Copyright © 2020 Biophysical Society. Published by Cell Press (Elsevier). All rights reserved. This is the accepted manuscript version of an article which has been published in final form at https://doi.org/10.1016/j.bpj.2020.08.009, made available on this repository under a Creative Commons CC BY-NC-ND attribution licence (https://creativecommons.org/licenses/by-nc-nd/4.0/).

1 *Final*
2 *08/10/2020*
3
4

5 # Allosteric priming of *E. coli* CheY by the flagellar
6 # motor protein FliM

7

8 **P. Wheatley[1], S. Gupta[2], A. Pandini[3,4], Y. Chen[5], C.J. Petzold[5], C.Y. Ralston[2],**
9 **D.F. Blair[1] & S. Khan[4,6*]**

10

11 *[1]Department of Biology, University of Utah, Salt Lake City, UT 84112. USA*
12 *[2]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National*
13 *Laboratory, Berkeley, CA 94720, USA.*
14 *[3]Department of Computer Science—Synthetic Biology Theme, Brunel University*
15 *London, Uxbridge, United Kingdom*
16 *[4]Computational Cell and Molecular Biology, the Francis Crick Institute, London,*
17 *United Kingdom*
18 *[5]Biological Systems and Engineering, Lawrence, Berkeley National Laboratory,*
19 *Berkeley, CA 94720, USA*
20 *[6]Molecular Biology Consortium, Lawrence Berkeley National Laboratory, Berkeley,*
21 *CA 94720, USA*
22

23 *[*]Corresponding Author*

25

26

27
Abstract

28

29 Phosphorylation of *Escherichia coli* CheY protein transduces chemoreceptor
30 stimulation to a highly cooperative flagellar motor response. CheY binds to the N-
31 terminal peptide of the FliM motor protein ($FliM_N$). Constitutively active D13K-Y106W
32 CheY has been an important tool for motor physiology. The crystal structures of
33 CheY and CheY·$FliM_N$ with and without D13K-Y106W have shown $FliM_N$ bound CheY
34 contains features of both active and inactive states. We used molecular dynamics
35 (MD) simulations to characterize the CheY conformational landscape accessed by
36 $FliM_N$ and D13K-Y106W. Mutual information measures identified the central features
37 of the long-range CheY allosteric network between D13K at the D57 phosphorylation
38 site and Y/W106 at the $FliM_N$ interface; namely the closure of the α4-β4 hinge and
39 inward rotation of Y/W106 with W58. We used hydroxy-radical foot-printing with
40 mass spectroscopy (XFMS) to track the solvent accessibility of these and other
41 sidechains. The solution XFMS oxidation rate correlated with the solvent-accessible
42 area of the crystal structures. The protection of allosteric relay sidechains reported
43 by XFMS confirmed the intermediate conformation of the native CheY·$FliM_N$
44 complex, the inactive state of free D13K-Y106W CheY and the MD-based network
45 architecture. We extended the MD analysis to determine temporal coupling and
46 energetics during activation. Coupled aromatic residue rotation was a graded rather
47 than a binary switch with Y/W106 sidechain burial correlated with increased $FliM_N$
48 affinity. Activation entrained CheY fold stabilization to $FliM_N$ affinity. The CheY
49 network could be partitioned into four dynamically coordinated sectors. Residue
50 substitutions mapped to sectors around D57 or the $FliM_N$ interface according to
51 phenotype. $FliM_N$ increased sector size and interactions. These sectors fused
52 between the substituted K13K-W106 residues to organize a tightly packed core and
53 novel surfaces that may bind additional sites to explain the cooperative motor
54 response. The community maps provide a more complete description of CheY
55 priming than proposed thus far.
56
57
58
Statement of Significance

59 CheY affinity for $FliM_N$, its binding target at the flagellar motor, is increased by
60 phosphorylation to switch rotation sense. Atomistic simulations based on CheY and
61 CheY·$FliM_N$ crystal structures with and without the phospho-mimetic double
62 substitution (D13K-Y106W) showed CheY compaction is entrained to increased $FliM_N$
63 affinity. Burial of exposed aromatic sidechains drove compaction, as validated by
64 tracking sidechain solvent accessibility with hydroxyl-radical foot-printing. The
65 substitutions were localized at the phosphorylation pocket (D13K) and $FliM_N$
66 interface (Y106W). Mutual information measures revealed these locations were
67 allosterically coupled by a specialized conduit when the conformational landscape of
68 $FliM_N$-tethered CheY was modified by the substitutions. Novel surfaces stabilized by
69 the conduit may bind additional motor sites, essential for the high cooperativity of
70 the flagellar switch.
71

72 # Introduction

73

74 *Escherichia coli* CheY is a founding member of a bacterial response
75 regulator superfamily that uses aspartate phosphorylation to regulate
76 diverse signal relays (1, 2). The CheY $\beta_5\alpha_5$ fold has structural homology with
77 small eukaryotic signal-transducing proteins (3). CheY phosphorylation
78 couples the occupancy of the chemoreceptor patch to the motile response in
79 bacterial chemotaxis. Previous studies of CheY have established it as a
80 model for fundamental design principles in protein allostery (4). Here, we
81 study *E. coli* CheY binding to the FliM N-terminal peptide ($FliM_N$) responsible
82 for its initial interaction with the flagellar switch complex.
83 CheY, fused with green fluorescent protein (GFP), is both
84 phosphorylated and dephosphorylated at the polar chemoreceptor patch,
85 generating pulsatile fluctuations in intracellular phosphorylated CheY
86 (CheY~P) level (5, 6). The CheY~P diffuses to the flagellar motor within the
87 flagellar basal body, interacting with its C-ring (a.k.a. the switch complex), a
88 multi-subunit assembly composed of the proteins FliG, FliM and FliN. In *E.*
89 *coli*, the interaction increases clockwise {CW} rotation *(7)*. Single-cell
90 measurements expressing GFP-CheY under conditions where CheY~P is the
91 dominant form have shown that motor rotational bias has a sigmoidal
92 dependence on CheY concentration (Hill coefficient > 10.5, K_D = 3 μM) (8),
93 implying highly cooperative action of the captured CheY molecules switching
94 flagellar rotation. More recently, GFP-CheY occupancy was estimated to be
95 about 1/3 and < 1/10 of the 34 FliM subunits present per motor (9) for single
96 CW and CCW rotating motors respectively (10). The occupancy and rotation
97 state were coupled within the image time resolution (20 ms); while GFP-CheY
98 motor dissociation times (70 ms) were faster than the response times to
99 attractant stimuli (11). The single-motor kinetics also imply cooperative
100 CheY-motor interactions.
101 Biochemical experiments coupled with mutagenesis, motility assays
102 and X-ray crystal structures have established that CheY is phosphorylated at
103 a single aspartate ($D57{\sim}PO_4$). The aspartyl phosphate is labile with a 22.8s
104 half-life at ambient temperature (12). The affinity for the $FliM_N$ motor binding
105 target of non-phosphorylated *E. coli* CheY (K_D = ca. 450 μM) is 15x weaker
106 than for CheY~P as measured by fluorescence quenching of CheY residue
107 W58 adjacent to D57 (13, 14)). The binding of CheY~P to isolated, native
108 CCW-locked flagellar switch complexes had K_D stronger than that for $FliM_N$,
109 but was non-cooperative (15), in contrast to the *in-vivo* measurements of
110 rotation bias (8) or motor localization (10) that sample both rotation states
111 (**Supporting Information Section A. Table S1**). The conundrum how
112 cooperative responses arise by CheY~P binding to $FliM_N$ alone is increased by
113 the fact that $FliM_N$ is separated from the rest of the C-ring by a flexible tether
114 (16). Thus, evidence that CheY interaction with the switch involves two
115 binding sites, initial interaction with $FliM_{N,}$ followed by a subsequent

116 interaction of the $FliM_N$ tethered CheY to FliN in *E. coli* (17) provides a
117 plausible resolution. It has remained unclear whether the $FliM_N$ tether
118 facilitates the second-stage binding step only by increasing CheY local
119 concentration, or whether structural changes also occur that prime CheY to
120 bind FliN (**Figure 1**).
121
122
123
124
125 **Figure 1: CheY interactions with the flagellar motor.** *CheY shade intensity*
126 *and size denote activation state and FliN binding probability respectively. Binding of*
127 *activated CheY (CheY*) to isolated switch complexes is not cooperative (H = 1), but*
128 *the change in flagellar CW/CCW rotation bias is highly cooperative (H > 10) with*
129 *CheY* concentration. 1st stage (1) binding to $FliM_N$ enables 2nd stage (2) binding to*
130 *FliN. The increased local concentration due to 1st stage binding and the multiple FliN*
131 *copies enhance 2nd stage binding probability. The Inactive CheY binds weakly,*
132 *reducing $FliM_N$-tethered CheY binding events with FliN below the critical threshold*
133 *for CW rotation. This study provides evidence for structural changes in CheY that*
134 *may supplement increased local concentration for 2nd stage binding.*
135
136 An atomic structure for CheY~P is not available given the lability of the
137 aspartyl phosphate. Therefore, atomic structures of phospho-mimetic CheY
138 proteins obtained by chemical modification (18, 19) or mutagenesis (20-23)
139 have been used to reconstruct the activation mechanism. While both
140 chemically modified and mutated proteins are used in vitro biochemical
141 assays. only the latter can be expressed and studied in vivo (12, 20-22, 24-
142 26). The activating substitutions D13K, Y106W are the most potent
143 modulators reported, thus far of $FliM_N$ binding in vitro (13, 14, 27) and motor
144 rotation bias in vivo (10, 28-30). The comparison of CheYD13KY106W
145 efficacy with CheY~P from both in vitro and in vivo assays, the substantial
146 knowledge of its effects on motor physiology and the availability of atomic
147 structures with and without $FliM_N$ (Supporting Information Table S1) make
148 CheYD13KY106W the logical first choice for the elucidation of the molecular
149 priming mechanism.
150 The crystal structures of D13K-Y106W CheY alone and in complex with
151 $FliM_N$ showed bound $FliM_N$ was required for the activated CheY conformation.
152 They established CheY residues K91, Y106 and K119 as part of the $FliM_{N,}$
153 binding surface (23). K91 and K119 formed salt bridges with $FliM_N$. The W106
154 sidechain moved in as $FliM_N$ bound to switch K109 bonding interactions with
155 T87, D57 and, via bound water, with D12 (23). The structure of the native
156 CheY·$FliM_N$ complex exhibited some features of inactive CheY and some
157 features of the active D13K-Y106W CheY·$FliM_N$ (31). Notably, the orientation
158 of the Y106 sidechain matched that for W106 in the D13K-Y106W CheY·$FliM_N$
159 complex. The “intermediate” conformation of the native CheY·$FliM_N$ structure
160 challenged two-state CheY allostery models that coupled Y/W106 rotamer
161 state to T87 motions (32). An NMR study on free CheY (33) reached a similar
162 conclusion. CheY has high conformational plasticity as seen by the

163 discrepancies between crystal structures of activated CheY proteins (19, 23,
164 34). The coverage of the conformational landscape by crystal structures is
165 too sparse to resolve the conformational trajectories for activation by
166 phosphorylation or binding targets such as $FliM_N$. Alteration of low-affinity
167 binding interfaces, a common occurrence in signal-transducing phospho-
168 protein complexes by crystal packing contacts Is an additional concern (35).
169 CheY conformational plasticity is not well-described by classical protein
170 allostery concepts of “induced fit” (KNF) (36) or “conformational selection”
171 (e.g. MWC (37)) but is accommodated by modern ideas of allostery (38)
172 where protein-protein interactions between flexible partners have been
173 described in terms of a folding funnel, where the funnel bottom has a
174 “rugged” landscape with multiple minima (39). Accordingly, molecular
175 dynamics (MD) simulations and solution measurements have supplemented
176 the X-ray crystallography of free CheY structures. MD of free CheY examined
177 the coupling between Y106 rotation and T87 movements triggered by
178 hydrogen bond formation (40), showing that the β4-α4 loop is an important
179 determinant of allosteric signaling affected by lysine acetylation (41) and
180 extracted common design principles between CheY and other response
181 regulators with correlation analyses (42, 43).
182
183 Here, we detail simulations and solution measurements to better
184 understand the differences between the native and D13K-Y106W CheY
185 crystal structures. We resolved the complex conformational landscapes by
186 MD simulations with mutual information measures to determine the coupling
187 between protein fragments. Protection experiments with XFMS (X-ray foot-
188 printing with mass spectroscopy) (44, 45), a technique that probed sidechain
189 solvent accessibility in contrast to deuterium exchange of backbone
190 hydrogen atoms, supported the $FliM_N$ requirement for D13K-Y106W CheY
191 activation reported by the crystal structures, and the MD allosteric network
192 model. XFMS has a more straight-forward physical rationale than
193 fluorescence quenching for reporting sidechain motions over time-resolved
194 windows and is not limited by the size of the protein assembly. Further
195 analysis of the MD trajectories resolved multiple CheY Y106 rotamer states.
196 Inward orientation was temporally coupled to stabilization of both the CheY
197 fold and the $FliM_N$ interface in the CheY·$FliM_N$ complex, but not in CheY alone.
198 The coupling increased in D13K-Y106WCheY·$FliM_N$. The formation of a distinct
199 module that orchestrates CheY dynamics to stabilize new surface topologies
200 for possible second-stage binding to FliN was the signature of the fully
201 activated D13K-Y106W CheY·$FliM_N$ state.

202 # Materials & Methods

203

204 ## 1. Structure Preparation.

205

206 Structures of *Escherichia coli* CheY (PDB ID: 3CHY. 1.7-angstrom resolution
207 (46)) and complexes of native (PDB ID: 2B1J. 2.8 angstrom resolution (31)) and
208 mutant (13DKY106W) CheY (PDB ID: 1U8T. 1.5 angstrom resolution (23)) with $FliM_N$
209 were downloaded from Protein Data Bank. The label CheY* will, henceforth,
210 specifically apply to CheY13DKY106W. The 1U8T unit cell was a tetramer with 2
211 CheY* and 2 CheY*·$FliM_N$ complexes. We generated the native CheY·$FliM_N$ complex
212 structure (1U8T_DY) by in silico mutagenesis (13 K->D, 106 Y->W) to base the
213 simulations on well-resolved atomic coordinates The reverse mutagenesis and
214 analyses of the crystal structures are detailed in **Supporting Information Section**
215 **B**.

216

217 ## 2. Molecular Simulations.

218

219 ### *(a) Molecular Dynamics*.
220 A set of 3 replicas of duration 1 μs each was generated for the mutant (1U8T)
221 and native (1U8T_DY) complexes using GROMACS 2016.2 with Amber ff99sb*-ILDNP
222 force-field (47). Another set of 3 replicas of 500ns duration each was generated for
223 the native CheY (3CHY). Each system was first solvated in an octahedral box with
224 TIP3P water molecules with a minimal distance between protein and box boundaries
225 of 12 Å. The box was then neutralized with Na^+ ions. Solvation and ion addition were
226 performed with the GROMACS preparation tools (**Supporting Information**
227 **Section B)**.
228 Collective motions were identified by PCA of the conformational ensembles.
229 PCs were generated by diagonalization of the covariance matrix of C^α positions in
230 GROMACS. The overlap (cumulative root mean square inner product) of the PCs
231 between replicas (48)) and the PC dot product matrix was computed with the
232 GROMACS g-anaeig function.
233 The conformational ensembles were clustered and mean structures
234 representing the major clusters (n>5) computed with the GROMACS g-cluster
235 function. The energy landscape was computed with PROPKA 3.0 (49). PROPKA
236 calculates the free energy difference (ΔG) between the folded and unfolded states
237 as the protein charge varies with pH (50). CheY has 37 ionizable groups (9D, 12E,
238 10K, 4R, 2Y) plus N and C termini that determine its net charge. The ΔG is computed
239 from the perturbation of residue pK values by the protein environment; namely the
240 dielectric-dependant de-solvation penalty, backbone and sidechain hydrogen bonds
241 and interactions with other charged residues. For the complexes, the ΔG was
242 computed for the complex (ΔG_T) as well as CheY alone with $FliM_N$ removed (ΔG_{CheY}).
243 The ΔG_{CheY} was the free energy of the CheY fold. The interfacial energy $\Delta G_{interface}$ =
244 ΔG_T - ΔG_{CheY}
245 ### *(b) tCONCOORD*
246 tCONCOORD utilizes distance constraints based on the statistics of residue
247 interactions in a crystal structure library (51, 52), to generate conformational
248 ensembles from one crystal structure with solvent modelled as an implicit
249 continuum. tCONCOORD runs compared conformational ensembles for native CheY

250 (3CHY) with double-mutant CheY, extracted from the heterogenous 1U8T unit cell
251 that contains structures both with and without $FliM_N$. Sets of 16^4 = 65,536
252 equilibrium conformations with full atom detail were typically generated for each
253 structure. The overlap between ensemble subsets was > 99% when the subset size
254 was < $1/4$ of this value (53)**.** The details are in **Supporting Information Section**
255 **B.**
256

257 ## 3. Network Analysis.
258
259 ### *(a) Structural alphabet.*
260 Coordinated CheY motions were examined using mutual-information analysis.
261 The mutual-information (nMI) matrix encodes correlations between conformational
262 states of different parts of the protein backbone (**Supporting Information**
263 **Section B**). The states are represented by a structural alphabet (SA), a set of
264 recurring four residue fragments encoding structural motifs derived from PDB
265 structures (54). Fragments are assigned an SA designation according to backbone
266 dihedral angles, allowing conformation to be specified as a 1-dimensional string
267 (54). The fragments are represented as network nodes, with the connectivity
268 (edges) between them representing their correlated dynamics over the MD
269 trajectory.
270 ### *(b) Eigenvector Analysis.*
271 Statistically significant correlations between columns were identified with
272 GSATools (55) and recorded as a correlation matrix. The correlation matrix was
273 used to generate a network model with the residues as nodes and the correlations
274 as edges. In vector notation, the overall connectivity of a given fragment is
275 reported by its eigenvector centrality, *E* (“centrality”). The contribution of a
276 node to the network was estimated by its *E*, calculated directly from the correlation
277 adjacency matrix:
278 $$E.\{M\}_{corr} = E.\lambda$$
279 where the $\{M\}_{corr}$ is the correlation matrix. The λ is the eigenvalue.
280 The nMI contributions of local fragment motions were computed for the top
281 PCs and superimposed on their RMSF profiles to evaluate the mechanical behavior
282 of the network nodes in driving collective motions. Ensemble conformations and MD
283 runs were averaged for computation of the nMI between fragment positions, with >
284 2σ thresholds for selected top couplings. Pearson’s correlations were used for
285 comparison. Significance limits were set in GSATools.
286 ### *(c) Community Analysis.*
287 The Girvan–Newman algorithm (56) was used to identify community
288 structure. Then the network was collapsed into a simplified graph with one node per
289 community, where the node size is proportional to the number of residues. Edge
290 weights represent the number of nMI couplings between communities (57).
291 Community analysis of correlation networks identifies relatively independent
292 communities that behave as semi-rigid bodies. Graphs were constructed with the
293 *igraph* library (58) in R (*https://cran.r-project.org/web/packages/igraph/*) and
294 visualized in Cytoscape (*http://www.cytoscape.org/).*
295
296 ## 4. Overexpression and purification of CheY proteins
297 The CheY-pET21b plasmids with *E. coli cheY* alone and fused with $FliM_N$ (17)
298 were modified to incorporate the double mutation D13K, Y106W. The native and
299 mutated plasmids were expressed in *E. coli* strain BL21/DE3. The expressed
300 proteins were purified with fast protein liquid chromatography (FPLC) The $FliM_N$.CheY
301 fusion interacts with FliN (17) and is more potent than CheY alone in potentiation of
302 CW rotation (P. Wheatley, unpublished). 3D models of the $FliM_N$.CheY fusions were
303 obtained with the I-Tasser suite (59). In all top five models, $FliM_N$ was docked in the
304 location seen in the crystal structures of the CheY·$FliM_N$ complexes. The top model

305 had cs = -1.08, RMSF = 7.2+4.2 angstroms (against CheY, $FliM_N$ crystal structures)
306 **Supporting Information Section C**.
307
308 **5. X-Ray Foot-printing (XF)**
309
310 Protein samples (CheY, $FliM_N$.CheY, CheY*and $FliM_N$CheY* were prepared in 10
311 mM potassium phosphate buffer (pH 7.2), 100 mM NaCl, and 10 mM $MgCl_2$.
312 Exposure range was determined empirically by adding Alexa488 to protein solutions
313 as previously described (60). Sample irradiation was conducted without Alexa488
314 dye using a microfluidic set-up with 100 mm and 200 mm ID tubing in combination
315 with a syringe pump as previously described (61). After exposure at ALS beamline
316 3.2.1, samples were immediately quenched with methionine amide to stop the
317 secondary oxidations and stored at -80 °C for LCMS analysis.
318 The oxidized fraction, F, for a single residue modification was given by the
319 equation
320 $$F=\{ Xi/ ¿$$
321 where Xi is the oxidized residue abundance of one of the monitored residues
322 in a trypsinized peptide and T is the unoxidized peptide.
323 Best fit first-order rates were calculated in Sigmaplot version 12. Protection
324 factors (PFs) were calculated as the ratio of the intrinsic residue reactivity over its
325 foot-printing rate (62). Its logarithm (log (PF)) was proportional to the SASA. The
326 relation assumes that the foot-printing rate was related to the activation energy
327 associated with the accessibility of the side-chain to hydroxy radicals and the initial
328 step of hydrogen abstraction It empirically gave the best-fit for proteolyzed peptides
329 on a model data set, extended here to single residues (62).
330
331 **6. Mass Spectrometry (MS) Analysis**

332 X-Ray exposed protein samples were digested by Trypsin and the resulted
333 peptide samples were analyzed in an Agilent 6550 iFunnel Q-TOF mass
334 spectrometer coupled to an Agilent 1290 LC system (Agilent Technologies, Santa
335 Clara, CA). Approximately 10 pmol of peptides were loaded onto the Ascentis
336 Peptides ES-C18 column (2.1 mm x 100 mm, 2.7 µm particle size; Sigma- Aldrich,
337 St. Louis, MO) at 0.400 mL/min flow rate and were eluted with the following
338 gradient: initial conditions were 95% solvent A (0.1% formic acid), 5% solvent B
339 (99.9% acetonitrile, 0.1% formic acid). Solvent B was increased to 35% over 5.5
340 min, and was then increased to 80% over 1 min, and held for 3.5 min at a flow rate
341 of 0.6 mL/min, followed by a ramp back down to 5% B over 0.5 min where it was
342 held for 2 min to re-equilibrate the column to original conditions. Peptides were
343 introduced to the mass spectrometer from the LC using a Jet Stream source (Agilent
344 Technologies) and spectra acquired with Agilent Mass Hunter Workstation Software
345 B.06.01. The peptide precursor peak intensities were measured in Mass Hunter
346 quantitative analysis software. Further details and data sets are given in
347 **Supporting Information Section C**.

Results

We analyzed conformational ensembles generated by MD to identify dynamic changes in CheY architecture, using loops and residues implicated in the allosteric relay (see Introduction) as markers. We used XFMS protection experiments to relate the crystal structures to the conformation landscape in solution and test the dynamics predicted by the MD simulations (**Figure 2A**).

1. CheY activating residue substitutions D13K/Y106W stabilize FliM$_N$ association.

Three MD replica runs each was performed for the native CheY structure (3CHY.pdb(46)), the activated D13K-Y106WCheY in complex with N-terminal FliM peptide (FliM$_N$) and alone (1U8T.pdb(23)), and a complex of native (non-activated) CheY with FliM$_N$ engineered in silico from 1U8T.pdb (Methods). The crystal structures showed residue Y106 was in the OUT conformation in CheY (3CHY), but in the IN conformation in CheY·FliM$_N$ (2B1J) and D13K-Y106WCheY·FliM$_N$ (1U8T). The Y/W106 rotamer state was correlated with the orientation of the W58 and F111 sidechains. The engineered complex was used instead of the crystal structure (2B1J.pdb) (31) since the latter, in addition to the lower resolution, had a systematic bias in its RMSF profile from the N- to C-terminus. The bias may be due to mosaicity in the crystal consistent with increased CheY-FliM$_N$ interfacial dynamics (**Supporting Information Figure S1**).

Henceforth, the 1U8T_DY CheY. FliM$_N$ will be referred to as the "native CheY complex" and CheY* FliM$_N$ as the "mutant CheY complex". The root-mean-square fluctuation (C$^\alpha$ RMSF) profile for each structure, averaged over three 1 μs runs, are shown in **Figure 2B.** The MD excluded the first three residues (M$_1$GD$_3$) of the FliM$_N$ sequence (M$_1$GDSILSQAEIDALL$_{16}$) as these were not resolved in the 1U8T structure. The CheY* FliM$_N$ complex had higher RMSF values for the α-helix 1 (residues 22-30) and connected β5-α5 loop (residues 109-114), but lower values for the α_4-β_4 loop (residues 88-96) relative to CheY, CheY·FliM$_N$. These flexibility differences were consistent with the altered bond arrangements between residues D12, K13 and K109 (α-helix 1 / β5-α5 loop) and bond formation between K91, (α_4-β_4 loop) and FliM$_N$ D3 seen in the crystal structures (23). The profiles are compared with B-factors for the X-ray structures. The B-factors were high relative to the MD-derived RMSF's, particularly in loop regions, reflecting conformational heterogeneity of these segments in the crystals (Supporting Information Figure S1).

The 3CHY MD trajectories revealed transitions of Y106 between the OUT and IN states, consistent with electron density observed for both states

392 in the crystal structure. $FliM_N$ secondary structure, the CheYK119-$FliM_N$D12
393 salt-bridge and Y106/W106 rotamer state were conserved between the 2B1J
394 and 1U8T crystal structures. However, the raw MD trajectories of the
395 complexes showed $FliM_N$ had higher mean C^α RMSF values when CheY was
396 wild type than when it carried the activating substitutions **(Supporting**
397 **Information Videos S1, S2, S3).** This difference was due to
398 association/dissociation of the $FliM_N$ N and C termini from native CheY. In
399 CheY*·$FliM_N$ trajectories, the peptide center was tethered by the CheYK119-
400 $FliM_N$D12 salt bridge. CheY* W106 was locked IN and part of the segment
401 with the lowest C^α RMSF together with K109 and F111. In CheY·$FliM_N$
402 trajectories, OUT excursions of Y106 cleaved this salt-bridge and weakened
403 interfacial attachments **(Figure 2C, Supporting Information Video S4)**.
404 Thus, the MD confirmed the suggestion from the CheY·$FliM_N$ crystal structure
405 that its $FliM_N$ interface was labile.
406

407 **Figure 2: Dynamics of CheY-$FliM_N$ association.** ***A.*** *Structure of CheY in complex*
408 *with $FliM_N$ (2B1J-AC.pdb). Colors indicate $FliM_N$ (yellow), tryptic CheY fragments*
409 *(blue), allosteric relay loops (green), sidechains (M (magenta), K (cyan), Y, W, F*
410 *(gold). D57 C^α (red asterisk), Mg^{2+} (magenta),* ***B.*** *MD RMSF profiles for the combined*
411 *replica trajectories for the three structures analyzed in this study. Bars mark CheY*
412 *loops α3-β3 (white) and α4-β4 (black). Asterisks mark residues Y/W106 (black), K109*
413 *(cyan) and F111 (yellow). $FliM_N$ residue D12 (red asterisk) forms a salt-bridge with*
414 *CheY K119.* **C.** *Snapshots of CheY (blue) Y106 (red) transitions in 1U8T_DY coupled*
415 *to internal and interfacial residues. $FliM_N$ (yellow). (i) T87 (lime). Supporting*
416 *Information Video S2. (ii) K119 (green), $FliM_N$ D12 (pink). Supporting Information*
417 *Video S4.*
418

419 **2. Two loops control CheY network dynamics**.
420

421 Previous MD simulations focused on the coupling between selected
422 residues implicated from genetic or biochemical data in the long-range
423 allosteric communication within CheY (40, 41, 43). Here, we develop a CheY
424 network model (**Figure 3A**), constructed in (42). for a quantitative
425 description of allostery within the entire protein from the MD conformational
426 ensembles. The model is based on three key concepts: the structural
427 alphabet (SA), the normalized mutual information (nMI) and the eigenvector
428 centrality (*E*). The 3D C^α conformation of four residue peptides is uniquely
429 specified by three bond and torsion angles with distinct conformational
430 clusters ("alphabets") resolved upon inspection of the PDB protein structure
431 database (54). First, the SA was used to convert the 3D CheY fold to a 1D N-
432 C terminal sequence of four-residue fragments and different conformations
433 in an MD trajectory represented as a sequence of 1D strings. Second, the
434 correlation between the confirmation of different CheY fragments within the
435 sequence was computed as the mutual information (MI). The normalized
436 mutual information (nMI) was the MI corrected for correlations expected by
437 chance and the estimated uncertainty due to the finite number of

438 conformations in the ensemble. The nMI couplings constituted the "edges" of
439 the CheY network with correlation strength denoted by the edge thickness.
440 The fragment positions within the sequence alignment formed the "nodes" of
441 the network. Third, the connectivity of the network was determined by *E*, a
442 measure of the influence of individual nodes in the network as reflected by the
443 coupling of their dynamics with other nodes in the nMI correlation matrix. A
444 node *E* limit value of 0 represented the case where its dynamics did not
445 affect other nodes in the network. An E limit of 1 represented the case where
446 its conformational fluctuations switched the entire network between discrete
447 structural states (see Materials & Methods for formal definitions).
448 First, we identified the central nodes in the CheY global network with
449 the highest connectivity (**Figure 3B**). The central CheY nodes were the loops
450 β3-α3 (D_{57}WNMPNMDG) and β4-α4 (T_{87}AEAKK). A third prominent node just
451 below the 1σ threshold was the short β_5-strand (Y_{106}VVKP). Second, we used
452 tCONCOORD, a computationally inexpensive method to generate
453 conformational ensembles for comparison of the CheY and CheY*
454 conformational landscapes.
455 Interpretation of differences between CheY and CheY* crystal
456 structures based on isolated landmarks (for example, Y106 rotamer state (IN/
457 OUT in CheY (3CHY.pdb) (46)) versus W106 (OUT in CheY*(23)) are
458 complicated by the CheY conformational plasticity. Analysis of the
459 tCONCOORD ensembles showed the central network nodes remained
460 unchanged, with both CheY Y106 and CheY* W106 sidechains restricted to a
461 limited OUT-orientation range (**Supporting Information Figure S2**).
462
463 We next examined the CheY and CheY* complexes with $FliM_N$ (**Figure
464 3C**). We split the CheY ensemble into four sub-populations to assess the
465 significance of differences observed between it and the complexes. The
466 network connectivity, as formalized by centrality plots, showed significant
467 changes in the complexes relative to the CheY protein alone. There was a
468 dramatic reduction in the centrality of loop β4-α4 and associated β-strand
469 $*_{106}$VVKP (* = Y (CheY·$FliM_N$), W (CheY*·$FliM_N$)) at the FliM binding surface.
470 Their roles as network nodes were reduced in CheY·$FliM_N$ and abolished in
471 CheY*·$FliM_N$. This trend contrasted with the conservation of these nodes for
472 CheY*. The centrality of α-helix 1 increased with its mobility (Figure 2C).
473
474 **Figure 3: CheY network dynamics. *A.*** *The global network has nodes*
475 *(residue fragments) and edges (mutual information weighted node interactions).* ***B.***
476 *Nodes containing residues that are part of the allosteric relay (W58, K91, Y106,*
477 *K109, F111) have high scores in the CheY network. In A, B, these residues and*
478 *control residues (M17, M60, M63) monitored by XFMS are highlighted (yellow*
479 *circles).* ***C.*** *Centrality profiles of the complexes. (**(i)** CheY/$FliM_N$ (green).* ***(ii)***
480 *CheY*·$FliM_N$ (red)) compared with the native CheY profile (mean $\pm$ s.e; blue lines).*
481 *The dotted line ((ii) red) plots the mutual information between the local loop*
482 *fragment dynamics and collective PC1 motions. Complex formation reduced the*
483 *centrality of the β4-α4 loop that together with the β3-α3 loop formed central nodes*
484 *in the CheY network. Activating mutations eliminated the β4-α4 loop as a node but*

485 *did not alter the contribution of the* β*3-*α*3 loop in CheY*·FliM*$_N$*. Horizontal bars*
486 *indicate* α*3-*β*3 (white) and* α*4-*β*4 (black) loops as in Figure 2B.*
487

488
489 **3. Immobilization of the α4-β4 loop modulates CheY collective**
490 **motions.**
491
492 We used Principal Component Analysis to characterize CheY collective
493 motions and their modulation by $FliM_N$ binding and the activating
494 substitutions. The Principal Components (PCs) are derived from the atomic-
495 coordinate covariance matrix and describe C^{α} backbone movements, ranked
496 according to the amplitude of the structural variation they explain. The
497 collective motions were described well by the first few PCs, as found for
498 other proteins. The first three principal components (PCI-PC2-PC3) accounted
499 for > 60% of all motions in each case. These three PCs comprise bending
500 and twisting modes organized around the β-sheet core. A core sub-
501 population of CheY conformations was observed in MD trajectories generated
502 by all three structures. When CheY is in complex with $FliM_N$, new sub-
503 populations comparable in size to the core were generated. These were
504 distinct in the CheY·$FliM_N$ and CheY*·$FliM_N$ complexes. Thus, new
505 conformational ensembles are accessed upon binding of $FliM_N$, with the
506 potential to produce binding surfaces for additional targets (**Supporting**
507 **Information Figure S3A, B**).
508 Loops act as hinge elements for collective motions. Their mechanics
509 give insight into the modules they control (39). We computed loop β3-α3 and
510 β4-α4 hinge flexibility by mapping their RMSF onto the PC1 that accounts for
511 > 40 % of the total amplitude of the PC motions. Flexibility scaled with the
512 magnitude of the loop RMSFs relative to the mean PC1 RMSF. We computed
513 hinge contribution to the PC1 as the nMI between PC1 variance and the local
514 loop fragment dynamics. The long β3-α3 loop partitioned into two segments.
515 The short D_{57}WN and the adjacent M_{60}PNMDG loop segments behaved as
516 rigid (low RMSF) and flexible (high RMSF) hinges respectively to control
517 native CheY PC1 dynamics. In the CheY*-$FliM_N$ complex, the β3-α3 loop hinge
518 was retained, but with inverted flexibility of the two segments. The transition
519 for loop β4-α4 was more dramatic from a flexible hinge in native CheY to a
520 closed hinge that acted as a rigid lever arm in CheY*-$FliM_N$. The reduced
521 flexibility decreased β4-α4 loop centrality and influence on PC1 motions
522 (**Supporting Information Figure S3C**),
523
524
525 **4. Protection experiments support the "intermediate" CheY·$FliM_N$**
526 **structure and the MD allosteric network.**

527 We studied homogenous solutions of CheY and $FliM_N$-CheY fusion
528 proteins (**Supporting Information Figure S4**), to measure the changes
529 predicted by the crystal structures and the MD network model. The fusions
530 were critical since the affinity of FliM for CheY is weak and that for the
531 inactive protein even weaker (Introduction). The crystal structures reported
532 that **(i)** Aromatic sidechain internalization in CheY was entrained to $FliM_N$

533 attachment, and **(ii)** The configurations of free CheY with or without the
534 D13K-Y106W substitutions were similar. The MD revealed **(iii)** $FliM_N$
535 attachment was more labile in the native versus D13K-Y106W complexes,
536 and **(iv)** generated a network model to discriminate between CheY
537 fragments that changed upon activation from those that did not. These
538 predictions were assessed by comparing the sidechain solvent accessibility
539 of allosteric relay residues Y106, W58, K91, K109, F111 and K119 by
540 hydroxyl radical foot-printing in the native and D13K-Y106W CheY proteins,
541 and their $FliM_N$-fusion constructs. The control residues predicted not to
542 change during activation were the β3-α3 loop residues M60 and M63 in
543 proximity to W58, and the M17 in proximity to D/K13.

544 Aromatic residues have high intrinsic sidechain reactivities with
545 hydroxyl radicals, exceeded only by methionine and cysteine (absent from *E.*
546 *coli* CheY) followed by the alkaline sidechains. Tryptic digestion partitioned
547 CheY into six separated peptides that were distinguished by mass
548 spectroscopy (MS) based on their characteristic m/z ratio, allowing oxidation
549 of these residues to be monitored. Dose-response curves were generated for
550 each of the four constructs (CheY, CheY*, CheY-$FliM_N$, CheY*-$FliM_N$). For each
551 residue examined, the curves from two independent experiments were
552 pooled (**Supporting Information Figure S5**).
553

554 CheY residues of the allosteric relay at the $FliM_N$ interface and distant
555 from it were designated “interface” and “core” residues respectively. The
556 oxidation of the interfacial residues (K119, Y/W106, K91) was reduced in the
557 complexes (**Figure 4 A. Ii-iii**). Importantly, oxidation of the core residues
558 also decreased with complex formation (**Figure 4 A. Ci-iii)**. In contrast,
559 there was no significant difference between oxidation rates for β3-α3 loop
560 control residues M60, M63 in the fusion proteins versus the free proteins,
561 while the oxidation of the control M17 in the fusions was comparable or
562 greater than in the corresponding free CheY proteins (**Figure 4B**).
563

564 Protection factors (PFs) were computed from the initial rates from the
565 single residue dose-response curves following protocols established by the
566 study of 24 peptides from 3 globular model proteins (62), with intrinsic
567 reactivities mostly determined thus far from measurements on small
568 peptides (63). We first evaluated the agreement between solvent
569 accessibility reported by the XFMS measurements and the crystal structures.
570 Protection factors read out the solvent-accessible surface area (SASA), with
571 some caveats (62), The log(PF)s were plotted against the residue SASA in the
572 crystal structures. The overall correlation was comparable to published
573 values for the peptide correlations for the model proteins (62), indicating
574 that the changes in the dose-response plots for the monitored residues are
575 due, in large part, to non-polar bulk solvent accessibility changes (**Figure**
576 **5A**). Outliers (M17, K109, F111) were restricted to a small CheY protein
577 volume in the structures (**Supporting Information Figure S6**). The crystal
578 structures may not reflect the solution conformation of this local region, but

579 bonding interactions may also contribute (Supporting Information Figure S6
580 legend). The correlation improved markedly (0.60 -> 0.86), without further
581 correction, if the outliers were excluded.
582
583 The PFs for CheY*, CheY·FliM$_N$ and CheY*·FliM$_N$ were then normalized
584 for each residue against the value obtained for CheY (**Figure 5B**). The
585 normalized (log (PF)s) provided a quantitative measure for the increase for
586 both interfacial and core residues in the CheY·FliM$_N$ and CheY*·FliM$_N$ fusions
587 relative to the values for CheY. These residues were significantly more
588 protected in CheY*·FliM$_N$ than CheY·FliM$_N$. In contrast, the protection of the
589 control residues in the fusions (CheY·FliM$_N$, CheY*·FliM$_N$) did not differ
590 significantly from that measured for CheY. The normalized PFs showed no
591 significant difference in protection for interfacial, core or control residues in
592 CheY* relative to CheY.
593
594
595

596 **Figure 4: XFMS Measurements.** *Dose-response curves for* ***A****. Relay**.** Interfacial*
597 *residues* ***(Ii)*** *Y/106W.* ***(Iii)*** *K119**.*** ***(Iiii)*** *K91. Core residues* ***(Ci)*** *K109.* ***(Cii)*** *F111*
598 ***(Ciii)*** *W58.* ***B****. Control residues.* ***(i)*** *M60, M63.* ***(ii)*** *M17. Initial rates (dashed lines)*
599 *were obtained from least-squares linear regression of the decrease in the un-*
600 *oxidized fraction with dose.*
601
602 **Figure 5. Single residue oxidations related to SASA.** ***A.*** *Log (PF)s plotted*
603 *against the side-chain solvent accessible surface area (SASA) calculated from the*
604 *crystal structures. Pearson correlation coefficients: 0.86 (minus M17 (rose), K109*
605 *(cyan). See text). Overall = 0.60 {CheY= (-)0.76; CheY*= (-)0.70; FliM$_N$.CheY=*
606 *(-)0.54; FliM$_N$.CheY*= (-)0.12}. Best-fit (black dashed line), 95% confidence limit*
607 *(blue lines), 95% prediction limit (red lines).* ***B.*** *Protection of interfacial (K119,*
608 *Y/W106, K91), core (F111, K109, W58) and control (M17, M60, M63) residues in*
609 *CheY*, CheY·FliM$_N$, CheY*FliM$_N$ relative to their protection in CheY. {Protection}$_{norm}$ =*
610 *Log {PF/PF$_{CheY}$}. Positive values indicate increased protection.*
611
612 The protection profiles showed that solvent accessibility for the
613 allosteric relay residues decreased in the order CheY < CheY·FliM$_N$ <
614 CheY*·FliM$_N$. The control residues either showed no changes or the opposite
615 trend. Changes in the solvent accessibility of CheY* relative to CheY were not
616 significant. Thus, in conclusion, the XFMS experiments validated the main
617 predictions of the crystal structures and the conformational ensembles
618 generated from them.
619
620

621 **5. Energetics of CheY stabilization by FliM$_N$ and D13K/Y106W**
622 **residue substitutions.**
623

624 The XFMS measurements correlated solution population shifts in
625 selected residue positions to each other and with the crystal structures. The
626 temporal couplings between these shifts could only be studied with MD. We
627 next analyzed the MD trajectories to extract this information.
628 We examined the temporal coupling between the electrostatic
629 stabilization of the interface and the CheY fold with the rotational states of
630 residue Y106 (106W in CheY*·FliM$_N$). CheY*·FliM$_N$ 106W sidechain was locked
631 IN (Supporting Information Video S3). In contrast, Y106 in CheY (Supporting
632 Information Video S1) and CheY·FliM$_N$ (Supporting Information Video S2)
633 made frequent OUT <-> IN excursions. Dwell times in the Y106 rotamer
634 states measured from the raw CheY trajectories were 107 ± 34 ns (OUT) and
635 15 ± 4 ns (IN)). The CheY·FliM$_N$ Y106 sidechain was predominantly in the IN
636 orientation, with mean dwell time 239 ± 123 ns, 15-fold greater than for free
637 CheY. The conformational ensembles in the MD trajectories were clustered
638 based on the C^{α} backbone dynamics {RMSF} The major clusters represented
639 distinct backbone conformational states accessed during the MD runs. The
640 average structures for these clusters were compared to each other and the
641 crystal structures with PROPKA. The mean ΔG values at pH 7.0 were CheY (-
642 4.8 ± 1.0 (n=7)) < CheY·FliM$_N$ (-5.8 ± 1.6 (n=4)) < CheY*·FliM$_N$ (-9.9 ± 2.2
643 (n=3)). All CheY clusters had Y106 in the OUT orientation (θ = 126.7 ± 3.8^{o})
644 indicating that CheY Y106 IN states were too short-lived to influence
645 backbone dynamics. CheY*·FliM$_N$ clusters had W106 in the IN orientation (θ =
646 54.1 ± 2.3^{o}). The CheY·FliM$_N$ clusters, in striking contrast, spanned the entire
647 Y106 rotamer range. Thus, the intermediate CheY·FliM$_N$ Y106 rotamer states
648 were sufficiently stable to affect backbone dynamics (**Figure 6**).
649 .
650
651

652

653 **Figure 6: Rotamer Y/W106 energetics.** ***A. Interface and CheY fold***
654 ***stabilization.*** *Interface (ΔG_{int}, triangle), CheY fold (ΔG_{CheY}, circle). Linear regressions*
655 *(interface (dashed), fold (solid).* ***(i)*** *CheY·FliM$_N$ (green). 2B1J crystal values (lime).*
656 *Vertical lines and rectangles show (CheY (cyan) and CheY*. FliM$_N$ (red) θ and ΔG_{core}*
657 *range respectively. Correlations: θ-$\Delta G_{interface}$ (R = 0.23, Pearson = 0.63); θ-ΔG_{CheY} (R =*
658 *0.43, Pearson = 0.21).* ***(ii)*** *CheY*·FliM$_N$ (red). 1U8T crystal values (purple).*
659 *Correlations: θ-$\Delta G_{interface}$ (R = 0.96, Pearson = 0.98); θ-ΔG_{CheY} (R = 0.85, Pearson =*
660 *0.33)* ***B.*** *CheY conformation and Y106 (green) sidechain rotamer orientation in*
661 *representatives of the major CheY·FliM$_N$ clusters.*
662

663 Next, we computed the activation energetics by measurement of
664 ionizable residue electrostatics with PROPKA. There was a weak stabilization
665 of the CheY FliM$_N$ interface and core with the internalization of the Y106
666 sidechain. The buried CheY*·FliM$_N$ W106 sidechain had a substantially more
667 restricted rotation range than the CheY·FliM$_N$ Y106 sidechain. However, the
668 correlation between side-chain orientation and stabilization of CheY* FliM$_N$

669 interface and CheY* core was stronger, consistent with a more-tightly packed
670 CheY*·FliM$_N$ complex. The stabilization of the interface by the D13K/Y106W
671 residue substitutions was consistent with the different FliM$_N$ binding affinities
672 measured in solution for active versus inactive CheY states. The novel result
673 was the coupled stabilization of the CheY fold for both CheY·FliM$_N$ and
674 CheY*·FliM$_N$.
675 The energetics computed for the 1U8T crystal structure was in line
676 with results from the MD conformational ensembles. In contrast, the values
677 computed for the 2B1J crystal structure were outliers reporting higher energy
678 states relative to the values obtained from the MD runs, an outcome that
679 may be linked to errors in atomic coordinate positions due to the increased
680 B-factor values around the 2B1J CheY·FliM$_N$ interface (Supporting Information
681 Figure S1) and/or deformation of the local volume around K109, FIII, M17 by
682 crystal packing contacts (Supporting Information Figure S6).
683
684 **6. An emergent sector orchestrates CheY* allosteric communication.**
685
686 We developed the network model for a comprehensive representation
687 of the temporal conformational couplings. The centrality analysis identified
688 network nodes with the dominant couplings but the non-nodal fragment
689 couplings that constituted (>95%) of the information available in the nMI
690 matrix were not well-represented. We used community analysis, a recently
691 developed tool for detection of higher-order organization of protein dynamics
692 (64, 65), Community networks are collapsed networks that reduce, partition
693 and map the protein into contiguous, semi-rigid bodies (“sectors”) that may
694 be schematized for a concise, comprehensive representation. The
695 schematics and their mapping onto the 3D structure will be henceforth
696 referred to as community “network” and “map” respectively.
697
698 Community analysis of native CheY revealed distinct sectors (n > 5)
699 displaying coordinated dynamics. The β3 strand $F_{53}VISD_{57}$ occupied a central
700 location in contact with all sectors. Sector A, organized around the D57
701 phosphorylation site coupled to the other sectors, particularly with sector-B,
702 organized around the FliM$_N$-binding surface. The tCONCOORD CheY*
703 community map, when compared against the corresponding CheY map,
704 showed a small increase in sector A relative to sector C interactions with
705 sector B (**Figure 7A)**. This result may indicate limited activation of CheY*
706 relative to CheY detectible with the more sensitive community versus global
707 network, but does not challenge the conclusion that CheY and CheY* have
708 similar dynamic architecture based on the retention of the α_3-β_3 and α_4-β_4
709 loops as network nodes.
710
711 The MD resolved the tCONCOORD sector C into two sectors (**Figure**
712 **7B**). Importantly, reported residue substitutions partitioned to sectors A and
713 B in the more detailed map according to phenotype (**Supporting**
714 **Information Section D. Table S2**). Positions, where these are known to

715 affect dephosphorylation kinetics (65), mapped to sector A. Residues known
716 to affect $FliM_N$ binding or rotation bias, such as sites of suppressor
717 substitutions for CW- or CCW-biasing FliM lesions (66), mapped to sector B.
718 Positions yielding mutations that affect interaction with the CheY-
719 phosphatase CheZ (67) were adjacent to Sector D, the smallest sector
720 obtained for CheY. Sector C, comparable in size to A, might be expected to
721 influence the overall stability and rigidity of the protein.
722

723 Changes in loop dynamics upon complex formation were reflected in
724 the networks (Figure 7B). The couplings between sectors A (phosphorylation)
725 and B ($FliM_N$ binding) were strengthened relative to the free protein. Sector B
726 expanded at the expense of sector C and coupled more strongly to sector A
727 in the CheY·$FliM_N$ network. The mutated residue D13 was part of a loop that
728 flipped from sector A to sector B. A fifth sector (E ($K_{45-48}N_{62}$-$L_{65-68}A_{101}$-$S_{104-107}F_{111-}$
729 $_{114}K_{119-123}$)) spanned by the substituted residues (K13, W106) formed in the
730 network of the activated-mutant CheY-$FliM_N$ complex (CheY*·$FliM_N$). The E-
731 sector fragments were drawn from sector A ($K_{45,}$ $N_{62,}$ K_{119}), sector B ($A_{101,}$ $S_{104,}$
732 F_{111}) or fragments adjacent to these sectors in the free CheY community
733 network. Sector E formed a surface-exposed ridge that connected the $FliM_N$
734 α-helix, via $S_{104-107}$ and $K_{119-123}$, to sector C residues E_{35} and (via K_{45}) E_{37},
735 (**Figure 7C, Supporting Information Video S5**). The top nMI couplings
736 connected sector E fragments within the central β3-α3 loop to the D57
737 phosphorylation site. These couplings were unchanged by complex
738 formation.

739

740 **Figure 7: Changes in community network architecture triggered by**
741 **D13K/Y106W substitutions and $FliM_N$ peptide.** *The reduced number of sectors*
742 *compared to single fragments as nodes provided a concise, quantitative readout of*
743 *the protein dynamics.* ***A. CheY and CheY* community maps.*** *Networks* ***(boxed***
744 ***insets)*** *from tCONCOORD runs show the reduction in the size of sector C relative to*
745 *sectors A and B in CheY* versus CheY.* ***B. CheY, CheY·$FliM_N$ and CheY*·$FliM_N$***
746 ***community architecture.*** *Networks* ***(top)*** *and maps* ***(bottom).*** *$FliM_N$ = yellow*
747 *(cartoon representation). The MD detected four dynamic sectors for CheY (A= cyan,*
748 *B = blue, C = orange, D = red). The sector C from the tCONCOORD runs is resolved*
749 *into two sectors (C and D) in the MD runs. Node size = sector membership; edge*
750 *thickness = weighted inter-sector interactions). Sectors A and B are built around the*
751 *phosphorylation site (D57 (red asterisks)) and the $FliM_N$ binding surface*
752 *respectively. They increase at the expense of sector C upon complex formation. The*
753 *presence of phospho-mimetic substitutions in the CheY*·$FliM_N$ complex creates an*
754 *additional sector E from sectors A and B, that orchestrates interactions with sectors*
755 *C and D.* ***C. CheY*·$FliM_N$ community map showing Sector E surface.*** *See*
756 *Supporting Information Video S5 for 3D perspective. Sidechains identify the*
757 *substituted residues (K13, W106) and $FliM_N$ binding residue K119, a part of Sector E.*
758 *Sectors are colored as in B. The strength of the top (>+2σ) nMI couplings (lines)*

couplings are reflected in their thickness and color (low (yellow) -> high (red)). D57 (red asterisk).

Discussion

The results of this study advance our understanding of CheY conformational plasticity and activation in important ways (**Figure 8)**.

Figure 8: Allosteric priming in E. coli CheY. A. *Reaction coordinate (x-axis) showing stabilization of the CheY fold coupled with CheY activation. The inward rotation of residue Y106 and the increased residue W58 fluorescence quenching due to its internalization, represented by red asterisk size, have been used to measure CheY activation and FliM$_N$ binding respectively. Horizontal bars indicate multiple local minima. CheY ensembles (blue) have large conformational heterogeneity, controlled by a flexible β4-α4 loop. They sample both Y106 IN and OUT rotamer states; but the IN state is too short-lived to generate CheY sub-populations with distinct backbone conformations. FliM$_N$ bound CheY ensembles (green) sample a conformational landscape with a large ΔG range, with prominent troughs among the local minima that track the progressive stabilization of the CheY fold and concerted Internalization of Y106 entrained to tighter FliM$_N$ attachment. FliM$_N$ bound to D13K-Y106W CheY (CheY*) confines the CheY fold to conformational space (red) around the global minimum. The β4-α4 loop is immobilized by the CheY.K91-FliM$_N$ salt-bridge and W106 plus W58 are locked IN a tightly-packed CheY core, with the emergence of a dedicated sector (E) for communication between the phosphorylation site and binding interface. This sector is central to the dynamics of the stabilized CheY core.*

1. FliM$_N$ as an allosteric effector.

X-ray crystallography in concert with behavioral and biochemical studies has built a valuable mechanistic framework based on visual inspection of structural landmarks, guided by chemical intuition. Examination of the native CheY·FliM$_N$ crystal structure led to the proposal that the complex was an intermediate between active and inactive state consistent with a flexible β4-α4 loop (31). The structure challenged existing two-state switch models; but puzzlingly the central element in the models, the Y106 rotamer state, was not in an intermediate conformation but the activated rotamer state and the decrease in FliM$_N$ affinity relative to the activated complex was difficult to discern. These issues have been resolved by the MD simulations and XFMS measurements reported in this study.

The CheY·FliM$_N$ conformational landscape generated by MD simulations of the reverse-engineered 1U8T_DY structure had prominent minima that reflected intermediate FliM$_N$ attachment entrained to Y106 rotation states that ranged between the dominant OUT state in free CheY and the W106 IN state in activated CheY*·FliM$_N$. XFMS determined solvent accessibility values for the CheY·FliM$_N$ allosteric relay sidechains that were intermediate between values obtained for inactive CheY and active CheY*·FliM$_N$. These values were correlated with the protection of the interfacial lysine residues that monitored FliM$_N$ attachment. The D13K-Y106W residue substitutions as seen in the crystal structures did not alter the CheY fold to any significant extent

808 in the absence of $FliM_N$; a result supported in this study by both simulation
809 and measurement. The MD clarified that $FliM_N$ stabilized CheY and
810 strengthened allosteric communication between its binding interface and the
811 D57 phosphorylation site due to formation, in part, of the CheY.K91-$FliM_N$.D3
812 salt-bridge. The salt-bridge decreased the flexibility of the β4-α4 hinge,
813 consistent with earlier studies (31, 41).
814
815 **2. The dynamics and energetics of activation.**
816
817 This study documents a broad, high-energy CheY conformational
818 landscape with shallow minima consistent with the high conformational
819 plasticity suggested by the CheY crystal structures and early MD studies
820 (Introduction). Network analysis, based on mutual information between short
821 protein fragments established that two loops (β3-α3, β4-α4) act as flexible
822 hinges to control the dynamics. The CheY MD trajectories revealed episodes
823 where the Y106 sidechain is buried (IN), but cluster analysis determined the
824 inward motions were too brief to influence backbone dynamics in contrast to
825 the case for CheY·$FliM_N$. The buried states of the Y106 sidechain have not
826 been visualized to our knowledge in inactive CheY crystal structures.
827 The CheY conformations of the major CheY·$FliM_N$ clusters were more
828 stable than the dominant CheY conformations reported by the MD or the
829 conformation in the 2B1J crystal structure. The lifetimes of the CheY Y106 IN
830 states in CheY·$FliM_N$ were substantially greater than in free CheY and
831 represented in the major clusters, There was a weak correlation between the
832 stability of the CheY fold, the $FliM_N$ interface and the position of the Y106
833 sidechain. The CheY ΔG values in the major CheY·$FliM_N$ clusters overlapped
834 with the values in the inactive CheY and activated CheY*·$FliM_N$ clusters.
835 The mean CheY*·$FliM_N$ ΔG value was more stable than for CheY·$FliM_N$.
836 This was also the case for the interfacial ΔG values. The position of the W106
837 sidechain was restricted to a narrow range. Nevertheless, the ΔG values for
838 both the CheY fold and its $FliM_N$ interface, as well as the rotamer position of
839 the W/Y106 sidechain, were similar for the dominant CheY*·$FliM_N$ and
840 CheY*·$FliM_N$ conformational clusters. The similarity may explain capture
841 during crystallization of the Y106 sidechain in the 2B1J structure in a position
842 superimposable with the W106 sidechain in the 1U8T structure. The better
843 correlation of W106 sidechain position, in the MD clusters and the 1U8T
844 structure, with the CheY fold and $FliM_N$ interface ΔG values, reflects the tight-
845 packing due to the D13K-Y106W substitutions. The ΔG and W106 rotation
846 angles of the CheY*·$FliM_N$ clusters had no overlap with values for CheY
847 clusters.
848 Allosteric communication may range from largely enthalpic, as in
849 lysozyme, to largely entropic with change in flexibility rather than shape
850 (68). Both energy terms contribute to CheY allosteric activation. CheY
851 activation has aspects that “invoke conformational selection”, namely the
852 selection of the global minimum from the multiple minima sampled by the

853 native CheY·$FliM_N$ conformational ensemble by the D13K-Y106W residue
854 substitutions. Other aspects, such as the formation of the allosteric relay
855 based on local changes in the loop and sidechain rotamer dynamics
856 triggered by $FliM_N$ attachment support "induced fit". Neither description is
857 complete.

858

859 **3. Community networks – a new measure for response regulator**
860 **signal transduction.**

861

862 It has long been recognized that two-state allosteric models have
863 heuristic value but that an analytical description is desirable (32). Many
864 conformational states, as suggested (69), may be essential to explain how
865 subtle changes in CheY sequence trigger diverse motile responses. In *B.*
866 *subtilis*, for example, CheY~P stimulates CCW rather than CW rotation in
867 contrast to *E. coli*, but remains critical for chemotaxis (70). In *Thermatoga*
868 *maritima,* the middle domain of FliM ($FliM_M$) could be the second-stage CheY
869 binding target *(16)*, The diverse sensory responses triggered by CheY
870 homologs even within one species (e.g. *Caulobacter crescentus* (71)), as well
871 as the variable signal transduction strategies employed by response
872 regulators (1), emphasize the need for a more complete description.
873 Community networks have been used previously (65) to identify jointly
874 moving regions that do not track backbone secondary structure but are
875 governed instead by side-chain motions. This work is the first application of
876 this approach to the response regulator superfamily.

877

878 Distinct protein sectors with correlated motions were identified in
879 community networks. The extensive library of CheY residue substitutions was
880 exploited for functional assignment of the sectors. Two sectors, namely the
881 neighborhood of the phosphorylation site (sector A) and the region of $FliM_N$
882 binding (sector B) had clear functional importance. Two other sectors lacked
883 strong, specific phenotypes and might have broader functions in maintaining
884 the overall CheY fold. The long β3-α3 loop influenced movements of the β3
885 strand that formed a sector junction, consistent with its central role in the
886 reported PC motions. Similar motions take place in other proteins that utilize
887 β-sheets for signal transduction (72).

888 $FliM_N$ attachment increased the size of sectors A and B in the CheY
889 community network. The CheY*·$FliM_N$ community network was distinguished
890 from the CheY and CheY·$FliM_N$ networks by a fifth sector (E), drawn from
891 sectors A and B, that formed a dedicated conduit between the
892 phosphorylation and $FliM_N$-binding sites to cement the allosteric linkage, with
893 the substituted residues K13 and W106 at its boundaries. The emergence of
894 sector E was tied to the closure of the β4-α4 hinge by the CheY.K91-$FliM_N$.D3
895 salt-bridge and "freeze-out" of W_{106}VVKP β-strand dynamics by the burial of
896 aromatic residues for a tightly packed core. This sector connects with all

897 other sectors and has a large surface profile. It may directly or indirectly
898 define a region important for binding to FliN.
899 The CheY* protein is impaired in its interactions for other chemotaxis
900 proteins, the CheA kinase and CheZ (6), that have CheY binding surfaces
901 that overlap with that for $FliM_N$ (69, 73). Sector E may also influence the
902 regulation of phosphorylation by these proteins. An important future goal
903 would be to apply the integrated approach presented here to detect how
904 CheY~P discriminates between these components of the chemotaxis circuity.
905
906 Rotamer reorientation of aromatic sidechains is a common theme in
907 phospho-proteins, but diverse strategies for coupling side-chain motions to
908 phosphorylation exist. In eukaryotic protein kinases, activation is controlled
909 by DFG motif loops. These loops take on multiple IN and OUT orientations,
910 with orientation correlated with activation. In Aurora kinase A,
911 phosphorylation triggers transition between distinct IN orientations, rather
912 than between IN and OUT states (74). In calcium calmodulin-dependent
913 kinase, IN and OUT DFG states are loosely coupled to kinase domain
914 phosphorylation (75). in CheY XFMS reported $D_{57}WN_{59}$ internalization was
915 coupled to protection at the $FliM_N$ interface. We envisage that XFMS will have
916 applications in other phospho-relays given ongoing developments in MS
917 sensitivity and high-throughput analyses since most amino acids are
918 modified by hydroxy radicals to a greater or lesser extent
919
920 The sparse sampling by crystal structures may miss high-energy states, such
921 as the intermediate states of the CheY 106 sidechain, that are important for
922 deciphering mechanism. MD simulations provide a much more detailed
923 sampling of the conformational landscape, but their challenge is to extract
924 the essential features from the large conformational ensembles obtained; a
925 challenge only partially met by standard PCA and RMSF analyses. Our study
926 shows that community maps provide a concise, comprehensive description
927 based on quantitative criteria for identification of the key features of CheY
928 allosteric activation. They could provide the optimal compromise for
929 mechanistic dissection of signal transduction strategies in the response
930 regulator superfamily.

931

932 **SUPPORTING CITATIONS:**

933 References (76-87) appear in the **Supporting Material**.

934 **AUTHOR CONTRIBUTIONS:**

935 **Paige Wheatley:** Investigation, Visualization, Writing - Review & Editing
936 **Sayan Gupta:** Methodology, Investigation
937 **Alessandro Pandini:** Conceptualization, Software, Formal analysis,
938 Validation, Visualization
939 **Yan Chen:** Methodology, Investigation, Formal analysis, Validation

940 **Christopher J Petzold:** Methodology, Supervision, Funding Acquisition,
941 Writing - Review & Editing
942 **Corie Y. Ralston:** Conceptualization, Supervision, Funding Acquisition,
943 Writing - Review & Editing
944 **David F. Blair:** Conceptualization. Supervision, Funding Acquisition, Writing -
945 Draft, Review & Editing
946 **Shahid Khan:** Conceptualization, Formal analysis, Writing - Draft, Review &
947 Editing, Project Administration
948

949 **ACKNOWLEDGEMENTS:**
950 We thank Dr. Robert Bourret for comments on the manuscript and Martin
951 Horvath for assistance with the FPLC analysis of CheY proteins. This study was
952 supported by National Institutes of Health grants 1R01GM126218 (to C.Y.R) and
953 R01GM46683 (to D.F.B.). The XFMS was conducted at the Advanced Light Source
954 beamline 3.2.1 and the Joint BioEnergy Institute, supported by the Office of Science,
955 Office of Biological and Environmental Research, of the U.S. DOE under contract DE-
956 AC02-05CH11231. The MD simulations described in this paper were executed on the
957 Crick Data Analysis and Management Platform (CAMP), provided by the Francis
958 Crick Institute. Other computations utilized the Molecular Biology Consortium
959 computer cluster.
960

961
REFERENCES

962 1. Gao, R., S. Bouillet, and A. M. Stock. 2019. Structural Basis of Response
963 Regulator Function. Annual review of microbiology 73:175-197.
964 2. Galperin, M. Y. 2010. Diversity of structure and function of response regulator
965 output domains. Current opinion in microbiology 13(2):150-159.
966 3. Schroeder, C. M., J. M. Ostrem, N. T. Hertz, and R. D. Vale. 2014. A Ras-like
967 domain in the light intermediate chain bridges the dynein motor to a cargo-
968 binding region. eLife 3:e03351.
969 4. Bourret, R. B. 2010. Receiver domain structure and function in response
970 regulator proteins. Current opinion in microbiology 13(2):142-149.
971 5. Lipkow, K. 2006. Changing cellular location of CheZ predicted by molecular
972 simulations. Plos Comput Biol 2(4):e39.
973 6. Terasawa, S., H. Fukuoka, Y. Inoue, T. Sagawa, H. Takahashi, and A. Ishijima.
974 2011. Coordinated reversal of flagellar motors on a single Escherichia coli
975 cell. Biophysical journal 100(9):2193-2200.
976 7. Bren, A., and M. Eisenbach. 2000. How signals are heard during bacterial
977 chemotaxis: protein-protein interactions in sensory signal propagation.
978 Journal of bacteriology 182(24):6865-6873.
979 8. Cluzel, P., M. Surette, and S. Leibler. 2000. An ultrasensitive bacterial motor
980 revealed by monitoring signaling proteins in single cells. Science (New York,
981 N.Y 287(5458):1652-1655.
982 9. Thomas, D. R., N. R. Francis, C. Xu, and D. J. DeRosier. 2006. The three-
983 dimensional structure of the flagellar rotor from a clockwise-locked mutant of
984 Salmonella enterica serovar Typhimurium. Journal of bacteriology
985 188(20):7039-7048.
986 10. Fukuoka, H., T. Sagawa, Y. Inoue, H. Takahashi, and A. Ishijima. 2014. Direct
987 imaging of intracellular signaling components that regulate bacterial
988 chemotaxis. Science signaling 7(319):ra32.
989 11. Sagawa, T., Y. Kikuchi, Y. Inoue, H. Takahashi, T. Muraoka, K. Kinbara, A.
990 Ishijima, and H. Fukuoka. 2014. Single-cell E. coli response to an
991 instantaneously applied chemotactic signal. Biophysical journal 107(3):730-
992 739.
993 12. Ganguli, S., H. Wang, P. Matsumura, and K. Volz. 1995. Uncoupled
994 phosphorylation and activation in bacterial chemotaxis. The 2.1-A structure of
995 a threonine to isoleucine mutant at position 87 of CheY. The Journal of
996 biological chemistry 270(29):17386-17393.
997 13. McEvoy, M. M., A. Bren, M. Eisenbach, and F. W. Dahlquist. 1999.
998 Identification of the binding interfaces on CheY for two of its targets, the
999 phosphatase CheZ and the flagellar switch protein fliM. Journal of molecular
1000 biology 289(5):1423-1433.
1001 14. Schuster, M., R. Zhao, R. B. Bourret, and E. J. Collins. 2000. Correlated switch
1002 binding and signaling in bacterial chemotaxis. The Journal of biological
1003 chemistry 275(26):19752-19758.
1004 15. Sagi, Y., S. Khan, and M. Eisenbach. 2003. Binding of the chemotaxis
1005 response regulator CheY to the isolated, intact switch complex of the
1006 bacterial flagellar motor: lack of cooperativity. The Journal of biological
1007 chemistry 278(28):25867-25871.

1008 16. Dyer, C. M., A. S. Vartanian, H. Zhou, and F. W. Dahlquist. 2009. A molecular
1009 mechanism of bacterial flagellar motor switching. Journal of molecular biology
1010 388(1):71-84.
1011 17. Sarkar, M. K., K. Paul, and D. Blair. 2010. Chemotaxis signaling protein CheY
1012 binds to the rotor protein FliN to control the direction of flagellar rotation in
1013 Escherichia coli. Proceedings of the National Academy of Sciences of the
1014 United States of America 107(20):9370-9375.
1015 18. Lee, S. Y., H. S. Cho, J. G. Pelton, D. Yan, R. K. Henderson, D. S. King, L.
1016 Huang, S. Kustu, E. A. Berry, and D. E. Wemmer. 2001. Crystal structure of an
1017 activated response regulator bound to its target. Nature structural biology
1018 8(1):52-56.
1019 19. Halkides, C. J., M. M. McEvoy, E. Casper, P. Matsumura, K. Volz, and F. W.
1020 Dahlquist. 2000. The 1.9 A resolution crystal structure of phosphono-CheY, an
1021 analogue of the active form of the response regulator, CheY. Biochemistry
1022 39(18):5280-5286.
1023 20. Schuster, M., W. N. Abouhamad, R. E. Silversmith, and R. B. Bourret. 1998.
1024 Chemotactic response regulator mutant CheY95IV exhibits enhanced binding
1025 to the flagellar switch and phosphorylation-dependent constitutive signalling.
1026 Molecular microbiology 27(5):1065-1075.
1027 21. Zhu, X., J. Rebello, P. Matsumura, and K. Volz. 1997. Crystal structures of
1028 CheY mutants Y106W and T87I/Y106W. CheY activation correlates with
1029 movement of residue 106. The Journal of biological chemistry 272(8):5000-
1030 5006.
1031 22. Jiang, M., R. B. Bourret, M. I. Simon, and K. Volz. 1997. Uncoupled
1032 phosphorylation and activation in bacterial chemotaxis. The 2.3 A structure of
1033 an aspartate to lysine mutant at position 13 of CheY. The Journal of biological
1034 chemistry 272(18):11850-11855.
1035 23. Dyer, C. M., M. L. Quillin, A. Campos, J. Lu, M. M. McEvoy, A. C. Hausrath, E. M.
1036 Westbrook, P. Matsumura, B. W. Matthews, and F. W. Dahlquist. 2004.
1037 Structure of the constitutively active double mutant CheYD13K Y106W alone
1038 and in complex with a FliM peptide. Journal of molecular biology 342(4):1325-
1039 1335.
1040 24. Lukat, G. S., B. H. Lee, J. M. Mottonen, A. M. Stock, and J. B. Stock. 1991.
1041 Roles of the highly conserved aspartate and lysine residues in the response
1042 regulator of bacterial chemotaxis. The Journal of biological chemistry
1043 266(13):8348-8354.
1044 25. Bourret, R. B., J. F. Hess, and M. I. Simon. 1990. Conserved aspartate residues
1045 and phosphorylation in signal transduction by the chemotaxis protein CheY.
1046 Proceedings of the National Academy of Sciences of the United States of
1047 America 87(1):41-45.
1048 26. Appleby, J. L., and R. B. Bourret. 1998. Proposed signal transduction role for
1049 conserved CheY residue Thr87, a member of the response regulator active-
1050 site quintet. Journal of bacteriology 180(14):3563-3569.
1051 27. Immormino, R. M., C. A. Starbird, R. E. Silversmith, and R. B. Bourret. 2015.
1052 Probing Mechanistic Similarities between Response Regulator Signaling
1053 Proteins and Haloacid Dehalogenase Phosphatases. Biochemistry
1054 54(22):3514-3527.
1055 28. Turner, L., A. D. Samuel, A. S. Stern, and H. C. Berg. 1999. Temperature
1056 dependence of switching of the bacterial flagellar motor by the protein
1057 CheY(13DK106YW). Biophysical journal 77(1):597-603.

1058 29. Korobkova, E. A., T. Emonet, H. Park, and P. Cluzel. 2006. Hidden stochastic
1059 nature of a single bacterial motor. Phys Rev Lett 96(5):058105.
1060 30. Wang, F., H. Shi, R. He, R. Wang, R. Zhang, and J. Yuan. 2017. Non-
1061 equilibrium effects in the allosteric regulation of the bacterial flagellar switch.
1062 Nature Physics 13:710-714.
1063 31. Dyer, C. M., and F. W. Dahlquist. 2006. Switched or not?: the structure of
1064 unphosphorylated CheY bound to the N terminus of FliM. Journal of
1065 bacteriology 188(21):7354-7363.
1066 32. Stock, A. M., and J. Guhaniyogi. 2006. A new perspective on response
1067 regulator activation. Journal of bacteriology 188(21):7328-7330.
1068 33. McDonald, L. R., J. A. Boyer, and A. L. Lee. 2012. Segmental motions, not a
1069 two-state concerted switch, underlie allostery in CheY. Structure 20(8):1363-
1070 1373.
1071 34. Cho, K. H., B. R. Crane, and S. Park. 2011. An insight into the interaction
1072 mode between CheB and chemoreceptor from two crystal structures of CheB
1073 methylesterase catalytic domain. Biochemical and biophysical research
1074 communications 411(1):69-75.
1075 35. Luo, J., Z. Liu, Y. Guo, and M. Li. 2015. A structural dissection of large protein-
1076 protein crystal packing contacts. Sci Rep 5:14214.
1077 36. Hammes, G. G., Y. C. Chang, and T. G. Oas. 2009. Conformational selection or
1078 induced fit: a flux description of reaction mechanism. Proceedings of the
1079 National Academy of Sciences of the United States of America
1080 106(33):13737-13741.
1081 37. Changeux, J. P. 2013. 50 years of allosteric interactions: the twists and turns
1082 of the models. Nat Rev Mol Cell Biol 14(12):819-829.
1083 38. Tsai, C. J., and R. Nussinov. 2014. A unified view of "how allostery works".
1084 Plos Comput Biol 10(2):e1003394.
1085 39. Kumar, S., B. Ma, C. J. Tsai, H. Wolfson, and R. Nussinov. 1999. Folding
1086 funnels and conformational transitions via hinge-bending motions. Cell
1087 Biochem Biophys 31(2):141-164.
1088 40. Ma, L., and Q. Cui. 2007. Activation mechanism of a signaling protein at
1089 atomic resolution from advanced computations. Journal of the American
1090 Chemical Society 129(33):10261-10268.
1091 41. Fraiberg, M., O. Afanzar, C. K. Cassidy, A. Gabashvili, K. Schulten, Y. Levin,
1092 and M. Eisenbach. 2015. CheY's acetylation sites responsible for generating
1093 clockwise flagellar rotation in Escherichia coli. Molecular microbiology
1094 95(2):231-244.
1095 42. Pandini, A., A. Fornili, F. Fraternali, and J. Kleinjung. 2012. Detection of
1096 allosteric signal transmission by information-theoretic analysis of protein
1097 dynamics. FASEB journal : official publication of the Federation of American
1098 Societies for Experimental Biology 26(2):868-881.
1099 43. Foster, C. A., and A. H. West. 2017. Use of restrained molecular dynamics to
1100 predict the conformations of phosphorylated receiver domains in two-
1101 component signaling systems. Proteins 85(1):155-176.
1102 44. Gupta, S., M. Guttman, R. L. Leverenz, K. Zhumadilova, E. G. Pawlowski, C. J.
1103 Petzold, K. K. Lee, C. Y. Ralston, and C. A. Kerfeld. 2015. Local and global
1104 structural drivers for the photoactivation of the orange carotenoid protein.
1105 Proceedings of the National Academy of Sciences of the United States of
1106 America 112(41):E5567-5574.

1107 45. Gupta, S., M. Sutter, S. G. Remesh, M. A. Dominguez-Martin, H. Bao, X. A.
1108 Feng, L. G. Chan, C. J. Petzold, C. A. Kerfeld, and C. Y. Ralston. 2019. X-ray
1109 radiolytic labeling reveals the molecular basis of orange carotenoid protein
1110 photoprotection and its interactions with fluorescence recovery protein. The
1111 Journal of biological chemistry 294(22):8848-8860.
1112 46. Volz, K., and P. Matsumura. 1991. Crystal structure of Escherichia coli CheY
1113 refined at 1.7-A resolution. The Journal of biological chemistry 266(23):15511-
1114 15519.
1115 47. Aliev, A. E., M. Kulke, H. S. Khaneja, V. Chudasama, T. D. Sheppard, and R. M.
1116 Lanigan. 2014. Motional timescale predictions by molecular dynamics
1117 simulations: case study using proline and hydroxyproline sidechain dynamics.
1118 Proteins 82(2):195-215.
1119 48. Skjaerven, L., A. Martinez, and N. Reuter. 2011. Principal component and
1120 normal mode analysis of proteins; a quantitative comparison using the GroEL
1121 subunit. Proteins 79(1):232-243.
1122 49. Olsson, M. H., C. R. Sondergaard, M. Rostkowski, and J. H. Jensen. 2011.
1123 PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical
1124 pKa Predictions. Journal of chemical theory and computation 7(2):525-537.
1125 50. Yang, A. S., and B. Honig. 1993. On the pH dependence of protein stability.
1126 Journal of molecular biology 231(2):459-474.
1127 51. Seeliger, D., J. Haas, and B. L. de Groot. 2007. Geometry-based sampling of
1128 conformational transitions in proteins. Structure 15(11):1482-1492.
1129 52. de Groot, B. L., D. M. van Aalten, R. M. Scheek, A. Amadei, G. Vriend, and H. J.
1130 Berendsen. 1997. Prediction of protein conformational freedom from distance
1131 constraints. Proteins 29(2):240-251.
1132 53. Pandini, A., J. Kleinjung, W. R. Taylor, W. Junge, and S. Khan. 2015. The
1133 Phylogenetic Signature Underlying ATP Synthase c-Ring Compliance.
1134 Biophysical journal 109(5):975-987.
1135 54. Pandini, A., A. Fornili, and J. Kleinjung. 2010. Structural alphabets derived
1136 from attractors in conformational space. BMC bioinformatics 11:97.
1137 55. Pandini, A., A. Fornili, F. Fraternali, and J. Kleinjung. 2013. GSATools: analysis
1138 of allosteric communication and functional local motions using a structural
1139 alphabet. Bioinformatics (Oxford, England) 29(16):2053-2055.
1140 56. Newman, M. E. 2006. Modularity and community structure in networks. Proc.
1141 Natl. Acad. Sci. U. S. A. 103(23):8577-8582.
1142 57. Newman, M. E. 2004. Analysis of weighted networks. Phys Rev E Stat Nonlin
1143 Soft Matter Phys 70(5 Pt 2):056131.
1144 58. Csardi, G., and T. Nepusz. 2006. The igraph software package for complex
1145 network research. InterJournal Complex Systems:1695.
1146 59. Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. 2015. The I-TASSER
1147 Suite: protein structure and function prediction. Nature methods 12(1):7-8.
1148 60. Gupta, S., M. Sullivan, J. Toomey, J. Kiselar, and M. R. Chance. 2007. The
1149 Beamline X28C of the Center for Synchrotron Biosciences: a national resource
1150 for biomolecular structure and dynamics experiments using synchrotron
1151 footprinting. J Synchrotron Radiat 14(Pt 3):233-243.
1152 61. Bohon, J., R. D'Mello, C. Ralston, S. Gupta, and M. R. Chance. 2014.
1153 Synchrotron X-ray footprinting on tour. J Synchrotron Radiat 21(Pt 1):24-31.
1154 62. Huang, W., K. M. Ravikumar, M. R. Chance, and S. Yang. 2015. Quantitative
1155 mapping of protein structure by hydroxyl radical footprinting-mediated

1156 structural mass spectrometry: a protection factor analysis. Biophysical journal
1157 108(1):107-115.
1158 63. Davies, M. J., Dean, R.T. 1997. Radical-Mediated Protein Oxidation: From
1159 Chemistry to Medicine. Oxford University Press.
1160 64. Kornev, A. P., and S. S. Taylor. 2015. Dynamics-Driven Allostery in Protein
1161 Kinases. Trends Biochem. Sci. 40(11):628-647.
1162 65. McClendon, C. L., A. P. Kornev, M. K. Gilson, and S. S. Taylor. 2014. Dynamic
1163 architecture of a protein kinase. Proceedings of the National Academy of
1164 Sciences of the United States of America 111(43):E4623-4631.
1165 66. Roman, S. J., M. Meyers, K. Volz, and P. Matsumura. 1992. A chemotactic
1166 signaling surface on CheY defined by suppressors of flagellar switch
1167 mutations. Journal of bacteriology 174(19):6247-6255.
1168 67. Sanna, M. G., R. V. Swanson, R. B. Bourret, and M. I. Simon. 1995. Mutations
1169 in the chemotactic response regulator, CheY, that confer resistance to the
1170 phosphatase activity of CheZ. Molecular microbiology 15(6):1069-1079.
1171 68. Tsai, C. J., A. del Sol, and R. Nussinov. 2008. Allostery: absence of a change in
1172 shape does not imply that allostery is not at play. J. Mol. Biol. 378(1):1-11.
1173 69. Schuster, M., R. E. Silversmith, and R. B. Bourret. 2001. Conformational
1174 coupling in the chemotaxis response regulator CheY. Proceedings of the
1175 National Academy of Sciences of the United States of America 98(11):6003-
1176 6008.
1177 70. Ward, E., E. A. Kim, J. Panushka, T. Botelho, T. Meyer, D. B. Kearns, G. Ordal,
1178 and D. F. Blair. 2019. Organization of the Flagellar Switch Complex of Bacillus
1179 subtilis. Journal of bacteriology 201(8).
1180 71. Nesper, J., I. Hug, S. Kato, C. S. Hee, J. M. Habazettl, P. Manfredi, S. Grzesiek,
1181 T. Schirmer, T. Emonet, and U. Jenal. 2017. Cyclic di-GMP differentially tunes
1182 a bacterial flagellar motor through a novel class of CheY-like regulators. eLife
1183 6.
1184 72. Fenwick, R. B., L. Orellana, S. Esteban-Martin, M. Orozco, and X. Salvatella.
1185 2014. Correlated motions are a fundamental property of beta-sheets. Nat
1186 Commun 5:4070.
1187 73. Zhu, X., K. Volz, and P. Matsumura. 1997. The CheZ-binding surface of CheY
1188 overlaps the CheA- and FliM-binding surfaces. The Journal of biological
1189 chemistry 272(38):23758-23764.
1190 74. Ruff, E. F., J. M. Muretta, A. R. Thompson, E. W. Lake, S. Cyphers, S. K.
1191 Albanese, S. M. Hanson, J. M. Behr, D. D. Thomas, J. D. Chodera, and N. M.
1192 Levinson. 2018. A dynamic mechanism for allosteric activation of Aurora
1193 kinase A by activation loop phosphorylation. eLife 7.
1194 75. Pandini, A., H. Schulman, and S. Khan. 2019. Conformational coupling by
1195 trans-phosphorylation in calcium calmodulin dependent kinase II. PLoS
1196 Comput Biol 15(5):e1006796.
1197 76. Scharf, B. E., K. A. Fahrner, and H. C. Berg. 1998. CheZ has no effect on
1198 flagellar motors activated by CheY13DK106YW. Journal of bacteriology
1199 180(19):5123-5128.
1200 77. Scharf, B. E., K. A. Fahrner, L. Turner, and H. C. Berg. 1998. Control of
1201 direction of flagellar rotation in bacterial chemotaxis. Proceedings of the
1202 National Academy of Sciences of the United States of America 95(1):201-206.
1203 78. Cavallo, L., J. Kleinjung, and F. Fraternali. 2003. POPS: A fast algorithm for
1204 solvent accessible surface areas at atomic and residue level. Nucleic acids
1205 research 31(13):3364-3366.

1206 79. Fornili, A., A. Pandini, H. C. Lu, and F. Fraternali. 2013. Specialized Dynamical
1207 Properties of Promiscuous Residues Revealed by Simulated Conformational
1208 Ensembles. Journal of chemical theory and computation 9(11):5127-5147.
1209 80. Motta, S., C. Minici, D. Corrada, L. Bonati, and A. Pandini. 2018. Ligand-
1210 induced perturbation of the HIF-2alpha:ARNT dimer dynamics. Plos Comput
1211 Biol 14(2):e1006021.
1212 81. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through
1213 velocity scaling. J. Chem. Phys 126:014101.
1214 82. Parrinello, M., and S. Rahman. 1981. Polymorphic transitions in single crystals
1215 A new molecular dynamics method. J. Appl. Phys 52:7182-7190.
1216 83. Pandini, A., F. Morcos, and S. Khan. 2016. The Gearbox of the Bacterial
1217 Flagellar Motor Switch. Structure 24(7):1209-1220.
1218 84. Roulston, M. S. 1999. Estimating the errors on measured entropy and mutual
1219 information. Physica D 125:285-294.
1220 85. Gonzalez Fernandez-Nino, S. M., A. M. Smith-Moritz, L. J. Chan, P. D. Adams, J.
1221 L. Heazlewood, and C. J. Petzold. 2015. Standard flow liquid chromatography
1222 for shotgun proteomics in bioenergy research. Front Bioeng Biotechnol 3:44.
1223 86. Perez-Riverol, Y., A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J.
1224 Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Perez, J. Uszkoreit, J.
1225 Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A.
1226 F. Jarnuczak, T. Ternent, A. Brazma, and J. A. Vizcaino. 2019. The PRIDE
1227 database and related tools and resources in 2019: improving support for
1228 quantification data. Nucleic acids research 47(D1):D442-D450.
1229 87. Eyal, E., R. Najmanovich, B. J. McConkey, M. Edelman, and V. Sobolev. 2004.
1230 Importance of solvent accessibility and contact surfaces in modeling side-
1231 chain conformations in proteins. Journal of computational chemistry
1232 25(5):712-724.

1233