JOURNAL OF BACTERIOLOGY, Jan. 2003, p. 254–261 Vol. 185, No. 1
0021-9193/03/$08.00+0 DOI: 10.1128/JB.185.1.254–261.2003

The Crystal Structure of the Phosphorylation Domain in PhoP Reveals a Functional Tandem Association Mediated by an Asymmetric Interface

Catherine Birck,[1]† Yinghua Chen,[2] F. Marion Hulett,[2] and Jean-Pierre Samama[1]*

Groupe de Cristallographie Biologique, IPBS-CNRS, 31077 Toulouse, France,[1] and Laboratory for Molecular Biology, Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois 60607[2]

Received 9 July 2002/Accepted 4 October 2002

PhoP from *Bacillus subtilis* belongs to the OmpR subfamily of response regulators. It regulates the transcription of several operons and participates in a signal transduction network that controls adaptation of the bacteria to phosphate deficiency. The receiver domains of two members of this subfamily, PhoB from *Escherichia coli* and DrrD from *Thermotoga maritima*, have been structurally characterized. These modules have similar overall folds but display remarkable differences in the conformation of the β4-α4 and α4 regions. The crystal structure of the receiver domain of PhoP (PhoPN) described in this paper illustrates yet another geometry in this region. Another major issue of the structure determination is the dimeric state of the protein and the novel mode of association between receiver domains. The protein-protein interface is provided by two different surfaces from each protomer, and the tandem unit formed through this asymmetric interface leaves free interaction surfaces. This design is well suited for further association of PhoP dimers to form oligomeric structures. The interprotein interface buries 970 Å² from solvent and mostly involves interactions between charged residues. As described in the accompanying paper, mutations of a single residue in one salt bridge shielded from solvent prevented dimerization of the unphosphorylated and phosphorylated response regulator and had drastic functional consequences. The three structurally documented members of the OmpR family (PhoB, DrrD, and PhoP) provide a framework to consider possible relationships between structural features and sequence signatures in critical regions of the receiver domains.

PhoR-PhoP in *Bacillus subtilis* comprises a prototypical two-component system activated in conditions under which inorganic phosphate is limiting (reviewed in reference 22). In this bacterium, the response regulator PhoP controls the expression of 31 genes in 10 operons, acting as a transcriptional activator or repressor depending on the promoter. Sequence similarities of the C-terminal effector domains identify PhoP as a member of the winged-helix OmpR subfamily, whose members occur frequently in bacterial species, as well as in some eukaryotes (8). PhoP belongs to the largest group of response regulators present in the *B. subtilis* genome according to the conservation of the residues forming the hydrophobic core of the DNA binding domain (16).

The structural similarities among the effector modules in OmpR, PhoB, and DrrD (24, 29, 38, 45) and among regulatory domains in general (2, 7, 13, 18, 19, 23, 32, 37, 41, 48, 49, 51, 52) are in sharp contrast to the different strategies that are used in the response regulators to regulate protein activity and to respond to the common phosphorylation event. The spatial extent and magnitude of the conformational changes in the α4-β5-α5 region that essentially sense phosphorylation of the aspartic acid vary in the few regulatory domains that have been characterized in both states (3, 10, 20, 28, 31). In two cases, the CheY-CheA complex (53) and the phosphorylated FixJN dimer (3), detailed comparisons can be made between the molecular species in the free and associated, phosphorylated and unphosphorylated states (19, 39, 40, 49, 52). The results illustrate the adaptability of the buried surfaces that are exploited for protein-protein interactions and show that rather small alterations in structure are sufficient to disrupt or foster the molecular recognition between partners.

In *Escherichia coli* CheB (14) and NarL (2), functional regions of the effector domains are masked by the regulatory domains prior to phosphorylation. A transcription activation surface in the effector domain of NarL contacts the regulatory domain in the unphosphorylated protein and prevents the tail-to-tail association of NarL to DNA (5). This may not be the case in *E. coli* PhoB, a response regulator that belongs to the same subfamily as *B. subtilis* PhoP and *Thermotoga maritima* DrrD. A comparison of the crystal structure of the effector domain of PhoB in complex with its target DNA sequence (4) with the crystal structure of full-length DrrD (7) suggested that none of the secondary elements involved in DNA binding are shielded by the regulatory domain in PhoB (4). It was proposed that only the position of this domain in the unphosphorylated protein prevents head-to-tail binding of the response regulator to the Pho box. The negative regulation essentially involves interactions mediated by helix α5, in agreement with genetic evidence (1).

A change in oligomeric state is another issue in meeting the specific regulatory needs. Homodimerization of the phosphorylated *Sinorhizobium meliloti* FixJ protein is a requirement for transcriptional activity (12). Dimerization is mediated by the

* Corresponding author. Present address: Département de Biologie et Génomique Structurales, IGBMC, 1 rue Laurent Fries, BP 10142, 67404 Illkirch, France. Phone: (33) 3 88 65 35 45. Fax: (33) 3 88 65 32 76. E-mail: samama@igbmc.u-strasbg.fr.

† Present address: Département de Biologie et Génomique Structurales, IGBMC, 67404 Illkirch, France.

receiver domains that associate through molecular twofold symmetry (3). Spo0A also dimerizes upon phosphorylation, and interestingly, unphosphorylated dimers of the full-length protein obtained through α5 exchange between two receiver domains activate transcription (32, 33). This observation indicates that the dimeric state favors protein binding to the 7-bp DNA consensus sequence of the 0A boxes and seems to be in agreement with the head-to-tail binding of Spo0A to two 0A boxes (5). A further step in diversity is provided by ArcA, which mediates the response of *E. coli* to anaerobic conditions. The unphosphorylated homodimer forms, upon phosphorylation, a tetramer of dimers containing phosphorylated and unphosphorylated proteins at a ratio of 1:1, an assembly that was shown to be a prerequisite for binding to DNA (26).

The X-ray structure of the regulatory domain of PhoP (PhoPN) described here and the functional studies described in the accompanying paper provide yet another example of protein-protein interactions between response regulators. PhoPN is a dimer in which the two protomers associate through an asymmetric interface that buries 970 Å^2 from solvent. The mode of association between receiver domains of PhoP has never been observed before and seems to be perfectly suited to allow further identical association of PhoPN tandems to form oligomeric species. The structural results are in line with the binding of the unphosphorylated dimeric PhoP species to the four 6-bp direct repeats that form the core binding region of the response regulator and with the cooperative and/or oligomerization process that occurs upon association of the protein with PhoP-regulated promoters (15, 35).

MATERIALS AND METHODS

Protein expression and purification. The PhoPN protein (residues 1 to 135) was cloned into the T7 promoter-based expression vector pET16b (Novagen), yielding pWL32M, and was expressed in *E. coli* BL21(DE3) as previously described (36). The cells were lysed by sonication in buffer containing 50 mM Tris (pH 7.8), 1 M NaCl, 10 mM imidazole, and 2 mM β-mercaptoethanol (buffer A) at 4°C. The lysate was clarified by centrifugation for 2 h at 9,000 × *g*. The soluble fraction was loaded onto a 5-ml Ni^{2+} affinity column (Amersham Biosciences), and the PhoPN protein was eluted by using a gradient of 30 to 300 mM imidazole in buffer A. The PhoPN fractions, identified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, were pooled and dialyzed overnight against 20 mM Tris (pH 7.8)–150 mM NaCl–2 mM β-mercaptoethanol. The dialyzed protein was digested with 2% (wt/wt) Xa factor (Biolabs) for 6 h at room temperature after addition of 2 mM $CaCl_2$. Phenylmethylsulfonyl fluoride (1 mM) was then added to the reaction mixture. After threefold dilution (by volume) in 20 mM Tris (pH 7.8)–0.1 mM EDTA–2 mM β-mercaptoethanol (buffer Q), PhoPN was purified by anion-exchange chromatography on a UnoQ6 column (Bio-Rad) and was eluted by using a gradient of 0 to 0.2 M NaCl in buffer Q. The PhoPN-containing fractions were then loaded on an S75 16/60 gel filtration column (Amersham Biosciences) in a solution containing 20 mM Tris (pH 7.8), 150 mM NaCl, 0.1 mM EDTA, and 1 mM dithiothreitol. This procedure yielded approximately 12 mg of pure PhoPN per liter of cell culture.

For production of the selenomethionine-substituted protein (Se-PhoPN), plasmid pWL32M was transformed into the methionine auxotroph strain B834(DE3). The culture was grown at 30°C in minimal medium supplemented with 17 amino acids, the bases for nucleic acids, various salts, $MgSO_4$, $FeSO_4$, glucose, and selenomethionine (21). Protein expression was induced by addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when the cell culture reached an optical density at 600 nm of 0.5. The cells were harvested by centrifugation after 10 h, and Se-PhoPN was purified by using a procedure identical to that described above for the wild-type protein except that all buffers were supplemented with 15 mM β-mercaptoethanol. Electrospray mass spectrometry of the purified Se-PhoPN established that the seven methionine residues were replaced by selenomethionine residues.

Crystallization. PhoPN in a solution containing 20 mM Tris (pH 7.8), 150 mM NaCl, 0.1 mM EDTA, and 1 mM dithiothreitol or Se-PhoPN in a solution containing 20 mM Tris (pH 7.8), 150 mM NaCl, 0.1 mM EDTA, and 10 mM tris-(2-carboxyethyl)phosphine hydrochloride was concentrated to 5 mg/ml by using Centricon 10 (Millipore). Crystals were obtained at 4°C by the hanging-drop vapor diffusion method. The initial crystallization conditions were identified by using the Crystal Screen sparse matrix (25). After optimization, crystals that were approximately 0.2 by 0.2 by 0.1 mm were obtained after 1 day by mixing 1 μl of the protein solution and 1 μl of a reservoir solution containing 23 to 25% polyethylene glycol 10000 (PEG 10000), 0.1 M sodium citrate (pH 5.8), and 3% PEG MME 550. The crystals were cryocooled in a nitrogen stream after a 5-s transfer in a solution containing 19% PEG 6000, 0.1 M sodium citrate (pH 5.8), 5% glycerol, and 15 mM β-mercaptoethanol.

The PhoPN-Mn^{2+} complex was obtained by soaking PhoPN crystals in a solution containing 23% PEG 10000, 0.1 M sodium cacodylate (pH 6.0), and 3% PEG MME 550, then in the same solution at pH 6.6, and finally in a solution containing 23% PEG 10000, 0.1 M sodium cacodylate (pH 7.0), 3% PEG MME 550, and 20 mM $MnCl_2$ for 23 h.

The crystals belong to the tetragonal space group $P4_12_12$ with cell parameters of $a = b = 45.704$ Å and $c = 134.811$ Å ($a = b = 45.476$ Å and $c = 133.528$ Å for the Se-PhoPN crystals), with one molecule per asymmetric unit.

Structure determination. Multiwavelength anomalous dispersion data were collected with a single crystal of Se-PhoPN. The data were collected at 100 K on the BW7A beam line of the European Molecular Biology Laboratory outstation at Deutsches Elektronen Synchrotron (Hamburg, Germany). The fluorescence spectrum recorded from the cryocooled crystal was used to select the wavelengths at the selenium K edge (λ1 = 0.9788 Å, minimum of f′), at the peak (λ2 = 0.9785 Å, maximum of f″), and at a high-energy remote wavelength (λ3 = 0.9611 Å). The data set for the Mn^{2+}-bound wild-type PhoPN crystal was collected at 100 K on the ID14-EH1 beam line of the European Synchrotron Radiation Facility (Grenoble, France).

The seven selenium atoms were identified by the MAD phasing program SOLVE (50) by using diffraction data to 1.6 Å. The initial phases (mean figure of merit, 0.50) were improved by density modification by using DM (11) and assuming a solvent content of 45% in the unit cell. The quality of the resulting 1.6-Å electron density map was excellent, and this map allowed tracing of almost the entire molecule by the warpNtrace procedure (47). The initial model was refined by using the maximum-likelihood method implemented in REFMAC (42).

The structure of the PhoPN-Mn^{2+} complex was solved by molecular replacement by using AmoRe (44) and the Se-PhoPN structure as the search model, which yielded an R factor and correlation coefficient of 0.39 and 77.7%, respectively, for data between 8 and 3 Å. Structure refinement was performed by using CNS (6) and REFMAC 5 (46) in the final stages of refinement. Statistics for the PhoPN-Mn^{2+} complex are shown in Table 1. The current model comprises 121 residues, 177 water molecules, two sodium ions, and one manganese ion. Residues 122 to 135 from the linker region and the additional histidine residue at the N terminus were not observed and were presumably disordered. Electron density was not observed for the side chain atoms of the M1, E11, E12, L17, E36, K83, D84, E85, and E86 residues, which were modeled as alanine residues. Dual side chain conformers were modeled for L80, and half occupancies were assigned to side chain atoms of residues 44, 46, 69, 70, 83, 84, 115, 119, and 120. The crystallographic R factor and R_{free} values were 0.186 and 0.233, respectively (Table 1). All residues except K58 are in the allowed regions of a Ramachandran plot, and 95.4% of them are in the most favored phi and psi angles, as defined by PROCHECK (30). The average B factors are 20.9 Å^2 for the protein atoms and 20.7 Å^2 for the manganese ion (PDB code 1MVO).

RESULTS

Overview of the PhoPN structure. The structure of PhoPN was solved by the MAD method from a single crystal of a selenomethionine derivative of PhoPN. The electron density after solvent flattening allowed automatic chain tracing of almost the entire molecule. The PhoPN-Mn^{2+} complex of the wild-type protein described in this paper was refined to 1.6-Å resolution with crystallographic R factor and R_{free} values of 0.186 and 0.233, respectively (Table 1). Binding of the Mn^{2+} ion in the active site induced no significant conformational changes in the receiver domain. PhoPN displays the five-strand doubly wound α/β sandwich fold shared by the regulatory modules in the response regulator family (Fig. 1). All residues

TABLE 1. Reflection data and structure refinement

Data set	Se-PhoPN			PhoPN-Mn^{2+}
	λ1 (edge)	λ2 (peak)	λ3 (remote)	
Wavelength (Å)	0.9788	0.9785	0.9611	0.934
Resolution (Å)	33.3-1.56	33.3-1.56	33.3-1.6	27.1-1.6
Observations	515,316	492,358	493,128	224,913
Unique reflections	20,664	20,770	20,710	18,249
Completeness (%) (overall/last shell)[a]	99.0/99.0	99.6/99.6	99.4/99.4	92.0/92.0
Redundancy	5.2/4.7	7.1/6.6	6.4/5.9	6.1/4.3
R_{sym} overall/last shell	0.053/0.284	0.045/0.280	0.055/0.292	0.091/0.187
Overall figure of merit		0.50 (SOLVE)		
Refinement				
Resolution				27.1-1.6
No. of molecules in the asymmetric unit				1
No. of protein atoms (including hydrogens)				1,925
No. of solvent molecules				177
No. of manganese ions				1
No. of sodium ions				2
No. of reflections (work/test)				17,225/878
Crystallographic R factor/R_{free}				0.186/0.233
Mean temperature factor ($Å^2$)				17.5

[a] The resolution ranges in the highest bin are 1.65 to 1.56 and 1.69 to 1.60 Å for Se-PhoPN and PhoPN-Mn^{2+}, respectively

except K58 from the β3-α3 loop adopt low-energy main chain dihedral angles (30). This may be because the ammonium group forms a salt bridge interaction with D34 (2.95 Å), the residue which caps the N terminus of helix α2 by forming two hydrogen bond interactions with the main chain nitrogen atoms of E36 and E37.

Protein-protein interface. Examination of the crystal packing revealed the occurrence of a major interface between the PhoP regulatory domains. This interface buries 970 $Å^2$ from solvent and involves 8% of the surface area of each monomer. The association between molecules does not involve rotational symmetry (i.e., twofold rotation about an axis through the interface) and is generated in the following way. Each regulatory domain offers two surfaces (surfaces A and B) for association with another molecule. Surface A is defined by α4, β5, loop β5-α5, and α5, and surface B is defined by α3, loop β4-α4, and α4. Two PhoPN protomers associate through an interface that involves surface A from one molecule and surface B from the other molecule (Fig. 1). This asymmetric tandem leaves one free surface A and one free surface B and may therefore associate with other tandems in a strictly identical way. This association occurs in the crystal, forming layers of multimers in the *ab* plane. In the orthogonal direction, the contact between layers involves two hydrogen bonds provided by the main chain atoms of residues 72 to 74 in PhoPN molecules.

The molecular association almost exclusively involves polar interactions provided by 10 acidic or basic side chains and four main chain carbonyl oxygen atoms (Fig. 2 and Table 2). The salt bridge between R113 (α5) and D60 (α3) is shielded from solvent and seems remarkable in view of its central position at

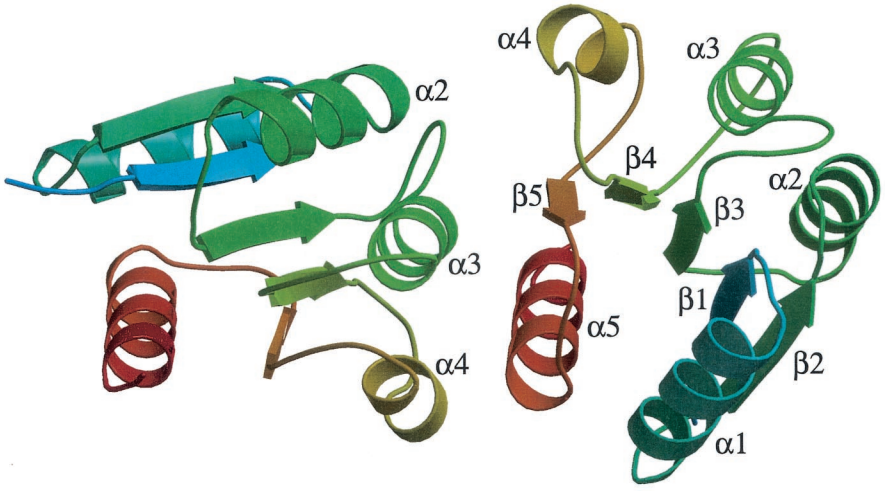

FIG. 1. Asymmetric interface between PhoPN protomers. Each protomer is represented by ribbons, and the color varies from light blue (N terminus) to red (C terminus). The two domains associate through surface A (α4, β5, loop β5-α5, and α5) from protomer 1 (right) and surface B (α3, loop β4-α4, and α4) from protomer 2 (left) (the protomers are arbitrarily numbered). This setting leaves surfaces B and A from protomers 1 and 2, respectively, free for association with other tandem units.

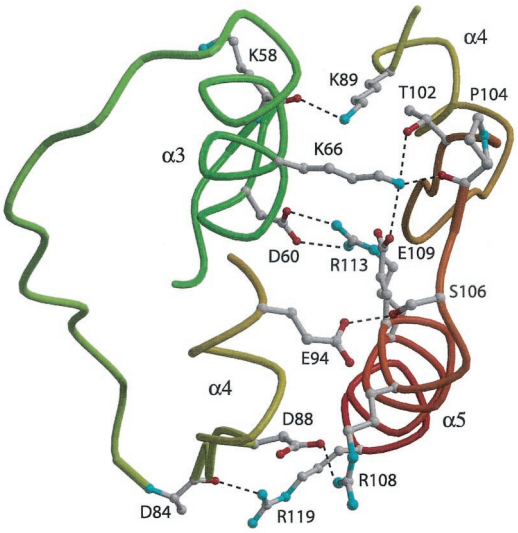

FIG. 2. Network of polar interactions at the PhoPN dimer interface. The hydrogen bonds are represented by dashed lines. The salt bridge interaction between D60 and R113 is shielded from solvent.

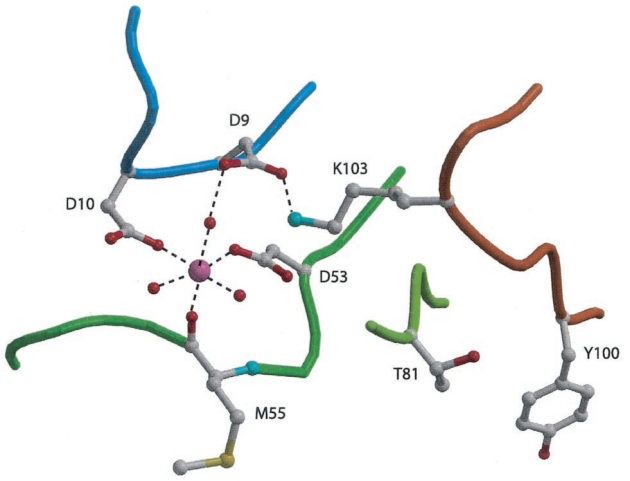

FIG. 3. Octahedral coordination of the manganese ion in the active site of PhoPN. The Mn^{2+} ion is represented by a large magenta sphere. Small red spheres and dashed lines indicate water molecules and polar interactions, respectively. Outside the active site, two conserved residues in receiver domains, T81 at the C-terminal edge of strand β4 and Y100 from strand β5, are also shown.

the interface. R113 appears to be the single positively charged residue within a group of negatively charged amino acids provided by the domain to which it belongs (D99 and E109) and by D60 and E63 from the associated molecule. We reasoned that a mutation of R113 would cancel the electrostatic balance in this area and be detrimental to the molecular association between regulatory domains. Mutation of this arginine residue to alanine or glutamate indeed changed the oligomeric state of the PhoP protein and had drastic consequences for the biological activity of the response regulator in vivo (9).

Active site and functionally conserved residues. The acidic pocket is defined by four loop regions at the C-terminal ends of strands 1, 3, 4, and 5 and has the invariant residues (D10, D53, and K103) in their typical positions in the active site. The manganese ion in the PhoPN-Mn^{2+} complex is coordinated to the carbonyl oxygen of M55 and to the carboxylate groups of D10 and D53, the site of phosphorylation. Three water molecules are liganded to the metal ion that displays perfect octahedral coordination (Fig. 3). The K103 side chain from the β5-α5 loop forms a salt bridge interaction with D9 (2.8 Å). The hydroxyl group of T81, an important residue in signal transduction, is oriented away from the acidic pocket. With respect to the active site geometry, PhoPN is similar to other receiver domains prior to phosphorylation.

PhoP was shown to form inactive disulfide-mediated dimers in the absence of a reducing agent (34). The single cysteine residue in the PhoP sequence (C65) belongs to helix α3 but is shielded from solvent in the crystal structure. Accessibility of the thiol group and disulfide bond formation require alterations of the conformation in the region encompassing helices α3 and α4. These structural changes likely affect the active site area, in line with the nonactivity of these dimers. However, the PhoP protein is a noncovalent dimeric species (36), and the association between molecules, mediated by the interactions between receiver domains described in this paper, involves the α3 and α4 helices as part of the B surface. This association shields and likely stiffens the α3-α4 area, and it seems likely that disulfide-mediated dimer formation is only an incidental event that occurs in vitro with a fraction of the free monomers. Nevertheless, this process indicates flexibility of the monomeric protein in this region, a feature that may contribute to the optimal fit of this surface in the protein-protein interface in which it is involved.

β4-α4 loop and α4 helix. The β4-α4 loop and α4 helix regions distinguish PhoPN from the other structurally characterized regulatory domains. The β4-α4 loop is much longer, with residues 82 to 85 forming a β-turn and residues 86 to 88 displaying an extended conformation (Fig. 4). The only observed interaction that may stabilize this fold arises from hydrogen bonds (2.7 Å) that link the carboxylate of D53 to the carbonyl oxygen of E85 through three water molecules. Helix α4 (residues 89 to 94) is less than two helix turns long. Its orientation and position are unrelated to those of the corresponding helix in the ortholog PhoB from *E. coli* and in DrrD from *T. maritima* (7, 48),which belong to the same subfamily of

TABLE 2. Interacting surfaces in PhoPN and interactions at the protein-protein interface

Type of contacts	Surface A		Surface B	
	Residue	Location	Residue	Location
Polar	K89	α4	K89 m.c. oxygen[a]	α3
	T102	β5	K66	α3
	T102 m.c. oxygen	β5	K66	α3
	P104 m.c. oxygen	β5-α5	K66	α3
	E109	α5	K66	α3
	S106	β5-α5	E94	α4
	R108	α5	D88	β4-α4
	R113	α5	D60	α3
	R119	α5	D84 m.c. oxygen	β4-α4
Apolar	A112	α5	I62	α3
	A112	α5	L91	α4

[a] m.c. oxygen, main chain oxygen.

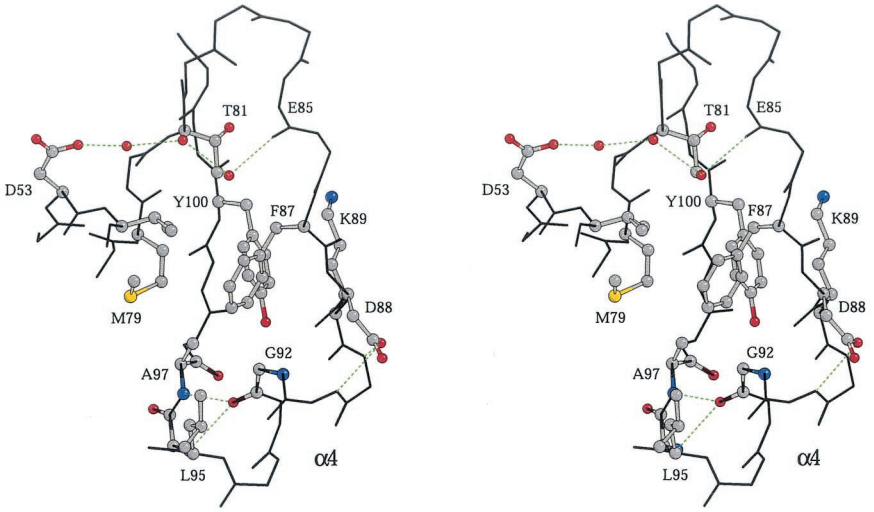

FIG. 4. Stereo view of the β4-α4 and α4 areas of PhoPN. The protein backbone is indicated by black lines; oxygen atoms are red, carbon atoms are grey, nitrogen atoms are blue, and sulfur atoms are yellow. Residue F87 from the β4-α4 loop contributes to the hydrophobic pocket and anchors this region to the protein core. The hydrogen bond (green dashed lines) between D53 and the main chain carbonyl atom of E85 involves three water molecules (small red spheres) and may stabilize the conformation of the loop.

response regulators. Compared to the directions in PhoB and DrrD, the direction of the α4 helix axis in PhoPN differs by 60° and 90°, respectively (Fig. 5).

In PhoPN, this helix is N capped by D88 and is stabilized at the C terminus by hydrogen bonds between the carbonyl oxygen atom of residue 92 and the main chain nitrogen atoms of residues 96 and 97 from the α4-β5 loop (Fig. 4). In contrast to other regulatory domains, there is no hydrophobic residue from α4 that anchors the helix to the protein core. It is the side chain of F87 from the β4-α4 loop that contributes to the hydrophobic pocket between strands β3, β4, and β5 and helices α3 and α4. This cavity in PhoPN also contains V54, M79, L95, and A97, as well as Y100, the conserved aromatic residue from strand β5. The aromatic side chain borders the hydrophobic pocket and is parallel to and at van der Waals distance from the hydrophobic moiety of the K89 side chain (Fig. 4). The phenol group points to the Cα carbon atom of G92 in the center of helix α4, and any side chain in that position would create a steric conflict with Y100. The special inward orientation of Y100 ($\chi_1 = -55°$) in PhoPN has also been observed in PhoB from *E. coli* and DrrD from *T. maritima* (7, 48).

Comparison of the PhoPN, PhoB, and DrrD structures. Phosphate assimilation and metabolism in *E. coli* and *B. subtilis* involve the PhoB and PhoP ortholog response regulators. The regulatory domains display 72% sequence similarity. Superposition of the central five-strand pleated sheet from PhoB (48) and PhoP indicates that helices α1 and α5 located on one side of the sheet occupy very similar positions in the two structures (root mean square deviation, rmsd, for all atoms = 0.88 Å), in contrast to helices α2, α3, and α4 and loops α3-β4 and β4-α4

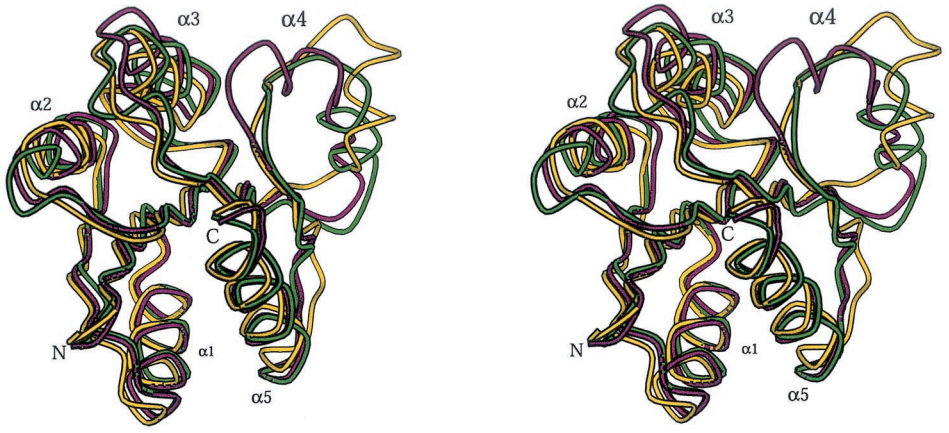

FIG. 5. Stereo view of the superimposed receiver domains of PhoP from *B. subtilis* (purple), PhoB from *E. coli* (pdb 1B00) (green), and DrrD from *T. maritima* (pdb 1KGS) (yellow). The superimposition of the structures was based on the best fit of the central five-strand sheet and revealed the similarities and differences in the protein folds. The conformations of the β4-α4 loops and α4 helices are not related in these three members of the OmpR family.

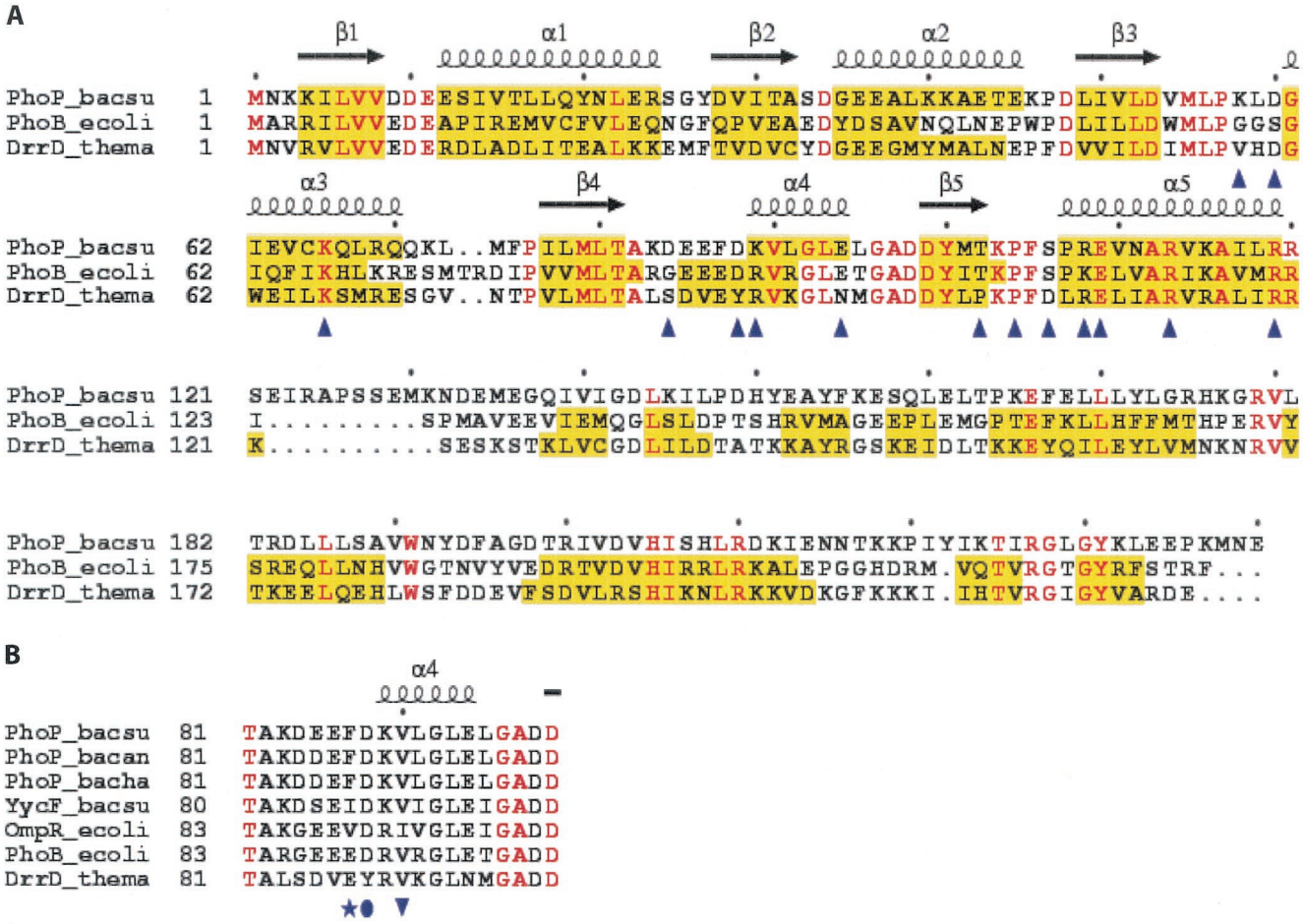

FIG. 6. Sequence alignment of PhoP from *B. subtilis* and other response regulators. (A) Structure-based sequence alignment of PhoP with PhoB from *E. coli* (pdb 1B00) and DrrD from *T. maritima* (pdb 1KGS). The secondary structure elements of PhoP are indicated by coils for α helices and arrows for β strands. The secondary structure elements in each receiver domain are indicated by a yellow background. Identical residues are indicated by red type. Residues that are involved in the PhoPN dimer interface are indicated by blue triangles. (B) Four response regulators displaying sequence signatures similar to that of PhoP in the region from position 81 to position 99: PhoP from *B. anthracis* (NP-658617), PhoP from *B. halodurans* (Q9K850), YycF from *B. subtilis* (P37478), and OmpR from *E. coli* (P03025). These proteins have a hydrophobic residue at position 87 (PhoP numbering). PhoB and DrrD have a glutamic acid residue at that position. Despite conservation of V90, substitution of D90 in PhoB by Y88 in DrrD induces a shift in register by two residues of the α4 anchor (7). PhoP_bacsu, *B. subtilis* PhoP; PhoB_ecoli, *E. coli* PhoB; DrrD_thema, *T. maritima* DrrD; PhoP_bacan, *B. anthracis* PhoP; PhoP_bacha, *B. halodurans* PhoP; YycF_bacsu, *B. subtilis* YycF; OmpR_ecoli, *E. coli* OmpR.

present on the other side of the sheet (rmsd = 2.60 Å) (Fig. 5). The same features were found when PhoPN was compared to the receiver domain of DrrD (7) (rmsd = 0.91 Å and rmsd = 3.44 Å for the similar and dissimilar surfaces of the sheet, respectively) (Fig. 5).

DISCUSSION

Sequence alignment of PhoP with PhoB and DrrD, the two other members of the OmpR subfamily with known structures (Fig. 6), indicated that the sequences from T81, the last residue from β4, to D99, the first residue from β5 (PhoPN numbering), involve the same number of residues and are closely related. Despite these similarities, the conformations of the three proteins in the region encompassing the β4-α4 loop, helix α4, and the α4-β5 loop are unrelated, and the positions of the invariant G92 residue between pairs of proteins differ by more than 8 Å. The difference between the orthologs PhoP and PhoB, from *B. subtilis* and *E. coli*, respectively, was striking in view of the identity of 14 of 19 residues in the region from position 81 to position 99 (Fig. 6B). The identical residues include V90, which in PhoB (V92) anchors helix α4 to the protein core (48). Sequence and structure analyses of these two proteins suggested that the occurrence of F87 in PhoP instead of E89 in PhoB might be responsible for the different geometries in this part of the structure. The situation in DrrD is different. In this protein, despite the valine conservation, a shift in register by two residues of the α4 anchor (Y88) results in a different orientation of the helix (7). These observations point to the extreme variability in conformation of this helix in the OmpR family. In the case of PhoP at least, this region plays a critical functional role in dimerization.

PhoP is a dimeric protein in vitro in both the phosphorylated

and unphosphorylated states (36), and we observed in the crystal structure a molecular interface between two PhoPN protomers that buries 970 Å^2 from solvent. This extent of interacting surface is in the range of the extents of interacting surfaces of the functionally relevant protein-protein interfaces that were described for CheY and CheA (53) and for two phosphorylated FixJN molecules in the homodimer (3). The association of PhoPN molecules does not involve twofold symmetry. Dimerization occurs by association of the A surface (α4, β5-α5, β5, α5) from one protomer with the B surface (α3, β4-α4, α4) from the other protomer (Fig. 1). This leaves free A and B surfaces on the tandem, which thus appears as an unprecedented unit that can oligomerize through identical associations with other dimers. This possibility matches well the established DNA binding properties of PhoP (15). DNA footprinting studies showed that there was cooperative binding between PhoP dimers at PhoP-activated promoters and indicated that both the unphosphorylated and phosphorylated species could bind to the four 6-bp direct repeats that form the core binding region on DNA of the response regulator. The in vivo results demonstrate the functional importance of the interface revealed by the crystal structure (9). These findings indicate that the unphosphorylated and phosphorylated PhoP dimers have structural relationships that allow similar binding of their C-terminal effector domains to the target DNA. The fact that only phosphorylated PhoP dimers can stimulate gene transcription or repression suggests that phosphorylation may favor cooperativity for DNA binding or interactions with other partners in the transcription complex (36).

Phosphorylation is not necessarily accompanied by large conformational changes or by changes in the oligomerization state of the proteins (10, 28, 31), and one of the striking examples in which no change in structure has been observed by X-ray diffraction is the active D13K mutant of CheY (27). The absence of negative regulation of the effector domain in unphosphorylated PhoP, the fact that phosphorylation does not change the dimeric state of the protein, and the functional importance of the tandem association illustrated in this study raise questions about the structural consequences of phosphorylation of PhoP.

Determination of the structure of PhoPN revealed some features that distinguish this protein from PhoB and DrrD. The proteins were examined in other response regulators. The focus was on the amino acid sequence of the region connecting β4 and β5 and on the residues involved in the protein-protein interface (Table 2). Among the 20 orthologs found in databases, PhoP from *Bacillus anthracis* and PhoP from *Bacillus halodurans* are the two response regulators that have sequence signatures identical to those of PhoP from *B. subtilis* (Fig. 6B). YycF, which has been shown to be part of an essential signal transduction pathway in *B. subtilis* (17), is also closely related to PhoP. Interestingly, the sequence of OmpR from *E. coli* in the critical regions mentioned above indicates that this protein is much more similar to PhoP from *B. subtilis* than to PhoB from *E. coli* or to DrrD from *T. maritima* (Fig. 6B).

In contrast, regardless of the conservation of the residues at the interprotein interface, it seemed from structure comparisons of the receiver domains that PhoB and DrrD may not associate in a similar way to PhoP because the topology of surface B, defined by α3, β4-α4, and α4 (Fig. 5), cannot accommodate the association of two protomers observed in PhoPN. Accordingly, it was found in the crystal structure determination of PhoB that two molecules were associated by a complementary hydrophobic surface that buries 560 Å^2 from solvent and involves helix α1, loop β5-α5, and the N terminus of helix α5 (48).

R113 is invariant in members of the OmpR family. The PhoP structure determination together with in vivo studies of the $PhoP_{R113E}$ mutant protein (9) demonstrated the critical functional contribution of R113 in this response regulator. A study of two OmpR mutants, E96A and R115S, documented that the proteins were phosphorylated by the kinase EnvZ but were defective in the oligomerization process observed in vitro with the wild-type protein (43). The fact that the phosphorylated $PhoP_{R113E}$ mutant also remains monomeric and the relatedness of OmpR and PhoP mentioned above raise the possibility that R113 may contribute to a similar association process.

ACKNOWLEDGMENTS

We thank the scientific staffs of the European Synchrotron Radiation Facility (Grenoble, France) and the European Molecular Biology Laboratory at Deutsches Elektronen Synchrotron (Hamburg, Germany) for providing excellent data collection facilities.

This work was supported by National Institutes of Health grant GM-33471 to F.M.H. and by grants from CNRS and Le Programme de Recherche en Microbiologie Fondamentale of the French Ministry of Research to J.-P.S.

ADDENDUM IN PROOF

The molecular structures of the protein-DNA complexes for Spo0A and NarL reported in reference 5 have been published (H. Zhao, T. Masadek, J. Zapf, J. Hoch, and K. Varughese, Structure **10:**1041–1050, 2002; A. E. Maris, M. R. Sawaya, M. Kaczor-Grzeskowiak, M. R. Jarvis, S. M. D. Bearson, M. L. Kopka, I. Schroder, R. P. Gunsalus, and R. E. Dickerson, Nat. Struct. Biol. **9:**771–778, 2002).

REFERENCES

1. **Allen, M. P., K. B. Zumbrennen, and W. R. McCleary.** 2001. Genetic evidence that the α5 helix of the receiver domain of PhoB is involved in interdomain interactions. J. Bacteriol. **183:**2204–2211.
2. **Baikalov, I., I. Schroder, M. Kaczor-Grzeskowiak, K. Grzeskowiak, R. P. Gunsalus, and R. E. Dickerson.** 1996. Structure of the *Escherichia coli* response regulator NarL. Biochemistry **35:**11053–11061.
3. **Birck, C., L. Mourey, P. Gouet, B. Fabry, J. Schumacher, P. Rousseau, D. Kahn, and J. P. Samama.** 1999. Conformational changes induced by phosphorylation of the FixJ receiver domain. Structure **7:**1505–1515.
4. **Blanco, A. G., M. Sola, F. X. Gomis-Rüth, and M. Coll.** 2002. Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. Structure **10:**701–713.
5. **Bourret, R. B., N. W. Charon, A. M. Stock, and A. H. West.** 2002. Bright lights, abundant operons—fluorescence and genomic technologies advance studies of bacterial locomotion and signal transduction: review of the BLAST meeting, Cuernavaca, Mexico, 14 to 19 January 2001 J. Bacteriol. **184:**1–17.
6. **Brünger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren.** 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. Sect. D **54:**905–921.
7. **Buckler, D. R., Y. Zhou, and A. M. Stock.** 2002. Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from *Thermotoga maritima.* Structure **10:**153–164.
8. **Chang, C., and R. C. Stewart.** 1998. The two-component system: regulation of diverse signaling pathways in prokaryotes and eukaryotes Plant Physiol. **117:**723–731.
9. **Chen, Y., C. Birck, J.-P. Samama, and F. M. Hulett.** 2003. Residue R113 is

essential for PhoP dimerization and function: a residue buried in the asymmetric PhoP dimer interface determined in the PhoPN three-dimensional crystal structure. J. Bacteriol. **185:**262–273.
10. **Cho, H. S., S. Y. Lee, D. Yan, X. Pan, J. S. Parkinson, S. Justus, D. E. Wemmer, and J. G. Pelton.** 2000. NMR structure of activated CheY. J. Mol. Biol. **297:**543–551.
11. **Collaborative Computational Project Number 4.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D **50:**760–763.
12. **Da Re, S., J. Schumacher, P. Rousseau, J. Fourment, C. Ebel, and D. Kahn.** 1999. Phosphorylation-induced dimerisation of the FixJ receiver domain. Mol. Microbiol. **34:**504–511.
13. **Djordjevic, S., P. N. Goudreau, Q. Xu, A. M. Stock, and A. H. West.** 1998. Structural basis for methylesterase CheB regulation by a phosphorylation-activated domain. Proc. Natl. Acad. Sci. USA **95:**1381–1386.
14. **Djordjevic, S., and A. M. Stock.** 1998. Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system. J. Struct. Biol. **124:**189–200.
15. **Eder, S., W. Liu, and F. M. Hulett.** 1999. Mutational analysis of the *phoD* promoter in *Bacillus subtilis*: implications for PhoP binding and promoter activation of Pho regulon promoters. J. Bacteriol. **181:**2017–2025.
16. **Fabret, C., V. A. Feher, and J. A. Hoch.** 1999. Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. J. Bacteriol. **181:**1975–1983.
17. **Fabret, C., and J. A. Hoch.** 1998. A two-component signal transduction system essential for growth of *Bacillus subtilis*: implications for anti-infective therapy. J. Bacteriol. **180:**6375–6383.
18. **Feher, V. A., J. W. Zapf, J. A. Hoch, F. W. Dahlquist, J. M. Whiteley, and J. Cavanagh.** 1995. 1H, 15N, and 13C backbone chemical shift assignments, secondary structure, and magnesium-binding characteristics of the *Bacillus subtilis* response regulator, Spo0F, determined by heteronuclear high-resolution. NMR Protein Sci. **4:**1801–1814.
19. **Gouet, P., B. Fabry, V. Guillet, C. Birck, L. Mourey, D. Kahn, and J. P. Samama.** 1999. Structural transitions in the FixJ receiver domain. Structure **7:**1517–1526.
20. **Halkides, C. J., M. M. McEvoy, E. Casper, P. Matsumura, K. Volz, and F. W. Dahlquist.** 2000. The 1.9 Å resolution crystal structure of phosphono-CheY, an analogue of the active form of the response regulator, CheY. Biochemistry **39:**5280–5286.
21. **Hendrickson, W. A., J. R. Horton, and D. M. LeMaster.** 1990. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. **9:**1665–1672.
22. **Hulett, F. M.** 2002. The Pho regulon, p. 193–201. *In* A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and its closest relatives: from genes to cells. ASM Press, Washington, D.C.
23. **Im, Y. J., S.-H. Rho, C.-M. Park, S.-S. Yang, J.-G. Kang, J. Y. Lee, P.-S. Song, and S. H. Eom.** 2002. Crystal structure of a cyanobacterial phytochrome response regulator. Protein Sci. **11:**614–624.
24. **Itou, H., and I. Tanaka.** 2001. The OmpR-family of proteins: insight into the tertiary structure and functions of two-component regulator proteins. J. Biochem. **129:**343–350.
25. **Jancarik, J., and S. H. Kim.** 1991. Sparse matrix sampling: a screening method for crystallization of proteins. J. Appl. Crystallogr. **24:**409–411.
26. **Jeon, Y., Y. S. Lee, J. S. Han, J. B. Kim, and D. S. Hwang.** 2001. Multimerization of phosphorylated and non-phosphorylated ArcA is necessary for the response regulator function of the Arc two-component signal transduction system. J. Biol. Chem. **276:**40873–40879.
27. **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. J. Biol. Chem. **272:**11850–11855.
28. **Kern, D., B. F. Volkman, P. Luginbuhl, M. J. Nohaile, S. Kustu, and D. E. Wemmer.** 1999. Structure of a transiently phosphorylated switch in bacterial signal transduction. Nature **40:**894–898.
29. **Kondo, H., A. Nakagawa, J. Nishihira, Y. Nishimura, T. Mizuno, and I. Tanaka.** 1997. *Escherichia coli* positive regulator OmpR has a large loop structure at the putative RNA polymerase interaction site. Nat. Struct. Biol. **4:**28–31.
30. **Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton.** 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26:**283–291.
31. **Lewis, R. J., J. A. Brannigan, K. Muchova, I. Barak, and A. J. Wilkinson.** 1999. Phosphorylated aspartate in the structure of a response regulator protein. J. Mol. Biol. **294:**9–15.
32. **Lewis, R. J., K. Muchova, J. A. Brannigan, I. Barak, G. Leonard, and A. J. Wilkinson.** 2000. Domain swapping in the sporulation response regulator Spo0A. J. Mol. Biol. **297:**757–770.
33. **Lewis, R. J., D. J. Scott, J. A. Brannigan, J. C. Ladds, M. A. Cervin, G. B. Spiegelman, J. G. Hoggett, I. Barak, and A. J. Wilkinson.** 2002. Dimer formation and transcription activation in the sporulation response regulator Spo0A. J. Mol. Biol. **316:**235–245.
34. **Liu, W.** 1997. Biochemical and genetic analyses establish a dual role for PhoP in *Bacillus subtilis* PHO regulation. Ph.D. thesis. University of Illinois at Chicago, Chicago.
35. **Liu, W., S. Eder, and F. M. Hulett.** 1998. Analysis of *Bacillus subtilis tagAB* and *tagDEF* expression during phosphate starvation identifies a repressor role for PhoP~P. J. Bacteriol. **180:**753–758.
36. **Liu, W., and F. M. Hulett.** 1997. *Bacillus subtilis* PhoP binds to the *phoB* tandem promoter exclusively within the phosphate starvation-inducible promoter. J. Bacteriol. **179:**6302–6310.
37. **Madhusudan, M., J. Zapf, J. A. Hoch, J. M. Whiteley, N. H. Xuong, and K. I. Varughese.** 1997. A response regulatory protein with the site of phosphorylation blocked by an arginine interaction: crystal structure of Spo0F from *Bacillus subtilis.* Biochemistry **36:**12739–12745.
38. **Martinez-Hackert, E., and A. M. Stock.** 1997. The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor. Structure **5:**109–124.
39. **McEvoy, M. M., D. R. Muhandiram, L. E. Kay, and F. W. Dahlquist.** 1996. Structure and dynamics of a CheY-binding domain of the chemotaxis kinase CheA determined by nuclear magnetic resonance spectroscopy. Biochemistry **35:**5633–5640.
40. **McEvoy, M. M., H. Zhou, A. F. Roth, D. F. Lowry, T. B. Morrison, L. E. Kay, and F. W. Dahlquist.** 1995. Nuclear magnetic resonance assignments and global fold of a CheY-binding domain in CheA, the chemotaxis-specific kinase of *Escherichia coli*. Biochemistry **34:**13871–13880.
41. **Müller-Dieckmann, H.-J., A. A. Grantz, and S.-H. Kim.** 1999. The structure of the signal receiver domain of the *Arabidopsis thaliana* ethylene receptor ETR1. Structure **7:**1547–1556.
42. **Murshudov, G. N., A. A. Vagin, and E. J. Dodson.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect. D **53:**240–255.
43. **Nakashima, K., K. Kanamaru, H. Aiba, and T. Mizuno.** 1991. Signal transduction and osmoregulation in *Escherichia coli*. A novel type of mutation in the phosphorylation domain of the activator protein, OmpR, results in a defect in its phosphorylation-dependent DNA binding. J. Biol. Chem. **266:**10775–10780.
44. **Navaza, J.** 1994. AMoRe: an automated package for molecular replacement. Acta Crystallogr. Sect. A **50:**157–163.
45. **Okamura, H., S. Hanaoka, A. Nagadoi, K. Makino, and Y. Nishimura.** 2000. Structural comparison of the PhoB and OmpR DNA-binding/transactivation domains and the arrangement of PhoB molecules on the phosphate box. J. Mol. Biol. **295:**1225–1236.
46. **Pannu, N. J., G. N. Murshudov, E. J. Dodson, and A. Read.** 1998. Incorporation of prior phase information strengthens maximum-likelihood structure refinement. Acta Crystallogr. Sect. D **54:**1285–1294.
47. **Perrakis, A., R. Morris, and V. S. Lamzin.** 1999. Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. **6:**458–463.
48. **Solà, M., F. Gomis, L. Serrano, A. Gonzalez, and M. Coll.** 1999. Three-dimensional crystal structure of the transcription factor PhoB receiver domain. J. Mol. Biol. **285:**675–687.
49. **Stock, A. M., E. Martinez-Hackert, B. F. Rasmussen, A. H. West, J. B. Stock, D. Ringe, and G. A. Petsko.** 1993. Structure of the Mg(2+)-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. Biochemistry **32:**13375–13380.
50. **Terwilliger, T. C., and J. Berendzen.** 1999. Automated MAD and MIR structure solution. Acta Crystallogr. Sect. D **55:**849–861.
51. **Volkman, B. F., M. J. Nohaile, N. K. Amy, S. Kustu, and D. E. Wemmer.** 1995. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. Biochemistry **34:**1413–1424.
52. **Volz, K., and P. Matsumura.** 1991. Crystal structure of *Escherichia coli* CheY refined at 1.7-A resolution. J. Biol. Chem. **266:**15511–15519.
53. **Welch, M., N. Chinardet, L. Mourey, C. Birck, and J. P. Samama.** 1998. Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. Nat. Struct. Biol. **5:**25–29.