4 domain.
1038 Shown here is one of four “dipping” events observed in the wild type unit cell simulations,
1039 leading to modified contacts between the CheA dimer and receptor TODs. Strong contacts
1040 between P3 and neighboring receptor dimers (D333/K390 shown here with licorice
1041 representation) are disrupted in favor of new contacts between P3/P4 and P4/receptor stabilizing
1042 the dipped state (R297/E397 and E390/R379 respectively, shown here with licorice
1043 representation). Related to Figure 4.

1044

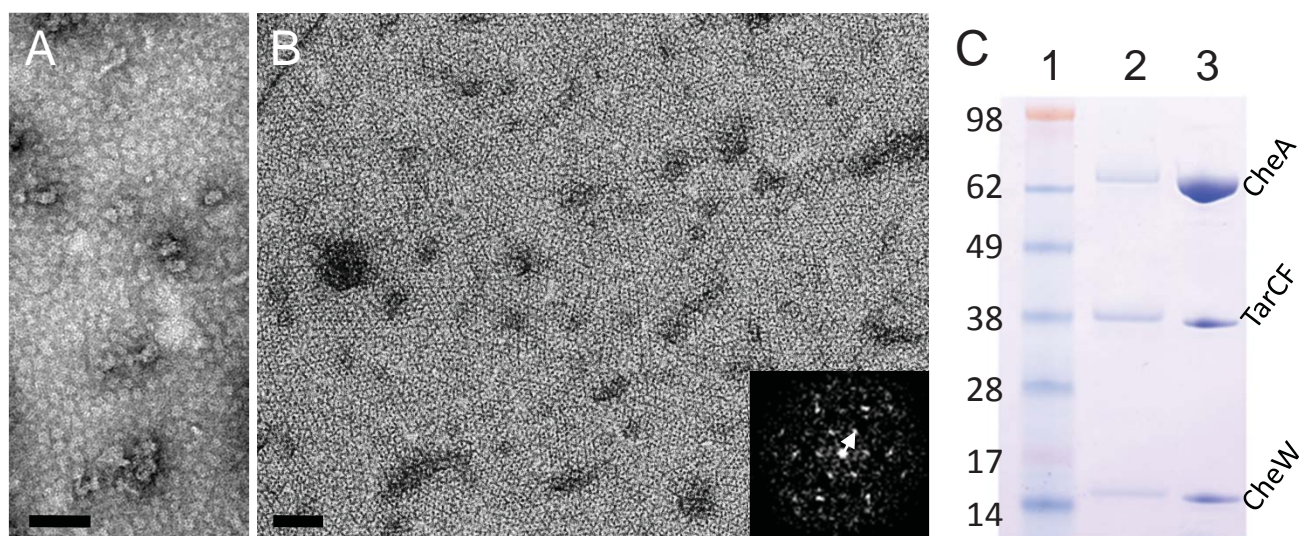


Figure 1

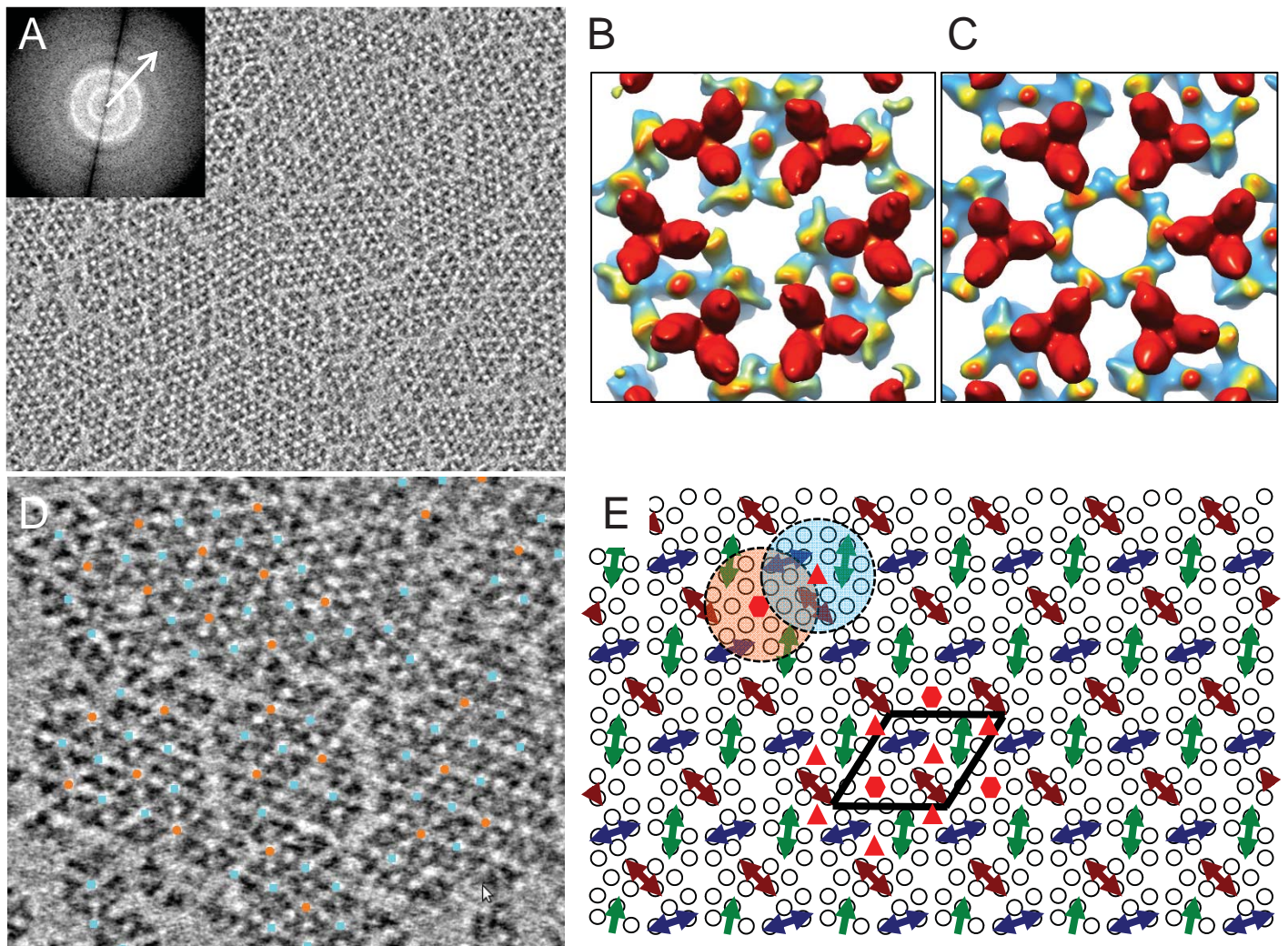


Figure 2

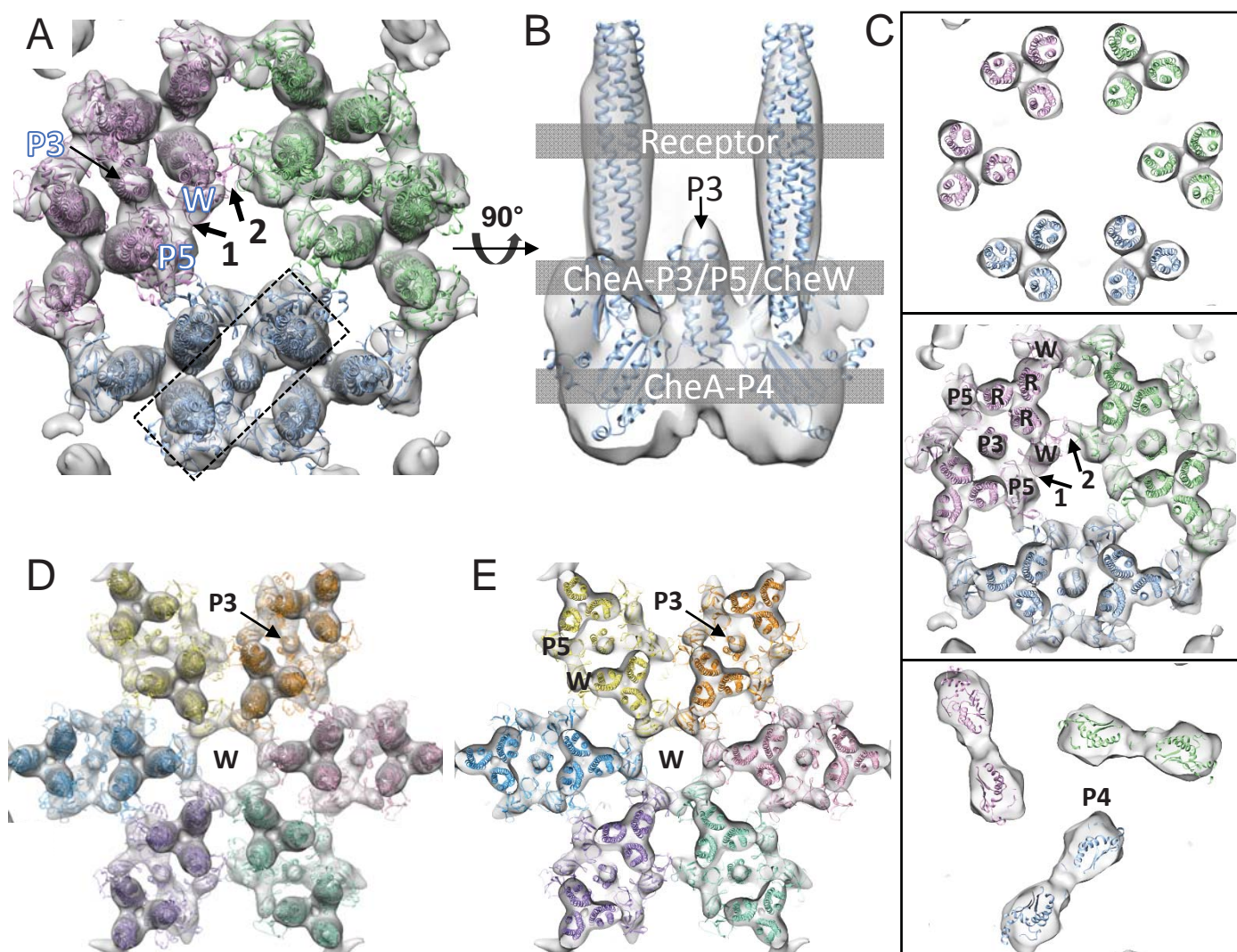


Figure 3

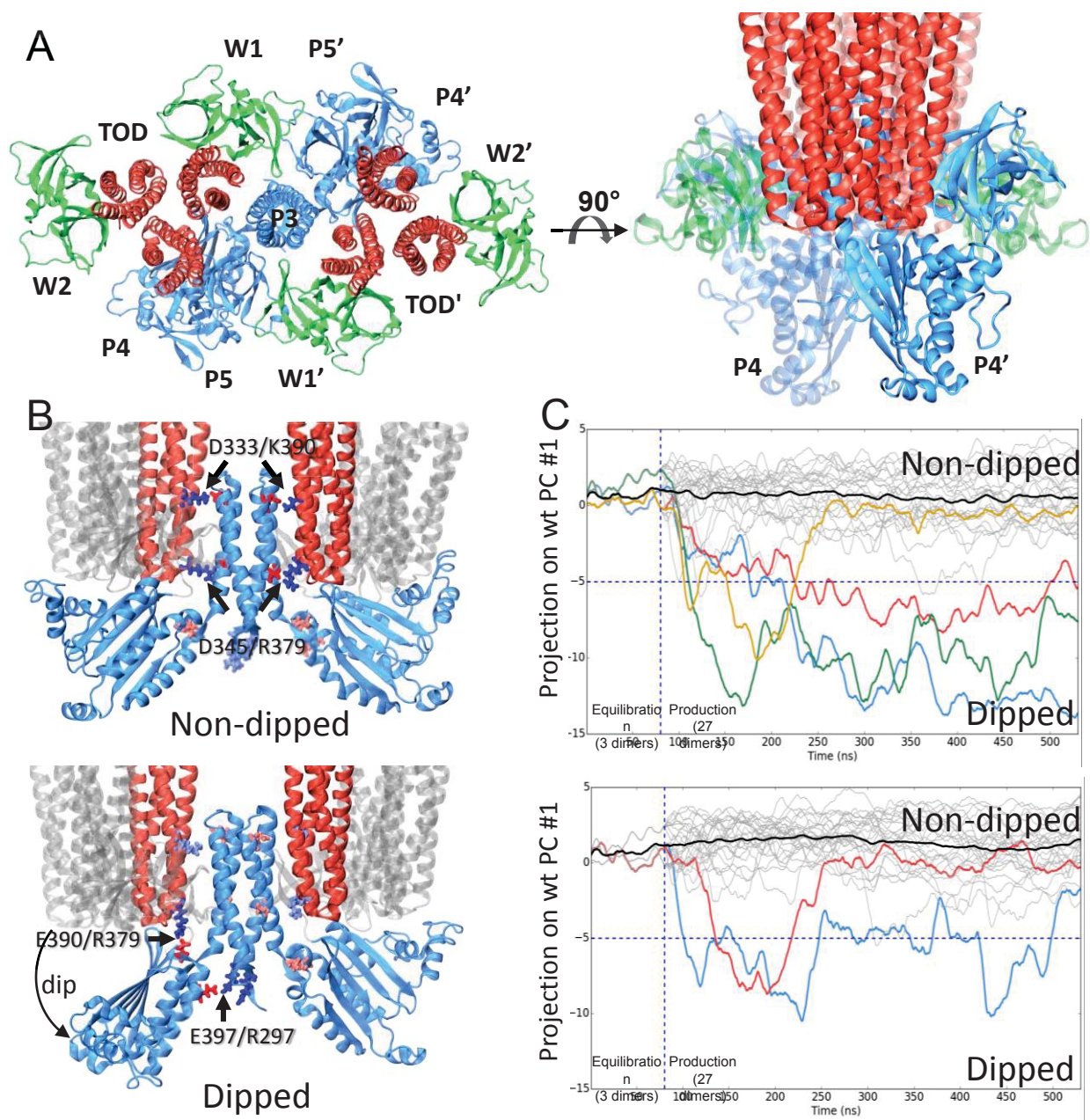


Figure 4

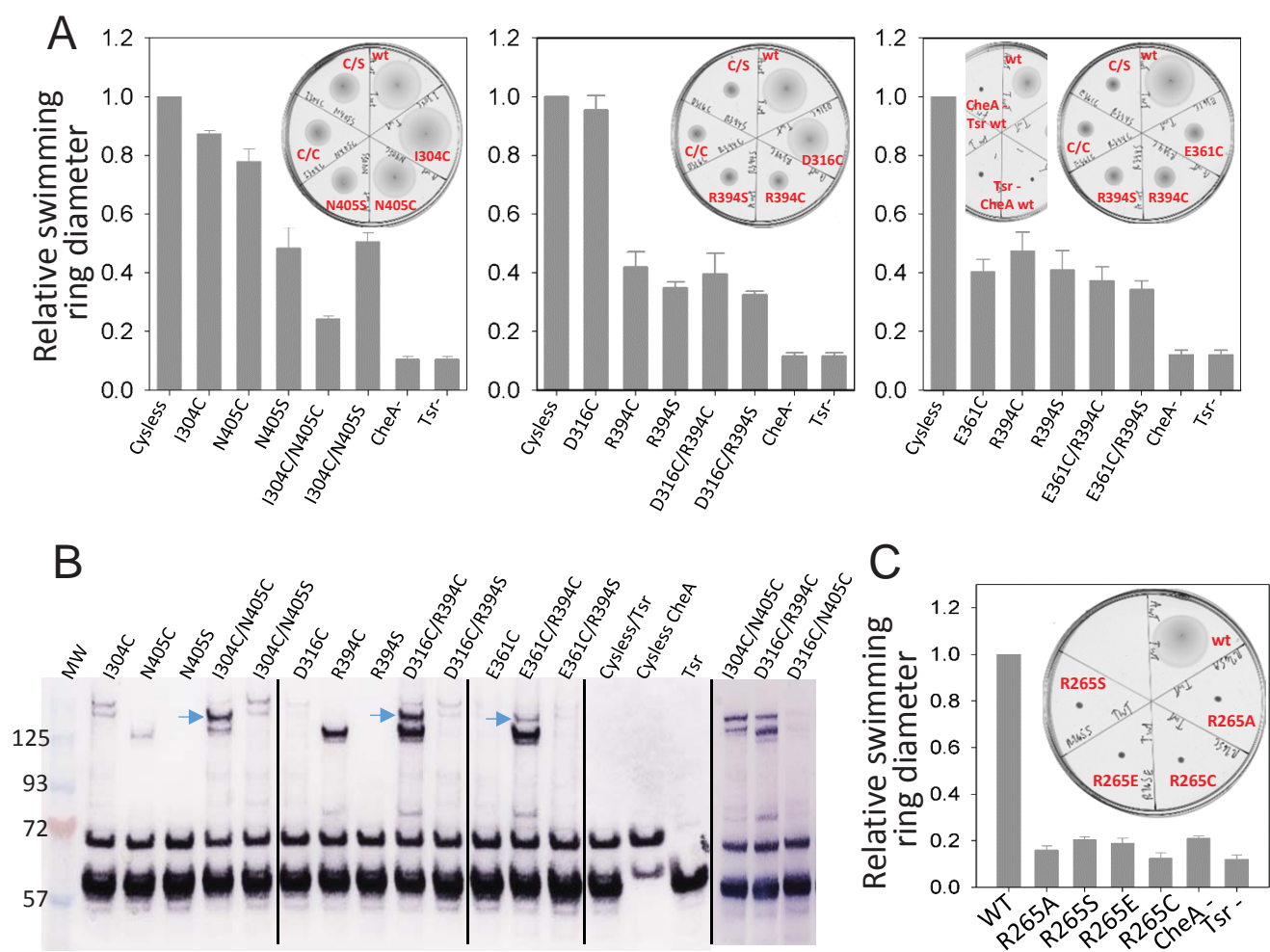


Figure 5