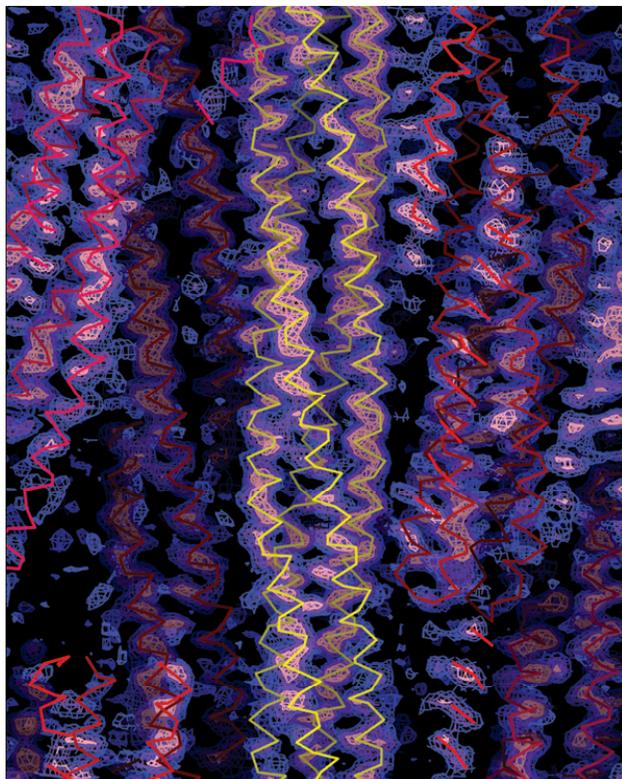
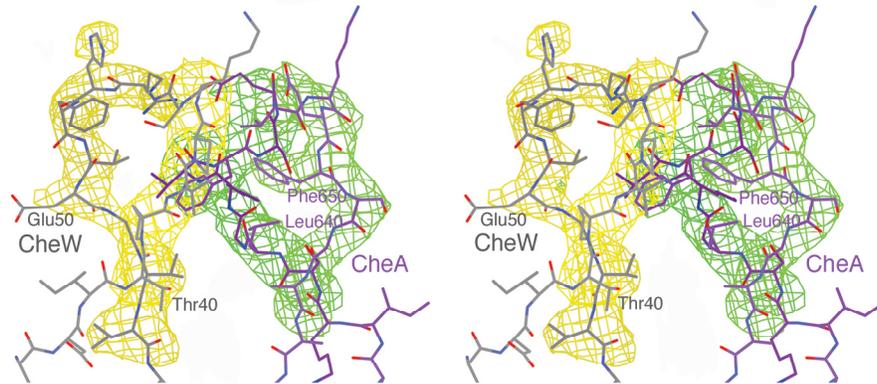


**Figure 1.** Initial electron-density map for the MCP<sub>1143c</sub> crystal structure.



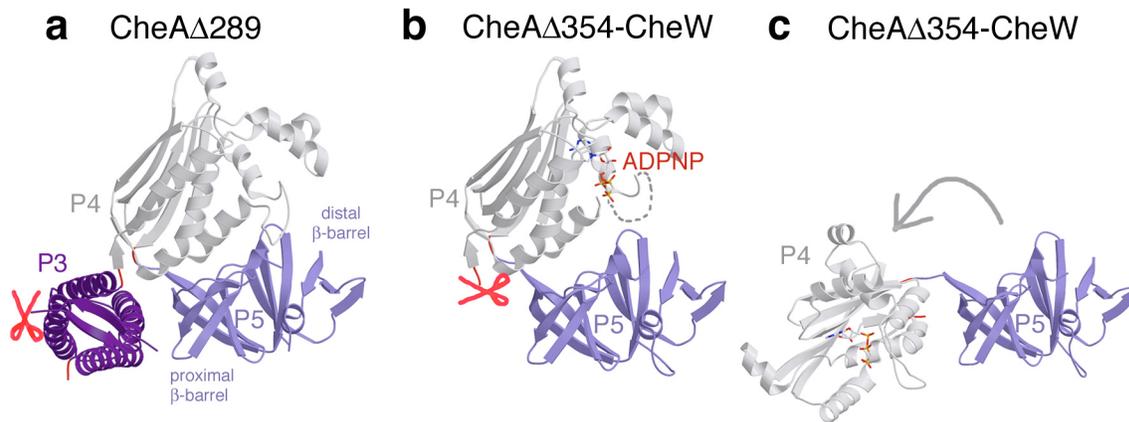
Experimental Pb-phased 2.5 Å electron density map contoured at 1 standard deviation ( $\sigma$ ) in blue and  $3\sigma$  in pink for the MCP<sub>1143c</sub> structure (yellow C $\alpha$  traces for part of one MCP<sub>1143c</sub> dimer, red traces for symmetry-related molecules).

**Figure 2.** Omit electron-density map for the CheA 354-CheW crystal structure.



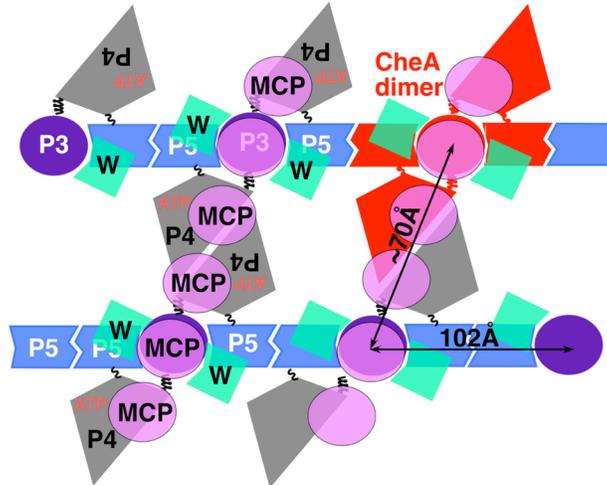
Stereoview of a 3.5 Å resolution  $F_o-F_c$  omit map contoured at  $3.5\sigma$  with either residues 40-53 from CheW (yellow density) or residues 639-653 of CheA (green density) removed from  $F_c$ . A portion of the interface between the two proteins is shown.

**Figure 3.** P4 domain orientations in the CheA 354-CheW crystallographic complex.



(a) Relative positions of P3 (purple), P4 (grey), and P5 (blue) in the crystal structure of dimeric CheA 289. Red scissors show point of N-terminal truncation. In the CheA 354:CheW structure, one molecule in the asymmetric unit (b) has a similar orientation of P5 and P4 compared to the CheA 289 structure, whereas the other (c) contains P4 in a position that would overlap with P3 in CheA 289.

**Figure 4.** The CheA-CheW-MCP array.



A schematic diagram for the kinase:receptor array to complement the detailed representation of **Fig. 4d** (same view). Close packing of the CheA layers generates contacts between the P4 domains and allows for two additional receptors to associate with each CheA-CheW complex. The receptors peripheral to the core complex may interact with the P4 domains. This model is based on the following assumptions:

- 1) three receptors bind one dimeric CheA-CheW, as defined by our structural studies
- 2) the interactions surfaces on CheW and the receptors defined by biochemistry and genetics must match,
- 3) the P5-P5 contacts we find in crystals and solution associate CheA molecules,
- 4) the receptor bundles interact in a closed pack fashion, as they do in the crystal, and
- 5) the modification sites must be accessible to the adaptation enzymes.

At higher receptor stoichiometries, additional receptor molecules could be accommodated between the CheA-CheW layers.

# METHODS

## A) Crystallography

### Crystallographic Structure Determination and Refinement

**MCP<sub>1143c</sub>**- Patterson analysis revealed two lead atoms per asymmetric unit in the MCP<sub>1143c</sub> crystals. Anomalous differences from the lead sites were weak, probably due to anisotropic diffraction (all helical bundles align in the crystal) and non-spherical spot shapes. Diffraction data from two separate multiwavelength anomalous diffraction (MAD) experiments were combined in SOLVE<sup>1</sup> to generate the initial experimental phases (Figure of Merit = 0.62). MAD data from Seleno-Methionine modified receptors did not contribute appreciably to the quality of the initial map, but anomalous difference Fourier maps made from anomalous differences taken at the peak wavelength of the Se-edge were essential to confirm the helical register of the model. With this information the model was built manually in XFIT<sup>2</sup> in a 3.5 Å SOLVE generated map, solvent-flattened in RESOLVE<sup>1</sup>. The final model (one MCP<sub>1143C</sub> dimer made of residues A225-A529 and B225-B528, two Pb sites, and 586 waters) was then refined against a 2.5 Å data set with CNS<sup>3</sup> to R-factor = 0.232 and R<sub>free</sub> = 0.279 for F > 2σ(F). All residues have backbone geometries within expected bounds of stereochemistry.

**CheA/CheW**: The position of the *T. maritima* CheA (P4-P5) and CheW were determined by molecular replacement using a P4 domain, a P5 domain (PDB code: 1B3Q)<sup>4</sup> and the NMR structure of CheW (PDB code: 1K0S)<sup>5</sup> as search models with PHASER<sup>6</sup>. Automated search routines of PHASER found one molecule of P4 and one molecule of P5 in the asymmetric unit. Map inspection followed by rigid body refinement, allowed placement of one CheW bound to the first P5 domain, and then, in a subsequent cycle, an additional P5-CheW complex. The second P4 domain could be placed only after model rebuilding and refinement. The final model has two CheA P4-P5 domains and two CheW proteins in the asymmetric unit and was refined with CNS to R-factor = 0.250 and R<sub>free</sub> = 0.288. All residues have backbone geometries within expected bounds of stereochemistry.

## B) Pulsed ESR

### Instrumentation

The 4-pulse DEER and 6-pulse DQC ESR techniques were executed at the 17.5 GHz frequency of Ku band (waveguide P-band). For DEER capability, a 2D-FT ESR spectrometer<sup>7</sup>

was modified to contain a dielectric resonator coupled to a Ku-band waveguide, all enclosed by a CF935 helium flow cryostat (Oxford Instruments, Ltd.). The microwave component was modified to include a microwave oscillator and a pulse-forming channel, followed by a Ku-band traveling-wave tube amplifier (TWTA). Pulses from this assembly were passed through a variable attenuator and injected into the main waveguide by means of the off-shelf 10 dB directional coupler. Use of two TWTA's eliminates instrumental cross-talk between pump and observation pulses in DEER, leaving only contributions from spectral excitation overlap to the remaining coupling. Pumping frequency  $\omega_B$  was positioned at the center maximum of the nitroxide spectrum 65 MHz above detection frequency,  $\omega_A$ ;  $\pi/2$  and  $\pi$  pulses were 16 and 32 ns in DEER and 3.2 and 6 ns in DQC, respectively. For most experiments, DEER was usually employed because of satisfactory signal-to-noise-ratio (SNR), reduced problems with deuterium modulation, and less signal attenuation in the presence of more than two spins compared to DQC. However, for short distances or low spin concentration, DQC improved SNR by a factor of 2-3.

### **Labeling Considerations**

Due to dimeric nature of CheA, single residue substitution to cysteine produced two labeling sites. Such mutations were used to obtain inter-subunit distances between the symmetry related-residues. Double mutations, however, resulted in two such pairs and 4 distances possible for symmetrical conformations. Mutations were made such that inter-subunit distances were significantly longer than the intra-subunit, inter-domain distance of interest. Signals then could be readily separated. Although all four (or generally six) distances could be in principle determined, protein conformational freedom and spin-label side chain flexibility produced complex signals that were difficult to deconvolute. Protein concentrations were typically in range of 15-150  $\mu\text{M}$ , but spin-labeling efficiency led to reduced concentration of double-labeled protein by a factor of 1.5-4, depending on labeling site. Complexes of double-labeled CheA with WT CheW or vice versa probed how complex formation perturbs local structure. In general, the effects were insignificant. The intra protein distances were all consistent with the expectations based on the X-ray and NMR structures and favored geometries of the nitroxide side-chain.

### **Metric Matrix Distance Geometry**

Metric matrix distance geometry<sup>8</sup> was applied with the distance information provided from spin labels at three positions on CheA $\Delta$ 289 (N553C, D579C and E646C – all on the P5

domain) and CheW (S15C, S72C, S80C). In all 11 out of 12 inter-protein distances were assigned numeric values (a distance between CheA $\Delta$ 289 S568C and CheW S15C was outside the ESR range). Distance constraints from S568C on CheA $\Delta$ 289 to all three CheW positions were intentionally left out of the matrix to cross-validate the solution.

With one CheW label site chosen arbitrarily as the origin, a 5x5 metric matrix was generated from the measured distances, and diagonalized with MATHEMATICA. All five eigenvalues showed positive values (7026, 543, 351, 209, 10) due to uncertainty in the measurements. The three largest eigenvalues were used to calculate the coordinates of the three label sites on CheA $\Delta$ 289 and two label sites on CheW; two possible mirror-image related coordinate sets were generated. Least square fitting between the three site coordinates calculated from distance geometry and the C $\alpha$  atoms of the labeled residues on the CheA $\Delta$ 289 structure (PDB code: 1B3Q) and the three CheW coordinates and their corresponding positions on a rigid CheW structure (taken from the NMR ensemble PDB code: 1K0S) produced the complex. Only one of the two mirror-related solutions showed a crash-free model of the association between CheA and CheW. The three distances left out of the calculation were predicted within a 5.1 Å rmsd by the model.

### **Rigid-body refinement of complexes under ESR constraints**

Coordinates of the P5-CheW complex determined by distance geometry were optimized by conjugate-gradient rigid-body refinement in CNS<sup>3</sup> under the restraints of the ESR distances. These distances were treated like NMR NOE restraints with lower limit errors set to 4 Å for distances < 40 Å, 6 Å for distances 40-50 Å and 10 Å for distances > 50 Å; and upper-limit errors set to 1 Å for distances < 50 Å, and 2 Å for distances > 50 Å. The structure was refined until the gradient converged. To test convergence, CheW was displaced manually in various directions and the complex refined. Within rigid-body displacements of ~15 Å and rotations of degrees 30 degrees, the same unique solution was found. These methods applied to the CheA dimer under restraints of the inter-subunit P5 distances suggest that in solution P5 moves slightly towards P3 compared to the structure of CheA $\Delta$ 289.

### **Inter-domain distances**

An extensive set of inter-domain, inter-subunit and inter-protein distance measurements were made to evaluate the structure of CheA $\Delta$ 289 and CheA $\Delta$ 289-CheW complex in dilute solution. In a few cases, heterodimers were constructed so that intra-subunit distances could be determined in absence of contaminating signals from labels related by the dimer symmetry.

A feature of CheA clearly evident in these experiments is the mobility of the individual domains about their respective hinge regions. P4 tends to be quite mobile (as evidence by the E387(A)/E387(B) distance and broad distributions between P4-label sites and those on other domains). P5 is less mobile due to contacts with P3.

We also find that in the absence of CheA, CheW will interact with itself at an affinity in the ranges of tens of micromolar (a far weaker interaction than the nM dissociation constant of *T. maritima* CheA $\Delta$ 289 and CheW). The tendency of crystalline CheA $\Delta$ 289 and CheA $\Delta$ 354 dimers to associate with each other via the P5 domains is also supported by CheA self-associations as the protein concentration increases in the ESR experiments. Distances presented in Table 1 were carried out under conditions of dilute solution where these additional interactions were minimized. Nevertheless, the tendency of the studied proteins to associate very likely had some impact (rather weak but non-negligible) on measured inter-domain distances. The crystal structure of CheA $\Delta$ 289<sup>4</sup> shows some asymmetry in the positions of P4 and P5, which are tethered by short mobile hinges. At 50-200  $\mu$ M we observe moderate mobility of these domains that is probably within the amplitude of change between subunits found in the asymmetric crystal structure. It should be noted that the ESR distances suggest that P5 (in both the free CheA $\Delta$ 289 dimer, and the CheW- $\Delta$ 289 complex) assumes an average position relative to P3 slightly different but within +/-10 Å of that predicted by the CheA $\Delta$ 289 crystal structure. The position of P5 may well be affected by P5/P5 contacts in higher oligomers. Increasing concentration above 500  $\mu$ M shows a larger network or aggregation of currently unknown nature; the local spin concentration, as indicated by the intermolecular part of the DEER signal, was tripled to that expected from uniform solution, qualitatively suggesting at least 3 CheA dimers per group.

## C) Figures

Figures were made with Molscript<sup>9</sup> and rendered with Raster3D<sup>10</sup>.

## D) References

1. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Cryst.* **D55**, 849-861 (1999).
2. McRee, D.E. XtalView: a visual protein crystallographic software system for X11/Xview. *J. Mol. Graph.* **10**, 44-47 (1992).
3. Brunger, A.T. et al. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**, 905-921 (1998).

4. Bilwes, A.M., Alex, L.A., Crane, B.R. & Simon, M.I. Structure of CheA, a signal-transducing histidine kinase. *Cell* **96**, 131-141 (1999).
5. Griswold, I.J. et al. The solution structure and interactions of CheW from *Thermotoga maritima*. *Nature Struct. Biol.* **9**, 121-125 (2002).
6. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. & Read, R.J. Likelihood-enhanced fast translation functions. *Acta Cryst.* **D61**, 458-464 (2005).
7. Borbat, P.P., Crepeau, R.H. & Freed, J.H. Multifrequency two-dimensional Fourier transform ESR: an X/Ku-band spectrometer. *J. Magn. Reson.* **127**, 155-67 (1997).
8. Crippen, G.M. & Havel, T.F. *Distance Geometry and Molecular Conformation*, (John Wiley & Sons, New York, NY, 1988).
9. Kraulis, P.J. Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946-950 (1991).
10. Merritt, E.A. & Murphy, M.E.P. Raster3D Version 2.0: a program for photorealistic molecular graphics. *Acta Crystallogr.* **D50**, 869-873 (1994).