NIH Public Access
Author Manuscript
J Mol Biol. Author manuscript; available in PMC 2007 April 24.

Published in final edited form as:
J Mol Biol. 2007 February 16; 366(2): 626–641.

The X-ray crystal structures of two constitutively active mutants of the *E. coli* PhoB receiver domain give insights into activation

Raquel Arribas-Bosacoma[1], **Soo-Ki Kim**[2,3], **Cristina Ferrer-Orta**[1], **Alexandre G. Blanco**[1,5], **Pedro J.B. Pereira**[1,4], **F. Xavier Gomis-Rüth**[1], **Barry L. Wanner**[2], **Miquel Coll**[1], and **Maria Solà**[1,*]

1 *Institut de Biologia Molecular de Barcelona (CSIC), c/Jordi Girona 18-26, 08034 Barcelona, Spain, and IRB - Parc Científic de Barcelona, c/Josep Samitier 1-5, 08028 Barcelona, Spain*

2 *Department of Biological Sciences; Purdue University; West Lafayette; Indiana 47907 USA*

Abstract

The PhoR/PhoB two-component system is a key regulatory protein network enabling *Escherichia coli* to respond to inorganic phosphate (Pi) starvation conditions by turning on Pho regulon genes for more efficient Pi uptake and use of alternative phosphorus sources. Under environmental Pi depletion, the response regulator (RR) component, PhoB, is phosphorylated at the receiver domain (RD), a process that requires Mg^{2+} bound at the active site. Phosphorylation of the RD relieves the inhibition of the PhoB effector domain (ED), a DNA-binding region that binds to Pho regulon promoters to activate transcription. The molecular details of the activation are proposed to involve dimerisation of the RD and a conformational change in the RD detected by the ED. The structure of the PhoB RD shows a symmetrical interaction involving α1, loop β5α5 and N-terminus of α5 elements, also seen in the complex of PhoB RD with Mg^{2+}, in which helix α4 highly increases its flexibility. PhoB RD in complex with Mg^{2+} and BeF_3^- (an emulator of the phosphate moiety) undergoes a dramatic conformational change on helix α4 and shows another interaction involving α4, β5 and α5 segments. We have selected a series of constitutively active PhoB mutants ($PhoB^{CA}$) that are able to turn on the Pho regulon promoters in the absence phosphorylation and, as they cannot be inactivated, should therefore mimic the active RD state of PhoB and its functional oligomerisation. We have analysed the $PhoB^{CA}$ RD crystal structures of two such mutants, Asp53Ala/Tyr102Cys and Asp10Ala/Asp53Glu. Interestingly, both mutants reproduce the homodimeric arrangement through the symmetric interface encountered in the unbound and magnesium-bound wild-type PhoB RD structures. Besides, the mutant RD structures show a modified active site organization as well as changes at helix α4 that correlate with repositioning of surrounding residues, like the active-site events indicator Trp54, putatively redifining the interaction with the ED in the full-length protein.

*Corresponding author: E-mail msvcri@ibmb.csic.es; Phone +34 93 400 61 50; Fax +34 93 204 59 04.
[3]Present address: Department of Animal Products and Environment; Konkuk University; 1 Hwayang-dong, Gwangin-gu; Seoul 143-701, Korea
[4]Present address: Grupo de Estrutura Molecular, IBMC; Rua do Campo Alegre, 823; 4150-180 Porto, Portugal
[5]Present address: Instrumentation Laboratory; 101 Hartwell Av., Lexington, MA 02421
Other authors' contact information: S.-K. Kim, sookikim@konkuk.ac.kr, B. L. Wanner, blwanner@purdue.edu, M. Coll, mcoll@ibmb.csic.es

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Keywords

PhoB; two-component signal transduction; receiver domain; X-ray crystal structure; asp-phosphorylation

Introduction

Fine-tuned regulatory systems are indispensable for adaptation and survival of bacteria to environmental changes. Two-component systems (TCSs) enable rapid response to stresses from changing growth conditions. Among others, TCS rule chemotaxis, quorum sensing, nutrient uptake, nitrogen fixation, osmoregulation, sporulation and pathogenic host invasion [1; 2; 3; 4]. TCSs are widespread in bacteria and bacterial pathways; for example, more than 30 different systems are estimated to operate in *Escherichia coli* and *Streptomyces coelicolor*[5]. TCSs are also present in fungi, plants and eukaryotic organelles [6; 7; 8; 9; 10; 11; 12; 13] and, as they are absent in mammals, they have been envisaged as valid targets for the development of new antimicrobials in order to increase the medical armamentarium available to evade the menace of increasing bacterial antibiotic resistance [14; 15]. TCSs comprise at least two paradigmatic elements, a signal-sensing histidine kinase and a response regulator (RR). The RR mediates the adaptive response, either through interaction with target proteins or with specific DNA sequences, in the latter cases acting as a transcriptional repressor or activator. The commonly accepted mechanism foresees that the transmembrane sensor kinase is autophosphorylated at a conserved His residue in its cytosolic kinase domain in response to an extra- or intra-cellular signal detected by its sensor module. Subsequently, the phosphoryl group is transferred in a magnesium-dependent manner to an Asp residue of the cognate intracellular receiver domain (RD) of the RR, resulting in the signal transduction [2; 9]. The latter may show additional domains, among them a C-terminal effector domain (ED) whose activity is modulated by Asp-phosphorylation of the RD. In lower eukaryotes and plants, the signalling cascade generally comprises at least a third element, a His-containing phosphotransfer protein that mediates signal transmission from the sensory histidine kinase to the RR [2]. Once the ED is activated, the RR binds to its target, changes its oligomeric state and/or exerts its enzymatic function. As soon as the triggering stimulus vanishes, the TCS is switched off through magnesium-dependent dephosphorylation of the RR, either by a specific phosphatase or catalysed by the response regulator itself [2; 16; 17; 18; 19].

Under environmental Pi-depleted conditions, the *E. coli* PhoR/PhoB TCS is involved in the assimilation of this essential nutrient from other sources. This process involves controlling the expression of more than 38 genes belonging to the phosphate (Pho) regulon that enables acquisition of Pi with higher affinity and to employ, for example, organic phosphates or phosphonates [20; 21; 22]. The first component of this environmental Pi detection TCS is the transmembrane histidine kinase, PhoR, which autophosphorylates and phosphorylates/dephosphorylates the second component, PhoB [23; 24; 25; 26; 27]. PhoB, a 229-amino acid transcription factor, consists of an N-terminal RD (residues 1–122) and a C-terminal ED (residues 131–229). The PhoB RD displays a nucleotide-binding fold characteristic of RR superfamily members [28]. Besides, the ED is a σ^{70} RNA-polymerase-subunit dependent transcription activator, and this groups PhoB in the OmpR/PhoB subfamily of RRs [23]. The structure of this domain is made up of a modified winged helix-turn-helix motif and DNA recognition is exerted in tandem [29]. The P_i signalling response involves three processes: activation, deactivation, and inhibition [30]. The latter two processes occur with PhoR unphosphorylated during the basal state under excess P_i levels. The signal that sparks off this system is a low level of extracellular P_i, detected either through the extracellular domain of PhoR or the high-affinity P_i-specific transporter Pst, an ATP-binding cassette transporter encoded by the *pstSCAB-phoU* operon [31; 32]. Under P_i starvation, PhoR is *trans*-

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

phosphorylated on a conserved His residue. Subsequently, the phosphate is transferred to an invariant on the RD, Asp53, in a magnesium-dependent manner. This activation leads the PhoB ED to bind to the Pho box with a 150-fold higher affinity [21] and to enhance DNA transcription [33] interacting with the σ^{70} subunit of RNA polymerase [24; 34]. There, a third player, PhoU, has been proposed to intervene. Based on its structural similarity with the Bcl2-associated athanogene (Bag) domain, a cofactor for the eukaryotic chaperone 70-kDa heat shock protein Hsp70 family, and the resemblance of the Hsc70 ATPase domain with the PhoR histidine kinase (30% sequence similarity), a matchmaker role for PhoU during deactivation has been suggested. Upon binding to PhoR and/or phosphorylated PhoB, PhoU is thought to stimulate PhoB dephosphorylation [35].

Biochemical data have demonstrated that PhoB is a dimer before and after phosphorylation [33; 36]. In addition, the N-terminal RD acts as an inhibitor of the C-terminal ED DNA binding, since its scission results in a polypeptide that binds DNA more tightly and activates transcription [27; 37]. Further data also show that ED removal stimulates RD dimerisation in an active form [38]. Dimerisation enables proper action of the ED, which binds DNA in tandem. The current model of activation postulates that phosphorylation switches interdomain rearrangements that release the inhibition and enables a dimer-mediated response. However, the structure of full-length PhoB is not available neither in the active or inactive state, possibly due to the flexible long linker between domains that might hamper crystallization. The full-length structure of another OmpR-family member, DrrB [39], shows that the interaction between the two domains involves a conserved aromatic residue from the RD (equivalent to PhoB Tyr102) and an aspartate from the ED, which is also present in PhoB. Moreover, the apo-PhoB RD structure shows a dimer in the asymmetric unit involving a symmetric α1-Lβ5α5 interface that, according to the biochemical data [38], could correspond to the contact surface within the dimer in the full-length RR structure following activation [28]. The same protein-protein interaction has been recently reported for the structure of PhoB RD in complex with magnesium, obtained in a different crystal form [40]. However, the complex with Mg^{2+} showed destabilization of the α4-β5-α5 surface, suggesting an initial conformational change that would lead to another interaction, through α4-β5-α5, also postulated to be the one for active PhoB [36]. In order to shed some light on this dichotomy between possible active surfaces, we have identified a series of mutants affecting PhoB RD that render a constitutively active phenotype ($PhoB^{CA}$) with respect to transcription of the Pho regulon. We have further studied the RD structure of two such constitutively active mutant proteins encompassing the mutations Asp53Ala/Tyr102Cys (DAYC), or Asp10Ala/Asp53Glu (DADE). The mutated residues lie within or near the active-side pocket and explain why they confer phosphorylation-independent activation of the ED response. Accordingly, they therefore contribute to a closer mimic of the active RD state.

Results

1. Constitutively active PhoB mutants

The two constitutively active PhoB mutants examined in this study, DADE and DAYC, have the WT phosphorylation site Asp53 substituted by Ala or Glu, respectively. Previous studies had shown that mutants bearing substitutions Asp53Ala, Asp53Glu or Asp53Asn are not phosphorylable and fail to elicit a PhoB ED response [41]. Our results show that the effect of these mutations is completely different if accompanied by other substitutions, that is, the combination of Asp10Ala with Asp53Glu (DADE) or Asp53Glu with Tyr102Cys (DAYC) results in constitutively active forms of PhoB ($PhoB^{CA}$). Moreover, the single mutation Asp10Ala also present in DADE also renders a non-phosphorylable inactive protein [41]. Therefore, the combination of two inactivating mutations, Asp10Ala and Asp53Ala, gives a highly active full-length protein. Residues Asp10 and Asp53, together with Glu9 and Glu11,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

confer a net electronegative charge to the active-site cavity [28], participate directly or indirectly to the coordination of the metal required for phosphorylation [36; 40], and their modification alters the metal-binding capacity and the activity of the protein [41]. The other position affected, Tyr102 in DAYC, is far from the acidic active site [28] and is highly conserved as an aromatic residue [42]; several genetic, mutational and structural studies suggest a post-phosphorylation role for this residue in the activation process [43; 44]. The same substitution Tyr to Cys at the equivalent position in the close orthologue OmpR also rendered a partially constitutively active protein [45]. The fact that in OmpR this single amino-acid substitution gives rise to an active protein whereas in PhoB the single Asp53Ala mutation yields an inactive one, suggests that the activating switch in DAYC is the aromatic exchange to cysteine.

2. Key residues and structural features of PhoB RD

The PhoB RD structure (PDB 1b00 [28], hereafter referred as WT-RD), like the other TCS RR members, shows a doubly-wound $(\alpha/\beta)_5$ fold consisting of a central five-stranded twisted parallel β-sheet flanked by two (α1 and α5) and three (α2, α3 and α4) helices on either side (Fig. 1a). The central β-sheet shows a β2β1β3β4β5 topology and β1, β3 and β4 form the protein core. Strands alternate with helices along the primary sequence. At the C-terminal edge of the β-sheet, the loops connecting strands β1, β3 and β5 with their subsequent helices α1, α3 and α5 (Lβ1α1, Lβ3α3 and Lβ5α5, respectively) surround a wide cavity that harbours the active site (Figs. 1a and b). Inside this, there is the invariable phosphorylation site, Asp53 [24; 46], which establishes a salt bridge with the highly conserved Lys105, as WT-RD shows. Residues Glu9, Asp10, Glu11 and Asp53 confer a net negative charge on the catalytic cavity that extends to nearby segments Lβ2α2 (Glu33 and Asp34), the N-terminus of helix α2 (Asp36) and helix α4 (Glu87, Glu88, Glu89 and Asp90) [28; 40]. WT-RD forms a homodimer comprised of WT-A and WT-B protomers, which show structural differences at helix α3 (as a consequence of crystal packing interactions of the preceding loop Lβ3α3 in WT-A), helix α4 and the loop Lβ4α4 [28]. Between helices α3 and α4 there is Trp54 (Fig. 1a), an indicator of active-site events [33; 41]. Although the relative position of helices α3 and α4 diverge in WT-A and –B, in either case Trp54 keeps the *trans(t)-105* (*t* refers to the χ^1 orientation of 180° [47], 105 to χ^2 orientation in degrees [48]) conformation. However, upon Mg^{2+} binding, helix α4 becomes highly flexible and Trp54 rotates its side chain to the *gauche(g)*$^+$*-90* rotamer [48] (g^- refers to the χ^1 orientation of −60° [47]) entering into the active site, where it performs several interactions [40]. Between helix α4 and Lβ4α4 there is Tyr102 (in β5, Fig. 1a), highly conserved as an aromatic residue and found in three rotamers among different RDs: a) g^- conformation, with the side chain buried between helix α4 and β5 pointing towards Lα4β5 at the C-terminus of the α-helix (Figs. 1a,c); b) g^+ conformation (in g^+, χ^1 orientation is of +60°), with the side chain pointing towards the bulk solvent or interacting with the ED, like in the OmpR-family member DrrB [39] or with another protein, as in the inactivated DctD dimer [49]; and c) *t* conformation, with the side chain pointing towards Lβ4α4 at the N-terminus of helix α4. This last orientation occurs upon phosphorylation at the active site and entails that Thr83 (Ser or Thr throughout RDs), at the end of β4 (Fig. 1a), is displaced towards the phosphate moiety to interact with its phosphoryl group; this pattern of correlated Tyr-Thr side chain movements has been observed in all activated RD structures, and has been termed "YT coupling" [50].

3. The interface in WT PhoB RD structure reappears in the constitutively active mutants

WT-RD shows a dimeric arrangement with a protein-protein interface observed only in the PhoB RD [28; 36; 40] but not in any other RR structure. The constituting protomers, WT-A and WT-B, are related by an almost perfect non-crystallographic two-fold axis (Fig. 1b), to which both β-sheet C-termini converge, joining the two active sites in a continuous electronegative depression [28; 40]. Underneath and perpendicularly, the WT-A and WT-B dimerisation surfaces are self-complementary and feature a hydrophobic core. The interacting segments comprise helix α1, Lβ5α5 and the N-terminus of helix α5 from both monomers (thus, “α1-Lβ5α5

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

interface", Fig. 1b). The interface area spans 560 $Å^2$, thus is rather small and points to a transient homodimeric contact (Table 3 and [51]), a notion supported by some experiments in solution [36]. The residues involved are Pro13, Ile14, Met17, Phe20, Val21, Pro106, Phe107 and Pro109 (Fig 1d). Among these, the conserved residues (Ile14, 59%; Pro106, 100%; Phe107, 97.9%) border the active site edge while the remaining residues, constituting the core of the interaction, are not conserved. Both DADE and DAYC RD structures show this dimer (Fig. 1b). The two molecules in the DADE crystal a.u., DADE-A and DADE-B, interact with each other through an interface of 556 $Å^2$ per protomer. Each molecule contributes with the same residues as for WT plus Gln24 (the stereochemical parameters are listed in table 3. Considering the DAYC mutant, the structure shows again the same interface under burial of 538 $Å^2$ involving Glu11, Pro13, Glu16, Met17, Phe20, Gln24, Arg85, Lys105 and Lys110 from either monomer (DAYC-A and DAYC-B), plus Ile14 and Val21 from DAYC-B. It is notable that Arg85 in DAYC-A introduces Lβ4α4 as a novel interface interacting segment (Fig. 1b). Arg85 makes a double salt bridge with Glu11 $O^{\varepsilon 2}$ which is rotated out from the active site. The stereochemical analysis of the DAYC interface (Table 3) shows again a small and planar surface, suggesting a transient interaction, yet constituted by a high percentage (> 60%) of apolar atoms consistent with a stable contact [73]. The α1-Lβ5α5 interface is also found in the structure of PhoB RD in complex with Mg^{2+} (from now on, Mg-RD), which shows this interaction either within the a.u. or formed by two crystallographically-related molecules [40]. Furthermore, the structure of the PhoB RD in complex with Mg^{2+} and activated with the phosphate analogue BeF_3^- (BeF_3-RD) [36] shows two interfaces, one of them involving the same segments as for the α1-Lβ5α5 surface although both protomers are rotated 90º with respect to each other if compared to the WT-RD. The prediction of stability (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver[52]) for the α1-Lβ5α5 interaction in all these structures predicts a stable contact in solution in all cases (solvation free energy gain upon formation of the assembly is negative, WT-RD, -13.1; DADE, -11.0; DAYC, -14.4; BeF_3-RD, -12.2 kcal/mol). Therefore, all these data suggest a high tendency of PhoB in establishing this interaction. Moreover, purification by analytical size exclusion chromatography showed in our hands that native PhoB RD elutes as a single peak corresponding to a molecular weight (MW) of 30.2 kDa molecule (data not shown). Hence, considering that the PhoB monomer is globular and has an estimated MW of 14.5 kDa, a stable dimer must exist in the buffer conditions assayed. A similar result was found for DADE, either for the full-length protein (elution peak corresponded to a MW of 42 kDa instead of 26 for a monomer) or the RD construct (elution peak corresponded to a MW 31.3 kDa instead of 14.3). It seems unlikely that a dimeric interaction strong enough to withstand size exclusion chromatography would include an interface distinct from the α1-Lβ5α5 found in the crystals.

4. DADE and WT RDs differ around helix α4

The DADE dimer is similar to the WT-RD one (*rmsd* of 0.43 Å, Table 2), as both are packed similarly in space group $P2_12_12_1$ (cell constants are isomorphous, Table 1) and show asymmetry between monomers (*rmsd* between DADE-A and DADE-B is 0.52 Å, Table 2). The best fitting (0.43 Å, Table 2) implies superposing DADE-A onto WT-A and DADE-B onto WT-B. Largest differences between DADE and WT dimers affect helix α4, in particular between DADE-A (the helix spans from Gly86 to Leu95) and WT-A (Gly86 to Gly94), and the flanking loops Lβ4α4 and Lα4β5 (Arg85A C^α diverge 1.85 Å from its position in WT-A; Glu96A C^α, 2.15 Å). The electron density map for DADE-A is well defined for the whole molecule chain (as an example, see Fig. 2a). In contrast, in DADE-B the regions involving helix α4 have been traced tentatively due to electron density map discontinuity, so any comparison with this protomer should be made with caution. Overall, we can conclude that the DADE dimer shows once again the intrinsic flexibility of helix α4, also reported for the WT-B structure [28], and for the Mg-RD protomers A (Mg-A) and C (Mg-C) structures [40]. Superposition of both DADE-A and -B protomers shows that the C^α atom of Ser60 is displaced

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2.65 Å due to, in DADE-A, an interaction with the Lys117 N^{ζ} atom from a crystallographically-related molecule, a fact also observed in WT-A. In DADE-B, the disorder observed at helix α4 propagates to the neighbouring residue Trp54 (first residue in Lβ3α3), whose side chain is in discontinuous density. Nonetheless, this residual density suggests that this residue is alternating between two conformations, *t-90* and *t-105* (according to [48]), the latter also found in DADE-A and both WT protomers. It is noteworthy that this mutant reproduces the correlation between the disorder of helix α4 and Trp54 flexibility, a fact also found in the three molecules of Mg-RD which has been related to activation of the protein [40]. The coupled *t* conformations of Thr83 and Tyr102 observed in activated RDs do not occur in any of the monomers of this constitutive mutant. Tyr102 is in both cases in the g^- conformation, and Thr83 is in the g^- and g^+ conformations in DADE-A and -B, respectively, as seen in WT RD. No metal was found in the active site, as we describe below.

5. DAYC RD results in an increased symmetry within the RD dimer and a large displacement of helix α4

The two DAYC monomers are rigid and well defined for their complete main chain (except for the first residue in DAYC-A), including the mutated sites (Fig. 2b,c,d), with no areas with weak electron density. Superposition of both monomers shows that they are highly symmetric, with a very low deviation value (*rmsd* of 0.2 Å, Table 2). However, compared to WT-RD, DAYC displays significant distortions that the crystal lattices ($P2_12_12_1$ in WT vs. $P2_1$ in DAYC) could account for it, but the structure in the monoclinic crystals is more rigid and well-defined although the packing is much looser than in the orthorhombic ones. Due to the high symmetry between DAYC-A and –B, superposition of DAYC and WT dimers yields similar *rmsd* values independently of the dimer orientations: of 0.62 Å if DAYC-A fits WT-A and DAYC-B WT-B, or 0.63 Å if the WT monomers are inverted (Table 2). Besides, 122 C^{α}-positions of either DAYC monomer are more similar to the WT-B ones (deviation values in Table 2) because in the formers Lβ3α3 is in a perfect antiparallel β-hairpin conformation similarly to WT-B. Again in both DAYC monomers, Glu11 fulfils the space freed by the new smaller sidechain Ala53, pulling the whole Lβ1α1 towards the active site cavity (in DAYC-A Glu11 moves 1.6 Å, and 1.0 Å in –B with respect to WT-B), as well as dragging the N-terminus of helix α1, which later accommodates to the WT-B helical trace.

However, these are minor changes if compared to the displacement performed by helix α4, together with Trp54 (Fig. 1c). In WT-RD, the side chain of Trp54 is in the *t-105* conformation [48] (like in DADE-A, Fig. 2a), immersed in a hydrophobic pillow between helices α3 and α4 as explained. In DAYC it changes to an unseen g^-*-95* rotamer [48] (Figs. 1c, 2b, 2d) invading the space liberated by the Tyr102Cys substitution between helix α4 and strand β4 (Fig. 1c). The side chains that interact with Trp54 are Met81 (β4), Thr83 (β4), Glu88 (α4) and Val92 (α4), the latter contacting the other face of the Trp54 rings with respect to as in WT (Fig. 1c). The Cys102 side chain adopts different rotamers, g^+ toward the solvent in DAYC-A (Fig. 2c) and g^- and g^+ double conformation in DAYC-B. But it does not adopt the *t* conformation that the aromatic residue at position 102 displays during the RD YT coupling activation. Likewise, Thr83 is not in the activated *t* conformation but in g^-, towards the core between helix α4 and β4 (Figs.1c,2b). This allows Thr83 $O^{\gamma 1}$ and Trp54 $N^{\varepsilon 1}$ to make a hydrogen bond (Fig. 1c). The structures of CheY co-crystallised with magnesium [53], or activated with Mg^{2+}/BeF_3^- and bound to its target FliM [50] show that the equivalent Trp58 ring adopts a plane parallel to that of the DAYC-A/B Trp54 rings, in the former induced by stacking interactions with Glu93 [53] (PhoB Glu89, which exposed to the solvent) or, in the latter, by steric hindrance with the methylene part of the Glu89 side chain [50] (Arg85, also exposed). Interestingly, in DAYC-A and –B, similar van-der-Waals interactions occur between Trp54 and Val92 (Fig. 1c). Regarding helix α4, the replacement of Tyr102 for a cysteine disrupts the van-der-Waals interaction that the tyrosine establishes with Arg93 in WT-RD (Fig. 1c). Therefore, in DAYC,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the arginine is exposed to the solvent, the last turn of helix α4 is unwinded so that this segment finishes at Val92 (Fig. 1c). Besides, helix α4 is largely displaced along its axis and laterally, towards the active-site groove, pushing out the whole Lβ3α3 and breaking hydrogen bonds between strands β3 and β4. Accordingly, residues in helix α4 are considerably displaced. For example, Glu87 and Asp90 are separated by 4.01 Å and 3.66 Å, respectively, ongoing from WT-A to DAYC-B. Within the dimer, the large helical movement brings the N-terminal ends of both helices α4 closer to each other (Fig. 1b) approaching A and B Gly86's from 21.50 Å (WT) to 11.93 Å (DAYC). Despite the changes in the helix, some interactions are unaffected. For example, in DAYC-A and –B, Glu88 makes, as in the WT-RD, interactions with the active site (see below).

6. Structural analysis of the active sites of DADE and DAYC RDs

The active site of PhoB shows two subsites which we previously referred to as site M (for Mg^{2+}) and site P (for phosphate) [40]. Residues at site M are involved in direct or indirect coordination to Mg^{2+} (Figs. 1e,f), and include the highly conserved triad Glu9, Asp10 and Glu11 (Lβ1α1) of PhoB [36; 40]. At the other end of the cavity, Thr83 O^{γ} and Ala84 N (β4) are directly involved in phosphoryl oxygen (or BeF_3^-) binding [36], thus delimiting the site P (Fig. 1e). The P- and the M- sites are bridged by the side chain of Asp53, whose $O^{\delta 1}$ atom coordinates a phosphate oxygen and $O^{\delta 2}$ atom binds the cation (Fig. 1e). In parallel, the backbone of Lβ3α3 is also partitioned, with the amide nitrogens of positions 54 and 55 binding the phosphate moiety and the carbonyl oxygen of residue 55 coordinating the cation. Besides, Lys105 N^{ζ} (nearly invariant among homologues) from Lβ5α5 binds a phosphate oxygen and contacts the carboxylate of Glu9 of site M. At the M-site, the interactions between residue atoms and water molecules are preserved in the WT-A, Mg-A and -B, and BeF_3-A and -B active sites [28; 36; 40]. Instead, in site P the WT or Mg-A, -B and -C protomers show different organization, as the area is filled either by Glu88 or by Trp54, with variations in the solvent content [40].

6.1 The DADE dimer shows asymmetry in the protomer active sites—In the DADE mutant, the electron density at the active site is better defined for protomer A (Fig. 2a) than for protomer B, therefore we will refer to the former. The substitution at position 53 introduces an additional methylene which orients the Glu53 side chain towards the M-site. This places Glu53 $O^{\varepsilon 2}$ 0.52 Å away from the position occupied by the Mg^{2+} cation in the Mg-B structure (0.70 Å in Mg-C or at 1.13 Å in the BeF_3-RD's position [36; 40]). The orientation observed for the Glu53 side chain is possible due to the mutation at position 10, from an aspartate to an alanine (Fig. 1e). In DADE-B, the same arrangement is found. Moreover, in DADE-A, the same glutamate carboxylate group interacts with water molecule W1 (W-118 in DADE-B), which is equivalent to WT-A W54, to Mg-C coordinating W21 and to BeF_3-RD W16 (Fig. 1e). Like in all these structures, W1 contacts Glu9 $O^{\varepsilon 2}$. The other *new* atom Glu53 $O^{\varepsilon 1}$ is close to where W1 is in WT-A and, likewise, binds the amide at position 10. In WT-A, W1 further binds to Glu9 $O^{\varepsilon 2}$, but in DADE Glu9 has been changed to alanine. Briefly, the active site of DADE shows that the negatively charged Glu53 $O^{\varepsilon 2}$ is close to where the bivalent cation should be and, despite the difference in charge between the atom and the cation, the former mimics the latter by performing similar interactions with surrounding water molecules and protein atoms. The location of Glu53 $O^{\varepsilon 2}$ is too distant from Lys105 N^{ζ} to make the interaction that Asp53 and Lys105 establish in WT-A. In DADE-A, Lys105 compensates this lack by bridging Glu53 $O^{\varepsilon 2}$ through a *new* water molecule, W13 (Fig. 1e), which in turn contacts Glu88 of helix α4 through two water bridges (involving W17, W57 and W58, Fig 1e). Therefore, in the DADE, W13 is a linker between the M- and P-sites. More interestingly, W13 is 0.80 Å from the position of the BeF_3^- fluoride atom F3 in BeF_3-RD (Fig. 2e) and, likewise, hydrogen bonds W58 similarly as to F3 interacts with W91. Besides, W17 locates close (1.0 Å) to where F2 is (Fig. 2e). Therefore, the position of W13 and W17 recalls the location of BeF_3^- fluorines. However, F2 and F3 and the two water molecules contact the P-site border in a different manner: F2

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

interacts with Ala84 N but W13 does not, and F2 contacts Thr83 O^{γ} but W17 does not either. Instead, W17 and Ala84 N interact with the Glu88 carboxylate, dragging Lβ4α4 towards the active site cavity (Arg85 C^{α} is displaced 2.10 Å). Furthermore, the similarity to the BeF_3-RD active sites is only true for DADE-A, as in DADE-B there are not hydrogen bonds connecting site M and P, no water molecule could be identified in this second P-site, and the Glu88 side chain has high temperature B-factors, indicating a certain disorder, in accordance with the poor definition of helix α4. The asymmetry between the two DADE active sites is also found in WT-RD, yet is much more marked in DADE.

6.2 The DAYC protomers display almost identical active sites—DAYC lacks the residue targeted during phosphorylation and one of the magnesium ligands. Nevertheless, both DAYC protomers show an electron density peak at the active site that cannot be attributed to a solvent molecule because of the distances to surrounding atoms would be too short (2.4 in average, see below for specific distances) and the coordination is not not tetrahedric but octahedric-minus-one. We considered the sodium or magnesium elements as an alternative because i) the former was omnipresent during the purification procedure and in the crystallization solution, ii) because magnesium and sodium are isoelectronic with solvent molecules and they can be distinguished by the surrounding ligands and the corresponding distances. Both sodium and magnesium typically display an octahedral co-ordination of six oxygen ligands, with a theoretical cation-oxygen distance of 2.46 Å for Na^+ and 2.07 Å for Mg^{2+} [54, 55]. During purification and, specially, crystallization the protein was exposed to a high concentration of Na^+ and rendered crystals whose structure (refined at 1.45 Å resolution, Fig. 2d, Table 1) display in both active sites a cation coordinated in an octahedral-minus-one or distorted trigonal-bipyramidal manner, as one equatorial position for metal binding disappears due to the side chain change at position 53. However, despite the mutation the metal coordination is maintained. The distances to the surrounding ligands range from 2.31 to 2.44 Å, and therefore we suggest it is a Na^+, as for Mg^{2+} the distances would be too long as well as this cation was not present during purification and crystallization. Seven oxygens coordinate the cation, which are provided by the following protein atoms (Fig. 2f): the ligands in the plane are Asp10 $O^{\delta 1}$, DAYC-A W2 (W12 in DAYC-B) and W3 (W13 in DAYC-B); the apical ligands are Met55 O and a further solvent molecule W1 (W11); the Na^{2+} metal coordination shell is therefore completed by three water molecules, which are at similar locations than in Mg^{2+}-RD and BeF_3-RD active sites [36; 40,] Fig. 2f. Thus, the sodium coordination atoms have the proper/active orientation, regardless a Na^+ is at the coordination site, as the key residues display quite conserved conformations when comparing all cation-containing PhoB RD structures. The rim-forming segment Trp54-Pro57, which changes Lβ3α3 into a perfect β-ribbon, is pushed laterally by Lβ4α4 and helix α4 as described above. Due to the displacement of Met55 O (Lβ3α3) and the absence of Asp53, the metal moves about 1.1 Å towards the Lβ3α3 edge in both DAYC-A and –B molecules, if compared to all three Mg-RD protomers [40], or 0.80 Å regarding BeF_3-RD [36]. The hydrogen bond network propagates from the M-site to the P-site by direct interaction of the metal-coordinating W3 (W13 in DAYC-B) to W6 (W9), being the latter a knot linking Lys105 N^{ζ}, Leu82 O and Glu88 $O^{\varepsilon 2}$ (Fig. 1f). This water tandem is located close to where the BeF_3^- molecule is in BeF_3-RD [36] (Fig. 1f) but, unlike in DADE-A, the water molecules are not close to the fluorine atoms (compare Figs. 1e,f). From the Mg^{2+}-bound structure, it was postulated that the cation attracts Trp54 to the active site, displacing Glu88 out from it and allowing the phosphate moiety to be bound [40]. Contrary, in DAYC the connection of Glu88 to the active site is not only maintained but intensified, as the carboxylate interacts with the metal sphere *via* W3 (W13 in DAYC-B), and is water-bridged to both the amide of Met55 (Lβ3α3), whose carbonyl is involved in Na^+ coordination, and to Lys105 N^{ζ} (Lβ5α5), which is implicated in phosphate binding in the WT protein (Fig. 2e,f). Concerning the Trp54 side chain, it is buried between helix α4 and β4 in all DAYC molecules and it therefore does not enter the active site.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Discussion

The asymmetry within the PhoB RD dimer might be a condition for PhoB activation

We have determined the three-dimensional structure of the RD from two constitutively active PhoB mutants, Asp10Ala /Asp53Glu (DADE) and Asp53Ala/Tyr102Cys (DAYC), in their unliganded (in the case of DADE), or sodium-bound (for DAYC) forms. In all cases, two protomers interact through the α1-Lβ5α5 interface, a contact also found in the apo- and Mg^{2+}-bound WT RD. In DADE, the two active sites of the protomers are not identical, in one the solvent content and some side chains are better defined than in the other. This asymmetry is again reminiscent of the apo and Mg^{2+}-bound WT RD dimers. Despite this asymmetry, the two active sites in DADE show the same orientation of Glu53 towards site M, where it performs interactions with surrounding water molecules mimicking the magnesium ion. In contrast, the P-site where the phosphate binds shows a different organization: in DADE-A Glu88 participates in an extensive active site hydrogen bond network that includes solvent molecules, resulting in a net of interactions which recalls the one found in the PhoB RD structure activated with BeF_3^- (BeF_3-RD). In contrast, in DADE-B the Glu88 side chain is less well defined and tentatively traced extruding from the active site and no solvent molecules fill the P-site. It has been shown that, upon Mg^{2+} binding, the Trp54 and Glu88 side chains alternate at the P-site [40]. In DADE-B, the Trp54 side chain is hardly identifiable in a weak density close to the cavity. The exact orientation of the indole ring cannot be ascertained, yet it is interesting to note once again the correlation between movements of the indole ring and the weakening of the definition of helix α4, as found in Mg-RD [40]. The differences between both active sites within the dimer especially affects the P-site and is more pronounced in DADE than in WT. Such asymmetry within the dimer, with one an active site bearing a poorer solvent content that correlates with a higher flexibility in helix α4, could be a condition needed for the activation process, which initiates in one of the two protomers, being DADE more prone to activation than the WT protein. In general, it is thought that RRs are activated by a conformational change that occurs within the RD upon phosphorylation, which is transmitted to the ED through the interaction surface between these domains. BeF_3-RD shows that upon activation the active site containing BeF_3 and Mg^{2+} is well defined although one edge of the cavity is deeply rearranged with respect to the non-bound RD: helix α4 is rotated and exposes to the surface residues that formerly were oriented towards the active site cavity. Despite we know this active PhoB RD form, we don't know what happens to the full-length protein, as neither the inactive or active structures of the whole molecule are available. Two structures of full-length RR from OmpR/PhoB family have been determined, i.e. DrrB [39] and DrrD [56]. These full-length structures are monomeric and show different arrangements for the RD and ED domains. In DrrD [56], a small interface between the C-terminal end of RD helix α5 is involved in a contact with the N-terminal four-stranded anti-parallel β-sheet of the ED. In contrast, in DrrB [39], a wider RD surface contacts the ED. This includes helix α4, β5 and the N-terminal part of helix α5 from the RD, which contact the ED residues from the N-terminal four-stranded anti-parallel β-sheet. It is particularly interesting to notice that Tyr97 (Tyr102 in PhoB), between helix α4 and β5 of the RD, interacts with the conserved Asp131 (Asp140 in PhoB) of the ED. A similar interaction could happen between PhoB domains, involving the Tyr102-Asp140 tandem, which would be disrupted upon phosporylation due to the rotamer change of Tyr102, therefore transmitting activating/inhibiting signals to the ED (comparisons between PhoB and DrrB have to be made with caution because the linker connecting RD with ED are of the different length and the residues at the N-terminal β-sheet of the ED are not conserved). In contrast to DADE, there is a clear increase in the rigidity of the whole DAYC molecule, with both protomers being highly symmetric, including helix α4 and the active site hydrogen bond network. Moreover, the DAYC RD structure shows that the mutation eliminates van-der-Waals interactions that stabilize helix α4, which is void of its last turn in the mutant. In case the Tyr102-Asp140 hydrogen bond was present in PhoB and played the same role as suggested for DrrB, the Tyr102Cys mutation

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

would directly cancel this interaction. The substitution at Tyr102 and the strong displacement of helix α4 towards the active site cavity may directly affect the association of the RD with the ED, giving rise to an activated protein bypassing magnesium-binding and phosphorylation and thus the requirements of an increase in of the overall flexibility. It is noteworthy that in OmpR a single change at the equivalent position of Tyr102 results in an activated protein [45].

Protein-protein contacts in PhoB involves different interfaces suggesting different models of activation

Although all RDs exhibit a similar fold, PhoB RD oligomers do not follow a canonical pattern of protein-protein interactions: helix α1, loop β5α5 and the N-terminus of helix α5 make up a two-fold related interface (α1-Lβ5α5 interface) not observed in other RDs. Structures of four different RD forms, namely WT [28; 36], Mg^{2+}-bound [40], DAYC and DADE have been crystallized in three different crystal lattices ($P2_12_12_1$, $P2_1$ and C2). Nevertheless, they show the same dimeric arrangement, which is not associated with any other relevant interaction in the crystal. In the Mg-RD structure, this dimer is even formed twice, related by a crystallographic and a local two-fold axis. Besides, size exclusion chromatography demonstrated that the RDs form stable dimers (our unpublished results) at least in the indicated buffer conditions. The fact that this interaction is systematically reproducible and the packing contacts are not equivalent in three of the four crystals strengthens the hypothesis that this is the actual PhoB dimerisation interface occurring when the effector domain is absent [33; 38]. As mentioned, the chemical stability prediction (employing the PISA server, http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver[52]) envisages this assembly to be formed in solution. Interestingly, almost all residues involved in the α1-Lβ5α5 contact have a similarity score of 0% within PhoB/OmpR family members, suggesting that the tendency of PhoB to dimerise through this interface might either be a particularity of this protein or, at least, if the family *relatives* do interact through this interface the contacts must be then very specific from case to case. The α1-Lβ5α5 dimerization makes possible a domains arrangement like in DrrB, with a contact between Tyr102 from the RD and Asp140 from the ED. This suggests a mechanism of activation. The change of rotamer of Tyr102 upon phosphorylation would disrupt the interaction with the ED and, together with the increment of flexibility of helix α4 that eliminated further contacts, would allow the two EDs from the same dimer to bind to two consecutive DNA sequences of the Pho Box motifs, as the interdomain link is long enough (6 residues) to permit the flexibility required for a head-to-head RD contacts compatible with a head-to-tail ED arrangement. A second model of activation is conceivable if consider the interaction through the α1-Lβ5α5 surface observed in BeF_3-RD [36], also predicted to be stable, where the two partners are rotated 90º with respect the WT (resulting in a 90-α1-Lβ5α5 interaction surface). In it, the RD C-terminal ends orientate to opposite directions, putatively placing the ED in the farthest possible distance from each other, thus making a simultaneous interaction with a DNA direct repeat very difficult. Nonetheless, the mentioned long length of the linker still should allow the ED to be independent from the RD arrangement and interact with a pho box. Therefore, accordingly to this second model, phosphorylation should convert the inactive α1-Lβ5α5 interaction to a 90-α1-Lβ5α5 active one triggered by a reciprocal rotation of the protomers. No other surface would be involved.

The alternative active interface proposed for PhoB [36; 57; 58] involves helix α4, strand β5 and helix α5 (α4-β5-α5 interface) and, in particular, the signalling residue Tyr102. The analysis of the α4-β5-α5 interface shows a high degree of conservation of the residues involved suggesting it is a common interface within the family, as shown by several PhoB homologous structures [36; 57; 58]. This surface is larger and less planar than the α1-Lβ5α5 interface although it contains a higher percentage of polar interactions (about 50% , Table 3) [36; 40], suggesting again a transient interface [51]. Analysis with the PISA server suggests that this surface is not stable in solution, as the protomers are more stable alone than in the dimer. The calculated free energy

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

values are of −122.4 and −118.1 kcal/mol for isolated B and C molecules (1ZES.pdb), but upon complex formation the free energy is slightly positive. Interestingly, the PhoB representative in *Bacillus subtilis*, $PhoP_N$ [59] shows the same functional [60] surface α4-β5-α5 although is packing in a head-to-tail manner against helices α3 and α4 and Lβ4α4 from the other protomer [59]. None of the PhoB packing contacts in any of the structures presented until now resemble the asymmetrical interaction found in $PhoP_N$. Therefore, despite the high conservation within the family, the strategies of the different molecules for oligomerisation seem to be heterogeneous. Each type of dimer potentially provides different orientations for the corresponding ED. In particular, almost all the segments participating in the α4-β5-α5 interface coincide with the ones that interact with the ED in the full-length structure of DrrB, and therefore both interactions cannot happen simultaneously. If the inactive protein has a DrrB-like interdomain contact while the active molecule performs a α4-β5-α5 interaction, a release of the ED by disruption of specific interactions must occur before the α4-β5-α5 surface interacts with another molecule. The advantage of the symmetrical arrangement α4-β5-α5 in the PhoB RD is the positioning of their C-terminal ends at the same edge of the dimer, suggesting a relative position of the DNA-binding structure elements in the full-length protein compatible with binding of a linear double-stranded DNA helix. However, none of the interactions envisaged for PhoB as to be active, α1-Lβ5α5, 90-α1-Lβ5α5 and α4-β5-α5 do not explain how a symmetrical (or two-fold) interaction of the RD elicits ED head-to-tail interactions in order to bind to direct repeat motifs in the DNA, as required for activation, unless the rationale also takes into account the length of the linker.

Materials and Methods

Mutant isolation

A genetic strategy was devised for isolation of $PhoB^{CA}$ mutants showing high level activation of PhoB in the absence of phosphorylation. The scheme took into account earlier observations that: (i) high level activation of the Pho regulon results in extremely slow growth under conditions of normal growth with Pi in excess [61], (ii) expression of *phnC*-to-*phnP* operon for use of phosphonates as an alternative phosphorus source requires high level activation by activated (phosphorylated) PhoB protein [62], and (iii) growth on phosphonates is slow [63]. Accordingly, standard allele-replacement methods [62] were used to construct an *E. coli* K-12 strain with a chromosomal *rhaBp-phoB* fusion in which PhoB synthesis requires induction by rhamnose [64], a functional *phnC*-to-*phnP* operon [62], and one of three mutations that block phosphorylation of PhoB (Asp53Ala, Asp53Gln, or Asp53Glu). The strain also had deletions of genes encoding the histidine kinases CreC and PhoR and genes for acetyl phosphate synthesis (*ackA* for acetate kinase and *pta* for phosphotransferase), so that it was incapable of phosphorylating PhoB by known pathways [65]. From this strain, spontaneous phosphonate utilizing mutants were selected on media containing rhamnose and methylphosphonate as sole phosphorus source. These mutants were tested for rhamnose-dependent growth on methylphosphonate, to test for dependency on synthesis PhoB (presumably a constitutively active PhoB not requiring phosphorylation), and for mutations genetically linked to the *rhaBp-phoB* fusion (encoding the sole copy of PhoB). Five independent mutants were found to carry $PhoB^{CA}$ mutations, as verified by cloning and DNA sequencing and by transfer of the mutant *phoB* allele into a new host. The $PhoB^{CA}$ DADE and DAYC showed high level activation of Pho regulon expression that is quantitatively similar to levels in a wild-type (*pho* (*BR*)+) host under conditions of Pi starvation. Further, transcriptional activation by the $PhoB^{CA}$ DADE and DAYC proteins was shown to be unaffected by PhoR, CreC, or acetyl phosphate (data not shown). Expression of these $PhoB^{CA}$ proteins, e. g., from native *phoBp* (phoB promoter), resulted in severe growth inhibition and rapid accumulation of secondary inactivating mutations, thus validating the use of such a strategy to obtain high level

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

PhoB^{CA} mutants. Details of the isolation and characterization of these and other PhoB^{CA} mutants will be described elsewhere (S.-K. Kim and B. L. Wanner, manuscript in preparation).

Recombinant protein overproduction and purification

PhoB RD DAYC and DADE were prepared and purified following a variation of the procedure used for the WT apo-PhoB RD preparation [66]. Accordingly, sequences encoding residues Met1 to Ala127 were PCR amplified, subcloned with *NcoI* and *HindIII* into the pBAT-4 vector [67], and verified by automated DNA sequencing. The resulting plasmids were transformed into *E. coli* strain BL21 (DE3) by heat-shock transformation, after which transformants were grown in LB medium containing 100 μg/ml ampicillin at 37ºC with aeration to early exponential phase (OD_{600nm} ~ 0.4–0.6). Cultures were cooled down to room temperature and 0.1 m*M* isopropyl-β-D-thiogalactopyranoside was added for induction. Recombinant protein production was carried out at ~ 25ºC in order to decrease the formation of inclusion bodies. After 13 h, cells were collected by centrifugation at 4.500 rpm. The cell pellet was resuspended in buffer A (50 m*M* NaCl, 1m*M* ethylenediamine tetraacetate (EDTA), 1 m*M* 1,4-dithio-D, L-threitol (DTT), 20m*M* 2-amino-2-hydroxymethyl-propane-1,3-diol (Tris-HCl, pH 8). Cells were disrupted by sonication, after which 100 mg/ml phenylmethanesulfonyl fluoride and 25 μl DNase I (Boehringer-Mannheim) were added per litre of initial culture. Soluble fractions were separated by ultracentrifugation. The supernatant was treated with polyethyleneimine at 0.1% (w/v) to precipitate DNA. After centrifugation and filtration, the sample was subjected to FPLC anion exchange chromatography employing a HiLoad Q-Sepharose column equilibrated with buffer B (same as buffer A without EDTA). Samples were eluted with an increasing NaCl gradient and analyzed by SDS-PAGE. The purest eluted fractions were further purified by gel filtration using a FPLC Superdex 75 HiLoad 26/60 column, equilibrated against buffer C (same as buffer B with 150 mM NaCl for DAYC or 200mM for DADE). Finally, after a polishing step via anion exchange Mono-Q column chromatography equilibrated with buffer B and elution with a gradient from 50 to 500 m*M* NaCl, PhoB RD DAYC purity reached more than 98% , as confirmed by SDS-PAGE, with a yield of 25 mg per litre of cell culture. Similarly, the purification procedure rendered 30 mg/l of PhoB RD DADE. Purified samples were dialysed against 100 mM NaCl, 10 mM Tris-HCl pH 8, flash frozen in liquid nitrogen and stored at −80ºC. In order to verify the elution behaviour of both native and DADE mutant, analytical size exclusion chromatography was performed using a Superdex 75 HR10/30 column (Amersham).

Protein crystallisation, data collection and processing

PhoB^{CA} DADE and DAYC crystals were obtained from hanging drops by the vapour-diffusion method at 20º C employing LINBRO crystallization plates and 0.5 ml of reservoir solution. For DAYC, drops contained 3 μl of protein solution at 5.5 mg/ml and 3 μl of reservoir solution (20% (w/v) polyethylene glycol 4000 (PEG 4K), 0.4 M sodium acetate, 0.1 M Tris.HCl (pH 8), 0.01 *M* DTT). Rectangular-shaped single crystals appeared after one week. Table 1 shows cell constants. There are two molecules in the asymmetric unit (a.u.; $V_m = 2.2$ Å^3 Da^{-1}, solvent content 43% [68]). Crystallization trials of the full-length DAYC protein were unsuccessful. In the case of DADE, drops containing 3 μl of protein solution at 5.6 mg/ml and 6μl a reservoir solution (22.5% PEG 4K, 0.4 *M* sodium acetate, 0.1 *M* Tris.HCl (pH 8)) rendered orthorhombic crystals ($P2_12_12_1$). These are isomorphous to the WT PhoB RD crystals [28; 66] (Table 1), with a dimer in the a.u. (V_m= 2.0 Å^3 Da^{-1}, solvent content 38%). In order to preserve DADE and DAYC crystals for transport and during data collection, they were flash cryo-cooled in liquid nitrogen after soaking them in a solution containing 35% PEG 4K for about 20 min. Complete X-ray diffraction data sets were collected from one single crystal each at 100K on CCD area detectors at synchrotron ESRF (Grenoble, France) beamlines ID14-EH2 and BM16 and at beamline BW7B at DESY (Hamburg, Germany). Data were indexed and integrated with

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

MOSFLM [69] and scaled, merged and reduced with SCALA [70]. Data collection and processing statistics are in Table 1.

Structure determination and refinement

The DAYC and DADE RD were both solved by molecular replacement [71] with AMoRe [72] using the WT PhoB RD X-ray structure as searching model (Protein Data Bank (PDB) accession code 1b00 [28]) and data between 15 and 3 Å resolution. In each case, the mutated residues were substituted by Ala in the searching model. No clear solution for DAYC was obtained when searching with a dimer. In contrast, a monomer yielded two clear solutions with a cumulative correlation coefficient in structure factor amplitudes (CC_F; [72]) of 46.1% (second highest unrelated peak, 32.4%) and a crystallographic residual R of 46.4% (second highest 51.8 %) for DAYC. Calculations performed with DADE structure factor amplitudes rendered a CC_F of 48.5% (second highest 25.1%) and an R of 45.0% (second highest 53.2%). The appropriately rotated and translated coordinates were refined with CNS v. 1.0 [73] and REFMAC5 [74] using all data and employing bulk-solvent and anisotropic temperature factor correction. The cross-validation R_{free} was monitored throughout with sufficient reflections (>500 75) not used during refinement. After initial simulated-annealing refinement, cycles of positional refinement and temperature factor refinement were alternated with manual model building using TURBO-FRODO [76] on a Silicon Graphics workstation. In the final steps of the refinement process, the cations were identified based on coordination geometry, ligand distances and definition in both σ_A-weighted $|2mF_{obs} - dF_{calc}|$- and $|mF_{obs}-dF_{calc}|$ -type Fourier maps at the expected coordination sites in both molecules A and B of each dimer. Solvent molecules were added at appropriate positions if a clear positive difference density was seen at 2σ. Anisotropic B-factor refinement did not improve the statistical parameters, therefore this option was not included in the refinement step. The final DAYC RD model consists of all atoms of residues 2A-126A and 2B-122B, two sodium ions at the magnesium-binding sites, one Tris molecule, and 303 solvent molecules. Alternate conformations were built for residues Met17A, Ser37A, Leu56A, Met120A, Cys102B, and Ile116B. The model of DADE RD includes residues 3A-123A and 2B-123B and 170 solvent molecules. Ramachandran plots generated with PROCHECK [77] locate no residues in disallowed regions in any of the structures. All structures show two *cis*-peptide bonds between residues 44 and 45 and 105 and 106. Final refinement statistics are in Table 1.

Miscellaneous

Least-squares superimpositions were carried out with LSQMAN [78]. In order to calculate distances between similar atoms of different monomers, only the cores (β-sheet strands) of the molecules were superimposed onto the A monomer of WT RD. Dimer superimpositions were calculated using all C^{α} atoms. The conservation values of residues were computed as follows: first, the thirty most similar sequences to PhoB were determined with BLAST [79] and subjected to multiple alignment using Clustal W [80]. The resulting sequence superposition was submitted to ESPript [81] which converted the similarity values to pseudo-temperature factors integrated to the PDB coordinates. Figures have been made with GRASP [82], MOLMOL [83] and TURBO-FRODO [76]. The interface area between molecules was calculated using CNS. Identification of residues involved in the protomer interactions was performed by visual inspection of the interfaces. The corresponding stereochemical analysis was done by submitting the coordinates at the protein-protein interactions server www.biochem.ucl.ac.uk/bsm/PP/server/index.html [84], as well as the stability of the interactions was analysed with PISA (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html52).

Acknowledgements

We are in debt with Jenny Colom for help during protein purification. R. A. B. was recipient of fellowships from the Generalitat de Catalunya and the Ministerio de Educación y Cultura (Spain). S.K.K. was supported by the Faculty

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Research Fund of Konkuk University. A.G.B. was recipient of a fellowship from the Ministerio de Sanidad y Consumo (Spain), and PJBP from FEBS (Long-term Fellowship). This work was supported by the Spanish Ministerio de Educación y Ciencia (grants BFU2005-512028, GEN2003-20642 and BIO2003-132) and the Generalitat de Catalunya (Centre de Referència en Biotecnologia and grant 2005SGR-00280). B.L.W. was supported by NIH GM62662. M.S. is a beneficiary of the "Ramon y Cajal" Program from the Ministerio de Educación y Ciencia from Spain. The fruitful interaction between the laboratories was enabled by a FULBRIGHT Research Grant from the Spanish-US American Program for Scientific and Technological Cooperation. The coordinates for the refined DADE and DAYC PhoB RD models have been deposited to the Protein Data Bank at www.ebi.ac.uk/msd, with accession codes WWWW and XXXX.

References

1. Hoch JA. Two-component and phosphorelay signal transduction. Curr Op Microbiol 2000;3:165–170.
2. Stock AM, Robinson VLNGP. Two-component signal transduction. Annu Rev Biochem 2000;69:183–215. [PubMed: 10966457]
3. Wang W, Zhang W, Chen H, Chiao J, Zhao G, Jiang W. Molecular and biochemical characterization of a novel two-component signal transduction system, amrA- amkA, involved in rifamycin SV production in *Amycolatopsis mediterranei* U32. Arch Microbiol 2002;178:376–386. [PubMed: 12375106]
4. Wolanin PM, Thomason PA, Stock JB. Histidine protein kinases : key signal transducers outside the animal kingdom. Genome Biol 2002;3:reviews3013.1–3013.8. [PubMed: 12372152]
5. Walsh, C. Antibiotics: Actions, origins, resistance. 1. ASM Press; Washington, D.C: 2003.
6. Forsberg J, Rosenquist M, Fraysse L, Allen JF. Redox signalling in chloroplasts and mitochondria : genomic and biochemical evidence for two-component regulatory systems in bioenergetic organelles. Biochem Soc Trans 2001;29:403–407. [PubMed: 11497997]
7. Santos JL, Shiozaki K. Fungal histidine kinases. Sci STKE 2001;2001:RE1. [PubMed: 11752677]
8. Grefen C, Harter K. Plant two-component systems: principles, functions, complexity and cross talk. Planta 2004;219:733–42. [PubMed: 15232695]
9. Saito H. Histidine phosphorylation and two-component signaling in eukaryotic cells. Chem Rev 2001;101:2497–2509. [PubMed: 11749385]
10. Lohrmann J, Sweere U, Zabaleta E, Baurle I, Keitel C, Kozma-Bognar L, Brennicke A, Schafer E, Kudla J, Harter K. The response regulator ARR2: a pollen-specific transcription factor involved in the expression of nuclear genes for components of mitochondrial complex I in Arabidopsis. Mol Genet Genomics 2001;265:2–13. [PubMed: 11370868]
11. Sweere U, Eichenberg K, Lohrmann J, Mira-Rodado V, Bäurle I, Kudla J, Nagy F, Schäfer E, Harter K. Interaction of the response regulator ARR4 with phytochrome B in modulating red light signaling. Science 2001;294:1108–1111. [PubMed: 11691995]
12. Yonekura-Sakakibara K, Kojima M, Yamaya T, Sakakibara H. Molecular characterization of cytokinin-responsive histidine kinases in maize. Differential ligand preferences and response to *cis*-zeatin. Plant Physiol 2004;134:1654–1661. [PubMed: 15064375]
13. Papon N, Clastre M, Gantet P, Rideau M, Chenieux JC, Creche J. Inhibition of the plant cytokinin transduction pathway by bacterial histidine kinase inhibitors in *Catharanthus roseus* cell cultures. FEBS Lett 2003;537:101–105. [PubMed: 12606039]
14. Barrett JF, Hoch JA. Two-component signal transduction as a target for microbial anti-infective therapy. Antimicrob Agents Chemother 1998;42:1529–1536. [PubMed: 9660978]
15. Stephenson K, Hoch J. Two-component and phosphorelay signal-transduction systems as therapeutic targets. Curr Op Pharmacol 2002;2:507–512.
16. Lukat GS, Stock AM, Stock JB. Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis. Biochemistry 1990;29:5436–5442. [PubMed: 2201404]
17. McCleary WR, Stock JB. Acetyl phosphate and the activation of two-component response regulators. J Biol Chem 1994;269:31567–31572. [PubMed: 7989325]
18. Aizawa SH, Harwood C, Kadner RJ. Signaling components in bactrial locomotion and sensory reception. J Bacteriol 2000;182:1459–1471. [PubMed: 10692349]
19. Falke JJ, Hazelbauer GL. Transmembrane signaling in bacterial chemoreceptors. Trends Biochem Sci 2001;26:257–265. [PubMed: 11295559]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

20. Shinagawa H, Makino K, Nakata A. Regulation of the *pho* regulon in *Escherichia coli* K-12. Genetic and physiological regulation of the positive regulatory gene. phoB J Mol Biol 1983;168:477–488.
21. Wanner BL. Gene regulation by phosphate in enteric bacteria. J Cell Biochem 1993;51:47–54. [PubMed: 8432742]
22. Wanner, BL. Phosphorous assimilation and control of the phosphate regulon. In: Neidhardt, FC.; Curtiss, RI.; Ingraham, JL.; Lin, EC.; Low, KBJ.; Magasanik, B.; Reznikoff, W.; Riley, M.; Schaechter, M.; Umbarger, HE., editors. Escherichia coli and Salmonella typhimurium cellular and molecular biology. American Society for Microbiology; Washington, D.C.: 1996. p. 1357-1381.
23. Stock JB, Ninfa AJ, Stock AM. Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol Rev 1989;53:450–490. [PubMed: 2556636]
24. Makino K, Shinagawa H, Amemura M, Yamasa M, Nakata A. Signal transduction in the phosphate regulon of *Escherichia coli* involves phosphotransfer between PhoR and PhoB protein. J Mol Biol 1989;210:551–559. [PubMed: 2693738]
25. Parkinson JS. Signal transduction schemes of bacteria. Cell 1993:857–871. [PubMed: 8098993]
26. Hackenbeck R, Stock JB. Analysis of two-component signal transduction system involved in transcriptional regulation. Meth Enzymol 1996;273:281–300. [PubMed: 8791619]
27. Ellison DW, McCleary WR. The unphosphorylated receiver domain of PhoB silences the activity of its output domain. J Bacteriol 2000;182:6592–6597. [PubMed: 11073900]
28. Solà M, Gomis-Rüth FX, Serrano L, González A, Coll M. Three-dimensional crystal structure of the transcription factor PhoB receiver domain. J Mol Biol 1999:675–687.
29. Blanco AG, Solà M, Gomis-Rüth FX, Coll M. Tandem DNA recognition by PhoB, a two-component signal transduction transcriptional activator. Structure 2002;10:701–713. [PubMed: 12015152]
30. Wanner, BL. Phosphate signaling and the control of gene expression in *Escherichia coli*. In: Silver, S.; Walden, W., editors. Metal ions in gene regulation. Chapman and Hall, Ltd.; Sterling, Va: 1997. p. 104-128.
31. Willsky GR, Malamy MH. Characterization of two genetically separable inorganic phosphate transport systems in *Escherichia coli*. J Bacteriol 1980;144:356–365. [PubMed: 6998957]
32. Steed PD, Wanner BL. Use of *rep* technique for allele replacement to construct mutants with deletions of the *pstSCAB-phoU* operon: evidence of a new role of the PhoU protein in the phosphate regulon. J Bacteriol 1993;175:6797–6809. [PubMed: 8226621]
33. McCleary WR. The activation of PhoB by acetylphosphate. Mol Microbiol 1996;20:1155–1163. [PubMed: 8809768]
34. Makino K, Amemura M, Kawamoto T, Kimura S, Shinagawa H, Nakata A, Suzuki M. DNA binding domain of PhoB and its interaction with RNA polymerase. J Mol Biol 1996;256:15–26. [PubMed: 8648643]
35. Oganesyan V, Oganesyan N, Adams PD, Jancarik J, Yokota HA, Kim R, Kim SH. Crystal structure of the "PhoU-like" phosphate uptake regulator from *Aquifex aeolicus*. J Bacteriol 2005;187:4238–4244. [PubMed: 15937186]
36. Bachhawat P, Swapna GV, Montelione GT, Stock AM. Mechanism of Activation for Transcription Factor PhoB Suggested by Different Modes of Dimerization in the Inactive and Active States. Structure (Camb) 2005;13:1353–63. [PubMed: 16154092]
37. Allen MP, Zumbrennen KB, McCleary WR. Genetic evidence that the alfa5 helix of the receiver domain of PhoB is involved in interdomain interactions. J Bacteriol 2001;183:2204–2211. [PubMed: 11244058]
38. Fiedler U, Weiss V. A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules. EMBO J 1995;14:3696–3705. [PubMed: 7641688]
39. Robinson VL, Wu T, Stock AM. Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. J Bacteriol 2003;185:4186–94. [PubMed: 12837793]
40. Solà M, Drew DL, Blanco AG, Gomis-Rüth FX, Coll M. The co-factor-induced pre-active conformation in PhoB. Acta Cryst D Biol Crystallogr 2006;62:1046. [PubMed: 16929106]
41. Zundel CJ, Capener DC, MacCleary WR. Analysis of the conserved acidic residues in the regulatory domain of PhoB. FEBS letters 1998;441:242–246. [PubMed: 9883892]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

42. Volz K. Structural conservation in the CheY superfamily. Biochem 1993;32:11741–11753. [PubMed: 8218244]
43. Zhu X, Amsler CD, Volz K, Matsumura P. Tyrosine 106 of CheY plays an important role in chemotaxis signal transduction in *Escherichia coli*. J Bacteriol 1996;178:4208–4215. [PubMed: 8763950]
44. Zhu X, Rebello J, Matsamura P, Volz K. Crystal structure of CheY mutants Y106W and T87I/Y106W. CheY activation correlates with movement of residue 106. J Biol Chem 1997;272:5000–5006. [PubMed: 9030562]
45. Kanamaru K, Mizuno T. Signal transduction and osmoregulation in *Escherichia coli*: a novel mutant of the positive regulator, OmpR, that functions in a phosphorylation-independent manner. J Biochem 1992;111:425–430. [PubMed: 1618729]
46. Hiratsu K, Nakata A, Shinagawa H, Makino K. Autophosphorylation and activation of transcriptional activator PhoB of *Escherichia coli* by acetyl phosphate in vitro. Gene 1995;161:7–10. [PubMed: 7642140]
47. IUPAC-IUB Commission on Biochemical Nomenclature. Abbreviations and symbols for the description of the conformation of polypeptide chains. J Mol Biol 1970;52:1–17. [PubMed: 5485910]
48. Lovell SC, Word JM, Richardson JS, Richardson DC. The penultimate rotamer library. Proteins 2000;40:389–408. [PubMed: 10861930]
49. Park S, Meyer M, Jones AD, Yennawar HP, Yennawar NH, Nixon BT. Two-component signaling in the AAA + ATPase DctD: binding Mg2+ and BeF3- selects between alternate dimeric states of the receiver domain. Faseb J 2002;16:1964–6. [PubMed: 12368235]
50. Lee SH, Cho HS, Pelton JG, Yan D, Berry EA, Wemmer DE. Crystal structure of activated CheY. Comparison with other activated receiver domains. J Biol Chem 2001;276:16425–16431. [PubMed: 11279165]
51. Nooren IM, Thornton JM. Structural characterisation and functional significance of transient protein-protein interactions. J Mol Biol 2003;325:991–1018. [PubMed: 12527304]
52. Krissinel, E.; Henrick, K. Detection of Protein Assemblies in Crystals. In: Berthold, MR.; Glen, R.; Diederichs, K.; Kohlbacher, O.; Fischer, I., editors. Computational Life Sciences: First International Symposium, CompLife 2005, Konstanz, Germany, September 25–27, 2005. Proceedings. 3695. Springer Berlin; Heidelberg, Berlin/Heidelberg: 2005. p. 163
53. Bellsolell L, Prieto J, Serrano L, Coll M. Magnesium binding to the bacterial chemotaxis protein CheY results on large conformational changes involving its functional surface. J Mol Biol 1994;238:489–495. [PubMed: 8176739]
54. Harding MM. The geometry of metal-ligand interactions relevant to proteins. Acta Crystallogr D Biol Crystallogr 1999;55 (Pt 8):1432–43. [PubMed: 10417412]
55. Harding MM. Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. Acta Crystallogr D Biol Crystallogr 2002;58:872–4. [PubMed: 11976508]
56. Buckler DR, Zhou Y, Stock AM. Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from *Thermotoga maritima*. Structure (Camb) 2002;10:153–164. [PubMed: 11839301]
57. Bent CJ, Isaacs NW, Mitchell TJ, Riboldi-Tunnicliffe A. Crystal structure of the response regulator 02 receiver domain, the essential YycF two-component system of Streptococcus pneumoniae in both complexed and native states. J Bacteriol 2004;186:2872–9. [PubMed: 15090529]
58. Toro-Roman A, Mack TR, Stock AM. Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: A symmetric dimer mediated by the α4-β5-α5 face. J Mol Biol 2005;349:11–26. [PubMed: 15876365]
59. Birck C, Chen Y, Hulett FM, Samama JP. The crystal structure of the phosphorylation domain in PhoP reveals a functional tandem association mediated by an asymmetric interface. J Bacteriol 2003;185:254–61. [PubMed: 12486062]
60. Chen Y, Birck C, Samama JP, Hulett FM. 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 2003;185:262–73. [PubMed: 12486063]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

61. Steed PM, Wanner BL. Use of the rep technique for allele replacement to construct mutants with deletions of the pstSCAB-phoU operon: evidence of a new role for the PhoU protein in the phosphate regulon. J Bacteriol 1993;175:6797–6809. [PubMed: 8226621]
62. Metcalf WW, Jiang W, Daniels LL, Kim SK, Haldimann A, Wanner BL. Conditionally replicative and conjugative plasmids carrying lacZ for cloning, mutagenesis, and allele replacement in bacteria. Plasmid 1996;35:1–13. [PubMed: 8693022]
63. Wanner BL, Boline JA. Mapping and molecular cloning of the phn (psiD) locus for phosphonate utilization in Escherichia coli. J Bacteriol 1990;172:1186–1196. [PubMed: 2155195]
64. Haldimann A, Daniels LL, Wanner BL. Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the Escherichia coli phosphate regulon. J Bacteriol 1998;180:1277–1286. [PubMed: 9495769]
65. Wanner BL, Wilmes-Riesenberg MR. Involvement of phosphotransacetylase, acetate kinase, and acetyl phosphate synthesis in the control of the phosphate regulon in Escherichia coli. J Bacteriol 1992;174:2124–2130. [PubMed: 1551836]
66. Solà M, Gomis-Rüth FX, Guasch A, Serrano L, Coll M. Overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the receiver domain of PhoB. Acta Crystallogr sect D 1998;54:1460–1463. [PubMed: 10089538]
67. Peränen J, Rikkonen M, Hyvonen M, Kääriäinen L. T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. Anal Biochem 1996;236:371–373. [PubMed: 8660525]
68. Matthews BW. Solvent content of protein crystals. J Mol Biol 1968;33:491–497. [PubMed: 5700707]
69. Leslie, AGW. Macromolecular data processing. In: Moras, D.; Podjarny, AD.; Thierry, JC., editors. Crystallographic computing V. Oxford University Press; Oxford: 1991. p. 27-38.
70. Evans, P. Data reduction. In: Sawyer, L.; Isaacs, N.; Bailey, S., editors. Data collection and processing - Proceedings of the CCP4 Study Weekend 29–30 January 1993. SERC Daresbury Laboratory; Warrington: 1993. p. 114-122.
71. Huber R. Die automatisierte Faltmolekülmethode. Acta Crystallogr sect A 1965;19:353–356.
72. Navaza J. AMoRe : an automated package for molecular replacement. Acta Crystallogr sect A 1994;50:157–163.
73. Brünger AT. Crystallographic phasing and refinement of macromolecules. Curr Op Struct Biol 1991;1:1016–1022.
74. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr sect D 1997;53:240–255. [PubMed: 15299926]
75. Brunger AT. Free R value: Cross-validation in crystallography. Methods Enzymol 1997;277:366–396.
76. Carranza, C.; Inisan, AG.; Mouthuy-Knoops, E.; Cambillau, C.; Roussel, A. AFMB Activity Report 1996–1999. Marseille: 1999. Turbo-Frodo; p. 89-90.CNRS-UPR 9039
77. Laskowski RA, McArthur MW, Moss DS, Thornton JM. PROCHECK : a program to check the stereochemical quality of a protein structure. J Appl Crystallogr 1993;26:283–291.
78. Kleywegt GJ, Jones TA. A super position. ESF/CCP4 Newsletter 1994;31:9–14.
79. Altschul SF, Koonin EV. Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. Trends Biochem Sci 1998;23:444–7. [PubMed: 9852764]
80. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80. [PubMed: 7984417]
81. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999;15:305–8. [PubMed: 10320398]
82. Nicholls A, Sharp K, Honig B. Graphical Representation and Analysis of Structural Proteins. PROTEINS, Structure, Function and Genetics 1991;11:281.
83. Koradi R, Billeter M, Wüthrich K. MOLMOL: A program for display and analysis of macromolecular structures. J Mol Graphics 1996;14:51–55.
84. Jones S, Thornton JM. Principles of protein-protein interactions. Proc Natl Acad Sci U S A 1996;93:13–20. [PubMed: 8552589]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Abbreviations

a.u
asymmetric unit

BeF_3^-
beryllium fluoride

BeF_3-RD
PhoB RD in complex with Mg^{2+}/ BeF_3^- structure

DADE
Asp10Ala and Asp53Glu

DAYC
Asp53Ala and yr102Cys

ED
effector domain

Mg-RD
Mg2+-bound PhoB RD structure, Mg-A, protomer A of Mg^{2+}-bound PhoB RD

MW
molecular weight

$PhoB^{CA}$
constitutively active PhoB mutants

RD
receiver domain

rmsd
root mean square deviation

RR
response regulator

WT
wild- type

WT-RD
PhoB RD structure

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

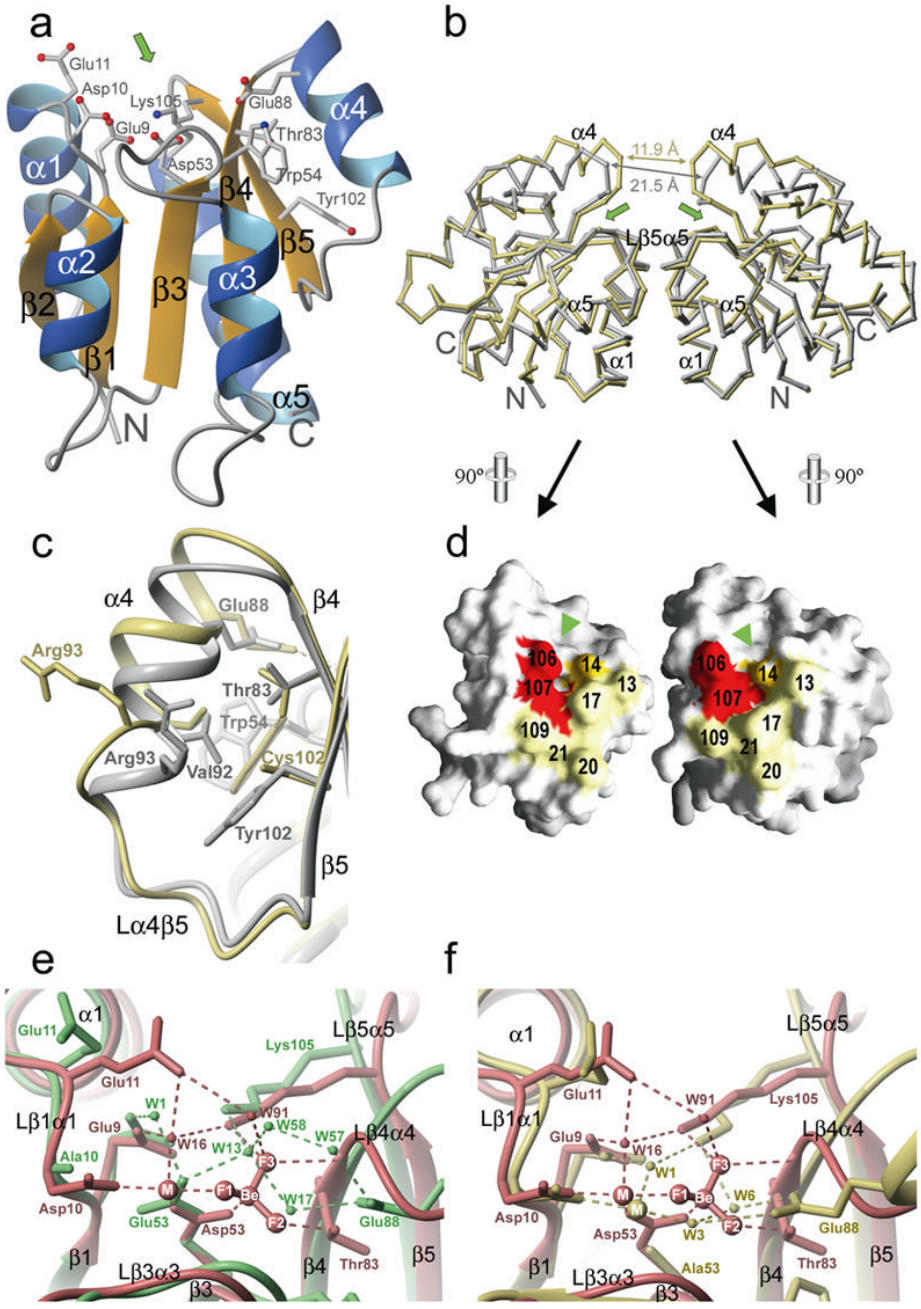

Figure 1.

a) *The structure of PhoB RD.* Richardson diagram of PhoB RD. Helices are shown as blue ribbons (α1 to α5) and β-strands as yellow arrows (β1 to β5). The N- and C-terminal ends are indicated (N and C, respectively). The green arrow points to the active site cavity. Key residue side chains mentioned throughout the text are highlighted. **b)** *Dimers interacting through the α1-Lβ 5α5 interface.* C^{α} plot showing the superimposition of WT RD (grey) and DAYC-Na (khaki) dimers. Secondary structure elements belonging to the interface are indicated; the distance between vicinal Lβ4α4 loops within a dimer is shown. The green arrows point to the active site cavities of each protomer within a dimer. The C- and N-terminal ends of the RDs are shown (N and C, respectively). In each case, molecules A are on the left and molecules B

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

on the right of the dimer. **c)** *Packing of WT (grey) and DAYC (khaki) helix α4 against the protein core*. Mutated residues or side chains that change the conformation are indicated twice, with colour coding according the molecule they belong to. α helices and β strands numbering is indicated. **d)** *Each constituting monomer within the WT RD dimer* (superimposed with its solid Connolly surface) has been rotated vertically 90º from the position occupied in (b) to grant insight into the interacting surface. The residues of each molecule participating in direct contacts are shown and labelled. The colour coding reflects conservation of each position, ranging from 0% (light yellow) to 100% (intense red). The approximate location of the active site cleft is indicated by a green arrowhead. e) *DADE (green) and BeF_3-RD [36] (coral) active sites are superimposed* (strands β1, β3 and β4 where used for an optimal superposition of residues in the active site). Mutated residues or side chains that change the conformation are indicated twice. The small spheres represent water (W) molecules, which are numbered according to the PDB. The M sphere represents the coordinated metal; the Be sphere the Beryllium and F1, F2 and F3 the fluorine atoms in the BeF_3^- molecule, an analogue of the activating phosphate moiety. Electrostatic and hydrogen bond interactions are represented by dashed lines and, together with water oxygen atoms, are coloured according to the PDB they belong to. W3 and the interaction between the cation and Met55 carbonyl are not shown for clarity. Secondary structure elements are indicated. f) *DAYC (khaki) and BeF_3-RD [36] (coral) are superimposed around their active sites*. The criteria of representation are the same as for e).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

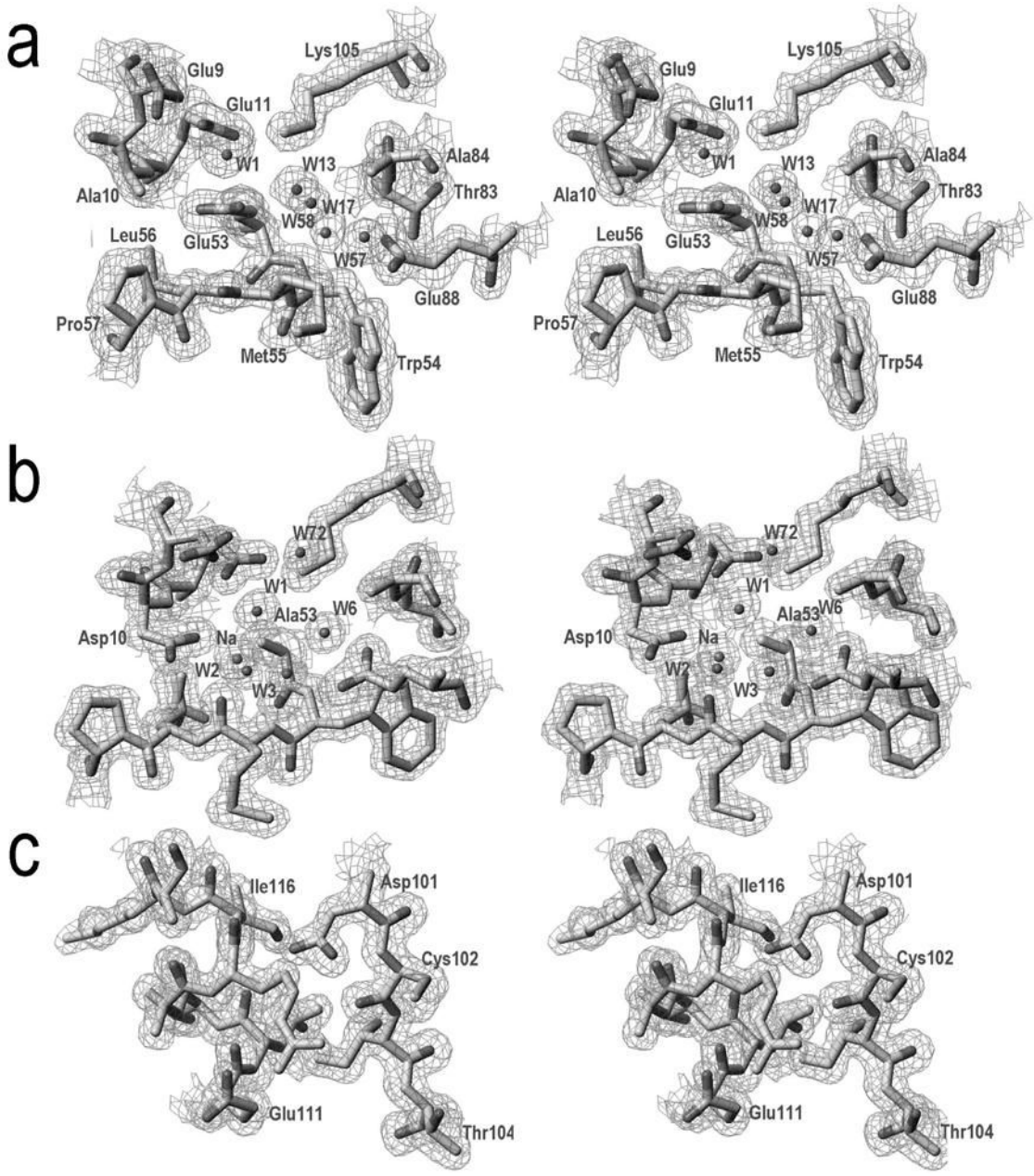

Figure 2. The active-site of PhoB RD

Stereo diagrams displaying the final models superimposed with their σ_A-weighted $(2mF_{obs}-dF_{calc})$-type electron density maps (contoured at 1σ). In each case, equivalent protein segments around the cation-binding site (positions 9 to 11; 53 to 57; 83 to 84; 88 and 105) of **(a)** DADE and **(b)** DAYC are shown. In **(c)** the final structure around the second mutation site of DAYC molecule A (position 102) is shown, including part of strand β5 and helix α5. In (a), all residues are labelled, whereas in (b) and (d) only those differing from (a) are denoted.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Data collection and refinement statistics.

	PhoB RD D53A+Y102C (DAYC)	PhoB RD D10A+D53E (DADE)
Wavelength (Å)	0.933	0.933
Space group	P 2_1	P $2_12_12_1$
Cell parameters (Å; º)	a=44.9, b=47.6, c=59.7, β = 99.3	a=32.9, b=60.2, c=116.5
No. of measurements	134,031	93,822
No. of unique reflections	41,837	24,245
Whole resolution range / last shell (Å)	58.7 – 1.45 / 1.53 – 1.45	32.9 – 1.70 / 1.79 – 1.70
Multiplicity	3.2	2.9
R_{merge} (%) [a]	6.7 (29.0)[b]	5.4 (34.7)
Completeness (%)	94.5 (70.3)	93.0 (96.6)
<I>/ σ(I)	15.2 (2.3)	14.3 (4.0)
Resolution range (Å)	50.0 – 1.45	32.9 – 1.70
Number of reflections (working set / test set)	41,049 / 770	23,577 / 638
R (%)[c]	18.9	22.3
R_{free} (%) [d]	20.4	25.6
Rmsd bonds (Å)	0.010	0.012
Rmsd angles (º)	1.31	1.94
No. of non-hydrogen protein atoms	1,977	1,951
Water molecules	303	170
Metal ions	2 (Na^+), 1 (Tris)	—

[a] $R_{merge} = \{\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)\}$x100, where $I_i(hkl)$ is the i^{th} intensity measurement of reflection *hkl* and <I(*hkl*)> is its a verage intensity.

[b] Values in parentheses correspond to the last shell.

[c] $R_{factor} = \{\sum_{hkl} ||F_{obs}| - k|F_{calc}||/\sum_{hkl}|F_{obs}|\}$x100

[d] Free R_{factor}: R_{factor} for a test set of re flections (> 500) not used during the refinement.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Root mean squares deviation between structures of protomers.

Dimer (protomer A / protomer B)	whole dimer[1]	monomers[2]	[3] WT-A/B against mutant-A/B	[4] WT B/A against mutant-A/B	[5] WT-A against mutant-A	WT-B against mutant-B	WT-B against mutant-A	WT-A against mutant-B
DADE-A / DADE-B	0.54 (240) 0.22 (240)	0.52 (120) 0.23 (120)	0.43 (242) 0.62 (229)	0.56 (241) 0.63 (225)	0.38 (121) 0.65 (118)	0.41 (120) 0.55 (120)	0.51 (121) 0.53 (119)	0.53 (120) 0.64 (118)
DAYC-A / DAYC-B								

The rmsd is calculated from the C^{α} atoms that devia te less than 3Å (number of a toms indicated in brackets). Values are given in Å.

[1] *rmsd* between C^{α} atoms of the whole dimer against itself (Mol1/Mol2 against Mol2/Mol1)

[2] *rmsd* between C^{α} atoms of the two monomers in a dimer (Mol1 against Mol2)

[3] *rmsd* between dimers, WT protomers A and B are superposed to mutant protomers A and B.

[4] *rmsd* between dimers, WT protomers B and A are superposed to mutant protomers A and B.

[5] The following 4 columns relate each WT protomer within a mutant protomer.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Interface stereochemical parameters between PhoB protomers.

Dimer In terface-type Area (Å2)	**WT α1-Lβ5α5 (560)***		**BeF_3 α4-β5- α5 (890)***	
Protein Interface Parameter&	**A**	**B**	**B**	**C**
In terface accessible surface area (ASA; in $Å^2$)	562	610	977	981
% Interface ASA of to tal surface	8.35	8.97	15.10	15.47
Planarity (*rmsd*)	1.34	1.53	3.23	3.26
Length/Breadth ratio	0.90	0.91	0.86	0.91
Segments at in terface	2	2	4	3
% Polar atoms in interface	27.9	27.2	47.0	51.6
% Non-polar atoms in interface	72.0	72.8	52.9	48.3
Gap volume index (Å)	1.75	1.75	1.41	1.41

* Total interaction surface divided by 2.

For definitions, see [61].