Submitted
09/22/2019

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*]

Abstract

Phosphorylation of *Escherichia coli* CheY couple's chemoreceptor output to flagellar motor response. The N-termini of FliM subunits (FliM$_N$) in the flagellar rotor prime CheY to bind FliN thereby inducing cooperative switching, an essential feature of bacterial chemotaxis. We analyzed molecular dynamics {MD} trajectories to identify networks of residues involved in the long-range allosteric activation of CheY by FliM$_N$. The CheY backbone was partitioned into four dynamically coordinated sectors, with activation tracked by changes in sector size and interactions. Bound FliM$_N$ closed the central α4-β4 loop hinge to strengthen correlations between sectors around its binding interface and D57 phosphorylation site. Inward W58 sidechain movements adjacent to the CheY D57 phosphorylation site were coupled to corresponding K91 and Y106 sidechain movements at the FliM$_N$ interface. Studies of the constitutively active CheY D13K-Y106W double mutant have related its structural changes with *in-vivo* signaling properties. The MD revealed that D13K-Y106W fused the phosphorylation site and FliM$_N$ binding sectors into a new surface-exposed sector and locked the 106W sidechain in the innermost rotamer configuration in CheY-FliM$_N$ complexes. X-ray foot-printing with mass spectroscopy exploited FliM$_N$-CheY fusion proteins to validate the concerted sidechain internalization of W58, K91 and Y106 triggered by bound FliM$_N$ and increased by D13K-Y106W. Oxidation rate was correlated with the solvent accessible surface area, with K109, another central element of the allosteric relay, an outlier likely due to hydrogen bonding. The measurements indicated the fusion proteins were an effective mimic of the crystallized complexes used for the MD simulations. In absence of the D13K-Y106W mutations, CheY Y106 sampled multiple inward rotamer states, but their coupling to backbone dynamics required bound FliM$_N$ to prolongation inward state lifetimes by bound FliM$_N$. Thus, as simulations have found for kinases, control of CheY activation by aromatic residue reorientation is more subtle than a binary ON-OFF switch.

Statement of Significance

The chemotaxis phospho-protein CheY is activated at the flagellar motor by the N-terminus (FliM$_N$) of FliM subunits. Crystal structures of FliM$_N$.CheY complexes with and without the phosphomimic D13K-Y106W double-mutation both have the residue 106 sidechain in the same IN orientation at the FliM$_N$ interface. Additional factors that explain activation were identified by atomistic simulations based on the crystal structures. Free CheY samples both IN and OUT Y106 rotamer states, but bound FliM$_N$ increases multiple IN-state lifetimes to alter backbone dynamics. D13K-Y106W triggers further alterations to select the W106 state seen in the crystals and create novel binding surfaces. Observed changes in sidechain position along the allosteric relay reported by the crystals were predicted and validated by X-ray foot-printing with mass-spectroscopy.

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 $\alpha\beta\alpha$ CheY fold has structural homology with small eukaryotic signal-transducing proteins (3). CheY phosphorylation by the CheA kinase is essential for coupling chemoreceptor array states to motor response in bacterial chemotaxis. Phospho-CheY produced within the chemoreceptor/CheA array diffuses to the flagellar motor, interacting with proteins of the C-ring (a.k.a. the switch complex) to increase clockwise {CW} rotation (4). In certain species such as *B. subtilis*, phospho-CheY exerts the opposite effect, stimulating CCW rather than CW rotation, but remains critical for chemotactic responses.

Classical ideas of protein allostery were developed on model cases such as lysozyme where the protein is flexible relative to the ligand (5). Conformational selection and sequential induced fit (6) arise as limiting cases. Ligand binding biases the conformational equilibrium to provide optimal solvation for the catalytic action. Allosteric signal transmission in a protein such as CheY poses new challenges because the protein is responsive to both phosphorylation and to the binding of a peptide (the N-terminus of FliM) that is also flexible. Conformational ensembles are generated when the ligand is also flexible (“soft”). This creates a complex conformational energy landscape with the potential to access novel conformations, which multiplies with each binding- or chemical-modification step. The recognition of such conformational complexity, informed by the protein-structure database, has led to the development of modern ideas of allostery (7). Energetics of the situation where both partners are flexible have been described using the analogy of the folding funnel, where the funnel bottom has a “rugged” landscape with multiple minima (8). Allosteric communication may range from largely enthalpic, as in lysozyme, to largely entropic, with change in flexibility rather than shape (9).

Here, we study allostery in CheY, with special attention to the consequences of its binding to the FliM N-terminal peptide (FliM$_N$ responsible for its initial interaction with the flagellar switch complex. Previous studies of CheY have spanned three decades (10), establishing this small protein as a model for fundamental design principles in protein allostery. Several crystal structures of *Escherichia coli* CheY have provided snapshots of the conformational changes driven by small-molecule phosphate mimics (11, 12), mutations (13-18) and FliM$_N$ (19, 20). CheY structures have shown that the phosphorylation site (D57) is coupled to changes at the FliM$_N$- binding surface by the residues Y106, T87, K91, K109 and W58. The Y106 side-chain, exposed in soluble CheY, moves in as FliM$_N$ binds, while T87 hydrogen bonds with K109. Binding of FliM$_N$ switches K109 bonding interactions. In the absence of FliM$_N$, K109 hydrogen bonds with D57 and, via bound water, with D12; on binding of FliM$_N$, the K109 side-chain is repositioned to bond with D12, with a water-mediated residual interaction with D57 (20). CheY residue K91 in the α4-β4 loop forms part of the FliM$_N$ binding surface (21).

The coverage of the accessible conformational landscape by crystal structures is too sparse to decipher all details of CheY structural modulation by phosphorylation and by binding of FliM$_N$. In addition, crystal-packing contacts can be confused with low-affinity binding interfaces (22), a common feature in signal-transducing phospho-relay systems. NMR and related spectroscopies, together with molecular dynamics simulations (MD), have previously been used to investigate the dynamics of CheY activation. NMR has shown that the phospho-aspartyl analog beryllium fluoride (BeF3) elicits global changes in backbone dynamics (23). Linear chemical shifts have argued for a two-state equilibrium of Y106 and segmental motions for other key residues implicated in allostery (24). Spin labels have measured distance changes between residues (25). MD has supported experimental indications that the β4-α4 loop is an important determinant of allosteric signaling (21) and examined a two-state allostery model based on the coupling between Y106 rotation and T87 movements triggered by hydrogen bond formation (26). Network analysis of MD trajectories based on mutual information can, in principle, provide a fuller description of the coordinated movements of various parts of a protein. This approach underpinned a recent comparative study of response regulators including CheY (27).

CheY has high conformational plasticity, harnessed in subtle ways by multiple binding targets (28, 29). Difference analyses of bound and unbound states essential for resolving the interplay between phosphorylation and the binding target of interest have been sparse by comparison with the many studies, experimental and computational, on free CheY conformational changes triggered by phosphate analogs, covalent modifications, and mutations. CheY structures with and without $\mathrm{FliM_N}$ are a notable exception (19, 20, 30, 31). The affinity of non-phosphorylated CheY for $\mathrm{FliM_N}$ (ca. 600 μM) is 20x weaker than for phosphorylated CheY (28). The activating mutations D13K, Y106W are potent modulators of motor rotation bias in vivo (32), but the structural changes triggered by these mutations do not map on changes triggered by chemical phospho-aspartyl analogues (12, 20, 23).

We present network analysis (33) of explicit (MD) and implicit (tCONCOORD (34)) solvent simulations for elucidation of allosteric communication. The analysis yielded quantitative measures for conformational selection and conserved residue-reaction pathways. As an adjunct to MD analysis, we used time-resolved X-ray footprinting (XFMS) (35, 36) to examine solution exposure of selected positions in $\mathrm{FliM_N}$-CheY fusion proteins, mimics of the $\mathrm{FliM_N}$-captured state of CheY and shown previously to bind to FliN (37). The use of fusion constructs simplified interpretation of the XFMS data by preventing solution heterogeneity that might otherwise dominate mixtures of weak-binding components. XFMS has been used previously to study protein global – local dynamics (38) and protein-protein interactions (39). The computational and experimental results together provide a useful description of cooperative structural changes that accompany activation of CheY by phosphorylation and by binding of $\mathrm{FliM_N}$, with implications for allosteric phenomena employed by phospho-relay systems.

Materials & Methods

1. Structure Preparation.

Structures of *Escherichia coli* CheY (PDB ID: 3CHY. 1.7 angstrom resolution) and complexes of native (PDB ID: 2B1J. 2.8 angstrom resolution) and mutant (13DKY106W) CheY (PDB ID: 1U8T. 1.5 angstrom resolution) with FliMn were downloaded from Protein Data Bank. The 1U8T unit cell was a tetramer with 2 CheY and 2 CheY. FliMn complexes. The 2B1J resolution was too low (2.8 angstrom) to be useful for molecular simulations. Instead reverse mutagenesis (13 K->D, 106 Y->W) was used to generate the native CheY.FliMn complex (1U8T_DY) for such purpose. Missing atoms were added in Swiss-PDB viewer (*www.expasy.org/spdbv*); missing loop segments were completed with Modeller (*https:/salilab.org/modeller*). Mutant substitutions were made in Pymol *(http://pymol.org*) then energy minimized in Modeller.

Contact residues, surfaces and energies were extracted from the PDB files with the sub-routines (*ncont, pisa*) available in CCP4 version 7 (*http://www.ccp4.ac.uk/*). Comparison with experimental B-factors and geometrical analyses were performed with GROMACS version 4.5.7 (*http://www.gromacs.org/*). Solvent accessible surface area (SASA) was computed with the POPS server (*https://mathbio.crick.ac.uk*) (40).

2. Molecular Simulations.

(a) Molecular Dynamics.

A set of 3 replicas of duration 1 µs each were generated for the mutant (1U8T) and native (1U8T_DY) complexes using GROMACS 2016.2 with Amber ff99sb*-ILDNP force-field (41). Another set of 3 replicas of 500 ns 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. A multistage equilibration protocol, modified from (42) as described in (43), was applied to all simulations to remove unfavourable contacts and provide a reliable starting point for the production runs including: steepest descent and conjugate gradient energy minimisation with positional restraints (2000 kJ mol^{-1} nM^{-2}) on protein atoms followed by a series of NVT MD simulations to progressively heat up the system to 300 K and remove the positional restraints with a finally NPT simulation for 250 ps with restraints lowered to 250 kJ mol^{-1} nM^{-2}. All the restraints were removed for the production runs at 300 K. In the NVT simulations temperature was controlled by the Berendsen thermostat with coupling constant of 0.2 ps, while in the NPT simulations the V-rescale thermostat (44) was used with coupling constant of 0.1 ps and pressure was set to 1 bar with the Parrinello-Rahman barostat and coupling constant of 2 ps (45). Time steps of 2.0 fs with constraints on all the bonds were used. The particle mesh Ewald method was used to treat the long-range electrostatic interactions with the cut-off distances set at 12 Å. The MD runs reached stationary root mean square deviation (RMSD) values within 3 ns.

(b) tCONCOORD

tCONCOORD utilizes distance constraints based on the statistics of residue interactions in a crystalstructure library (34, 46), to generate conformational ensembles from one structure. The 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$_N$. The random atom displacemnts were limited to 2 nm^3, the iteration limit for generation of a new structure was 500 and a threshold solvaation score of 2.2 (34). 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 (47).

3. Network Analysis.

(a) Structural alphabet.

The mutual information $I(X;Y)$ between two variables(X) and (Y) is

$$I(X;Y) = H(X) + H(Y) - H(X,Y);$$

where $H(X,Y)$ is the joint probability distribution;

The normalized mutual information, $nMI(X;Y) = (I(X;Y) - \varepsilon(X;Y))/(H(X,Y));$ $H(X)$ is a measure of the entropy $\Delta S(X)$ that is related to the number of microstates and their probability. k_B Is the Boltzmann constant

$$\Delta S(X) = k_{B.}\ln(W_X) = k_B.\sum_{i=1}^{n} p(X_i).\ln(\mathrm{p}(X_i))$$

$$H(X) = \Delta S(X)/k_{B.}$$

$\varepsilon(X;Y)$ is the expected, finite-size error.

The finite-size error estimated as in earlier publications (e.g. (27, 48)) corrects for the effects of finite data and quantization on the probability distribution (49). The nMI couplings are detected as correlated changes in fragment dynamics, after spatial filtration to isolate long-range couplings (27). The SA (50), is a set of recurring four residue fragments encoding structural motifs derived from PDB structures. There is no need for discretisation and / or optimisation of parameters as the fragment set is pre-calculated.

(b) Eigenvector Analysis.

Statistically significant correlations between columns were identified with GSATools (33) 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. The contribution of a node to the network was estimated by the eigenvector centrality, *E*, calculated directly from the correlation matrix:

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

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

Collective motions were identified by PCA of the conformational ensembles. PCs were generated by diagonalization of the covariance matrix of C$^\alpha$ positions in GROMACS 4.5.7. The overlap (cumulative root mean square inner product) of the PCs between replicas (51)) and the PC dot product matrix was computed with the GROMACS g-anaeig function. The nMI contribution of local fragment motions was 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 threshold for selected top couplings. Pearson's correlations were used for comparison. Significance limits were set in GSATools.

(c) Community Analysis.

The Girvan–Newman algorithm (52) 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 (53). Community analysis of correlation networks identifies relatively independent communities that behave as semi-rigid bodies. Graphs were constructed with the *igraph* library (54) 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

E. coli strain BL21/DE3 with CheY-pET21b plasmid (37) was grown in 15 mL LB + 100 ug/mL ampicillin at 32°C for 15-18 hours. The culture was diluted 1:100 into 1 L LB + 100 ug/mL ampicillin + 250 uM IPTG and grown at 32°C until OD600 0.5-0.7. Cells were collected by centrifugation at 5000g for 20 minutes at 4°C . The cell pellet was resuspended in 8 mL lysis buffer (1 mg/ml lysozyme, 1 mM PMSF, 50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole, pH 8) and rocked on ice for 2 hours. 5 ug/mL DNAse I and 2.5 mM $MgCl_2$ were added and the lysed solution was rocked on ice for an additional 20 minutes, or until the lysate was no longer viscous. Lysed samples were sonicated for 30 pulses (Branson Sonifier 250, duty cycle 50, output 3) to completely lyse cells. The lysate was centrifuged at 10,000g for 30 minutes at 4°C to remove cellular debris from the HIS-tagged CheY protein. The lysate supernatant was combined with 4 mL Ni-sepharose beads and mixed gently at 4°C for 1 hour. The CheY-HIS-Ni-sepharose beads were washed twice with 10 mL 50 mM NaH_2PO4, 300 mM NaCl, 20 mM imidazole, pH 8; twice with 10 mL 50 mM NaH2PO4, 300 mM NaCl, 30 mM imidazole, pH 8; and twice with 10 mL 50 mM NaH2PO4, 300 mM NaCl, 40 mM imidazole, pH 8. CheY protein was washed off the beads 7-8 times with 1mL 50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8 . 4 mL of purified CheY protein was dialyzed into XF (X-ray footprinting) buffer (10 mM potassium phosphate buffer (pH 7.2), 100 mM NaCl, and 10 mM MgCl2) at 4°C.

Purified CheY, $FliM_N$.CheY, CheY*, and $FliM_N$.CheY* proteins were analyzed by fast protein liquid chromatography (FPLC) on an AKTA Superdex-75 10/300 GL column in X-ray footprinting buffer at 22°C (conductivity 28.9 mS/cm). Carbonic anhydrase (29 kDa) and ribonuclease A (13.6 kDa) were also analyzed on the same Superdex-75 column in X-ray footprinting buffer as molecular weight standards **(Supporting Information Figure S1)**.

3D models of the $FliM_N$.CheY fusions were obtained with the I-Tasser suite (55). I-Tasser uses threading and template-matching to determine secondary structure segments. Tertiary 3D topology is generated by *ab initio* folding of undetermined regions by replica-exchange Monte-Carlo dynamics and refined to remove steric clashes and optimize hydrogen bonding. The final models are assigned a confidence score (cs = -5.0 -> 1.0) based on multi-parametric comparison against Protein Data Bank structures. In all top 5 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, RMSD = 7.2+4.2 angstroms (against CheY, $FliM_N$ crystal structures).

5. X-Ray Footprinting (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 (56). Sample irradiation was conducted without Alexa488 dye using a microfluidic set-up with 100 mm and 200mm ID tubing in combination with a syringe pump as previously described (57). 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

$$F = \{Xi/(T + (\sum Xi))\}$$

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 (58). Its logarithm (log(PF)) was proportional to the SASA. The relation assumes that the footprinting 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 (58).

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) operating in positive-ion mode (3,500 V). Source parameters employed gas temp (250°C), drying gas (14 L/min), nebulizer (35 psig), sheath gas temp (250°C), sheath gas flow (11 L/min), VCap (3,500 V), fragmentor (180 V), OCT 1 RF Vpp (750 V). The data were acquired with Agilent MassHunter Workstation Software B.06.01 operating in either full MS mode or Auto MS/MS mode whereby the 20 most intense ions (charge states, 2–5) within 300-1,400 m/z mass range above a threshold of 1,500 counts were selected for MS/MS analysis. CheY native and oxidized peptides were identified by searching MS/MS data against *E. coli* protein database with Mascot search engine version 2.3.02 (Matrix Science). Based on the information of accurate peptide m/z value and retention time, the peptide precursor peak intensities were measured in MassHunter quantitative analysis software.

Results

1. Activating mutations D13K/Y106W stabilize FliM$_N$ association and lock in W106.

Three MD replica runs each were performed for the wild-type CheY structure (3CHY.pdb(59)), the activated CheY double mutant D13K/Y106W in complex with N-terminal FliM peptide (FliM$_N$)(1U8T.pdb(20)), and a complex of wild-type (non-activated) CheY with FliM$_N$ engineered in silico from 1U8T.pdb (Methods). The crystal structure of the native CheY/FliM$_N$ complex (2B1J.pdb) was not used, because its resolution is relatively low and exhibits a systematic upward trend in B-factors from the N- to C-terminus (**Supporting Information Figure S2**).

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. Henceforth, the double mutant CheYD13K/Y106W will be referred to as “CheY*”; FliM$_N$.CheY as the “wild-type CheY complex” and FliM$_N$.CheY* as the “mutant CheY complex”.The root-mean-square fluctuation (C$^\alpha$ RMSD) profiles for each structure, averaged over three 1 μs runs, are shown in **Figure 1** and are compared with B-factors for the X-ray structures in Figure S1. The B-factors were high relative to the MD-derived RMSD’s, particularly in loop regions, possibly reflecting conformational heterogeneity of these segments in the crystal.

In the crystal structure of CheY (3CHY.pdb) , the side-chain of Y106 exhibits two rotamer states, IN and OUT(59). 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. In MD trajectories of the FliM$_N$-CheY complexes, FliM$_N$ had higher mean C$^\alpha$ RMSD values when CheY was wild type than when it carried the activating mutations. Inspection of the raw MD trajectories **(Supporting Information Videos S1, S2, S3)** showed this difference was due to transient association/dissociation of the FliM$_N$ N and C termini from wild-type CheY. The peptide centre was tethered by a CheY-K119 to FliM$_N$-D12 salt bridge. In the non-mutated CheY, OUT excursions of Y106 cleaved this salt-bridge and weakened interfacial attachments **(Supporting Information Video S4)**. When CheY carried the activating mutations, FliM$_N$ did not dissociate and W106 was locked in the IN orientation in the MD trajectories (20). FliM$_N$ secondary structure and the salt-bridge are conserved in crystal structures of both the wild-type and activated-mutant CheY complexes, but B-factors of the FliM$_N$ termini are higher in the wild-type complex owing to fewer hydrogen bonds and loss of a CheY-K91 to FliM$_N$-D3 salt-bridge (19). Thus, the upward trend of B-values in the wild-type CheY complex may be due to a smeared-out electron density at the interface formed between FliM$_N$ and CheY C-terminal residues including K119. In conclusion, the MD supported and refined observations of the FliM$_N$ interface reported by the crystal structures. The differences in the CheY core are analyzed below.

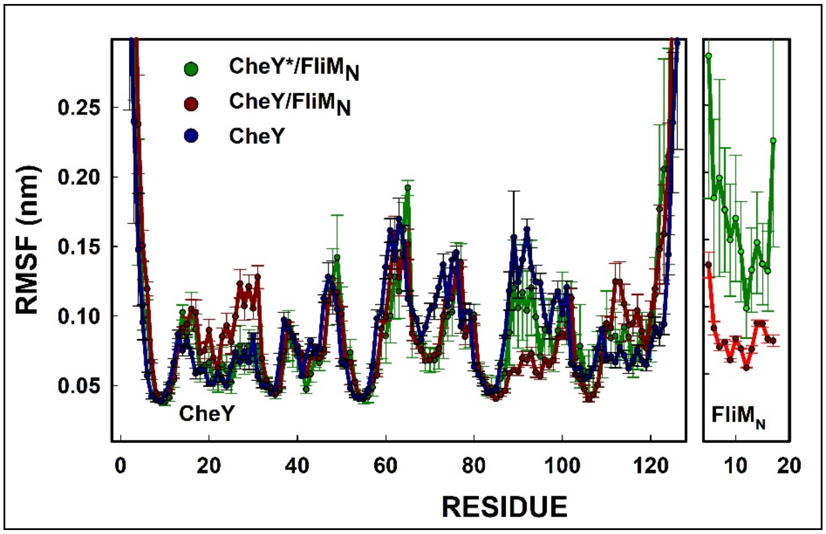


Figure 1: Dynamics of the CheY-FliM$_N$ association. *MD RMSF profiles for the combined replica trajectories for the three structures analyzed in this study. FliM$_N$ residue D12 forms a salt-bridge with CheY K119. Asterisk denotes double-mutant CheY complex.*

2. FliM$_N$ association and activating mutations generate novel CheY conformational ensembles.

We used Principal Component Analysis to determine whether CheY adopts new conformations as a result of FliM$_N$ binding and activating mutations. The Principal Components (PCs) are derived from the atomic-coordinate covariance matrix and describe the coordinated (co-varying) backbone movements that most efficiently account for the MD trajectories. Principal components are ranked according to the amount of structural variation they explain. The collective motions of the CheY backbone are described well by the first few PCs, as has also been found for other proteins (**Supporting Information Figure S3**). PCI-PC2-PC3 plots define a conformational space and provide a snapshot of conformations sampled by the various proteins (**Figure 2**). A core sub-population of CheY conformations was observed in MD trajectories generated by all three structures. This core set contained the bulk of the conformations obtained for free CheY (3CHY). When CheY is in complex with FliM$_N$, new sub-populations comparable in size to the core were generated, and were distinct in the wild-type and activated-mutant complexes (Figure 2). Thus, new conformational ensembles are accessed upon binding of FliM$_N$, with the potential to produce binding surfaces for additional targets. Inspection of the first three PCs reveals that these consist largely of bending and twisting modes organized around the β-sheet core **(Supporting Information Videos S1, S2, S3)**. Movements of this kind are known to take place in other proteins that utilize β-sheets for signal transduction (60) .

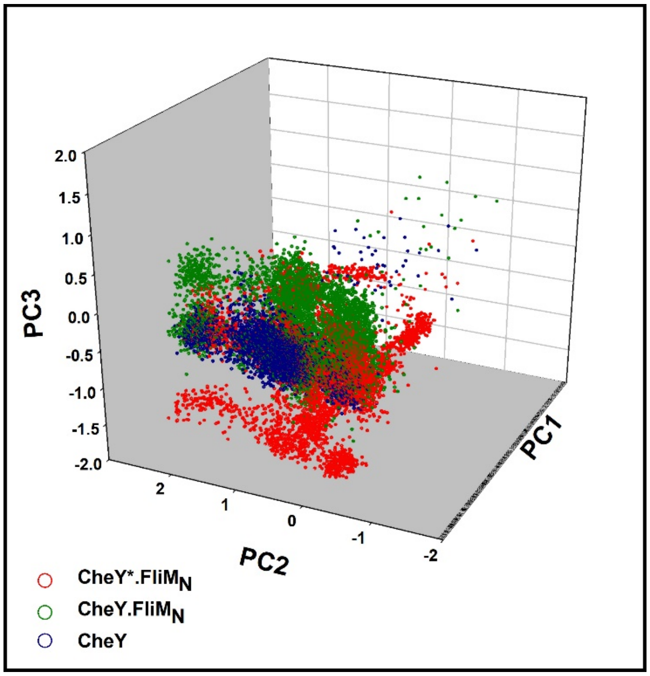


Figure 2: Conformational ensembles populated by FliM peptide and activating mutations. *The MD trajectories of the three structures (CheY, CheY.FliM$_N$, CheY*.FliM$_N$) were merged based on a common atom set. The member structures were projected onto the first three PCs. Non-overlapping structures indicate new conformational space made accessible by the binding of FliM$_N$. Distinct CheY conformational ensembles are sampled by free CheY, CheY.FliM$_N$, and CheY*.FliM$_N$.*

3. The dynamic architecture of CheY.

.

Coordinated motions in CheY were further examined using mutual-information analysis. The mutual-information (nMI) matrix encodes correlations between conformational states of different parts of the protein backbone, represented in terms of a structural alphabet (SA) derived from cluster analysis of protein structures in the pdb. Four-residue fragments are assigned an SA designation according to backbone dihedral angles, allowing conformation to be specified in terms of a 1-dimensional string (50). The segments can then be represented as network nodes,with the connectivity (edges) between them representing their correlated dynamics over the MD trajectory (see Methods for details). In vector notation, the overall extent of connectivity of a given fragment is reported by its eigenvector centrality ("centrality"). The centrality plotted against fragment number identified the central network nodes with the highest connectivity (**Figure 3A**). These nodes for free CheY were the loops 2 (β3-α3; D$_{57}$WNMPNMDG) and 4 (β4-α4; T$_{87}$AEAKK). A third prominent node just below the 1σ threshold was the short β-sheet (Y$_{106}$VVKP).

We mapped the dynamics onto the 3D structure in two ways. First, the top (>+2σ) nMI couplings from the global network were identified, revealing "hot-spots" engaged in the most strongly-correlated motions. Most (>95%) of the dynamic-coupling information in the nMI matrix is not utilized in this approach. Therefore, we used a second method, community analysis (61, 62), to extract and map this additional information. In brief, community networks are collapsed networks that partition and map the protein into semi-rigid body sub-volumes, based on correlation dynamics (see Methods for details). The two measures may be superimposed into a composite map, henceforth referred to as a "community map". We will refer to the sub-volumes in community maps as "sectors".

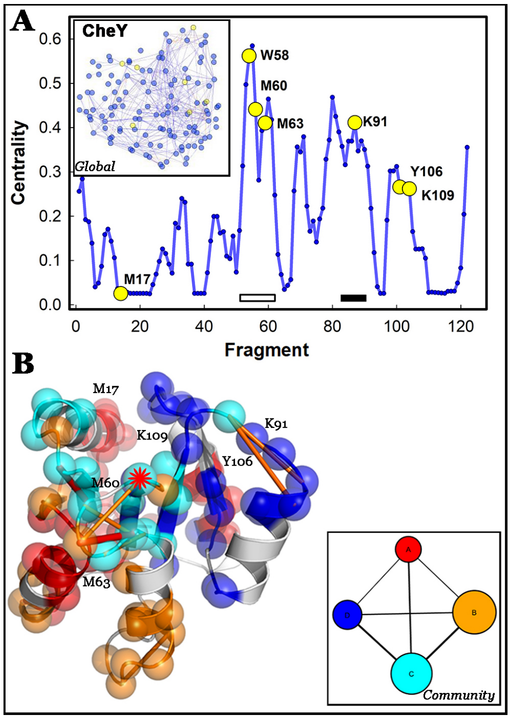


Figure 3: Network analysis. ***A.*** *The global network has nodes (residue fragments) and edges (mutual information weighted node interactions). The connectivity of the network is determined by the eigenvector centrality, a measure of the influence of individual nodes in the network. Nodes containing W58, K91, Y106, or K109 have high scores. Residues monitored by XFMS (M17, M60, M63) are also highlighted (yellow circles). Horizontal bars indicate* $\alpha3$*-*$\beta3$ *(white) and* $\alpha4$*-*$\beta4$ *(black) loop fragments respectively.* ***B.*** *Wild-type (non-activated) CheY community map. The strength of the top nMI couplings (lines) is reflected in their thickness and color (low (yellow) -> high (red)). CheY is composed of four sectors (A= red, B = orange, C = cyan, D = blue). The phosphorylation site (D57 (red asterisk)) and the* FliM_N *binding surface are part of sectors C and D respectively. Side-chain labels identify residues monitored by XFMS.* ***Box Inset:*** *Schematic of sector size and sector interactions. The reduced number of sectors compared to single fragments as nodes provides an effective quantitative readout of the matrix architecture (node size = sector residue membership; edge line thickness = weighted node interactions). The largest sectors (B and C) have comparable size.*

Community analysis of wild-type CheY reveals four major sectors displaying coordinated dynamics (**Figure 3B**). The β3 strand F$_{53}$VISD$_{57}$ occupies a central location in contact with all four sectors. Sector C, organized around the D57 phosphorylation site, is most prominent. It showed the strongest couplings to the other sectors, in particular to sector D, organized around the FliM$_N$-binding surface. Sector B has comparable size to C; it might be expected to influence the overall stability and rigidity of the protein. Known mutations were mapped onto the sectors (**Table 1**). Positions where mutations are known to affect dephosphorylation kinetics (63) mapped to sector C, while residues known to harbor suppressor mutations for CW- or CCW-biasing FliM mutations (64) mapped to sector D. Positions yielding mutations that affect interaction with the CheY-phosphatase CheZ (65) were adjacent to Sector A, the smallest sector obtained from the community analysis of wild-type CheY.

RESIDUE	SECTOR	MUTATION	CITATION
D57	D	D -> N	[16]
I95	D	I -> V	[14]
Y106	D	Y ->	[15]
T87	D	T -> I	[15]
A90	D	FliM^{sup}	[44]
T112	D	FliM^{sup}	[44]
K91	D	K -> R	[21]
V108	D$^{\#}$	FliM^{sup}	[44]
F111	D$^{\#}$	FliM^{sup}	[44]
D117	#	FliM^{sup}	[44]
D12	C	D ->	[16]
D13	C	D -> N, K	[16,43]
S56	C	S -> F	[44]
N59	C	F -> *	[43]
E89	C	F -> *	[43]
F14	B	F -> *	[43]
D27	A	FliM^{sup}	[44]
K26	A$^{\#}$	K -> E	[45]
N23	#	N -> D	[45]

Table 1: Mapping of CheY mutations to sectors identified in the community analysis. *Asterisks denote cases where the mutated residue does not map to a sector. Superscripted asterisks mark sectors < 3 positions (< 1 fragment) removed from an unassigned mutated residue.*

Comparison of the CheY and CheY* structures (20) shows that the structures are similar in absence of FliM$_N$. We used tCONCOORD to study whether the similarity extended to conformational flexibility. tCONCOORD is a computationally inexpensive method for rapid generation of conformational ensembles from a single structure (34, 46). In brief, the structure was rebuilt by random displacement of its atoms within limits, followed by iterative correction to eliminate bond violations until a new structure that satisfies all bonding constraints is obtained. The process was repeated to generate an ensemble, with simulation parameters listed in Methods. Full atom detail is preserved, but solvent is not modelled for increased computational speed. Instead, tCONCOORD has a solvation parameter that estimates the distance-dependent probability for a water molecule next to a particular atom to allow predicton of unstable hydrogen bonds (34) – solvent is "implicit". The cost is reduced spatio-temporal detail due to loss of long-range and bound-water interactions.

We first checked the tCONCOORD ensemble for wild-type CheY against its MD trajectory. The community map derived from the tCONCOORD ensemble showed three distinct sectors. Sectors C and D were as in the community map obtained from the MD, but sectors A and B were merged. The absence of short-lived Y106 rotamer IN states also reflected the reduced accuracy of tCONCOORD relative to MD. Nevertheless, the comparison of the tCONCOORD ensembles for native and activated-mutant CheY, our primary rationale for use of the method, established importantly that both the native Y106 and the activated mutant W106 side chains remained in the OUT orientation, a diagnostic of the inactive conformation, as seen in crystal structures. This agreement is noteworthy since the single mutant CheY Y106W crystal structure showed W106 in the IN state (15). Furthermore, the centrality and community architecture of wild-type CheY and CheY* were remarkably similar (**Supporting Information Figure S4**).

4. Loop dynamics modulate the link between D57 and the FliM$_N$-binding surface.

We next examined the wild-type and mutant CheY in complex with FliM$_N$ by the same methodologies used for free CheY. We split the free CheY data-set to evaluate its variance and thereby assess the significance of differences observed in the complexes. The network connectivity, as formalized by centrality plots, showed significant changes in the complexes relative to the CheY protein alone (**Figure 4A)**. There was a dramatic reduction in the centrality of loop β4-α4 and associated β-r106 at the FliM binding surface. Their roles as network nodes were reduced in the wild-type complex and abolished completely in the mutant complex.

The changes in loop dynamics were reflected in the community maps (**Figure 4B**). The couplings between sector C (phosphorylation) and sectors D (FliM$_N$ binding) and B were strengthened relative to the un-complexed protein. In the map of the native CheY complex, sector A expanded at the expense of sector B while increasing its coupling to sector D. A fifth sector (sector E, consisting of $K_{45}N_{62}L_{65}A_{101}S_{104}F_{111}K_{119}$) formed in the map of the activated-mutant CheY in complex with FliM$_N$. The E-sector segments were drawn from sector B (K_{45}, N_{62}, K_{119}) or sector D (A_{101}, S_{104}) or from regions adjacent to sector B (L_{65}) or sector D (F_{111}) in the free CheY community map. Sector E formed a surface-exposed ridge that connected the FliM$_N$ α-helix, via S_{104} andK_{119}, to sector C residues E_{35} and (via K_{45}) E_{37}, as may be appreciated from the 3D perspective provided by **Supporting Information Video S5**. The centrality of the highly connected loop β3-α3, and the top nMI couplings that linked fragments within it to the phosphorylation site, were unchanged by complex formation.

Loops act as hinge elements and their mechanics give insight into the modules they control (8). Loops β3-α3 and β4-α4 were treated as hinges consistent with their position as network nodes. We computed hinge flexibility by mapping their RMSF onto the principal collective motion (PC1). PC1 accounts for > 40 % of the total motion amplitude (Supporting Information Figure S2). 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 its variance and the local loop fragment dynamics (**Figure 5**). For this analysis, the long β3-α3 loop was 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 wild-type CheY PC1 dynamics. In the activated CheY-FliM_N complex, hinge function for the β3-α3 loop was retained, but with flexibility of the two segments inverted. The transition for loop β4-α4 was more dramatic. This segment functioned as a flexible hinge in wild-type CheY, but showed reduced flexibility, decreased centrality, and only marginal influence on PC1 motions in the CheY*-FliM_N complex. (Figure 4A). These changes are consistent with its transition from a flexible hinge to a rigid lever arm and have the effect of fusing elements of sectors B and D to drive creation of sector E, thereby increasing connectivity between sector C (phosphorylation) and sector D (FliM_N interface).

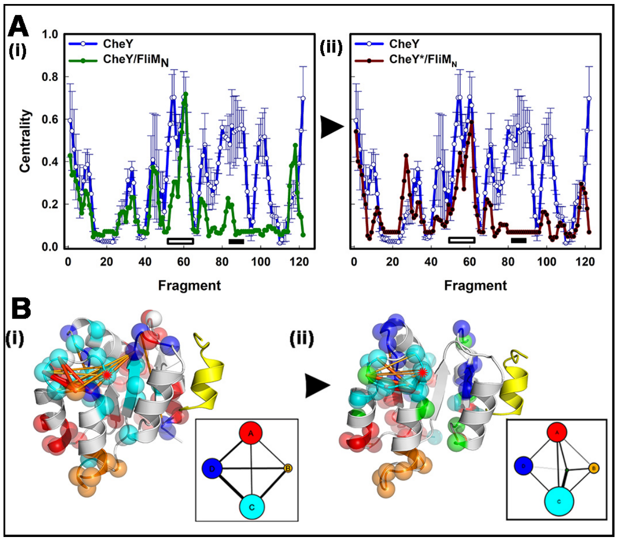


Figure 4: Changes in network architecture triggered by FliM_N peptide. ***A.*** *Centrality profiles of the FliM_N complexes ((**i**) CheY/FliM_N (green). (**ii**) CheY*.FliM_N (red)) compared with the native CheY profile (mean±s.e; blue lines). The standard error (s.e) was obtained from the variance of ensembles obtained by partitioning the summed CheY MD trajectories into four sub-populations. Capture of FliM peptide reduced the centrality of the β4-α4 loop that together with the β3-α3 loop forms central nodes in the 3CHY centrality profile. Bars indicate the loops as in Figure 3A. Activating mutations in the complex eliminated the influence of the β4-α4 loop as a central node; while the contribution of the β3-α3 loop remained unchanged.* ***B.*** *Community maps for (**i**) CheY.FliM_N. (**ii**) CheY*.FliM_N. FliM_N = yellow (cartoon). The top NMI couplings and*

sectors are named / color coded as in Fig 3B. D57 (red asterisks). ***Box Insets****: Schematics for sector interactions. Sector radii and edge line thickness as in Fig 3B*

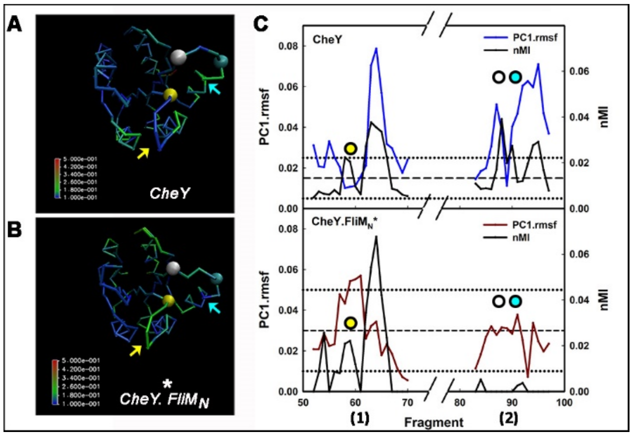


Figure 5: Hinge characterization. *PC1 RMSFs for their complete MD ensembles were mapped onto the structures.* ***A.*** *CheY (3CHY).* ***B.*** *CheY*.FliM$_N$ (1U8T). Color coded bars indicate RMSF values. Spheres denote W58 (yellow), K91 (cyan) and T87 (white) residues. Arrows identify β3-α3 (yellow) and β4-α4 (cyan).* ***C.*** *PC1 superimposed fragment nMI and RMSF values for the loops β3-α3 (1) and β4-α4 (2). The horizontal reference lines are CheY PC1 RMSF (mean (dashed)$\pm\sigma$(dotted)) values. The values are 0.015±0.01(3CHY); 0.03±0.02 (1U8T). Fragments with W58, T87 and T91 are marked (circles color-coded as in A,B).*

5. Frequency modulation by FliM$_N$ couples Y106 side-chain rotation to backbone dynamics.

The static crystal structures showed residue Y106 was in the OUT conformation in CheY (3CHY), but in the IN conformation in CheY.FliM$_N$ (2B1J) and CheY*.FliM$_N$ (1U8T). The conformational ensembles in the MD trajectories were clustered based on the C$^\alpha$ backbone dynamics {RMSD} to assess their coupling with the rotational states of residue Y106 (106W in CheY*.FliM$_N$). The major clusters represent distinct backbone conformational states accessed during the MD runs. Average conformations for the major clusters were compared to each other and the crystal structures on a 2D plot with axes representing two separate measures of protein fold-stability, the folding free energy {ΔG_{app}} and optimum pH {pH$_{st}$} (66). Structural alignment of representatives from the largest clusters confirmed that wild-type CheY Y106 was found predominantly in the OUT orientation, while CheY.FliM$_N$ Y106 predominantly superimposed with CheY*.FliM$_N$ 106W in the IN orientation (**Figure 6A)**. The major clusters for CheY (n =7) exhibited a large spread in pH$_{st}$ (4.5 -> 10) with ΔG_{app} values in the 0 -> -10 kcal/mol range. In contrast, C$^\alpha$ RMSD clusters (n = 3) for the activated CheY*.FliM$_N$ complex had a small pH$_{st}$ spread (8.4 -> 9.6) and were more stable (-10 -> -20 kcal/mol ΔG_{app}). The major clusters of the CheY.FliM$_N$ complex (n =4) showed pH$_{st}$ spread similar to CheY clusters and intermediate ΔG_{app} (-5 -> -15 kcal/mol) (**Figure 6B**).

The raw MD trajectories had revealed that while the CheY*.FliM$_N$ 106W side-chain was locked IN (Supporting Information Video S3), Y106 in both CheY (Supporting Information Video S1) and CheY.FliM$_N$ (Supporting Information Video S2) made frequent OUT <-> IN excursions during the MD runs. Dwell times in the two Y106 rotamer states measured from the raw CheY trajectories were 107$\underline{+}$34 ns (OUT) and 15$\underline{+}$4 ns (IN)). None of the seven CheY clusters had Y106 in the IN orientation. We conclude, therefore, that dwell times for the Y106 IN state are too short to influence backbone dynamics.

CheY.FliM$_N$ C$^\alpha$ backbone conformations overlapped with both CheY and CheY*.FliM$_N$ states. While the CheY.FliM$_N$ Y106 side-chain was predominantly in the IN orientation, with mean dwell time 239$\underline{+}$123 ns, 15-fold greater than for the uncomplexed CheY, multiple rotamer states with dwell-times > 25 ns were accessed during the MD runs and reported by the four major clusters. This result allowed us to investigate the coupling along the allosteric relay between the D57 phosphorylation site and residue Y106 at the FliM$_N$ interface. We monitored three residues {K91, K109, and W58} whose roles were noted in the introduction, in addition to Y106. Side-chain positions of both Y106 and K91 were correlated in a graded manner with the interfacial separation of the CheY and FliM$_N$ C$^\alpha$ backbones {**Figure 6Ci). ** Movements of K109 were small and uncorrelated with those of Y106. Although residue W58 belongs to a different sector (D) than the other residues (sector C), it is in proximity to and showed tightly coupled movements with Y106. (**Figure 6Cii**). We conclude that the transition between CheY.FliM$_N$ Y106 IN and OUT states is coupled to concerted conformational changes governed by multiple local minima reported by the four major clusters.

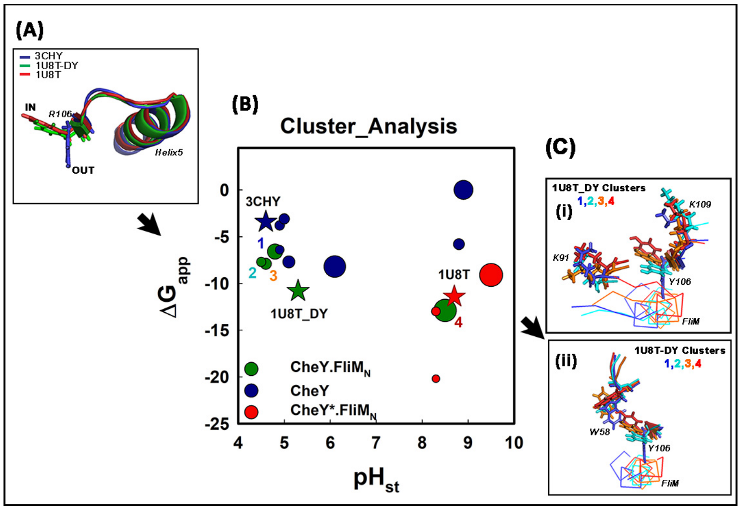


Figure 6: Coupling of Y106 rotamer orientation with CheY backbone dynamics. ***A. Match with crystal structures.*** *Y106/W rotamer orientation for the dominant cluster for each structural ensemble.* ***B. Representative structures of the major clusters*** *(circles; diameter indicates cluster size). C$^\alpha$ RMSD values for the MD trajectories were mapped on a 2D grid (the free energy change to achieve the stable configuration (ΔG_{app} (abscissa)) at optimal pH (pH_{st} (ordinate)). The asterisks represent the X-ray crystal structures.* ***C. 106 rotamer orientation – backbone***

dynamics coupling. (i) *Multiple Y106 rotamer orientations correlate with FliM peptide bound state in the native complex. Inward movement of K91 accompanies inward Y106 rotation, while K109 shows little change.* ***(ii)*** *The Y106 IN to OUT transition is coupled to coupled inward W58 movements.*

6. Oxidation of residues Y106, W58, and K91, but not K109, correlates with solvent accessibility.

The affinity of FliM for activated CheY is relatively weak and that for the inactive protein even weaker, as noted (Introduction). Therefore, we studied homogenous solutions of CheY and FliM$_{N}$-CheY fusion proteins to measure the changes brought about by the FliM$_{N}$ peptide in the allosteric signal relay as predicted by the MD simulations. We analyzed the side-chain solvent accessibility of residues Y106, W58, K91, K109 and K119 by hydoxyl radical foot-printing in the wild-type and activated-mutant CheY proteins, and their FliM$_{N}$-fusion constructs. Previous work documented the interaction of the FliM$_{N}$-CheY fusion with FliN, another protein component in the flagellar motor C-ring (37). Bioinformatic analysis (Methods) of the FliM$_{N}$.CheY fusions indicated that the flexible FliM$_{N}$-CheY linker docks FliM$_{N}$ onto CheY and CheY* as reported by the crystal structures.

Aromatic residues have high intrinsic reactivities with hydroxyl radicals, exceeded only by methionine and cysteine (absent from *E. coli* CheY), followed by the alkaline side chains. Thus, in addition to the five residues in the allosteric relay characterized by our MD simulations (Y106, K91, W58, K109, K119), we monitored oxidation of M60 and M63 in proximity to W58, as well as M17 four residues downstream from the mutated residue D13 as controls. Tryptic digestion partitioned CheY into six separated peptides that could be distinguished by mass spectroscopy (MS) based on their characteristic m/z ratio, allowing oxidation of these residues to be monitored with ease (**Figure 7**).

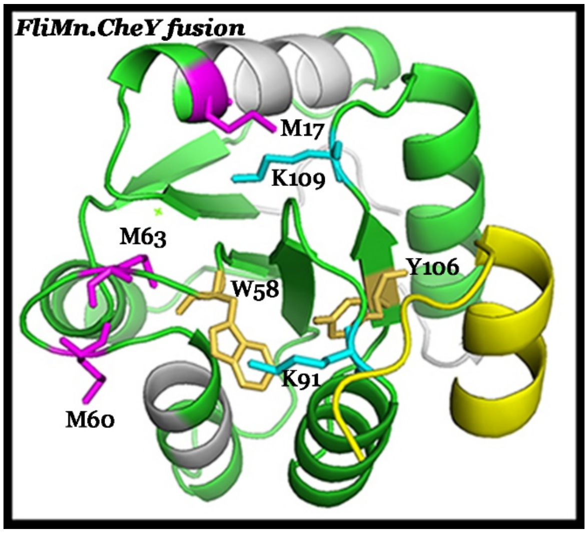


Figure 7: Residue positions examined by XFMS. *The residues are mapped on the modeled structure of the FliM$_N$.CheY (or FliM$_N$.CheY*) fusions tested in the XFMS experiments. Green segments indicate the tryptic peptides analyzed by MS after XF. FliM$_N$ peptide (yellow).*

Dose-response curves were generated for each of the four constructs (CheY, CheY*, CheY-FliM$_N$, CheY*-FliM$_N$) (**Figure 8**). For each residue examined, the curves from two independent experiments were pooled (**Supporting Information Figure S5**).

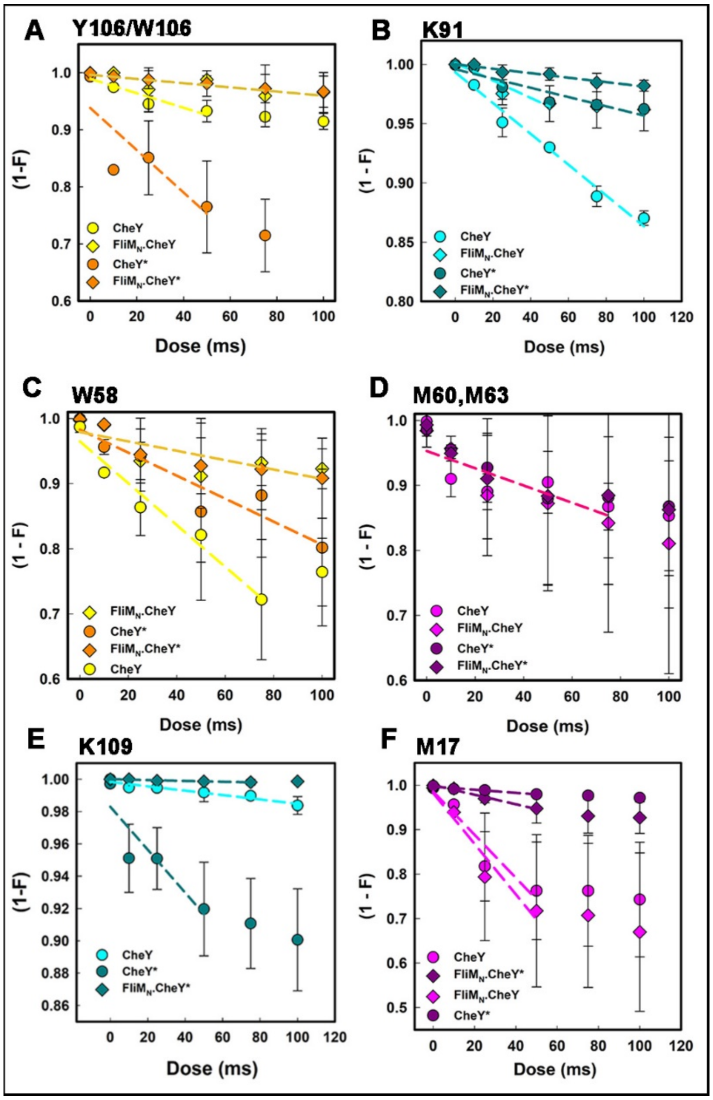


Figure 8: XFMS Measurements. *Dose response curves for* **A.** *Y/106W.* **B.** *K91* **C.** *W58.* **D.** *M60.M63.* **E.** *K109.* **F.** *M17. Initial rates (dashed lines) were obtained from linear best-fits.*

Oxidation of the Y106 or W106 side-chains was slowed in the FliM$_N$-fusions relative to the uncomplexed proteins, consistent with the expectation that binding of FliM$_N$ will increase the population fraction in the IN configuration (**Figure 8A**). Oxidation of W106 in activated CheY* was greater than that of the native CheY Y106 possibly due, in part, to greater (1.45x) intrinsic reactivity. In addition, it may be sterically more difficult to accommodate the bulkier tryptophan side chain in the interior of the protein in absence of the binding free energy supplied by FliM$_N$. In any case, the XFMS result is consistent with X-ray crystallographic evidence (20) that W106 is exclusively OUT in free CheY*.

The other interfacial residues are K119 and K91. The salt-bridge of K119 with FliM$_N$ D12 is important for complex formation (Figure 1, Supporting Information Video S4). K119 exhibited no detectable oxidation in the FliM$_N$-fusion proteins consistent with this premise. Some oxidation of K119 was observed in both the uncomplexed CheY and CheY* proteins, but the oxidation varied between experiments (Supporting Information Figure S4). Absence of K119 oxidation in the fusions is difficult to explain solely on the basis of reduced solvent exposure, since W106, whose side-chain is buried based on both MD and the crystal structures, showed detectable oxidation. K91 oxidation was reduced in the fusions relative to the uncomplexed CheY proteins (**Figure 8B)**. K91 oxidation occurred at a lower rate in the activated FliM$_N$.CheY* versus FliM$_N$.CheY and may reflect formation of a salt-bridge between K91 with FliM residue D3 in the former, but not the latter (20). The K119 and K91 oxidation data suggest that salt-bridge formation may protect against hydroxy radicals. More generally, the Y106/W, K119 and K91 data together indicate that the FliM$_N$ binding interaction in the fusion proteins is as predicted by the crystal structures of the complexes .

Residue W58 and the adjacent M60, M63 residues are part of the β3-α3 loop (D$_{57}$WNMPNMDG), the central node that organizes CheY dynamic architecture. W58 oxidation was reduced in the CheY.FliM$_N$ fusion proteins relative to the uncomplexed proteins, consistent with X-ray crystallography and with the MD prediction that W58 moves inward in the activated complex (**Figure 8C).** Like K91, W58 had a lower rate of oxidation in the mutant (CheY*, FliM$_N$.CheY*) versus wild-type (CheY, FliM$_N$.CheY) proteins. The data are consistent with the idea that the activating mutations in CheY* might facilitate side-chain internalization to a limited degree, to potentiate further burial upon complex formation. Residues M60 and M63 are in the solvent-exposed segment of the long loop that functions as a hinge to modulate CheY dynamics in both the uncomplexed proteins and the FliM$_N$ complexes. The oxidation rates of both methionine residues were insensitive to FliM$_N$ association, consistent with their surface location and the loop flexibility seen by MD (**Figure 8D)**.

K109 is more susceptible to oxidation in CheY* versus wild-type CheY (**Figure 8E)**. Reorientation of the D57 side-chain, observed in a single-mutant (D13K) crystal structure, eliminates its hydrogen bond with the K109 ε-amino group (18). The XFMS result is in line with the crystal structure if hydrogen bonding confers resistance to hydroxyl radical attack, as suggested above for salt-bridges based on the K91 data. Oxidation of K109 in the complexes with FliM$_N$ was very limited (FliM$_N$.CheY*) or undetectable (FliM$_N$.CheY). The crystal structures and MD both indicate that the K109 side-chain is more buried in FliM$_N$.CheY complexes relative to free CheY.

In conclusion, decreased oxidation rates in the fusions relative to the free proteins is the common theme for all residues in the allosteric relay (Y/W106, K91, K109, K119 and W58) that we have monitored by XFMS. Increased W106 and K109 oxidation in CheY* relative to wild-type CheY appear to be reflective of accessibility and bonding differences observed in the crystal structures.

M17 in helix α1 is distant from the FliM$_N$ binding surface and its low centrality and lack of membership in CheY*.FliM$_N$ community sectors determined by MD predict that it is not allosterically coupled to the FliM$_N$ interface. Consistent with prediction, oxidation of M17 in the FliM$_N$-fusions is comparable to or greater than in the uncomplexed CheY proteins (**Figure 8F).** An observed difference between M17 oxidation in the wild-type (CheY,FliM$_N$.CheY) versus mutant (CheY*,FliM$_N$.CheY*) proteins is unlikely to be due to local C$^\alpha$ flexibility differences, since M17 is within a helix. It could be due to perturbation of the local dielectric and bound waters caused by the D13K mutation, since the crystal structure indicates an altered hydrogen bonding network around the phosphorylation site in the D13K mutant (18).

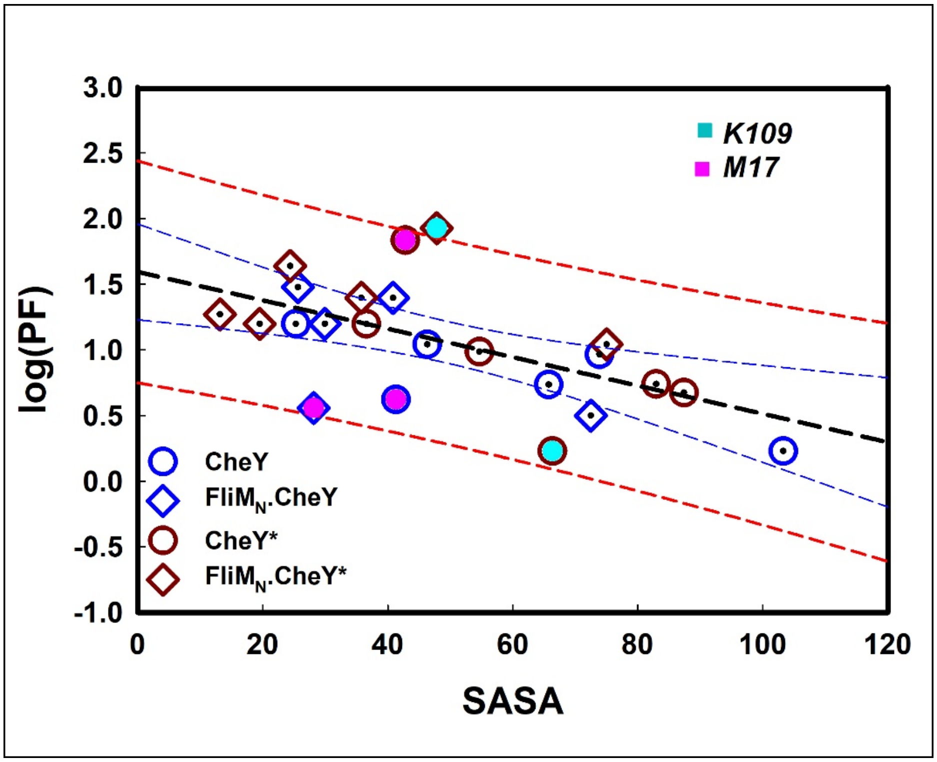


Figure 9. Single residue oxidations related to SASA. *Protection factors derived from the dose-response rates plotted against the side-chain solvent accessible surface area (SASA) calculated from the crystal structures. Pearson correlation coefficients: 0.86 (minus M17 (rose), K109 (cyan). See text). 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).*

The initial rates from the dose-response curves were used to compute protection factors (PFs) (**Figure 9**), following protocols established by the study of 24 peptides from 3 globular model proteins. This study showed that protection factors provide a read-out of the solvent accessible surface area (SASA) (58), using intrinsic reactivities mostly determined thus far from measurements on small peptides (67). We used single residue MS-analysis to determine PFs.

Outliers beyond the 95% confidence limit were found exclusively for cases involving K109 or M17. We speculate that the CheY* K109 outlier reflects the loss of hydrogen bonding interactions reported for the D13K mutant crystal structure (18) and an accompanying increase in the susceptibility of the K109 side-chain to oxidation. An even larger deviation, but in the opposite direction from the norm, is seen for $FliM_N$.CheY* K109. K109 oxidations in the fusions will be reduced to a greater degree than predicted by SASA alone, if ionic/hydrogen bonding confers resistance to oxidation as argued above, bond strength will increase for more buried locations due to the lower dielectric of the environment. Weighting of the M17 SASA with the intrinsic reactivity (58) or variability documented by clustering of the MD trajectories did not significantly improve alignment with the overall relation. Local perturbation of the bound water in the phosphorylation site, caused by a disrupted hydrogen-bonding network as found for CheY D13K (18), may affect oxygen radical reactivity of the inwardly oriented M17 side-chain (68). In addition, methionines are particularly susceptible to secondary solvent reactions (67). Some of these reactions could be influenced in ways that are not presently understood. It has been reported that SASA weighting by a contact parameter (69) improved its correlation with log(PF) (58). Our results suggest that closer study of the modulation of polar-residue reactivities by local dielectric and bonding interactions would be well-merited.

The overall correlation was comparable to published values for the model-protein peptide correlations. The correlation improved markedly (0.60 -> 0.86), without further correction, if the aberrant K109 and M17 values were excluded based on the arguments outlined above. Our relation importantly establishes that the changes in the dose-response plots for the monitored residues are in large part due to apolar bulk solvent accessibility changes.

Discussion

Although substantial advances have been made in understanding the mechanism of CheY-mediated switching of rotation sense, details of the allosteric linkages within the protein remain unclear**.** Because the phosphoryl aspartate is labile, mutants that mimic aspects of the activation mechanism have been used to examine conformational changes in CheY and their effect on its interaction with a the N-terminus of the flagellar rotor protein FliM, which has been proposed to capture CheY prior to a CheY-FliN interaction (37). The CheY mutant CheY D13K Y106W in particular has been well studied and is known to cause a strong CW rotational bias in vivo. The binding of activated CheY to isolated *E. coli* flagellar basal bodies is non-cooperative (70), presumably reflecting an initial interaction with FliM$_N$. In vivo, however,rotational bias displays a highly sigmoidal dependence on CheY concentration (Hill coefficient > 10.5) (71) implying highly cooperative action of the captured CheY molecules switching flagellar rotation from CCW to CW. It is not clear whether the FliM$_N$ binding step serves only to tether CheY to the flagellar rotor to increase the likelihood of a CheY-FliN interaction, or whether FliM$_N$ binding also triggers structural changes in CheY that prime CheY to bind FliN. In any case, the measured Hill coefficient implies that a significant proportion of the FliM subunits, which number around 34 (72), must be occupied with CheY for CW rotation to occur. Studies with *Thermatoga maritima* CheY indicate that its interaction with the switch involves two binding sites, with an initial interaction with FliM$_{N,}$ and a subsequent interaction involving the middle domain of FliM (FliM$_M$), as opposed to FliN in *E. coli* (30). Binding studies have not been reported in *B. subtilis*, but the architecture of the switch in this gram-positive model species is expected to resemble that in *T. maritima*, with CheY most likely interacting with FliM$_M$ rather than FliN. In this context, we note that CheY potentiates CCW rather than CW rotation in *B. subtilis*. The FliM$_{N.}$CheY fusion protein exploited in this study was engineered to mimic the state of CheY after it has been captured by FliM$_N$. The FliM$_{N.}$CheY fusion protein has been shown to interact with FliN (37) and is more potent than simple CheY in potentiation of CW rotation (P. Wheatley, unpublished).

In addition to FliM$_N$, CheY interacts with the kinase CheA that catalyzes CheY phosphorylation, and with the protein CheZ that accelerates CheY dephosphorylation. Binding sites for these three partners overlap; CheZ in particular contains a segment that is homologous to FliM$_N$ and that binds to CheY in a similar way. While X-ray crystallography has provided snapshots of various states of the protein, and diverse spectrocopies and computer simulations have identified elements important for allosteric signaling, the allosteric mechanism(s) that regulate CheY conformational state and its various targets remain unclear. A key aspect of this process, and a major focus of the present study, is the binding of CheY to FliM$_N$. Here, we have sought to elucidate changes in protein dynamics associated with FliM$_N$ binding, using recently developed network-analysis methods in conjunction with time-resolved footprinting as a probe of residue accessibilities.

CheY conformational plasticity. On binding to FliM$_N$, CheY takes on a distinct set of conformational states that can be differentiated on the basis of their conformational heterogeneity, as revealed by cluster analysis, and dynamics reported by principal collective motions and local community architecture. Wild-type and activated-mutant CheY give rise to distinct conformational populations with FliM$_N$ bound; these FliM$_N$-induced conformational differences appear large enough to influence subsequent binding steps, in particular the affinity of the protein for FliN. We speculate that this is important to the ability of activated CheY to potentiate switching.

CheY conformational selection. The MWC model (73) provides the simplest example of conformational selection and has been commonly utilized in models of CheY allostery. While this analytical approach has provided important insights, the richness of CheY conformational space revealed in the present study indicates the need for a more explicit approach based on extensive simulation of CheY dynamics. We found, based on cluster analysis of the MD trajectories, that conformational selection does occur upon FliM$_N$ binding, though less completely than in classical models. Clusters of activated mutant CheY*.FliM$_N$ complexes did not overlap with uncomplexed CheY clusters, occupied a more confined conformational landscape, and were more stable. Clusters of native CheY.FliM$_N$ complexes overlapped with both inactive CheY and activated CheY*.FliM$_N$ clusters, with backbone dynamics coupled to Y106 rotation state to delineate a possible pathway for CheY activation, as detailed below. Comparison of the implicit solvent ensembles generated for CheY and mutant CheY confirmed that the double mutation did not trigger activation in and of itself.

Further analyses utilized mutual information to identify community networks, which could be mapped onto the structure to identify dynamic sectors exhibiting correlated dynamics. This approach has been used previously (62) to identify jointly moving regions that do not track backbone secondary structure but are frequently governed instead by side-chain motions. Two of the communities identified here correspond roughly to sectors of clear functional importance, namely the neighborhood of the phosphorylation site and the region of FliM$_N$ binding. Two other communities are not as simply characterized, owing to the lack of mutations giving strong, specific phenotypes. These communities might have broader functions in maintaining the overall fold and conformational flexibility of CheY. A fifth, relatively-small community (sector E) that arose upon binding of FliM$_N$ to the activated-mutant, but not the wild-type, CheY protein is potentially of greatest interest. Residues in this community fall in a region between the phosphorylation and FliM$_N$-binding sites, and may represent an allosteric linkage between these sites. The residues involved are largely surface exposed, and it may also (or alternatively) define a region important for binding to FliN.

Loop dynamics Our conclusions, while broadly in line with previous studies, extend them in important ways. The integrated eigenvector and community network analyses provides a clearer view of the role of CheY loop regions in allosteric communication. The long β3-α3 loop is a key element underpinning the dynamics of the protein, influencing movements of the β3 “nexus” fragment. Although the β3-α3 loop is within four residues of the phosphorylation site (D57), its role appears similar in uncomplexed CheY and the FliM$_N$-CheY complexes, suggesting that it is not critically involved in allosteric signaling *per se*. Allosteric linkage appears instead to involve changes in β4-α4 and β-r106 dynamics, which are constrained upon binding of FliM$_N$. The flexibility of the β4-α4 hinge is coupled to the association of FliM$_N$ and the position of the residue-106 sidechain.

Side-chain dynamics: Two-state allosteric models of CheY have drawn on the fact that the conformation of the Y106 side-chain is bistable, apparently taking on one of two (IN or OUT) conformational states. The Y106 OUT and mutant W106 IN rotamer state were associated with distinct conformational clusters in CheY and CheY*.FliM$_N$ respectvely. In addition to the dominant OUT state in uncomplexed CheY and the IN state in activated CheY*.FliM$_N$, the CheY.FliM$_N$ Y106 sidechain displayed two intermediate states, with transitions between the states reflecting the strength of the FliM$_N$ association. K91 side-chain accessibility also tracked association strength. These movements were also coupled to side-chain motions of W58, a residue which, by virtue of its proximity to D57, appears likely to contribute to allosteric linkages between the phosphorylation site and the binding site for FliM$_N$. Binding of FliM$_N$ was found to affect the dynamics of interconversion of the Y106 rotamers; transitions are rapid in uncomplexed CheY but are slower upon binding of FliM$_N$. We suggest that this "freezing in" of conformational states of Y106 is an important aspect of the allosteric linkage within CheY, enabling a rigidification of the binding interface that contributes to the increase in FliM$_N$-binding affinity. **Figure 10** summarizes our results **(A)** and their implications for CheY allostery and flagellar motor switch cooperativity.

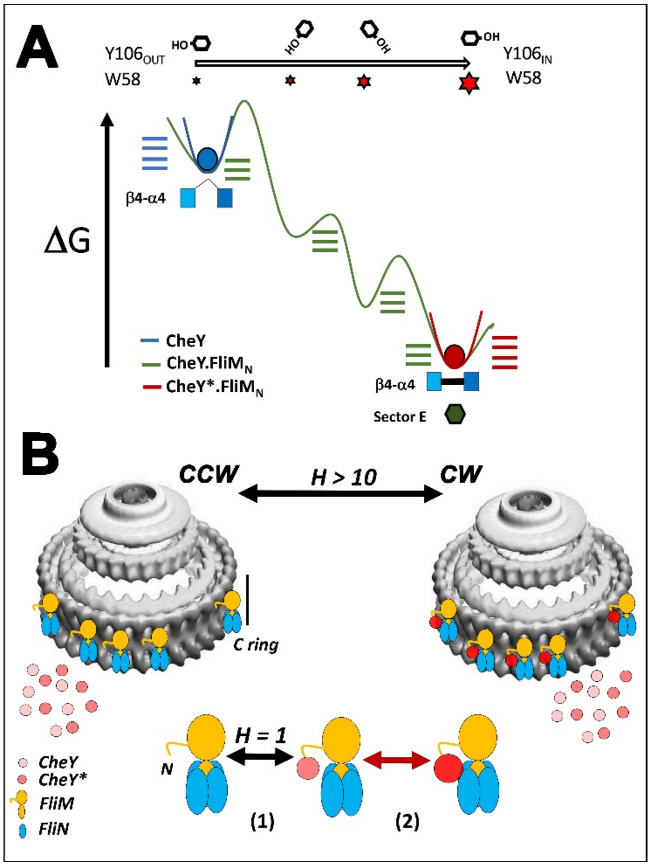


Figure 10: CheY activation and motor reactions in E. coli. ***A.*** *Activation is coupled to FliM*$_N$ *binding and rotamer transitions of CheY residue 106. (1) CheY ensembles (blue) have large conformational heterogeneity and sample both Y106 IN and OUT rotamer states, but the IN state is too short-lived to influence backbone dynamics with or without activating mutations. (2) FliM*$_N$ *bound to native CheY (green) generates multiple, Y106 IN rotamer states allosterically coupled via backbone dynamics with the D57 phosphorylation site monitored via W58. (3) FliM*$_N$ *bound to the phospho-mimic CheY (CheY*) locks residue 106 in the IN state in all major conformational clusters. (4) Activation is driven by rigidification of the* β*4-*α*4 loop that limits conformational heterogeneity and creates a new surface exposed dynamic sector to link the phosphorylation site with the FliM*$_N$ *binding interface. (5) Concerted movements of Y106, W58 and K91 (as probe for FliMN occupancy seen in simulations were supported by XFMS.* ***B.*** *Binding of CheY* to isolated basal bodies with C-rings is not cooperative (H = 1), but the change in CW/CCW rotation bias is highly cooperative (H > 10) with in vivo CheY* concentration. Evidence indicates that 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. In addition, as argued by this study and the X-ray crystallographic data on which it is based, allosteric activation by FliM*$_N$ *stabilizes bound CheY*.FliM*$_N$ *complexes creating new surface topology to potentially enhance 2nd stage binding. Stabilization of bound CheY.FliM*$_N$ *complexes is not as great, substantially reducing the likelihood that the critical threshold of bound CheY.FliN complexes to switch to CW rotation will be reached. CheY shade intensity represent activation state and FliN binding probability respectively.*

Rotamer reorientation of aromatic side-chains is a common theme in phospho-proteins, but there exist diverse strategies for coupling side-chain motions to phosphorylation. In eukaryotic protein kinases, activation is controlled by loops with the DFG motif. These loops take on multiple IN and OUT orientations, with orientation only loosely 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 contrast, the multiple Y106 orientations in the wild-type FliM$_N$-CheY complex are tightly linked to long-range changes in backbone dynamics and to the state of the D57 phosphorylation site as monitored by W58.

Most amino acids are modified by hydroxy radicals to a greater or lesser extent. Therefore it is feasible, given ongoing developments in MS sensitivity and high throughput analyses, to envisage fuller characterization of surface details in the near future. Here, we have used XFMS as a method for focussed measurement of concerted aromatic and associated side-chain motions predicted by simulations. The method complements methods such as fluorescence quenching for solution measurement of side-chain motions, has a more straight-forward physical rationale, provides simultaneous measurements over time-resolved windows and is not limited to the size of the protein assembly. We anticipate it will find increasing use in the mechanistic dissection of allostery in other macromolecular assemblies.

.

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

ACKNOWLEDGMENTS:

We thank 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. Galperin, M. Y. 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. Journal of bacteriology 188(12):4169-4182.
2. Gao, R., S. Bouillet, and A. M. Stock. 2019. Structural Basis of Response Regulator Function. Annual review of microbiology.
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. 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.
5. Warshel, A., and M. Levitt. 1976. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. Journal of molecular biology 103(2):227-249.
6. 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.
7. Tsai, C. J., and R. Nussinov. 2014. A unified view of "how allostery works". Plos Comput Biol 10(2):e1003394. Research Support, N.I.H., Extramural.
8. 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.
9. 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. Journal of molecular biology 378(1):1-11.
10. Bourret, R. B. 2010. Receiver domain structure and function in response regulator proteins. Current opinion in microbiology 13(2):142-149.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.

18. 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.
19. 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. Research Support, N.I.H., Extramural.
20. 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. Research Support, U.S. Gov't, P.H.S.
21. 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.
22. Luo, J., Z. Liu, Y. Guo, and M. Li. 2015. A structural dissection of large protein-protein crystal packing contacts. Sci Rep 5:14214.
23. 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.

24. 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.
25. Merz, G. E., P. P. Borbat, A. R. Muok, M. Srivastava, D. N. Bunck, J. H. Freed, and B. R. Crane. 2018. Site-Specific Incorporation of a Cu(2+) Spin Label into Proteins for Measuring Distances by Pulsed Dipolar Electron Spin Resonance Spectroscopy. The journal of physical chemistry. B 122(41):9443-9451.
26. 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.
27. Pandini, A., A. Fornili, F. Fraternali, and J. Kleinjung. 2012. Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. FASEB J. 26(2):868-881.
28. 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.

29. McEvoy, M. M., A. C. Hausrath, G. B. Randolph, S. J. Remington, and F. W. Dahlquist. 1998. Two binding modes reveal flexibility in kinase/response regulator interactions in the bacterial chemotaxis pathway. Proceedings of the National Academy of Sciences of the United States of America 95(13):7333-7338.
30. 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.

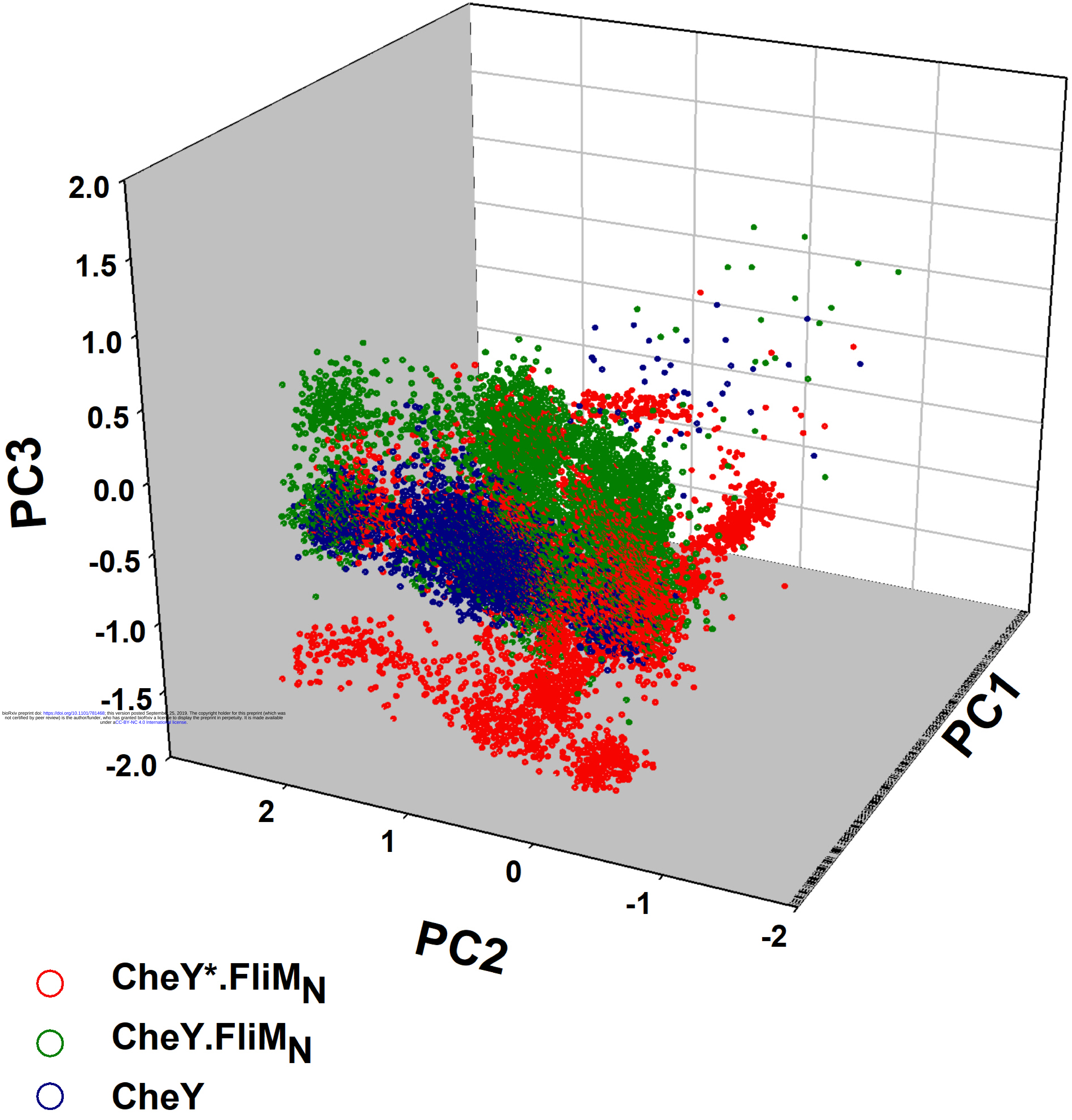
31. Ahn, D. R., H. Song, J. Kim, S. Lee, and S. Park. 2013. The crystal structure of an activated Thermotoga maritima CheY with N-terminal region of FliM. International journal of biological macromolecules 54:76-83.
32. Turner, L., A. D. Samuel, A. S. Stern, and H. C. Berg. 1999. Temperature dependence of switching of the bacterial flagellar motor by the protein CheY(13DK106YW). Biophysical journal 77(1):597-603.
33. 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. Research Support, Non-U.S. Gov't.

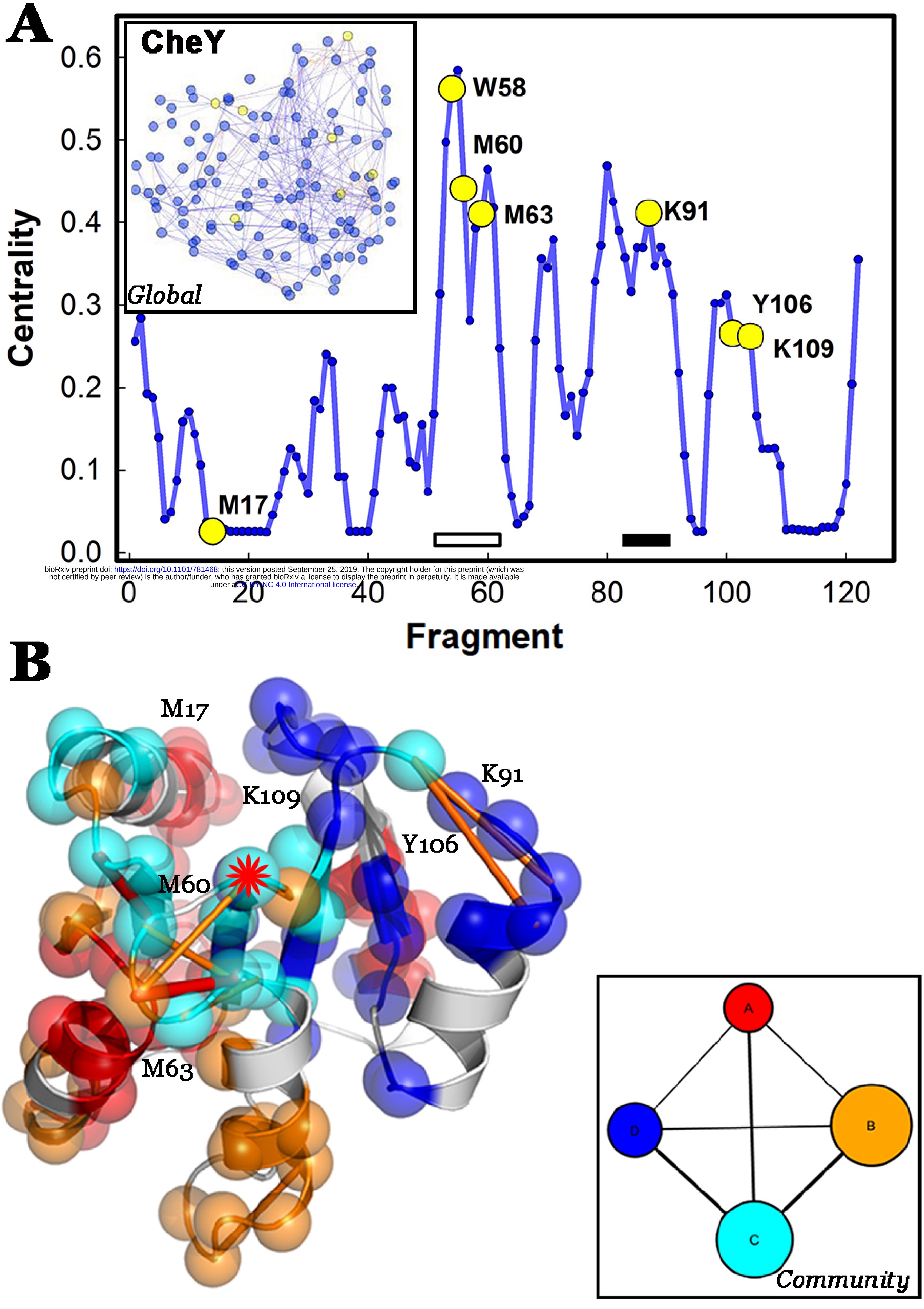
34. Seeliger, D., J. Haas, and B. L. de Groot. 2007. Geometry-based sampling of conformational transitions in proteins. Structure 15(11):1482-1492.
35. Morton, S. A., S. Gupta, C. J. Petzold, and C. Y. Ralston. 2019. Recent Advances in X-Ray Hydroxyl Radical Footprinting at the Advanced Light Source Synchrotron. Protein Pept Lett 26(1):70-75.
36. Xu, G., and M. R. Chance. 2007. Hydroxyl radical-mediated modification of proteins as probes for structural proteomics. Chemical reviews 107(8):3514-3543.
37. 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. Research Support, N.I.H., Extramural.
38. 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.
39. 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.
40. 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.
41. 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.
42. 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.
43. 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.
44. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. J. Chem. Phys. 126(1):014101.
45. Parrinello, M., and A. Rahman. 1981. Polymorphic transitions in single crystals. A new molecular dynamics method. J. Appl. Phys 52(9):7182-7190.
46. 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.
47. 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.
48. Pandini, A., F. Morcos, and S. Khan. 2016. The Gearbox of the Bacterial Flagellar Motor Switch. Structure 24(7):1209-1220.
49. Roulston, M. S. 1999. Estimating the errors on measured entropy and mutual information. Physica D 125:285-294.
50. Pandini, A., A. Fornili, and J. Kleinjung. 2010. Structural alphabets derived from attractors in conformational space. BMC bioinformatics 11:97. Research Support, Non-U.S. Gov't.
51. 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.
52. Newman, M. E. 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. U. S. A. 103(23):8577-8582.

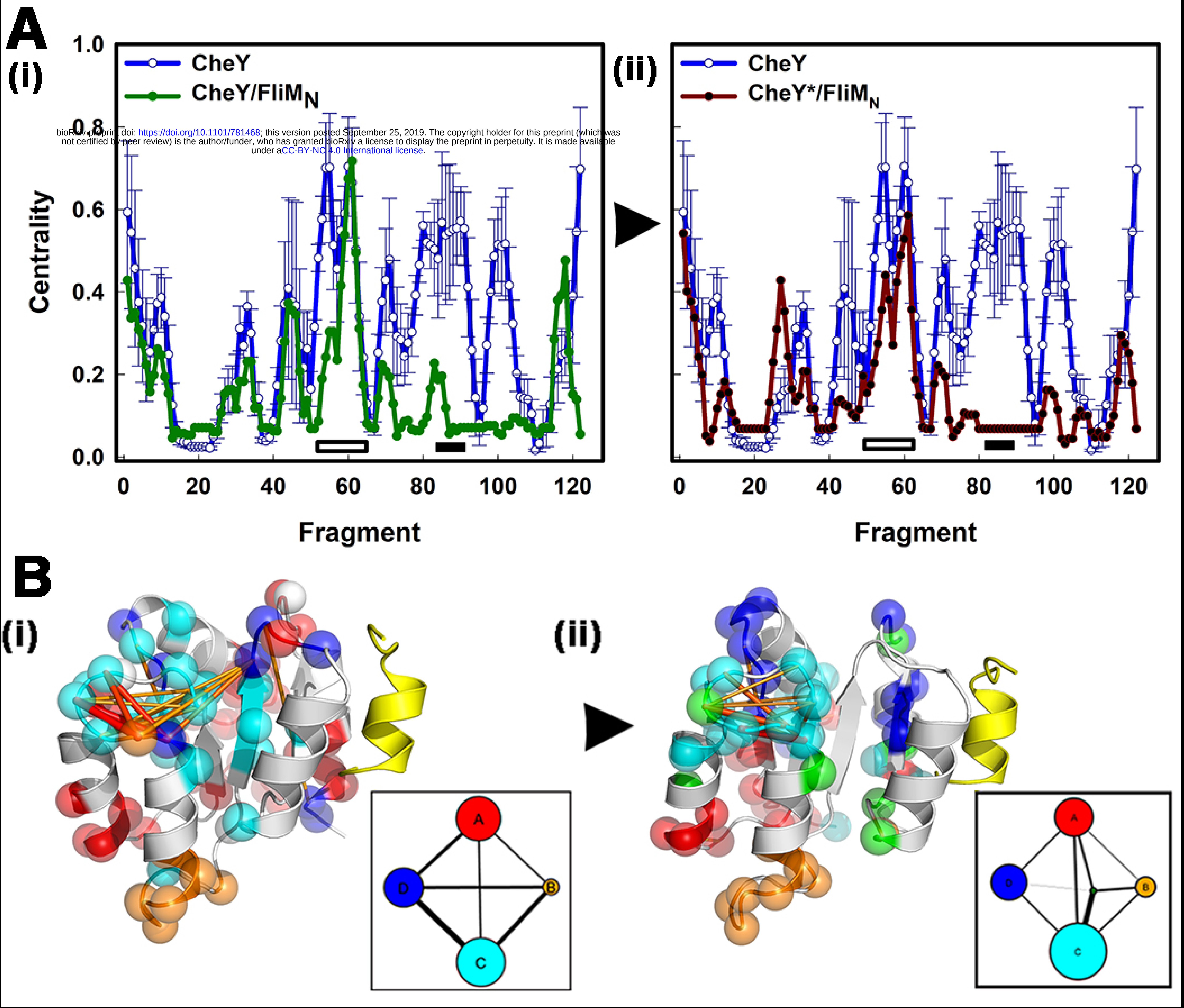
53. Newman, M. E. 2004. Analysis of weighted networks. Phys Rev E Stat Nonlin Soft Matter Phys 70(5 Pt 2):056131.
54. Csardi, G., and T. Nepusz. 2006. The igraph software package for complex network research. InterJournal Complex Systems:1695.
55. Yang, J., R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. 2015. The I-TASSER Suite: protein structure and function prediction. Nature methods 12(1):7-8.
56. 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.
57. 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.
58. 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.
59. 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.
60. 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.
61. Kornev, A. P., and S. S. Taylor. 2015. Dynamics-Driven Allostery in Protein Kinases. Trends in biochemical sciences 40(11):628-647.
62. 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.
63. Pazy, Y., A. C. Wollish, S. A. Thomas, P. J. Miller, E. J. Collins, R. B. Bourret, and R. E. Silversmith. 2009. Matching biochemical reaction kinetics to the timescales of life: structural determinants that influence the autodephosphorylation rate of response regulator proteins. Journal of molecular biology 392(5):1205-1220.
64. 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.
65. 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.
66. Rostkowski, M., M. H. Olsson, C. R. Sondergaard, and J. H. Jensen. 2011. Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. BMC Struct Biol 11:6.
67. Davies, M. J., Dean, R.T. 1997. Radical-Mediated Protein Oxidation: From Chemistry to Medicine. Oxford University Press.
68. Gupta, S., R. D'Mello, and M. R. Chance. 2012. Structure and dynamics of protein waters revealed by radiolysis and mass spectrometry. Proceedings of the National Academy of Sciences of the United States of America 109(37):14882-14887.
69. 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.
70. 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.
71. 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.

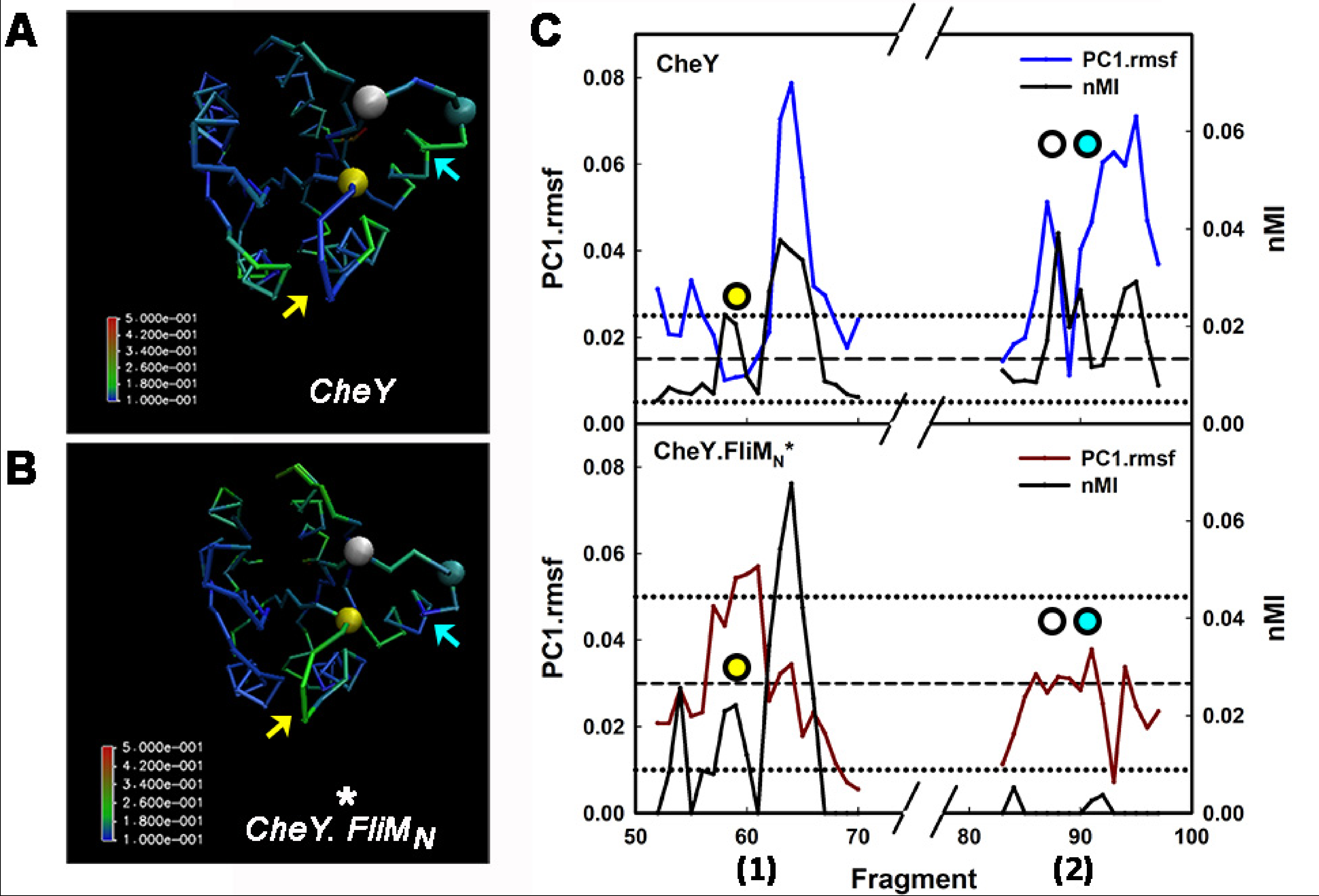
72. Thomas, D. R., N. R. Francis, C. Xu, and D. J. DeRosier. 2006. The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of Salmonella enterica serovar Typhimurium. Journal of bacteriology 188(20):7039-7048.
73. Changeux, J. P. 2013. 50 years of allosteric interactions: the twists and turns of the models. Nat Rev Mol Cell Biol 14(12):819-829.
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.

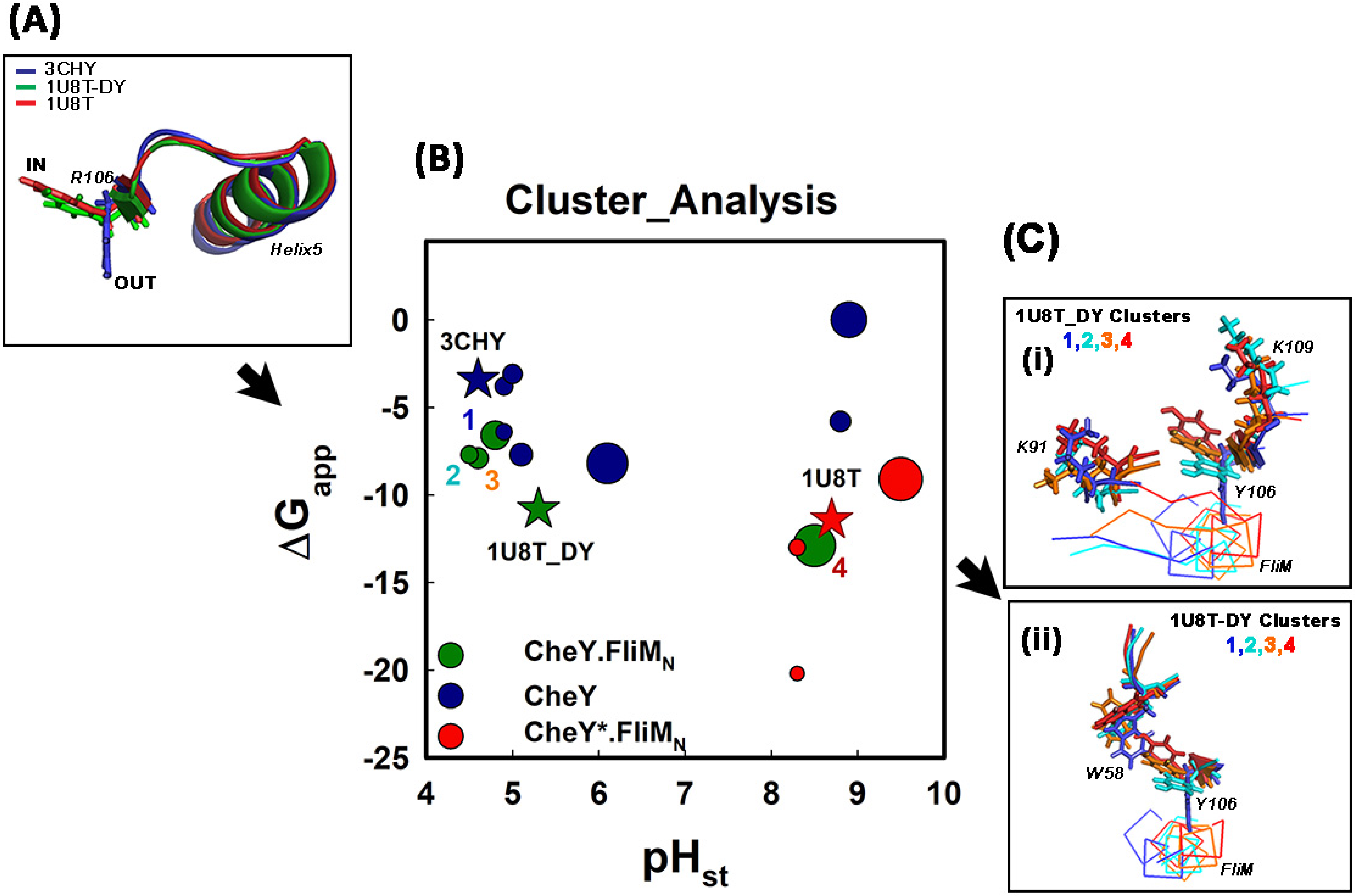

● CheY*/FliM$_N$
● CheY/FliM$_N$
● CheY
RMSF (nm)
0.25
0.20
0.15
0.10
0.05
CheY
FliM$_N$
0
20
40
60
80
100
120
10
20
RESIDUE

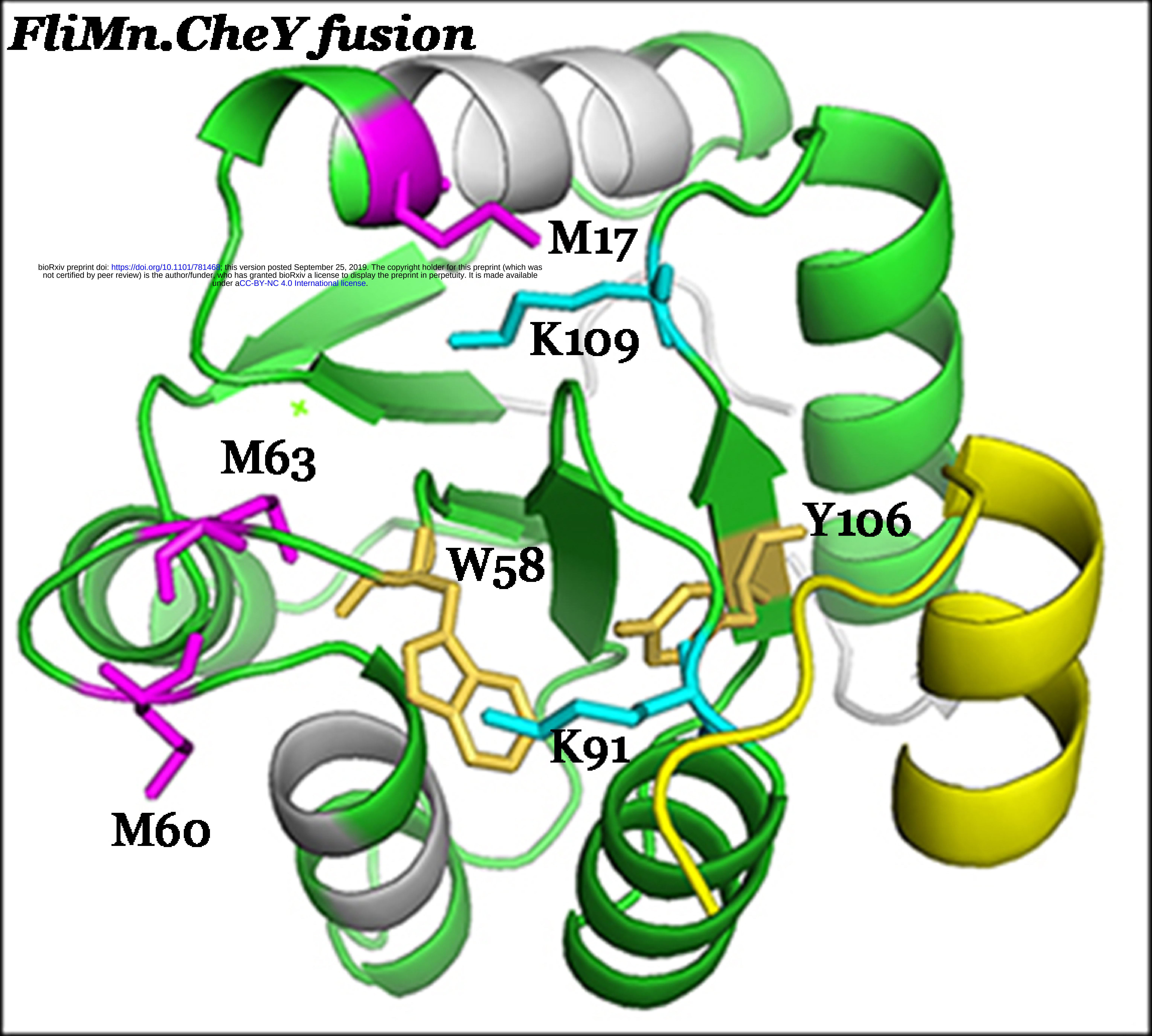

PC3
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
PC2
2
1
0
-1
-2
PC1
○ CheY*.FliM$_N$
○ CheY.FliM$_N$
○ CheY

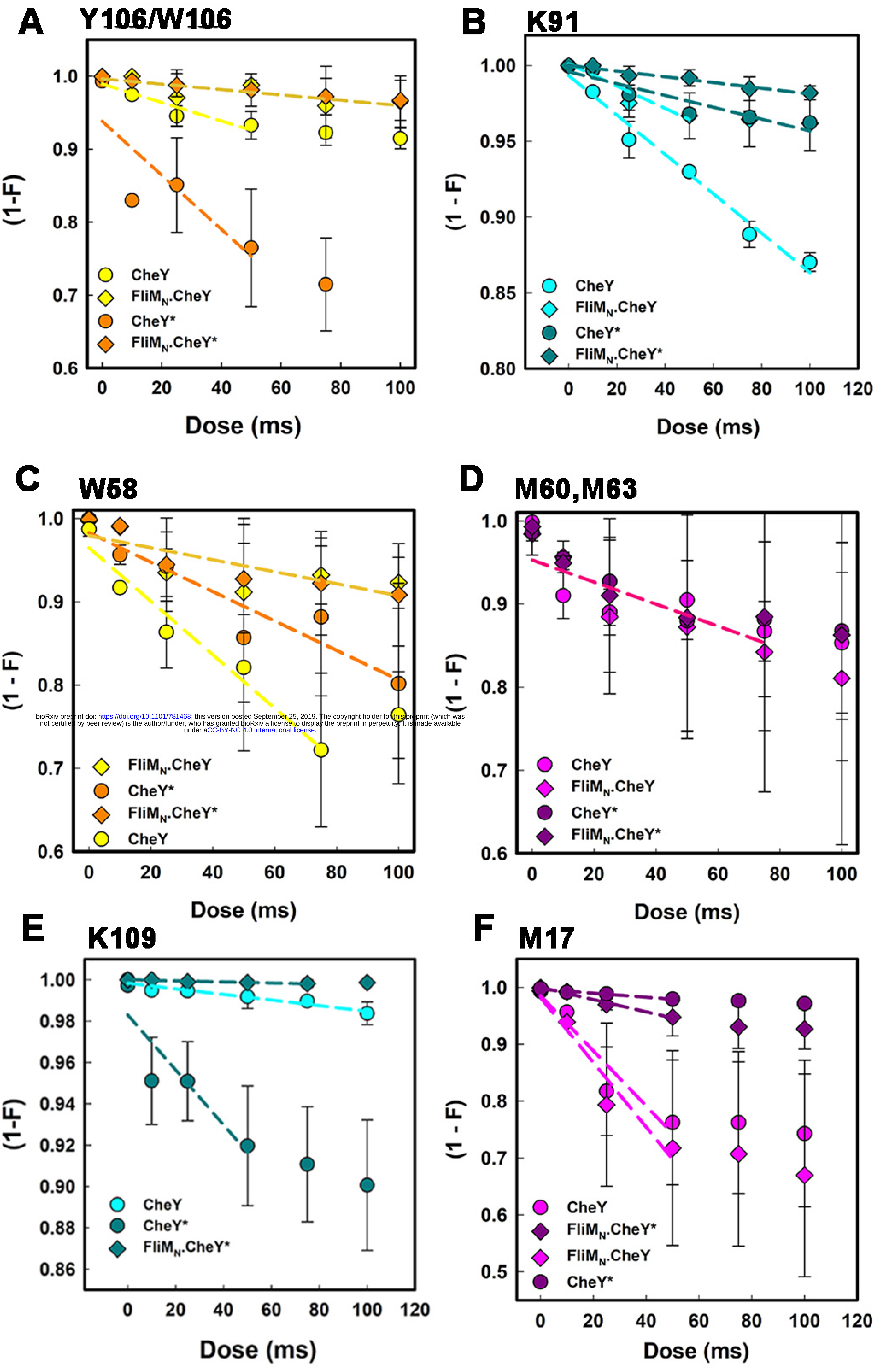

A
CheY
Global
Centrality
0.6
0.5
0.4
0.3
0.2
0.1
0.0
W58
M60
M63
K91
Y106
K109
M17
0
20
40
60
80
100
120
Fragment
B
M17
K109
K91
Y106
M60
M63
A
B
C
D
Community


A
(i)
CheY
CheY/FliM$_N$
Centrality
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
Fragment
(ii)
CheY
CheY*/FliM$_N$
0
20
40
60
80
100
120
Fragment
B
(i)
A
B
C
D
(ii)
A
B
C
D

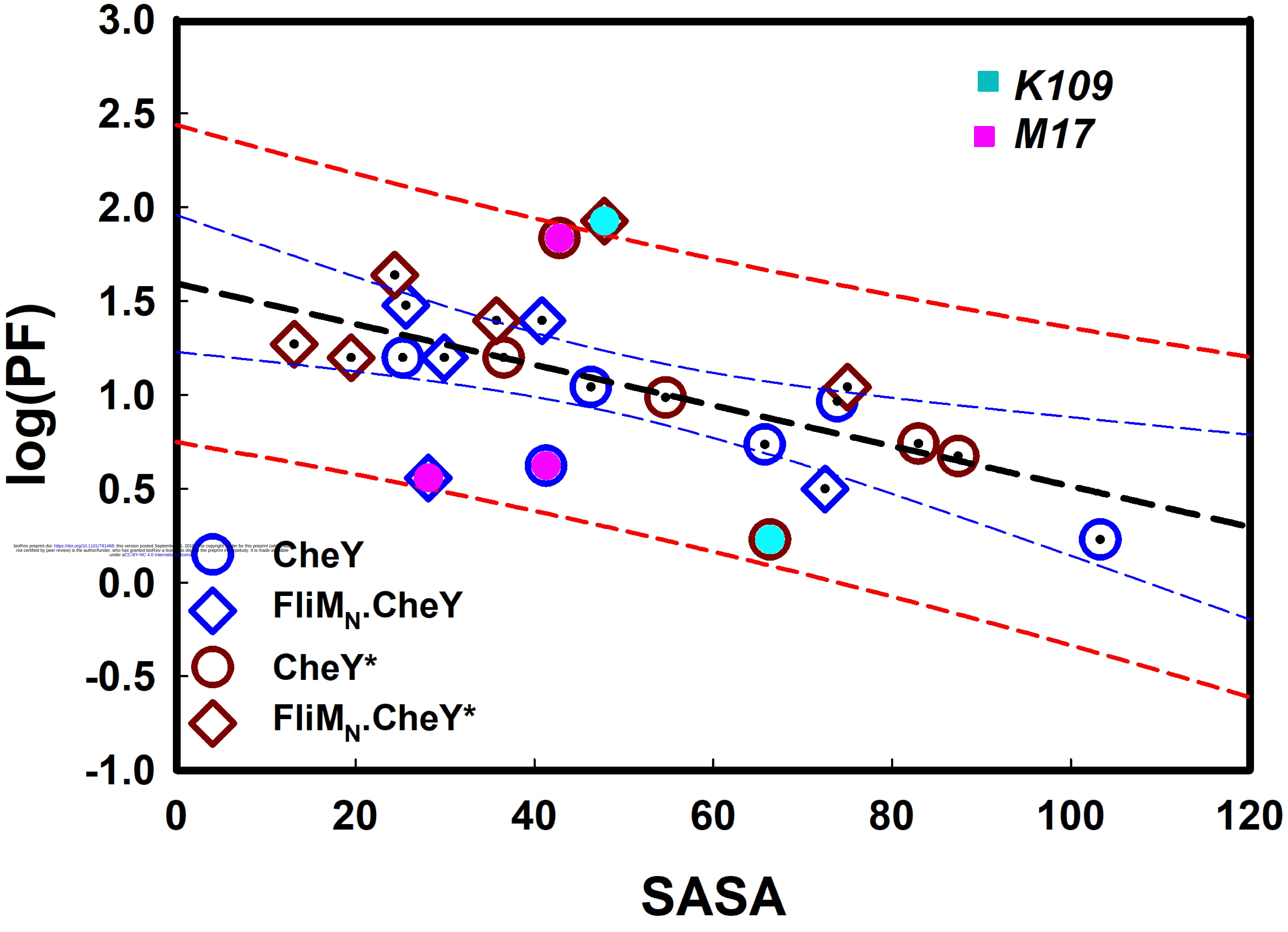

A
5.000e-001
4.200e-001
3.400e-001
2.600e-001
1.800e-001
1.000e-001
CheY
B
5.000e-001
4.200e-001
3.400e-001
2.600e-001
1.800e-001
1.000e-001
*
CheY. FliM$_N$
C
CheY
PC1.rmsf
nMI
0.08
0.06
0.04
0.02
0.00
0.06
0.04
0.02
0.00
CheY.FliM$_N$*
PC1.rmsf
nMI
0.08
0.06
0.04
0.02
0.00
0.06
0.04
0.02
0.00
PC1.rmsf
PC1.rmsf
nMI
nMI
50
60
70
80
90
100
(1)
Fragment
(2)

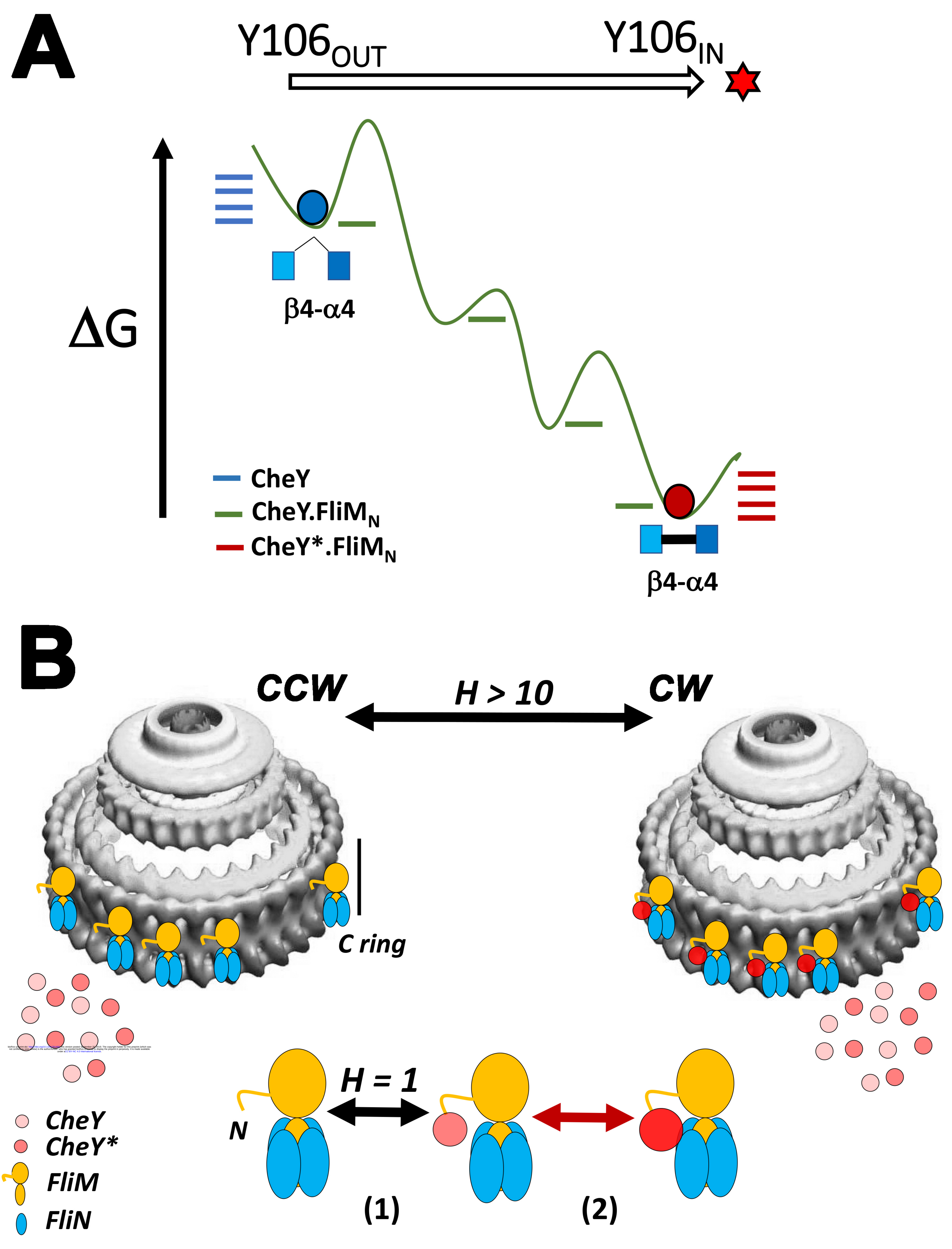

(A)
3CHY
1U8T-DY
1U8T
IN
R106
Helix5
OUT
(B)
Cluster_Analysis
3CHY
1
2
3
1U8T_DY
1U8T
4
CheY.FliM$_N$
CheY
CheY*.FliM$_N$
ΔG$_{app}$
pH$_{st}$
0
-5
-10
-15
-20
-25
4
5
6
7
8
9
10
(C)
1U8T_DY Clusters
1,2,3,4
(i)
K109
K91
Y106
FliM
(ii)
1U8T-DY Clusters
1,2,3,4
W58
Y106
FliM


FliMn.CheY fusion
M17
K109
M63
Y106
W58
K91
M60



A
Y106/W106
(1-F)
1.0
0.9
0.8
0.7
0.6
0
20
40
60
80
100
Dose (ms)
CheY
FliM$_N$.CheY
CheY*
FliM$_N$.CheY*
B
K91
(1 - F)
1.00
0.95
0.90
0.85
0.80
0
20
40
60
80
100
120
Dose (ms)
CheY
FliM$_N$.CheY
CheY*
FliM$_N$.CheY*
C
W58
(1 - F)
1.0
0.9
0.8
0.7
0.6
0
20
40
60
80
100
Dose (ms)
FliM$_N$.CheY
CheY*
FliM$_N$.CheY*
CheY
D
M60,M63
(1 - F)
1.0
0.9
0.8
0.7
0.6
0
20
40
60
80
100
Dose (ms)
CheY
FliM$_N$.CheY
CheY*
FliM$_N$.CheY*
E
K109
(1-F)
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0
20
40
60
80
100
120
Dose (ms)
CheY
CheY*
FliM$_N$.CheY*
F
M17
(1 - F)
1.0
0.9
0.8
0.7
0.6
0.5
0
20
40
60
80
100
120
Dose (ms)
CheY
FliM$_N$.CheY*
FliM$_N$.CheY
CheY*


K109
M17
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
log(PF)
CheY
FliM$_N$.CheY
CheY*
FliM$_N$.CheY*
0
20
40
60
80
100
120
SASA


A
Y106$_{OUT}$
Y106$_{IN}$
ΔG
β4-α4
β4-α4
CheY
CheY.FliM$_N$
CheY*.FliM$_N$
B
CCW
H > 10
CW
C ring
H = 1
N
(1)
(2)
CheY
CheY*
FliM
FliN