Title

Allosteric Priming of E. coli CheY by the Flagellar Motor Protein FliM

Journal

Biophysical Journal, 119(6)

ISSN

0006-3495

Authors

Wheatley, Paige

Gupta, Sayan

Pandini, Alessandro

et al.

Publication Date

2020-09-01

DOI

10.1016/j.bpj.2020.08.009

Peer reviewed

Final
08/10/2020

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

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

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

[*]*Corresponding Author*

Abstract

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

Statement of Significance

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

Introduction

Escherichia coli CheY is a founding member of a bacterial response regulator superfamily that uses aspartate phosphorylation to regulate diverse signal relays (1, 2). The CheY $\beta_5\alpha_5$ fold has structural homology with small eukaryotic signal-transducing proteins (3). CheY phosphorylation couples the occupancy of the chemoreceptor patch to the motile response in bacterial chemotaxis. Previous studies of CheY have established it as a model for fundamental design principles in protein allostery (4). Here, we study *E. coli* CheY binding to the FliM N-terminal peptide ($FliM_N$) responsible for its initial interaction with the flagellar switch complex.

CheY, fused with green fluorescent protein (GFP), is both phosphorylated and dephosphorylated at the polar chemoreceptor patch, generating pulsatile fluctuations in intracellular phosphorylated CheY (CheY~P) level (5, 6). The CheY~P diffuses to the flagellar motor within the flagellar basal body, interacting with its C-ring (a.k.a. the switch complex), a multi-subunit assembly composed of the proteins FliG, FliM and FliN. In *E. coli*, the interaction increases clockwise {CW} rotation *(7)*. Single-cell measurements expressing GFP-CheY under conditions where CheY~P is the dominant form have shown that motor rotational bias has a sigmoidal dependence on CheY concentration (Hill coefficient > 10.5, K_D = 3 μM) (8), implying highly cooperative action of the captured CheY molecules switching flagellar rotation. More recently, GFP-CheY occupancy was estimated to be about 1/3 and < 1/10 of the 34 FliM subunits present per motor (9) for single CW and CCW rotating motors respectively (10). The occupancy and rotation state were coupled within the image time resolution (20 ms); while GFP-CheY motor dissociation times (70 ms) were faster than the response times to attractant stimuli (11). The single-motor kinetics also imply cooperative CheY-motor interactions.

Biochemical experiments coupled with mutagenesis, motility assays and X-ray crystal structures have established that CheY is phosphorylated at a single aspartate ($D57{\sim}PO_4$). The aspartyl phosphate is labile with a 22.8s half-life at ambient temperature (12). The affinity for the $FliM_N$ motor binding target of non-phosphorylated *E. coli* CheY (K_D = ca. 450 μM) is 15x weaker than for CheY~P as measured by fluorescence quenching of CheY residue W58 adjacent to D57 (13, 14)). The binding of CheY~P to isolated, native CCW-locked flagellar switch complexes had K_D stronger than that for $FliM_N$, but was non-cooperative (15), in contrast to the *in-vivo* measurements of rotation bias (8) or motor localization (10) that sample both rotation states (**Supporting Information Section A. Table S1**). The conundrum how cooperative responses arise by CheY~P binding to $FliM_N$ alone is increased by the fact that $FliM_N$ is separated from the rest of the C-ring by a flexible tether (16). Thus, evidence that CheY interaction with the switch involves two binding sites, initial interaction with $FliM_N$, followed by a subsequent

interaction of the $FliM_N$ tethered CheY to FliN in *E. coli* (17) provides a plausible resolution. It has remained unclear whether the $FliM_N$ tether facilitates the second-stage binding step only by increasing CheY local concentration, or whether structural changes also occur that prime CheY to bind FliN (**Figure 1**).

Figure 1: CheY interactions with the flagellar motor. *CheY shade intensity and size denote activation state and FliN binding probability respectively. Binding of activated CheY (CheY*) to isolated switch complexes is not cooperative (H = 1), but the change in flagellar CW/CCW rotation bias is highly cooperative (H > 10) with CheY* concentration. 1st stage (1) binding to $FliM_N$ enables 2nd stage (2) binding to FliN. The increased local concentration due to 1st stage binding and the multiple FliN copies enhance 2nd stage binding probability. The Inactive CheY binds weakly, reducing $FliM_N$-tethered CheY binding events with FliN below the critical threshold for CW rotation. This study provides evidence for structural changes in CheY that may supplement increased local concentration for 2nd stage binding.*

An atomic structure for CheY~P is not available given the lability of the aspartyl phosphate. Therefore, atomic structures of phospho-mimetic CheY proteins obtained by chemical modification (18, 19) or mutagenesis (20-23) have been used to reconstruct the activation mechanism. While both chemically modified and mutated proteins are used in vitro biochemical assays. only the latter can be expressed and studied in vivo (12, 20-22, 24-26). The activating substitutions D13K, Y106W are the most potent modulators reported, thus far of $FliM_N$ binding in vitro (13, 14, 27) and motor rotation bias in vivo (10, 28-30). The comparison of CheYD13KY106W efficacy with CheY~P from both in vitro and in vivo assays, the substantial knowledge of its effects on motor physiology and the availability of atomic structures with and without $FliM_N$ (Supporting Information Table S1) make CheYD13KY106W the logical first choice for the elucidation of the molecular priming mechanism.

The crystal structures of D13K-Y106W CheY alone and in complex with $FliM_N$ showed bound $FliM_N$ was required for the activated CheY conformation. They established CheY residues K91, Y106 and K119 as part of the $FliM_N$, binding surface (23). K91 and K119 formed salt bridges with $FliM_N$. The W106 sidechain moved in as $FliM_N$ bound to switch K109 bonding interactions with T87, D57 and, via bound water, with D12 (23). The structure of the native CheY·$FliM_N$ complex exhibited some features of inactive CheY and some features of the active D13K-Y106W CheY·$FliM_N$ (31). Notably, the orientation of the Y106 sidechain matched that for W106 in the D13K-Y106W CheY·$FliM_N$ complex. The “intermediate” conformation of the native CheY·$FliM_N$ structure challenged two-state CheY allostery models that coupled Y/W106 rotamer state to T87 motions (32). An NMR study on free CheY (33) reached a similar conclusion. CheY has high conformational plasticity as seen by the

discrepancies between crystal structures of activated CheY proteins (19, 23, 34). The coverage of the conformational landscape by crystal structures is too sparse to resolve the conformational trajectories for activation by phosphorylation or binding targets such as $FliM_N$. Alteration of low-affinity binding interfaces, a common occurrence in signal-transducing phospho-protein complexes by crystal packing contacts Is an additional concern (35).

CheY conformational plasticity is not well-described by classical protein allostery concepts of “induced fit” (KNF) (36) or “conformational selection” (e.g. MWC (37)) but is accommodated by modern ideas of allostery (38) where protein-protein interactions between flexible partners have been described in terms of a folding funnel, where the funnel bottom has a “rugged” landscape with multiple minima (39). Accordingly, molecular dynamics (MD) simulations and solution measurements have supplemented the X-ray crystallography of free CheY structures. MD of free CheY examined the coupling between Y106 rotation and T87 movements triggered by hydrogen bond formation (40), showing that the $\beta4$-$\alpha4$ loop is an important determinant of allosteric signaling affected by lysine acetylation (41) and extracted common design principles between CheY and other response regulators with correlation analyses (42, 43).

Here, we detail simulations and solution measurements to better understand the differences between the native and D13K-Y106W CheY crystal structures. We resolved the complex conformational landscapes by MD simulations with mutual information measures to determine the coupling between protein fragments. Protection experiments with XFMS (X-ray foot-printing with mass spectroscopy) (44, 45), a technique that probed sidechain solvent accessibility in contrast to deuterium exchange of backbone hydrogen atoms, supported the $FliM_N$ requirement for D13K-Y106W CheY activation reported by the crystal structures, and the MD allosteric network model. XFMS has a more straight-forward physical rationale than fluorescence quenching for reporting sidechain motions over time-resolved windows and is not limited by the size of the protein assembly. Further analysis of the MD trajectories resolved multiple CheY Y106 rotamer states. Inward orientation was temporally coupled to stabilization of both the CheY fold and the $FliM_N$ interface in the CheY·$FliM_N$ complex, but not in CheY alone. The coupling increased in D13K-Y106WCheY·$FliM_N$. The formation of a distinct module that orchestrates CheY dynamics to stabilize new surface topologies for possible second-stage binding to FliN was the signature of the fully activated D13K-Y106W CheY·$FliM_N$ state.

Materials & Methods

1. Structure Preparation.

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

2. Molecular Simulations.

(a) Molecular Dynamics.

A set of 3 replicas of duration 1 μs each was generated for the mutant (1U8T) and native (1U8T_DY) complexes using GROMACS 2016.2 with Amber ff99sb*-ILDNP force-field (47). Another set of 3 replicas of 500ns duration each was generated for the native CheY (3CHY). Each system was first solvated in an octahedral box with TIP3P water molecules with a minimal distance between protein and box boundaries of 12 Å. The box was then neutralized with Na^+ ions. Solvation and ion addition were performed with the GROMACS preparation tools (**Supporting Information Section B)**.

Collective motions were identified by PCA of the conformational ensembles. PCs were generated by diagonalization of the covariance matrix of C^α positions in GROMACS. The overlap (cumulative root mean square inner product) of the PCs between replicas (48)) and the PC dot product matrix was computed with the GROMACS g-anaeig function.

The conformational ensembles were clustered and mean structures representing the major clusters (n>5) computed with the GROMACS g-cluster function. The energy landscape was computed with PROPKA 3.0 (49). PROPKA calculates the free energy difference (ΔG) between the folded and unfolded states as the protein charge varies with pH (50). CheY has 37 ionizable groups (9D, 12E, 10K, 4R, 2Y) plus N and C termini that determine its net charge. The ΔG is computed from the perturbation of residue pK values by the protein environment; namely the dielectric-dependant de-solvation penalty, backbone and sidechain hydrogen bonds and interactions with other charged residues. For the complexes, the ΔG was computed for the complex (ΔG_T) as well as CheY alone with $FliM_N$ removed (ΔG_{CheY}). The ΔG_{CheY} was the free energy of the CheY fold. The interfacial energy $\Delta G_{interface} = \Delta G_T - \Delta G_{CheY}$

(b) tCONCOORD

tCONCOORD utilizes distance constraints based on the statistics of residue interactions in a crystal structure library (51, 52), to generate conformational ensembles from one crystal structure with solvent modelled as an implicit continuum. tCONCOORD runs compared conformational ensembles for native CheY

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

3. Network Analysis.

(a) Structural alphabet.

Coordinated CheY motions were examined using mutual-information analysis. The mutual-information (nMI) matrix encodes correlations between conformational states of different parts of the protein backbone (**Supporting Information Section B**). The states are represented by a structural alphabet (SA), a set of recurring four residue fragments encoding structural motifs derived from PDB structures (54). Fragments are assigned an SA designation according to backbone dihedral angles, allowing conformation to be specified as a 1-dimensional string (54). The fragments are represented as network nodes, with the connectivity (edges) between them representing their correlated dynamics over the MD trajectory.

(b) Eigenvector Analysis.

Statistically significant correlations between columns were identified with GSATools (55) and recorded as a correlation matrix. The correlation matrix was used to generate a network model with the residues as nodes and the correlations as edges. In vector notation, the overall connectivity of a given fragment is reported by its eigenvector centrality, *E* ("centrality"). The contribution of a node to the network was estimated by its *E*, calculated directly from the correlation adjacency matrix:

$$E.\{M\}_{corr} = E.\lambda$$

where the $\{M\}_{corr}$ is the correlation matrix. The λ is the eigenvalue.

The nMI contributions of local fragment motions were computed for the top PCs and superimposed on their RMSF profiles to evaluate the mechanical behavior of the network nodes in driving collective motions. Ensemble conformations and MD runs were averaged for computation of the nMI between fragment positions, with > 2σ thresholds for selected top couplings. Pearson's correlations were used for comparison. Significance limits were set in GSATools.

(c) Community Analysis.

The Girvan–Newman algorithm (56) was used to identify community structure. Then the network was collapsed into a simplified graph with one node per community, where the node size is proportional to the number of residues. Edge weights represent the number of nMI couplings between communities (57). Community analysis of correlation networks identifies relatively independent communities that behave as semi-rigid bodies. Graphs were constructed with the *igraph* library (58) in R (*https://cran.r-project.org/web/packages/igraph/*) and visualized in Cytoscape (*http://www.cytoscape.org/).*

4. Overexpression and purification of CheY proteins

The CheY-pET21b plasmids with *E. coli cheY* alone and fused with $FliM_N$ (17) were modified to incorporate the double mutation D13K, Y106W. The native and mutated plasmids were expressed in *E. coli* strain BL21/DE3. The expressed proteins were purified with fast protein liquid chromatography (FPLC) The $FliM_N$.CheY fusion interacts with FliN (17) and is more potent than CheY alone in potentiation of CW rotation (P. Wheatley, unpublished). 3D models of the $FliM_N$.CheY fusions were obtained with the I-Tasser suite (59). In all top five models, $FliM_N$ was docked in the location seen in the crystal structures of the CheY·$FliM_N$ complexes. The top model

had cs = -1.08, RMSF = 7.2+4.2 angstroms (against CheY, $FliM_N$ crystal structures) **Supporting Information Section C**.

5. X-Ray Foot-printing (XF)

Protein samples (CheY, $FliM_N$.CheY, CheY*and $FliM_N$CheY* were prepared in 10 mM potassium phosphate buffer (pH 7.2), 100 mM NaCl, and 10 mM $MgCl_2$. Exposure range was determined empirically by adding Alexa488 to protein solutions as previously described (60). Sample irradiation was conducted without Alexa488 dye using a microfluidic set-up with 100 mm and 200 mm ID tubing in combination with a syringe pump as previously described (61). After exposure at ALS beamline 3.2.1, samples were immediately quenched with methionine amide to stop the secondary oxidations and stored at -80 °C for LCMS analysis.

The oxidized fraction, F, for a single residue modification was given by the equation

$$F = \{ Xi / ¿$$

where Xi is the oxidized residue abundance of one of the monitored residues in a trypsinized peptide and T is the unoxidized peptide.

Best fit first-order rates were calculated in Sigmaplot version 12. Protection factors (PFs) were calculated as the ratio of the intrinsic residue reactivity over its foot-printing rate (62). Its logarithm (log (PF)) was proportional to the SASA. The relation assumes that the foot-printing rate was related to the activation energy associated with the accessibility of the side-chain to hydroxy radicals and the initial step of hydrogen abstraction It empirically gave the best-fit for proteolyzed peptides on a model data set, extended here to single residues (62).

6. Mass Spectrometry (MS) Analysis

X-Ray exposed protein samples were digested by Trypsin and the resulted peptide samples were analyzed in an Agilent 6550 iFunnel Q-TOF mass spectrometer coupled to an Agilent 1290 LC system (Agilent Technologies, Santa Clara, CA). Approximately 10 pmol of peptides were loaded onto the Ascentis Peptides ES-C18 column (2.1 mm x 100 mm, 2.7 µm particle size; Sigma- Aldrich, St. Louis, MO) at 0.400 mL/min flow rate and were eluted with the following gradient: initial conditions were 95% solvent A (0.1% formic acid), 5% solvent B (99.9% acetonitrile, 0.1% formic acid). Solvent B was increased to 35% over 5.5 min, and was then increased to 80% over 1 min, and held for 3.5 min at a flow rate of 0.6 mL/min, followed by a ramp back down to 5% B over 0.5 min where it was held for 2 min to re-equilibrate the column to original conditions. Peptides were introduced to the mass spectrometer from the LC using a Jet Stream source (Agilent Technologies) and spectra acquired with Agilent Mass Hunter Workstation Software B.06.01. The peptide precursor peak intensities were measured in Mass Hunter quantitative analysis software. Further details and data sets are given in **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^α RMSF) profile for each structure, averaged over three 1 μs runs, are shown in **Figure 2B.** The MD excluded the first three residues (M_1GD_3) of the $FliM_N$ sequence ($M_1GDSILSQAEIDALL_{16}$) 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 $\beta5$-$\alpha5$ 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 / $\beta5$-$\alpha5$ 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

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

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

2. Two loops control CheY network dynamics.

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

conformations in the ensemble. The nMI couplings constituted the “edges” of the CheY network with correlation strength denoted by the edge thickness. The fragment positions within the sequence alignment formed the “nodes” of the network. Third, the connectivity of the network was determined by *E*, a measure of the influence of individual nodes in the network as reflected by the coupling of their dynamics with other nodes in the nMI correlation matrix. A node *E* limit value of 0 represented the case where its dynamics did not affect other nodes in the network. An E limit of 1 represented the case where its conformational fluctuations switched the entire network between discrete structural states (see Materials & Methods for formal definitions).

First, we identified the central nodes in the CheY global network with the highest connectivity (**Figure 3B**). The central CheY nodes were the loops β3-α3 (D_{57}WNMPNMDG) and β4-α4 (T_{87}AEAKK). A third prominent node just below the 1σ threshold was the short β_5-strand (Y_{106}VVKP). Second, we used tCONCOORD, a computationally inexpensive method to generate conformational ensembles for comparison of the CheY and CheY* conformational landscapes.

Interpretation of differences between CheY and CheY* crystal structures based on isolated landmarks (for example, Y106 rotamer state (IN/OUT in CheY (3CHY.pdb) (46)) versus W106 (OUT in CheY*(23)) are complicated by the CheY conformational plasticity. Analysis of the tCONCOORD ensembles showed the central network nodes remained unchanged, with both CheY Y106 and CheY* W106 sidechains restricted to a limited OUT-orientation range (**Supporting Information Figure S2**).

We next examined the CheY and CheY* complexes with $FliM_N$ (**Figure 3C**). We split the CheY ensemble into four sub-populations to assess the significance of differences observed between it and the complexes. The network connectivity, as formalized by centrality plots, showed significant changes in the complexes relative to the CheY protein alone. There was a dramatic reduction in the centrality of loop β4-α4 and associated β-strand $*_{106}$VVKP (* = Y (CheY·$FliM_N$), W (CheY*·$FliM_N$)) at the FliM binding surface. Their roles as network nodes were reduced in CheY·$FliM_N$ and abolished in CheY*·$FliM_N$. This trend contrasted with the conservation of these nodes for CheY*. The centrality of α-helix 1 increased with its mobility (Figure 2C).

Figure 3: CheY network dynamics. ***A.*** *The global network has nodes (residue fragments) and edges (mutual information weighted node interactions).* ***B.*** *Nodes containing residues that are part of the allosteric relay (W58, K91, Y106, K109, F111) have high scores in the CheY network. In A, B, these residues and control residues (M17, M60, M63) monitored by XFMS are highlighted (yellow circles).* ***C.*** *Centrality profiles of the complexes. (**(i)** CheY/$FliM_N$ (green). **(ii)** CheY*·$FliM_N$ (red)) compared with the native CheY profile (mean $\pm$ s.e; blue lines). The dotted line ((ii) red) plots the mutual information between the local loop fragment dynamics and collective PC1 motions. Complex formation reduced the centrality of the β4-α4 loop that together with the β3-α3 loop formed central nodes in the CheY network. Activating mutations eliminated the β4-α4 loop as a node but*

485 *did not alter the contribution of the β3-α3 loop in CheY*·FliM$_N$. Horizontal bars*
486 *indicate α3-β3 (white) and α4-β4 (black) loops as in Figure 2B.*
487

3. Immobilization of the α4-β4 loop modulates CheY collective motions.

We used Principal Component Analysis to characterize CheY collective motions and their modulation by $FliM_N$ binding and the activating substitutions. The Principal Components (PCs) are derived from the atomic-coordinate covariance matrix and describe C^α backbone movements, ranked according to the amplitude of the structural variation they explain. The collective motions were described well by the first few PCs, as found for other proteins. The first three principal components (PCI-PC2-PC3) accounted for > 60% of all motions in each case. These three PCs comprise bending and twisting modes organized around the β-sheet core. A core sub-population of CheY conformations was observed in MD trajectories generated by all three structures. When CheY is in complex with $FliM_N$, new sub-populations comparable in size to the core were generated. These were distinct in the CheY·$FliM_N$ and CheY*·$FliM_N$ complexes. Thus, new conformational ensembles are accessed upon binding of $FliM_N$, with the potential to produce binding surfaces for additional targets (**Supporting Information Figure S3A, B**).

Loops act as hinge elements for collective motions. Their mechanics give insight into the modules they control (39). We computed loop β3-α3 and β4-α4 hinge flexibility by mapping their RMSF onto the PC1 that accounts for > 40 % of the total amplitude of the PC motions. Flexibility scaled with the magnitude of the loop RMSFs relative to the mean PC1 RMSF. We computed hinge contribution to the PC1 as the nMI between PC1 variance and the local loop fragment dynamics. The long β3-α3 loop partitioned into two segments. The short D_{57}WN and the adjacent M_{60}PNMDG loop segments behaved as rigid (low RMSF) and flexible (high RMSF) hinges respectively to control native CheY PC1 dynamics. In the CheY*-$FliM_N$ complex, the β3-α3 loop hinge was retained, but with inverted flexibility of the two segments. The transition for loop β4-α4 was more dramatic from a flexible hinge in native CheY to a closed hinge that acted as a rigid lever arm in CheY*-$FliM_N$. The reduced flexibility decreased β4-α4 loop centrality and influence on PC1 motions (**Supporting Information Figure S3C**),

4. Protection experiments support the "intermediate" CheY·$FliM_N$ structure and the MD allosteric network.

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

attachment, and **(ii)** The configurations of free CheY with or without the D13K-Y106W substitutions were similar. The MD revealed **(iii)** $FliM_N$ attachment was more labile in the native versus D13K-Y106W complexes, and **(iv)** generated a network model to discriminate between CheY fragments that changed upon activation from those that did not. These predictions were assessed by comparing the sidechain solvent accessibility of allosteric relay residues Y106, W58, K91, K109, F111 and K119 by hydroxyl radical foot-printing in the native and D13K-Y106W CheY proteins, and their $FliM_N$-fusion constructs. The control residues predicted not to change during activation were the $\beta3$-$\alpha3$ loop residues M60 and M63 in proximity to W58, and the M17 in proximity to D/K13.

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

CheY residues of the allosteric relay at the $FliM_N$ interface and distant from it were designated “interface” and “core” residues respectively. The oxidation of the interfacial residues (K119, Y/W106, K91) was reduced in the complexes (**Figure 4 A. Ii-iii**). Importantly, oxidation of the core residues also decreased with complex formation (**Figure 4 A. Ci-iii)**. In contrast, there was no significant difference between oxidation rates for $\beta3$-$\alpha3$ loop control residues M60, M63 in the fusion proteins versus the free proteins, while the oxidation of the control M17 in the fusions was comparable or greater than in the corresponding free CheY proteins (**Figure 4B**).

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

bonding interactions may also contribute (Supporting Information Figure S6 legend). The correlation improved markedly (0.60 -> 0.86), without further correction, if the outliers were excluded.

The PFs for CheY*, CheY·$FliM_N$ and CheY*·$FliM_N$ were then normalized for each residue against the value obtained for CheY (**Figure 5B**). The normalized (log (PF)s) provided a quantitative measure for the increase for both interfacial and core residues in the CheY·$FliM_N$ and CheY*·$FliM_N$ fusions relative to the values for CheY. These residues were significantly more protected in CheY*·$FliM_N$ than CheY·$FliM_N$. In contrast, the protection of the control residues in the fusions (CheY·$FliM_N$, CheY*·$FliM_N$) did not differ significantly from that measured for CheY. The normalized PFs showed no significant difference in protection for interfacial, core or control residues in CheY* relative to CheY.

Figure 4: XFMS Measurements. *Dose-response curves for **A**. Relay**.** Interfacial residues **(Ii)** Y/106W. **(Iii)** K119**. (Iiii)** K91. Core residues **(Ci)** K109. **(Cii)** F111 **(Ciii)** W58. **B**. Control residues. **(i)** M60, M63. **(ii)** M17. Initial rates (dashed lines) were obtained from least-squares linear regression of the decrease in the un-oxidized fraction with dose.*

Figure 5. Single residue oxidations related to SASA. *A.* *Log (PF)s plotted against the side-chain solvent accessible surface area (SASA) calculated from the crystal structures. Pearson correlation coefficients: 0.86 (minus M17 (rose), K109 (cyan). See text). Overall = 0.60 {CheY= (-)0.76; CheY*= (-)0.70; $FliM_N$.CheY= (-)0.54; $FliM_N$.CheY*= (-)0.12}. Best-fit (black dashed line), 95% confidence limit (blue lines), 95% prediction limit (red lines).* ***B.*** *Protection of interfacial (K119, Y/W106, K91), core (F111, K109, W58) and control (M17, M60, M63) residues in CheY*, CheY·$FliM_N$, CheY*$FliM_N$ relative to their protection in CheY. $\{Protection\}_{norm}$ = Log $\{PF/PF_{CheY}\}$. Positive values indicate increased protection.*

The protection profiles showed that solvent accessibility for the allosteric relay residues decreased in the order CheY < CheY·$FliM_N$ < CheY*·$FliM_N$. The control residues either showed no changes or the opposite trend. Changes in the solvent accessibility of CheY* relative to CheY were not significant. Thus, in conclusion, the XFMS experiments validated the main predictions of the crystal structures and the conformational ensembles generated from them.

5. Energetics of CheY stabilization by $FliM_N$ and D13K/Y106W residue substitutions.

The XFMS measurements correlated solution population shifts in selected residue positions to each other and with the crystal structures. The temporal couplings between these shifts could only be studied with MD. We next analyzed the MD trajectories to extract this information.

We examined the temporal coupling between the electrostatic stabilization of the interface and the CheY fold with the rotational states of residue Y106 (106W in CheY*·FliM$_N$). CheY*·FliM$_N$ 106W sidechain was locked IN (Supporting Information Video S3). In contrast, Y106 in CheY (Supporting Information Video S1) and CheY·FliM$_N$ (Supporting Information Video S2) made frequent OUT <-> IN excursions. Dwell times in the Y106 rotamer states measured from the raw CheY trajectories were 107$\pm$34 ns (OUT) and 15$\pm$4 ns (IN)). The CheY·FliM$_N$ Y106 sidechain was predominantly in the IN orientation, with mean dwell time 239$\pm$123 ns, 15-fold greater than for free CheY. The conformational ensembles in the MD trajectories were clustered based on the C$^\alpha$ backbone dynamics {RMSF} The major clusters represented distinct backbone conformational states accessed during the MD runs. The average structures for these clusters were compared to each other and the crystal structures with PROPKA. The mean ΔG values at pH 7.0 were CheY (-4.8$\pm$1.0 (n=7)) < CheY·FliM$_N$ (-5.8$\pm$1.6 (n=4)) < CheY*·FliM$_N$ (-9.9$\pm$2.2 (n=3)). All CheY clusters had Y106 in the OUT orientation (θ = 126.7$\pm$3.8°) indicating that CheY Y106 IN states were too short-lived to influence backbone dynamics. CheY*·FliM$_N$ clusters had W106 in the IN orientation (θ = 54.1$\pm$2.3°). The CheY·FliM$_N$ clusters, in striking contrast, spanned the entire Y106 rotamer range. Thus, the intermediate CheY·FliM$_N$ Y106 rotamer states were sufficiently stable to affect backbone dynamics (**Figure 6**).

.

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

Next, we computed the activation energetics by measurement of ionizable residue electrostatics with PROPKA. There was a weak stabilization of the CheY FliM$_N$ interface and core with the internalization of the Y106 sidechain. The buried CheY*·FliM$_N$ W106 sidechain had a substantially more restricted rotation range than the CheY·FliM$_N$ Y106 sidechain. However, the 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 $\beta 3$ 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

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

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

Figure 7: Changes in community network architecture triggered by D13K/Y106W substitutions and $FliM_N$ peptide. *The reduced number of sectors compared to single fragments as nodes provided a concise, quantitative readout of the protein dynamics.* ***A. CheY and CheY* community maps.*** *Networks* ***(boxed insets)*** *from tCONCOORD runs show the reduction in the size of sector C relative to sectors A and B in CheY* versus CheY.* ***B. CheY, CheY·$FliM_N$ and CheY*·$FliM_N$ community architecture.*** *Networks* ***(top)*** *and maps* ***(bottom).*** *$FliM_N$ = yellow (cartoon representation). The MD detected four dynamic sectors for CheY (A= cyan, B = blue, C = orange, D = red). The sector C from the tCONCOORD runs is resolved into two sectors (C and D) in the MD runs. Node size = sector membership; edge thickness = weighted inter-sector interactions). Sectors A and B are built around the phosphorylation site (D57 (red asterisks)) and the $FliM_N$ binding surface respectively. They increase at the expense of sector C upon complex formation. The presence of phospho-mimetic substitutions in the CheY*·$FliM_N$ complex creates an additional sector E from sectors A and B, that orchestrates interactions with sectors C and D.* ***C. CheY*·$FliM_N$ community map showing Sector E surface.*** *See Supporting Information Video S5 for 3D perspective. Sidechains identify the substituted residues (K13, W106) and $FliM_N$ binding residue K119, a part of Sector E. 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 $\beta4$-$\alpha4$ 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 $\beta4$-$\alpha4$ 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 $\beta4$-$\alpha4$ 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

in the absence of $FliM_N$; a result supported in this study by both simulation and measurement. The MD clarified that $FliM_N$ stabilized CheY and strengthened allosteric communication between its binding interface and the D57 phosphorylation site due to formation, in part, of the CheY.K91-$FliM_N$.D3 salt-bridge. The salt-bridge decreased the flexibility of the β4-α4 hinge, consistent with earlier studies (31, 41).

2. The dynamics and energetics of activation.

This study documents a broad, high-energy CheY conformational landscape with shallow minima consistent with the high conformational plasticity suggested by the CheY crystal structures and early MD studies (Introduction). Network analysis, based on mutual information between short protein fragments established that two loops (β3-α3, β4-α4) act as flexible hinges to control the dynamics. The CheY MD trajectories revealed episodes where the Y106 sidechain is buried (IN), but cluster analysis determined the inward motions were too brief to influence backbone dynamics in contrast to the case for CheY·$FliM_N$. The buried states of the Y106 sidechain have not been visualized to our knowledge in inactive CheY crystal structures.

The CheY conformations of the major CheY·$FliM_N$ clusters were more stable than the dominant CheY conformations reported by the MD or the conformation in the 2B1J crystal structure. The lifetimes of the CheY Y106 IN states in CheY·$FliM_N$ were substantially greater than in free CheY and represented in the major clusters, There was a weak correlation between the stability of the CheY fold, the $FliM_N$ interface and the position of the Y106 sidechain. The CheY ΔG values in the major CheY·$FliM_N$ clusters overlapped with the values in the inactive CheY and activated CheY*·$FliM_N$ clusters.

The mean CheY*·$FliM_N$ ΔG value was more stable than for CheY·$FliM_N$. This was also the case for the interfacial ΔG values. The position of the W106 sidechain was restricted to a narrow range. Nevertheless, the ΔG values for both the CheY fold and its $FliM_N$ interface, as well as the rotamer position of the W/Y106 sidechain, were similar for the dominant CheY*·$FliM_N$ and CheY*·$FliM_N$ conformational clusters. The similarity may explain capture during crystallization of the Y106 sidechain in the 2B1J structure in a position superimposable with the W106 sidechain in the 1U8T structure. The better correlation of W106 sidechain position, in the MD clusters and the 1U8T structure, with the CheY fold and $FliM_N$ interface ΔG values, reflects the tight-packing due to the D13K-Y106W substitutions. The ΔG and W106 rotation angles of the CheY*·$FliM_N$ clusters had no overlap with values for CheY clusters.

Allosteric communication may range from largely enthalpic, as in lysozyme, to largely entropic with change in flexibility rather than shape (68). Both energy terms contribute to CheY allosteric activation. CheY activation has aspects that “invoke conformational selection”, namely the selection of the global minimum from the multiple minima sampled by the

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

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

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

Distinct protein sectors with correlated motions were identified in community networks. The extensive library of CheY residue substitutions was exploited for functional assignment of the sectors. Two sectors, namely the neighborhood of the phosphorylation site (sector A) and the region of $FliM_N$ binding (sector B) had clear functional importance. Two other sectors lacked strong, specific phenotypes and might have broader functions in maintaining the overall CheY fold. The long $\beta3$-$\alpha3$ loop influenced movements of the $\beta3$ strand that formed a sector junction, consistent with its central role in the reported PC motions. Similar motions take place in other proteins that utilize β-sheets for signal transduction (72).

$FliM_N$ attachment increased the size of sectors A and B in the CheY community network. The CheY*·$FliM_N$ community network was distinguished from the CheY and CheY·$FliM_N$ networks by a fifth sector (E), drawn from sectors A and B, that formed a dedicated conduit between the phosphorylation and $FliM_N$-binding sites to cement the allosteric linkage, with the substituted residues K13 and W106 at its boundaries. The emergence of sector E was tied to the closure of the $\beta4$-$\alpha4$ hinge by the CheY.K91-$FliM_N$.D3 salt-bridge and “freeze-out” of W_{106}VVKP β-strand dynamics by the burial of aromatic residues for a tightly packed core. This sector connects with all

other sectors and has a large surface profile. It may directly or indirectly define a region important for binding to FliN.

The CheY* protein is impaired in its interactions for other chemotaxis proteins, the CheA kinase and CheZ (6), that have CheY binding surfaces that overlap with that for $FliM_N$ (69, 73). Sector E may also influence the regulation of phosphorylation by these proteins. An important future goal would be to apply the integrated approach presented here to detect how CheY~P discriminates between these components of the chemotaxis circuity.

Rotamer reorientation of aromatic sidechains is a common theme in phospho-proteins, but diverse strategies for coupling side-chain motions to phosphorylation exist. In eukaryotic protein kinases, activation is controlled by DFG motif loops. These loops take on multiple IN and OUT orientations, with orientation correlated with activation. In Aurora kinase A, phosphorylation triggers transition between distinct IN orientations, rather than between IN and OUT states (74). In calcium calmodulin-dependent kinase, IN and OUT DFG states are loosely coupled to kinase domain phosphorylation (75). in CheY XFMS reported $D_{57}WN_{59}$ internalization was coupled to protection at the $FliM_N$ interface. We envisage that XFMS will have applications in other phospho-relays given ongoing developments in MS sensitivity and high-throughput analyses since most amino acids are modified by hydroxy radicals to a greater or lesser extent

The sparse sampling by crystal structures may miss high-energy states, such as the intermediate states of the CheY 106 sidechain, that are important for deciphering mechanism. MD simulations provide a much more detailed sampling of the conformational landscape, but their challenge is to extract the essential features from the large conformational ensembles obtained; a challenge only partially met by standard PCA and RMSF analyses. Our study shows that community maps provide a concise, comprehensive description based on quantitative criteria for identification of the key features of CheY allosteric activation. They could provide the optimal compromise for mechanistic dissection of signal transduction strategies in the response regulator superfamily.

SUPPORTING CITATIONS:

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

AUTHOR CONTRIBUTIONS:

Paige Wheatley: Investigation, Visualization, Writing - Review & Editing
Sayan Gupta: Methodology, Investigation
Alessandro Pandini: Conceptualization, Software, Formal analysis, Validation, Visualization
Yan Chen: Methodology, Investigation, Formal analysis, Validation

Christopher J Petzold: Methodology, Supervision, Funding Acquisition, Writing - Review & Editing

Corie Y. Ralston: Conceptualization, Supervision, Funding Acquisition, Writing - Review & Editing

David F. Blair: Conceptualization. Supervision, Funding Acquisition, Writing - Draft, Review & Editing

Shahid Khan: Conceptualization, Formal analysis, Writing - Draft, Review & Editing, Project Administration

ACKNOWLEDGEMENTS:

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

REFERENCES

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

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

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

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

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

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