

In different organisms, the mode of interaction between two signaling proteins is not necessarily conserved

Sang-Youn Park*, Bryan D. Beel†, Melvin I. Simon†, Alexandrine M. Bilwes*, and Brian R. Crane**

*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14850; and †Division of Biology, California Institute of Technology, Pasadena, CA 91125

Edited by Janet Thornton, European Bioinformatics Institute, Cambridge, United Kingdom, and approved June 24, 2004 (received for review February 13, 2004)

Although interfaces mediating protein–protein interactions are thought to be under strong evolutionary constraints, binding of the chemotaxis histidine kinase CheA to its phosphorylation target CheY suggests otherwise. The structure of *Thermotoga maritima* CheA domain P2 in complex with CheY reveals a different association than that observed for the same *Escherichia coli* proteins. Similar regions of CheY bind CheA P2 in the two systems, but the CheA P2 domains differ by an $\approx 90^\circ$ rotation. CheA binds CheY with identical affinity in *T. maritima* and *E. coli* at the vastly different temperatures where the respective organisms live. Distinct sets of P2 residues mediate CheY binding in the two complexes; conservation patterns of these residues in CheA and compensations in CheY delineate two families of prokaryotic chemotaxis systems. A protein complex that has the same components and general function in different organisms, but an altered structure, indicates unanticipated complexity in the evolution of protein–protein interactions and cautions against extrapolating structural data from homologs.

It is generally accepted that residues involved in functional protein–protein interactions are more stringently conserved across protein families than other surface residues that are not under the constraint of maintaining molecular recognition (1–5). Structure-informed sequence analyses (1, 2, 4–9) and experiment (10–13) demonstrate that interaction surfaces of protein partners coevolve. It is then reasonable to expect that related proteins will conserve a mode of association with their conserved binding targets. Indeed, this is often the case when comparing structures of complexes composed from homologs (7, 9).

In *Escherichia coli* chemotaxis, the most extensively characterized chemotaxis system, transient protein–protein interactions propagate changes in extracellular chemical concentrations to flagellar activity (for reviews, see refs. 14–18). At the center of this so-called “two-component” signaling pathway is the cytoplasmic histidine kinase CheA and its substrate, the response regulator CheY. In response to changes in receptor occupancy, homodimeric CheA uses ATP to autophosphorylate a conserved histidine residue on its phosphotransferase domain. The phosphoryl group is then transferred from this histidine to a conserved aspartyl residue on the response-regulator protein, CheY. Phosphorylated CheY interacts directly with the FlhM protein of the flagellar motor and thereby affects swimming behavior. In addition to CheY, CheA may also phosphorylate the response regulator domain of the methyltransferase CheB. CheB functions in mediating adaptation to constant levels of attractant or repellent.

Of the five domains that comprise CheA (P1–P5) (19, 20), P1 and P4 produce the kinase activity that transfers phosphate from ATP to the P1 histidine. P2 docks response regulators for interactions with P1. P3 dimerizes the CheA subunits and P5 mediates CheW binding and receptor coupling. Herein we report the crystallographic structure of the complex between

Thermotoga maritima CheA P2 domain and CheY at 1.9-Å resolution. The interface formed between the *T. maritima* proteins differs considerably from that previously seen for the homologous *E. coli* proteins (21–23). This is surprising given that CheA and CheY are well conserved throughout a wide range of chemotactic bacteria and that their association is essential for signaling (24).

Methods

Protein Preparation. The genes encoding *T. maritima* full-length CheA, domains P1–P2 (residue 1–264), and CheY were PCR cloned into the vector pET28a (Novagen) and expressed with a 6-Histidine tag in *E. coli* strain BL21 (DE3) (Novagen). Using Terrific Broth (Difco) as the medium and kanamycin (25 $\mu\text{g}/\text{ml}$), the proteins were purified as described (19). His tags were removed by thrombin digestion. After concentration by centrifugation (Amicon), the CheA and CheY proteins were mixed at an $\approx 1:2$ ratio (dimeric CheA:CheY) and run on a Superdex200 26/60 sizing column. The elution profile gave two major peaks that represented the CheA–CheY complex and free CheY. The high molecular weight complex was concentrated to ≈ 30 mg/ml in gel filtration buffer (50 mM Tris, pH 7.5/150 mM NaCl).

Crystallization and Data Collection. Initial conditions for growing of the CheA P2–CheY complex crystal were found in commercial screening solutions (Hampton Research, Riverside, CA) after 1 month at room temperature. Crystals of the CheA P2–CheY complex grew by vapor diffusion against a reservoir of 25–30% polyethylene glycol 12 K and 0.1 M HEPES, pH 7.5. SDS/PAGE confirmed CheY had crystallized with a 17-kDa proteolytic fragment of full-length CheA. Diffraction data were collected under a 100 K nitrogen stream at the Cornell High Energy Synchrotron Source beamline (F2) on a charge-coupled device Quantum detector, by using a cryo protectant of reservoir solution and 5% glycerol. Crystals belong to the space group $P2_12_12$ and contain one CheA P2 domain complexed to one CheY molecule per asymmetric unit. Data were processed by DENZO and SCALEPACK (25).

Structure Determination and Refinement. The position of the *T. maritima* CheY protein was determined by molecular replacement with free CheY as a search model [Protein Data Bank (PDB) code: 1TMY] by using EPMR (26) and diffraction data from 4.0- to 25.0-Å resolution (correlation coefficient = 0.416,

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ITC, isothermal titration calorimetry; ΔC_p , change in heat capacity; EP2, *Escherichia coli* CheA P2 domain; TP2, *Thermotoga maritima* CheA P2 domain; TCheY, *T. maritima* CheY; PDB, Protein Data Bank; ECheY, *E. coli* CheY.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1U05).

†To whom correspondence should be addressed. E-mail: bc69@cornell.edu.

© 2004 by The National Academy of Sciences of the USA

Table 1. Data collection and refinement statistics

Data collection	
Space group	P2 ₁ 2 ₁ 2
	(a=88.6 Å, b=91.4 Å, c=31.7 Å)
Resolution range (highest shell), Å	30–1.9 (2.0–1.9)
Unique reflections	20,209
Completeness overall (highest shell), %	99.7 (99.7)
R _{merge} * (highest shell)	0.097 (0.273)
Wilson B, Å ²	22.7
Refinement	
R factor† (R _{free})‡	0.227 (0.253)
No. of atoms (no. of residues, waters)	1,604 (204, 201)
rms deviation bonds, Å; angles, °	0.006, 1.2
Average B factor main chain (side chain, water), Å ²	24.4 (26.3, 42.7)

*R_{merge} = $\sum \sum_j |I_j - \langle I \rangle| / \sum \sum_j I_j$.

†R factor = $\sum (|F_{obs}| - |F_{calc}|) / \sum |F_{obs}|$.

‡R factor for 10% of the reflections selected at random and excluded from refinement.

R factor = 0.518). The automated refinement software ARP/WARP (27) improved the electron-density map, into which the CheA P2 domain was constructed. The final model (CheY residues 2–119 and P2 residues 176–260) was built manually with XFIT (28) and further refined with CNS at a resolution of 1.9 Å (Table 1).

Isothermal Titration Calorimetry (ITC). To determine precise extinction coefficients (ϵ), BCA assay (Pierce) and reductant compatible/denatured compatible assay (Bio-Rad) were used with BSA and cytochrome *c* standards for P1–P2 ($M_r = 32.9$ kDa) and CheY ($M_r = 13.2$ kDa). ϵ_{280} values for P1–P2 domain ($9,700 \text{ M}^{-1}\text{cm}^{-1}$) and CheY ($6,500 \text{ M}^{-1}\text{cm}^{-1}$) were taken as averages from both assays. Calorimetric measurements were carried out by using a VP-ITC titration calorimeter (Microcal, Amherst, MA) at temperatures of 18–70°C. Before titration, samples of P1–P2 and CheY were dialyzed against gel filtration buffer supplemented with 10 mM DTT. At temperatures between 25°C and 35°C, we observed a minor secondary interaction site of TP2

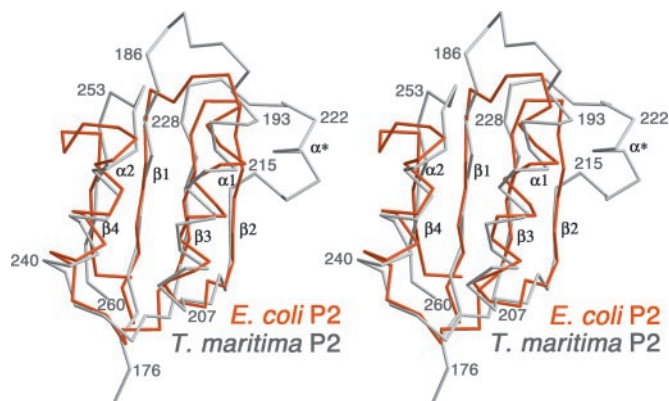


Fig. 1. Stereoview of C_α trace superposition for the P2 domains of *E. coli* (orange) and *T. maritima* CheA (gray). Secondary structural elements are reasonably conserved. However, *T. maritima* P2 has a helical insertion and a larger loop between β2 and β3, as well as a larger loop between β1 and α1 (rms deviation of 2.0 Å for 57 of 67 C_αs from EP2 with 57 of 86 C_αs from TP2). Figs. 1, 3, 5, and 6 were rendered with BOBSCRIPT (51).

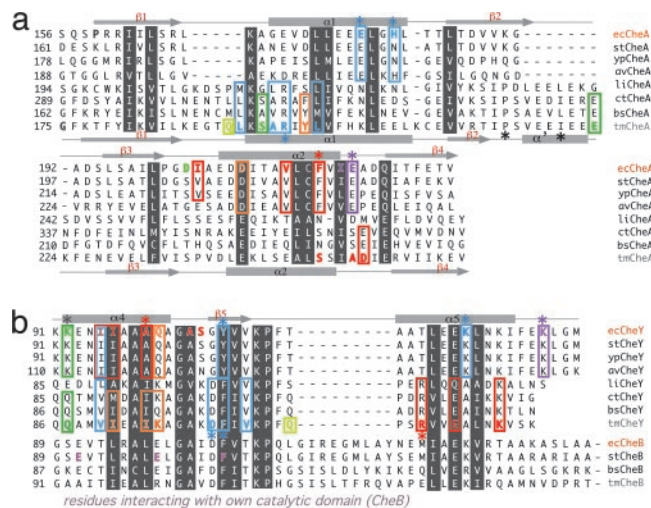


Fig. 2. Aligned amino acid sequences of CheA P2 (a) and CheY and CheB (b). (a) Residues mediating the P2–CheY interaction (colored blue, red, orange, purple, green, and yellow to define groups of interactions) delineate two structural families exemplified by the *E. coli* and *T. maritima* complexes. Conserved residues boxed with the same color in CheA and CheY contact each other through side chains. Colored residues that are not boxed interact through their main chain only. Colored stars highlight interface residues that are highly conserved in one family but not in the other and hence provide signatures for the two types of interactions. The four archaea genomes classified under family 2 have CheA P2 sequences closely related to *T. maritima* than *E. coli*, but their CheY sequences (not shown) contain the “GY” motif instead of “DF” (*T. maritima* residues 100 and 101). Black stars identify invariant residues within the *T. maritima* family likely important for the folding of helix α'. Four examples from each family are given. (b) CheB proteins conserve key residues involved in the *T. maritima* TP2–TcheY interaction. Residues involved in the interface between the response regulator and catalytic domains of *Salmonella typhimurium* CheB are shown in pink. Family 1, *E. coli* (ec), *Shigella flexneri*, *Yersinia pestis* (yp), *Burkholderia fungorum*, *Salmonella typhimurium* (st), *Ralstonia metallidurans*, *Ralstonia solanacearum*, and *Azotobacter vinelandii* (av). Family 2, *T. maritima* (tm), *Clostridium tetani* (ct), *Bacillus subtilis* (bs), *Clostridium acetobutylicum*, *Bacillus halodurans*, *Clostridium thermocellum*, *Thermoanaerobacter tengcongensis*, *Bacillus cereus*, *Listeria innocua*, *Archaeoglobus fulgidus*, *Pyrococcus horikoshii*, *Pyrococcus abyssi*, *Leptospira interrogans* (li), *Methanosarcina mazei*, and *Desulfitobacterium hafniense*. Bacterial genomes were searched with BLAST for organisms that contained P2 domains within CheA sequences and also a response regulator protein with high sequence similarity to CheY.

on *T. maritima* CheY (TcheY). In this temperature range, the reported thermodynamic values are for the major interaction site, which has thermodynamic parameters that correspond to the only site at physiological temperatures. Thermodynamic parameters were determined with the ORIGIN software package (Microcal). A positive change in heat capacity (ΔC_p) was estimated in a limited temperature range from the slope of change in enthalpy (ΔH) versus temperature.

Results and Discussion

P2 Domain Structures. NMR and crystallographic studies have shown the *E. coli* CheA P2 domain (EP2) to be an open-face β-sandwich with four antiparallel β-strands packed against two antiparallel helices (Fig. 1) (21–23, 29). The *T. maritima* CheA P2 domain (TP2) shares the topology of EP2, but TP2 has a helical insertion α' and a larger loop between β2 and β3, as well as a larger loop between β1 and α1 (Figs. 1 and 2a). Despite significant overall sequence identity between the *T. maritima* and *E. coli* full-length CheAs (34%), there is little sequence identity ($\approx 10\%$) between the two P2 domains. In fact, P2 is the least conserved domain of the CheA protein (30). Nevertheless, the *E. coli* and *T. maritima* proteins show significant structural

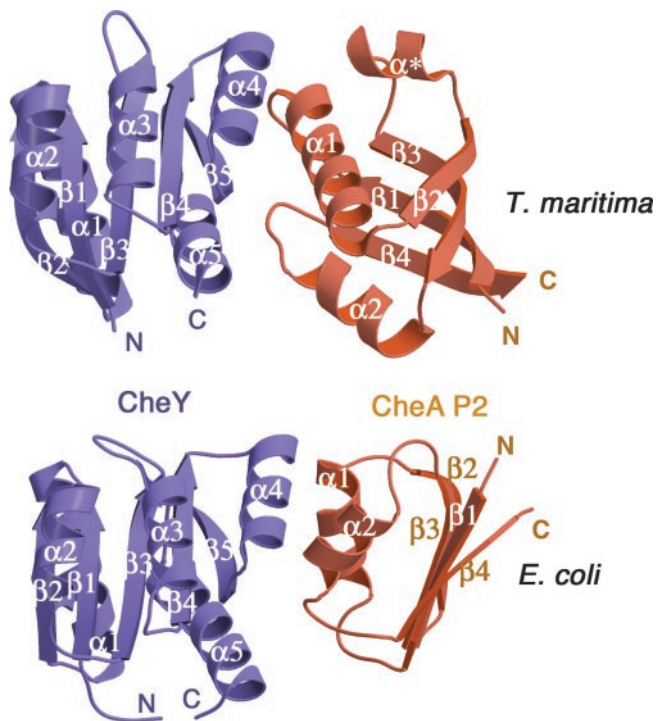


Fig. 3. CheY binds P2 differently in *T. maritima* than in *E. coli*. CheY (blue ribbons) binds CheA P2 domain (orange ribbons) in orientations that differ by $\approx 90^\circ$ for the proteins from *T. maritima* and *E. coli* [note positions of N (amino) and C (carboxyl) termini].

equivalence in the peptide backbone of their central β -sheets and two α -helices (Fig. 1).

CheY Structures. The structure of TCheY, which has been previously determined for the free response regulator (31) (PDB code 4TMY), is very similar to that of *E. coli* CheY (ECheY), for which there are now numerous structures in free, activated, and various complexed forms (refs. 22 and 32 and references therein). Association with TP2 results in little structural change in TCheY (rms deviation of 1.3 Å for all C_α s). There are also only modest differences between TCheY and ECheY within the P2–CheY complexes, and they result mainly from crystal contacts in the loop comprising residues 56–60 of TCheY [comparing TCheY in our structure to ECheY in PDB coordinates 1EAY (21) and 1A0O (23) gives rms deviations of 1.3–2.0 Å for 119 C_α atoms]. The greater sequence and structural conservation of the CheY proteins from *E. coli* and *T. maritima* (29% sequence identity, Fig. 2) compared to the P2 domains may result from additional conserved interactions of CheY with the CheA P1 domain and the flagellar switch protein FlIM (33).

Structure of the P2–CheY Complex. The complexes formed between CheA P2 and CheY from *E. coli* and *T. maritima* involve the same faces on each protein and similar regions on CheY, but the respective P2 proteins differ by an $\approx 90^\circ$ rotation in orientation (Figs. 3 and 4). Consequently, the P2 residues binding CheY in the two complexes differ in sequence, position, and type (Fig. 2). In contrast, two CheY residues central to both interfaces have the same sequence position but are of different type: ECheY Gly-105–Tyr-106 and TCheY Asp-100–Phe-101 (Fig. 2). In both complexes, the two proteins associate along their respective helical faces to orient CheY so that its phosphate-accepting Asp-57 residue projects away from the P2 interface for interaction with phosphorylated P1 (Fig. 3). However, in the *E. coli*

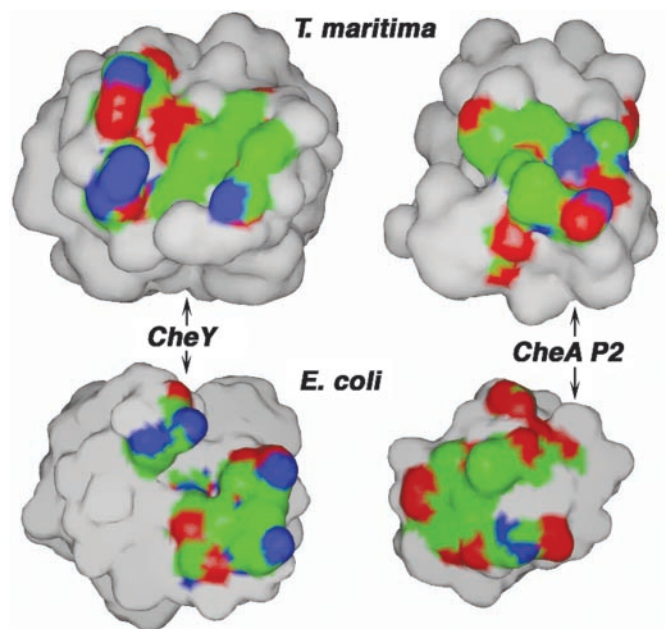


Fig. 4. Footprint of the P2–CheY interfaces for *T. maritima* and *E. coli*. Buried surface areas in the P2–CheY interfaces colored by residue type (carbon surface area, green; nitrogen, blue; oxygen, red). Solvent accessible surface area colored in gray. Rotating the molecules toward the center closes each interface and reconstitutes the complexes.

complex, the $\alpha 4$ – $\beta 5$ – $\alpha 5$ region of CheY runs perpendicular to P2 helices $\alpha 1$ and $\alpha 2$ (21, 23), whereas in the *T. maritima* complex, CheY $\beta 5$ inserts between and runs parallel to P2 helices $\alpha 1$ and $\alpha 2$ (Fig. 3).

The two interfaces also differ in the details of their intermolecular contacts, yet retain similar overall chemical compositions (Fig. 4). The P2–CheY complex of *T. maritima* buries more solvent-accessible surface area (1,506 Å²) than the *E. coli* complex [1,214 Å², PDB code: 1FFG (22)], primarily due to the rotated orientation and larger size of the TP2 domain. The compositions and areas of both buried surfaces are typical of those found in transient heterodimers (2, 3). Interprotein interactions include 5 hydrogen bonds, 7 hydrophobic contacts and one π stacking for the *T. maritima* complex, and 11 hydrogen bonds and 6 hydrophobic contacts for that of *E. coli* (Fig. 4). The amount of hydrophobic buried surface area in the TP2–TCheY complex (852 Å², 56%) is larger than, but the same proportion as, that buried in EP2–ECheY (685 Å², 56%). The higher surface complementarity (34) of the TP2–TCheY interface ($S_c = 0.82$) compared to that of EP2–ECheY ($S_c = 0.73$) even exceeds complementarity of interfaces formed by high-affinity protein/protein inhibitor complexes ($S_c = 0.72$ – 0.76) (34). Such complete meshing of TP2 with TCheY may relate to the thermostability of this interaction (see below).

Central hydrophobic contacts and peripheral hydrogen bonding interactions stabilize the interfaces of both complexes (Fig. 5). At the center of each interface, the P2 proteins provide a pocket for the insertion of a conserved aromatic residue: Tyr-106 in ECheY and Phe-101 in TCheY (Fig. 4). ECheY phosphorylation reduces the affinity of CheY for P2 by favoring rotation of Tyr-106 away from the P2 interface and into an internal conformation (35). Within the EP2–ECheY complex, Tyr-106 hydrogen bonds to EP2 $\alpha 1$ residues Glu-178, and His-181. ECheY Gly-105 allows Tyr-106 to contact EP2 Phe-214 (EP2 $\alpha 2$) edge on (Fig. 4b). In TCheY, Phe-101 replaces ECheY Tyr-106 and also may serve to couple changes in CheY phosphorylation to destabilization of the TP2 contact (31). Four residues on TP2

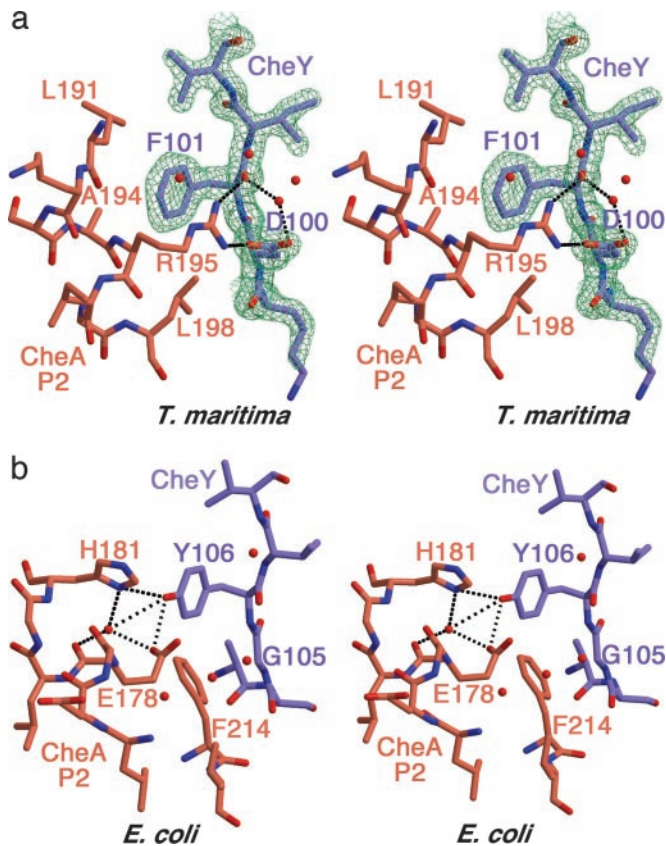


Fig. 5. Stereoviews of important interactions in the CheY-P2 interfaces of *T. maritima* (a) and *E. coli* (b). In CheY (blue), F101 (*T. maritima*), and Y106 (*E. coli*) bind into pockets provided by the respective P2 domains (orange). (a) In *T. maritima*, four residues on P2 α 1 encapsulate F101: R195, A194, L191 and L198. (b) In *E. coli*, Y106 hydrogen bonds to P2 E178 and interacts with H181. $2F_{\text{obs}} - F_{\text{calc}}$ electron density map shown for CheY at 1.9-Å resolution (green contours at 1.5 σ).

α 1 encapsulate TCheY Phe-101 and fix the rotated orientation of TP2 relative to TCheY: Ala-194 and Arg-195 stack with the phenyl ring, whereas Leu-191 and Leu-198 enclose the ring edges. A buried salt bridge between TCheY Asp-100 and TP2 Arg-195 anchors the insertion of Phe-101 into the pocket provided by TP2 α 1. In both the *E. coli* and *T. maritima* complexes, hydrophobic interactions between CheY α 4 and the side of P2 α 1 further sequester the binding pocket for Phe-101/Tyr-106. Interprotein hydrogen bonds peripheral to these central interactions primarily involve the C-terminal region of P2 α 2 but have different residue composition in the two complexes (Fig. 2).

Interestingly, the interaction between TCheY and TP2 α 1 mimics the interaction of ECheY with a 16-residue helical peptide from N terminus of the *E. coli* flagellar switch protein FliM (33). Superposition of the TCheY-TP2 complex with the ECheY-FliM peptide complex reveals that TP2 α 1 and the helical FliM peptide bind the same α 4- β 5 region of TCheY and ECheY, respectively (Fig. 6). This site on ECheY contains the Phe-101 (Tyr-106) residue that changes from an external to an internal position in response to CheY phosphorylation (35). Due to sequence differences on the TP2 and FliM helices, TP2 α 1 binds unphosphorylated TCheY with Phe-101 externalized, whereas the FliM peptide binds phosphorylated ECheY with Tyr-106 internalized. *T. maritima* FliM (FliM) conserves 14 of the 16 residues found in the *E. coli* FliM peptide. Thus, in *T. maritima* and related organisms, binding targets of CheY may contain a relatively simple structural determinant on a con-

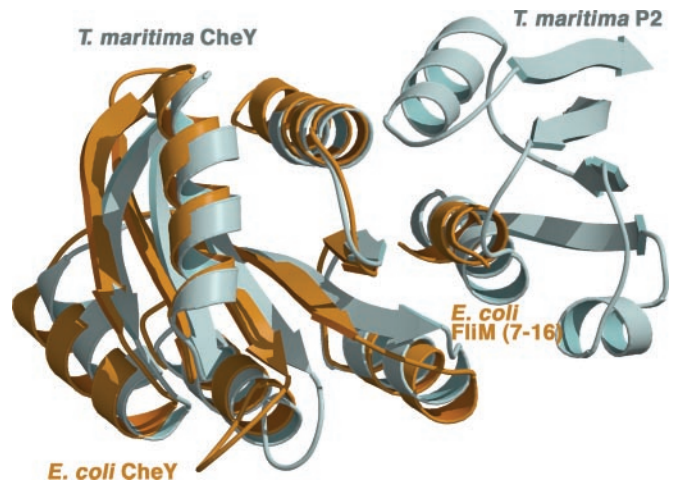


Fig. 6. *T. maritima* P2 binds the same site on CheY as does an *E. coli* FliM fragment. Superposition of TCheY complexed with CheA TP2 (gray) and activated ECheY complexed with FliM residues 7–16 (1F4V, orange) (33). Only CheY C $_{\alpha}$ atoms were superimposed.

served helical motif to discriminate phosphorylated from unphosphorylated CheY.

Thermodynamics of CheA-CheY Binding. ITC was used to investigate whether the altered orientation of the *T. maritima* CheA P2-CheY complex correlates with a change in the thermodynamics of binding. As in previous experiments with the *E. coli* proteins (36), binding titrations were carried out between CheY and a monomeric fragment composed of CheA domains P1 and P2. At 28°C, the dissociation constants of the *E. coli* and *T. maritima* complexes (1.3 and 0.2 μ M, respectively) are similar to each other and are lower than the cellular CheY concentration in *E. coli* [8 μ M (37)]. The comparable affinities at 28°C were unexpected because there is an $\approx 40^\circ$ difference in the temperatures where the two sets of proteins function and both binding constants are likely to be temperature dependent. A positive value for ΔC_p is expected with an increase in exposed hydrophobic surface area upon complex dissociation (38). In fact, the temperature dependence of the change in enthalpy (ΔH) $\approx 25^\circ\text{C}$ indicates that both P2-CheY interactions have a similar positive ΔC_p for dissociation [0.51 kcal \cdot mol $^{-1}\cdot$ K $^{-1}$ for TP2-TCheY and 0.22 kcal \cdot mol $^{-1}\cdot$ K $^{-1}$ for EP2-ECheY (36)]. These values agree reasonably well with empirical calculations based on the differences in exposed hydrophobic and hydrophilic surface areas in the complexes and free proteins (39) (0.25 kcal \cdot mol $^{-1}\cdot$ K $^{-1}$ for TP2-TCheY and 0.18 kcal \cdot mol $^{-1}\cdot$ K $^{-1}$ for EP2-ECheY). Nevertheless, the TP2-TCheY K_d is relatively temperature independent because of significant temperature dependence in ΔC_p itself (Fig. 7). This results in essentially equivalent affinities of CheA for CheY in *E. coli* and *T. maritima* at the respective temperatures where each organism lives (37°C and 80°C, Table 2). We tested the binding of TP2 to ECheY and to ECheY in which the two central interface residues were mutated to their correspondents in TCheY (ECheY Tyr-106-Phe, Gly-105-Asp) and saw no detectable interaction in either case.

The similarity in CheA P2-CheY binding affinities for *T. maritima* and *E. coli* at all temperatures investigated rules out the possibility that the structural divergence between them derived from an adaptation to a different temperature range. Furthermore, most of the P2 residues involved in CheY binding are conserved by a subfamily of chemotactic bacteria, many of which are not thermophiles.

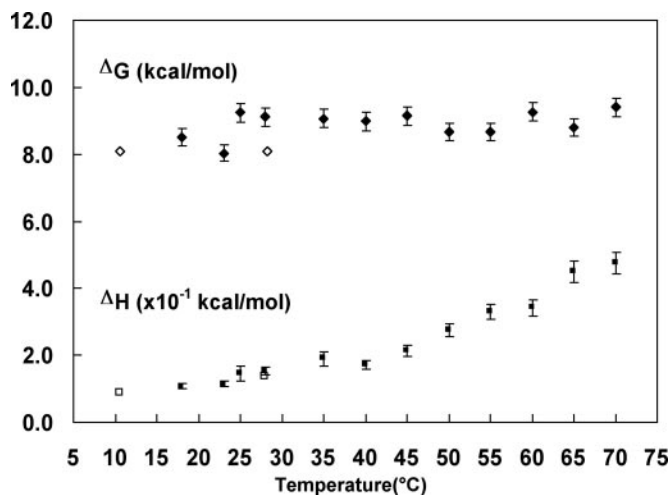


Fig. 7. Temperature dependence of the free energy and enthalpy associated with dissociation of the *T. maritima* P1P2-CheY complex. ITC derived ΔG (kcal·mol⁻¹, \blacklozenge) and change in enthalpy (ΔH) ($\times 10^{-1}$ kcal·mol⁻¹, \blacksquare) values for TP2-CheY dissociation from 10°C to 70°C. Previously reported *E. coli* values (24, \square) are shown for comparison (36). Linear changes in ΔH from 18°C to 35°C were used to estimate a ΔC_p of 0.51 (± 0.1) kcal·mol⁻¹·K⁻¹ at this range.

P2-CheY Interactions Delineate Two Families of Chemotaxis Proteins.

PSI-BLAST searches (40) of the available bacterial genome sequences for P2 homologs reveal at least two distinct families of CheA proteins. The first family, exemplified by *E. coli* CheA, conserves P2 residues involved in binding CheY with the orientation shown by the *E. coli* CheA P2-CheY structure (Fig. 2). The second family of P2 proteins, exemplified by *Bacillus subtilis* and *T. maritima*, conserve residues involved in binding CheY with the orientation shown by the *T. maritima* CheA P2-CheY structure. In all, 23 organisms were found by nested BLAST searches that contained both CheA P2 and CheY homologs: 8 in the *E. coli* family and 15 in the *B. subtilis* family (Fig. 2). The CheY partners in these organisms also conserve the residues required to distinguish the two orientations of binding. From this collection of sequences, the *B. subtilis* family appears to be more widespread and includes examples from Archaea. Importantly, organisms that are not thermophiles have CheA P2 domains predicted by sequence conservation to bind CheY in the *T. maritima* mode. (Fig. 2).

Despite conserving many chemotaxis proteins between them, *E. coli* and *B. subtilis* show differences in their mechanisms of chemotaxis (41). In both organisms, phosphorylated CheY (CheY-P) directly affects translocation by mediating changes in the direction of flagellar rotation, and an increase in attractant concentration increases the probability of smooth swimming. However, this response results from a decrease in CheY-P in one case (*E. coli*) and an increase in CheY-P in the other (*B. subtilis*) (42). *T. maritima* belongs in the *B. subtilis* chemotaxis family based on its complement of chemotaxis proteins and the extent of sequence similarity of these proteins to their homologs in *B. subtilis*. Given that CheA appears to have opposite regulation

with respect to attractant in *E. coli* and *B. subtilis*, it may be significant that the P2 domain of CheA binds CheY differently in *T. maritima* (a member of the *B. subtilis* family) than in *E. coli*. *T. maritima* and *B. subtilis* also contain CheC, a protein that resembles the flagellar motor protein FlhM, yet has been shown to bind CheA in *B. subtilis* (43). Thus, additional interactions with other chemotaxis proteins may also relate to the differences between the two families of CheA kinases.

CheB Binding to CheA in *T. maritima*. In *E. coli*, many of the CheY residues that interact with P2 are not conserved by the homologous response regulator domain of CheB (21), despite CheY and CheB having similar affinities for P2 (36). This suggested that P2 binds CheB differently than CheY (21). In support of an alternative binding mode for CheB, the structure of *Salmonella typhimurium* CheB revealed that the CheB catalytic domain blocks the surface of the CheB response regulator domain that corresponds to the CheY P2-binding site (44). Phosphorylation of the CheB regulatory domain destabilizes this interaction with the catalytic domain and exposes the active site in a mechanism that could parallel release of P2 by phosphorylated CheY (35, 44, 45). Interestingly, the response regulator domains of CheB proteins from both the *E. coli* and *B. subtilis* subfamilies conserve residues that are key for binding P2 in the orientation observed in the *T. maritima* complex. For example, the Asp-Phe sequence on the CheB equivalent of $\beta 5$ is conserved by all CheB sequences, as are the aliphatic hydrophobic residues on $\alpha 4$ that are important for interaction with TP2 $\alpha 1$ (Fig. 2) (44). Thus we tested, with ITC, the ability of full-length *T. maritima* CheB and its isolated N-terminal CheY-like regulatory domain to bind TP2 and full-length *T. maritima* CheA. Surprisingly, unlike EP2, which has an affinity for full-length CheB comparable to that for CheY (36), no binding was observed between TP2 or full-length CheA with either full-length CheB or its regulatory domain alone (within the detection limit of ITC: $K_d < 1$ mM). This suggests that either the interaction between *T. maritima* CheA and CheB may be solely mediated by phosphorylated P1 or the role of CheB in *T. maritima* chemotaxis is different from in *E. coli*.

Implications for the Evolution of Protein-Protein Interactions. If protein A binds protein B, the homologs A' and B' usually conserve the A-B mode of interaction, as long as it is constrained by function. P2 may solely serve to increase the effective concentration of CheY in the vicinity of the P1 domain (24). With the P2 domain removed, CheA can still phosphorylate P1, but the resulting reduction in phosphotransfer is likely to significantly affect chemotaxis (24). Thus, the exact form of the P2-CheY complex may not matter, provided it renders the phospho-accepting Asp accessible to the P1 phospho-donor. Nevertheless, how the two CheA-CheY interactions diverged is puzzling because to maintain binding affinity, mutations on one partner must be matched by mutations in the other. Attached domains are not under such constraints. For example, the homologous transcription-factor response regulators DrrB and DrrD have evolved different contacts and juxtapositions between their regulatory and effector domains (46), but domain

Table 2. Dissociation thermodynamic data of CheA P1-P2 domain and CheY for *E. coli* and *T. maritima*

Organism	Temperature, °C	N [†]	K _a , M ⁻¹	K _d , M	ΔH, kcal·mol ⁻¹	ΔG, kcal·mol ⁻¹	ΔS, cal mol ⁻¹ ·K ⁻¹
<i>E. coli</i> *	28.3	0.93	0.8 ± 0.1	1.3 ± 0.2	12.6 ± 0.5	8.1 ± 0.1	15 ± 2
<i>T. maritima</i>	28.0	0.73	4.3 ± 0.4	0.23 ± 0.02	15.2 ± 0.9	9.14 ± 0.05	20 ± 3
<i>T. maritima</i>	70.0	0.63	1.0 ± 0.2	1.0 ± 0.2	48 ± 2	9.4 ± 0.2	111 ± 6

*Values averaged from ref. 31.

[†]Binding stoichiometry.

fusion ensures proximity even in the presence of destabilizing mutations. In contrast, noncovalent protein interactions tend to conserve binding modes. In fact, *in vitro* selection experiments demonstrated convergence to similar interactions between non-covalent partners (47). Furthermore, compensating mutations within interfaces can identify protein/protein interactions from sequence (4–6). The CheA–CheY complexes from *E. coli* and *T. maritima* represent an extreme example of how divergent binding modes result from central residue substitution and peripheral residue conformational change. Smaller, but similar, perturbations in a transient protein complex have resulted from mutagenesis studies (10). Perhaps the robust nature of the chemotaxis system that insulates it against changes of individual parameters (48) allowed the divergence of the CheA–CheY interaction through states where binding was weakened.

On the other hand, the divergence of the CheA–CheY interaction could also derive from a divergence in mechanism of action. Varied binding modes for the cytokine receptors with their hormone ligands (e.g., growth hormone, erythropoietin, IFN- γ) allow specificity and tuning of cross talk between signaling pathways (49, 50). These associations, although somewhat different among partners, involve similar surfaces on related proteins. More variation is seen in the homodimeric assemblies of the homologous chemokines IL-8 (PDB code: 3IL8) and MCP-1 (PDB code: 1DOM). In these proteins, structurally

related subunits associate differently into homodimers and presumably enable different signaling properties (2). Likewise, the SH2 and SH3 domains of signaling kinases interact with each other in three very different ways (9). We do not know of another example like CheA and CheY where such a drastic change in assembly occurs between two partner proteins that have the same function in different organisms. It remains to be seen whether the two subclasses of P2-CheY complexes relate to the opposite effects of attractant on CheA activation in *E. coli* and *B. subtilis*.

The finding that interacting proteins may not necessarily conserve the mode of association found for their functionally equivalent homologs has important implications for predicting protein–protein interactions from patterns of residue conservation, gene organization, and known structures. Recent statistical analyses indicate that below 30–40% sequence identity, pairs of interacting proteins can show different assembly orientations (9). The varied interactions of CheA and CheY provide a striking example of the potential pitfalls of extrapolating functional understanding from the structures of homologs when their sequence similarity is not high.

We thank Cristian Gradinaru for assistance and the Cornell High Energy Synchrotron Source for data collection facilities. This work was supported by National Institutes of Health Grant GM066775 (to B.R.C.).

1. Valdar, W. S. J. & Thornton, J. M. (2001) *Proteins* **42**, 108–124.
2. Nooren, I. M. A. & Thornton, J. M. (2003) *J. Mol. Biol.* **325**, 991–1018.
3. Lo Conte, L. L., Chothia, C. & Janin, J. (1999) *J. Mol. Biol.* **5**, 2177–2198.
4. Valencia, A. & Pazos, F. (2002) *Curr. Opin. Struct. Biol.* **12**, 368–373.
5. Lichtarge, O. & Sowa, M. E. (2002) *Curr. Opin. Struct. Biol.* **12**, 21–27.
6. Yao, H., Kristensen, D. M., Mihalek, I., Sowa, M. E., Shaw, C., Kimmel, M., Kavraki, L. & Lichtarge, O. (2003) *J. Mol. Biol.* **326**, 255–261.
7. Nooren, I. M. A. & Thornton, J. M. (2003) *EMBO J.* **22**, 3486–3492.
8. Valdar, W. S. J. & Thornton, J. M. (2001) *J. Mol. Biol.* **313**, 399–416.
9. Aloy, P., Ceulemans, H., Stark, A. & Russell, R. B. (2003) *J. Mol. Biol.* **332**, 989–998.
10. Atwell, S., Ulsch, M., DeVos, A. M. & Wells, J. A. (1997) *Science* **278**, 1125–1128.
11. Jespers, L., Lijnen, H. R., Vanwetswinkel, S., Van Hoef, B., Brepoels, K., Collen, D. & De Maeyer, M. (1999) *J. Mol. Biol.* **290**, 471–479.
12. Goh, C. S., Bogan, A. A., Joachimiak, M., Walther, D. & Cohen, F. E. (2000) *J. Mol. Biol.* **299**, 283–293.
13. Moyle, W. R., Campbell, R. K., Myers, R. V., Bernard, M. P., Han, Y. & Wang, X. Y. (1994) *Nature* **368**, 251–255.
14. Parkinson, J. S. & Kofoed, E. C. (1992) *Annu. Rev. Genet.* **26**, 71–112.
15. Borkovich, K. A. & Simon, M. I. (1990) *Cell* **63**, 1339–1348.
16. Appleby, J. L., Parkinson, J. S. & Bourret, R. B. (1996) *Cell* **86**, 845–848.
17. Goudreau, P. N. & Stock, A. M. (1998) *Curr. Opin. Microbiol.* **1**, 160–169.
18. Wolanin, P. M. & Stock, J. B. (2003) in *Histidine Kinases in Signal Transduction*, ed. M. Inouye, R. D. (Academic, San Diego), pp. 74–123.
19. Bilwes, A. M., Alex, L. A., Crane, B. R. & Simon, M. I. (1999) *Cell* **96**, 131–141.
20. Bilwes, A. M., Park, S. Y., Quezada, C. M., Simon, M. I. & Crane, B. R. (2003) in *Histidine Kinases in Signal Transduction*, eds. Inouye, M. & Dutta, R. (Academic, San Diego), pp. 48–74.
21. McEvoy, M. M., Hausrath, A. C., Randolph, G. B., Remington, S. J. & Dahlquist, F. W. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 7333–7338.
22. Gouet, P., Chinarde, N., Welch, M., Guillet, V., Cabantous, S., Birck, C., Mourey, L. & Samama, J. P. (2001) *Acta Crystallogr. D* **57**, 44–51.
23. Welch, M., Chinarde, N., Mourey, L., Birck, C. & Samama, J. P. (1998) *Nat. Struct. Biol.* **5**, 25–29.
24. Stewart, R. C., Jahreis, K. & Parkinson, J. S. (2000) *Biochemistry* **39**, 13157–13165.
25. Otwinowski, A. & Minor, W. (1997) *Methods Enzymol.* **276**, 307–325.
26. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999) *Acta Crystallogr. D* **55**, 484–491.
27. Perrakis, A., Morris, R. M. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463.
28. McRee, D. E. (1992) *J. Mol. Graphics* **10**, 44–47.
29. McEvoy, M. M., Zhou, H., Roth, A. F., Lowry, D. F., Morrison, T. B., Kay, L. E. & Dahlquist, F. W. (1995) *Biochemistry* **34**, 13871–13880.
30. Alexandre, G. & Zhulin, I. B. (2003) *J. Bacteriol.* **185**, 544–552.
31. Usher, K. C., De la Cruz, A. F., Dalquist, F. W., Swanson, R. V., Simon, M. I. & Remington, S. J. (1998) *Protein Sci.* **7**, 403–412.
32. West, A. H. & Stock, A. M. (2001) *Trends Biochem. Sci.* **26**, 369–376.
33. Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D. L., Henderson, R. K., King, D. S., Huang, L. S., Kustu, S., Berry, E. A. & Wemmer, D. E. (2001) *Nat. Struct. Biol.* **8**, 52–56.
34. Lawrence, M. C. & Colman, P. M. (1993) *J. Mol. Biol.* **234**, 946–950.
35. Stock, A. M. & West, A. H. (2003) in *Histidine Kinases in Signal Transduction*, eds. Inouye, M. I. & Dutta, R. (Academic, San Diego), pp. 237–271.
36. Li, J., Swanson, R. V., Simon, M. I. & Weis, R. M. (1995) *Biochemistry* **34**, 14626–14636.
37. Kuo, S. C. & Koshland, D. E. J. (1987) *J. Bacteriol.* **169**, 1307–1314.
38. Makhatazde, G. & Privalov, P. L. (1995) *Adv. Protein Chem.* **47**, 307–420.
39. Park, S. Y., Quezada, C. M., Bilwes, A. M. & Crane, B. R. (2004) *Biochemistry* **43**, 2228–2240.
40. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, W. (1997) *Nucleic Acids Res.* **25**, 3389–3402.
41. Armitage, J. P. (1999) *Adv. Microb. Physiol.* **41**, 229–289.
42. Garrity, L. F. & Ordal, G. W. (1997) *Microbiology* **143**, 2945–2951.
43. Kirby, J. R., Kristich, C. J., Saulmon, M. M., Zimmer, M. A., Garrity, L. F., Zhulin, I. B. & Ordal, G. W. (2001) *Mol. Microbiol.* **42**, 573–585.
44. Djordjevic, S., Goudreau, P. N., Xu, Q. P., Stock, A. M. & West, A. H. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 1381–1386.
45. Anand, G. S., Goudreau, P. N., Lewis, J. K. & Stock, A. M. (2000) *Protein Sci.* **9**, 898–906.
46. Robinson, V. L., Wu, T. & Stock, A. M. (2003) *J. Bacteriol.* **185**, 4186–4194.
47. DeLano, W. L., Ulsch, M. H., de Vos, A. M. & Wells, J. A. (2000) *Science* **287**, 1279–1283.
48. Alon, U., Surette, M. G., Barkai, N. & Liebler, S. (1999) *Nature* **397**, 168–171.
49. Grotzinger, J. (2002) *Biochim. Biophys. Acta* **1592**, 215–223.
50. Frank, S. J. (2002) *Endocrinology* **143**, 2–10.
51. Esnouf, R. M. (1997) *J. Mol. Graphics* **15**, 133–138.