The X-ray Crystal Structures of Two Constitutively Active Mutants of the *Escherichia coli* PhoB Receiver Domain Give Insights into Activation

Raquel Arribas-Bosacoma¹, Soo-Ki Kim², Cristina Ferrer-Orta¹
Alexandre G. Blanco¹, Pedro J.B. Pereira¹, F. Xavier Gomis-Rüth¹
Barry L. Wanner², Miquel Coll¹ and Maria Solà¹*  

¹Institut 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  
²Department of Biological Sciences, Purdue University West Lafayette, IN 47907, USA  

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 Mg2+ 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 dimerization 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 β5a5 and N terminus of α5 elements, also seen in the complex of PhoB RD with Mg2+ in which helix α4 highly increases its flexibility. PhoB RD in complex with Mg2+ and BeF3− (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 (PhoBCA) 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 PhoBCA 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.

© 2006 Elsevier Ltd. All rights reserved.

*Corresponding author

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

Present addresses: S.-K. Kim, Department of Animal Science and Environment; Konkuk University; 1 Hwayang-dong, Gwangin-gu; Seoul 143-701, Korea; P. J.B. Pereira, Grupo de Estrutura Molecular, IBMC; Rua do Campo Alegre, 823; 4150-180 Porto, Portugal; A. G. Blanco, Instrumentation Laboratory; 101 Hartwell Av., Lexington, MA 02421, USA.

Abbreviations used: TCS, two-component system; a.u., asymmetric unit; BeF3−, beryllium fluoride; BeF3-RD, PhoB RD in complex with Mg2+/BeF3− structure; DADE, Asp10Ala and Asp53Glu; DADY, Asp53Ala and Tyr102Cys; ED, effector domain; Mg-RD, Mg2+-bound PhoB RD structure, Mg-A, protomer A of Mg2+-bound PhoB RD; MW, molecular weight; PhoBCA, constitutively active PhoB mutants; RD, receiver domain; RR, response regulator; WT, wild-type; WT-RD, PhoB RD structure.

E-mail address of the corresponding author: msvcri@ibmb.csic.es

0022-2836/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
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, TCSs rule chemotaxis, quorum sensing, nutrient uptake, nitrogen fixation, osmoregulation, sporulation and pathogenic host invasion.1–4 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 organelles6–13 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 antimicrobials in order to increase the medical armamentarium available to evade the menace of increasing bacterial antibiotic resistance.14,15 TCSs armamentarium available to evade the menace of antimicrobials in order to increase the medical envisaged as valid targets for the development of new and, as they are absent in mammals, they have been envisaged as valid targets for the development of new antimi...
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<sup>C5-A</sup>) 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**

**Constitutively active PhoB mutants**

The two constitutively active PhoB mutants examined here, DADE and DAYC, have the wild-type (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.1 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 Asp53Ala with Tyr102Cys (DAYC) results in constitutively active forms of PhoB (PhoBCA). Moreover, the single mutation Asp10Ala also present in DADE also renders a non-phosphorylatable 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, confer a net electro-negative charge to the active-site cavity,28 and their modification alters the metal-binding capacity and the activity of the protein.31 The other position affected, Tyr102 in DAYC, is far from the acidic active site28 and is highly conserved as an aromatic residue;32 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.

**Key residues and structural features of PhoB RD**

The PhoB RD structure (PDB, 1b00; hereinafter referred as WT-RD), like the other TCS RR members, shows a doubly-wound (α/β)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 (Figure 1(a)). The central β-sheet shows a β2|1|3|4|5 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 (Lβ1x1, L3αx3 and L4|5x5, respectively) surround a wide cavity that harbours the active site (Figure 1(a) and (b)). Inside this, there is the invariant 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 terminal of helix α2 (Asp36) and helix α4 (Glu87, Glu88, Glu89 and Asp90).23,46 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 Lβ2α4 in WT-A), helix α4 and the loop Lβ4α4.5,28 Between helices α3 and α4 there is Trp54 (Figure 1(a), 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

**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-Lβ5α5 interface. Cα 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 for clarity. 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 BeF3-RD49 (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 the coordinated metal; the Be sphere the Beryllium and F1, F2 and F3 the fluorine atoms in the BeF3 molecule, an analogue of the activating phosphate moiety. Electrostatic and hydrogen bond interactions are represented by broken 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 BeF3-RD49 (coral) superimposed around their active sites. The criteria of representation are the same as for (e).
other through an interface of 556 Å² per protomer. However, upon Mg²⁺ binding, helix α4 becomes highly flexible and Thr54 rotates its side chain to the gauche(g)°-90 rotamer.17 (g° refers to the χ’ orientation of −60°)36 entering into the active site, where it performs several interactions.62 Between helix α4 and Lβ4eα4 there is Tyr102 in (β5; Figure 1(a)), 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 (Figure 1(a) and (c)); (b) g+ conformation (in g°, χ’ orientation is of +60°), with the side-chain pointing towards the bulk solvent or interacting with the ED, like in the OmpR-family member DrrB39 or with another protein, as in the inactivated DetD dimer40, and (c) l conformation, with the side-chain pointing towards Lβ4eα4 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 (Figure 1(a)), 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”.30

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 RD28,36,40 but not in any other RR structure. The constituting protomers, WT-A and WT-B, are related by an almost perfect non-crystallographic 2-fold axis (Figure 1(b)), 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, Lβ5eα5 and the N terminus of helix α5 from both monomers (thus, “α1-Lβ5eα5 interface”; Figure 1(b)). The interface area spans 560 Å², thus is rather small and points to a transient homodimeric contact1 (Table 1), a notion supported by some experiments in solution.36 The residues involved are Pro13, Ile14, Met17, Phe20, Val21, Pro106, Phe107 and Pro109 (Figure 1(d)). 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 (Figure 1(b)). The two molecules in the DADE crystal asymmetric unit (a.u.), DADE-A and DADE-B, interact with each other through an interface of 556 Å² per protomer. Each molecule contributes with the same residues as for WT plus Gln24 (the stereochemical parameters are listed in Table 1). Considering the DAYC mutant, the structure shows again the same interface under burial of 538 Å² 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 Lα4β5 as a novel interface interacting segment (Figure 1(b)). Arg85 makes a salt bridge with Glu11 Oε2, which is rotated out from the active site. The stereochemical analysis of the DAYC interface (Table 1) 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-Lβ5eα5 surface although both protomers are rotated 90° with respect to each other if compared to the WT-RD. The prediction of stability15 for the α1-Lβ5eα5 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; BeF2-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

| Table 1. Interface stereochemical parameters between PhoB protomers* |
|--------------------------|--------------------------|
| Dimer interface-type area (Å²) | WT α1-Lβ5eα5 | BeF₂ α1-Lβ5eα5 |
| Protomer | A | B | C | A | B | C |
| Interface accessible surface area | 562 | 610 | 977 | 981 |
| ASA (in Å²) | 560b | 560b | 890b |
| Interface ASA of total 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 interaction | 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 |

* For definitions, see Jones & Thornton.84

† Total interaction surface divided by 2.61

http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/  

piserver
Table 2. Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>PhoB RD D53A+Y102C (DAYC)</th>
<th>PhoB RD D10A+D35E (DADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.933</td>
<td>0.933</td>
</tr>
<tr>
<td>Space group</td>
<td>P 2₁</td>
<td>P 2₁</td>
</tr>
<tr>
<td>Cell parameters (Å; °)</td>
<td>a = 44.9, b = 47.6, c = 59.7, β = 99.3</td>
<td>a = 32.9, b = 60.2, c = 116.5</td>
</tr>
<tr>
<td>No. of measurements</td>
<td>134,031</td>
<td>93,822</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>41,837</td>
<td>24,245</td>
</tr>
<tr>
<td>Whole resolution range/last shell (Å)</td>
<td>58.7–1.45/1.53–1.45</td>
<td>32.9–1.70/1.79–1.70</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>41,049/770</td>
<td>23,577/638</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.5 (70.3)</td>
<td>93.0 (96.6)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>3.2</td>
<td>2.9</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>6.7 (29.0)</td>
<td>5.4 (34.7)</td>
</tr>
<tr>
<td>Complete merge (%)</td>
<td>94.5 (70.3)</td>
<td>93.0 (96.6)</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50.0–1.45</td>
<td>32.9–1.70</td>
</tr>
<tr>
<td>R free (%)</td>
<td>20.4</td>
<td>25.6</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>18.9</td>
<td>22.3</td>
</tr>
<tr>
<td>No. of non-hydrogen protein atoms</td>
<td>1,977</td>
<td>1,951</td>
</tr>
<tr>
<td>No. of measurements</td>
<td>134,031</td>
<td>93,822</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>41,049/770</td>
<td>23,577/638</td>
</tr>
<tr>
<td>Water molecules</td>
<td>303</td>
<td>170</td>
</tr>
<tr>
<td>Metal ions</td>
<td>2 (Na⁺), 1 (Tris)</td>
<td>–</td>
</tr>
</tbody>
</table>

* Rmerge = Σhkl | Fobs | – | Fcalc | / Σhkl | Fcalc | x100, where I(hkl) is the ith intensity measurement of reflection hkl and <I(hkl)> is its average intensity.
* Values in parentheses correspond to the last shell.
* Rmerge = Σhkl | Fobs | – | k | Fcalc | / Σhkl | Fcalc | x100.
* Free Rmerge, Rmerge for a test set of reflections (>500) not used during the refinement.

The DADE dimer is similar to the WT-RD one (rmsd of 0.43 Å; Table 3), as both are packed similarly in space group P2₁2₁2₁ (cell constants are isomorphous; Table 2) and show asymmetry between monomers (rmsd between DADE-A and DADE-B is 0.52 Å; Table 3). The best fitting (0.43 Å; Table 3) 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 Gly96 to Leu95) and WT-A (Gly86 to Gly94), and the flanking loops Lβ4α4 and Lα4β5 (Arg85A Cα diverge 1.85 Å from its position in WT-A; Glu96A Cα, 2.15 Å). The electron density map for DADE-A is well defined for the whole molecule chain (as an example, see Figure 2(a)). 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 Cα atom of Ser60 is displaced 2.65 Å due to, in DADE-A, an interaction with the Lys117 Nε 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 Lβ3α3c), 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 Lovell et al.28), the latter also found in DADE-A and both

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DADE-A/DADE-B</td>
<td>0.54 (240)</td>
<td>0.52 (120)</td>
<td>0.43 (242)</td>
<td>0.56 (241)</td>
<td>0.38 (121)</td>
<td>0.41 (120)</td>
<td>0.51 (121)</td>
</tr>
<tr>
<td>DAYC-A/DAYC-B</td>
<td>0.22 (240)</td>
<td>0.23 (120)</td>
<td>0.62 (229)</td>
<td>0.63 (225)</td>
<td>0.65 (118)</td>
<td>0.55 (120)</td>
<td>0.53 (119)</td>
</tr>
</tbody>
</table>

The rmsd is calculated from the Cα atoms that deviate less than 3 Å (number of atoms indicated in brackets). Values are given in Å.

a. rmsd between Cα atoms of the whole dimer against itself (Mol1/Mol2 against Mol1/Mol2).
b. rmsd between Cα atoms of the two monomers in a dimer (Mol1 against Mol2).
c. rmsd between dimers, WT protomers A and B are superposed to mutant protomers A and B.
d. rmsd between dimers, WT protomers B and A are superposed to mutant protomers A and B.
e. The following four columns relate each WT protomer within a mutant protomer.
WT protomers. It is noteworthy that this mutant reproduces the correlation between the disorder of helix $\alpha_4$ and Trp54 flexibility, a fact also found in the three molecules of Mg-RD which has been related to activation of the protein. 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 $g^+\beta$ conformations in DADE-A and -B, respectively, as seen in WT-RD. No metal was found in the active site, as we describe below.

**DAYC RD results in an increased symmetry within the RD dimer and a large displacement of helix $\alpha_4$**

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 (Figure 2), 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 3). However,
compared to WT-RD, DAYC displays significant
distortions that the crystal lattices ($P_2_12_12_1$ in WT
versus $P_2_1$ in DAYC) could account for it, but the
structure in the monoclinic crystal is more rigid and
well defined although the packing is much looser
than in the orthorhombic one. 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 3). Besides,
$\alpha$-positions of either DAYC monomer are more
similar to the WT-B ones (deviation values in Table
3) because in the formers $\beta$s is in a perfect
antiparallel $\beta$-hairpin conformation similarly to WT-
B. Again in both DAYC monomers, Glu11 fulfills the
space freed by the new smaller side-chain Ala53,
pulling the whole $\alpha$L1a1 towards the active site
cavity (in DAYC-A Glu11 moves 1.6 Å, and 1.0 Å in
$\alpha$-B with respect to WT-B), as well as dragging the
N-terminus of helix $\alpha$1, which later accommodates to
the WT-B helical trace.

However, these are minor changes if compared to
the displacement performed by helix $\alpha$4, together
with Trp54 (Figure 1(c)). In WT-RD, the side-chain of
Trp54 is in the $t$-105 conformation 38 (like in DADE-
A; Figure 2(a)), immersed in a hydrophobic pillow
between helices $\alpha$3 and $\alpha$4 as explained. In DAYC it
takes to an unseem $g^{\prime}$-95 rotamer 39 (Figures 1(c),
2(b)) invading the space liberated by the Tyr102Cys
substitution between helix $\alpha$4 and strand $\beta$4 (Figure
1(c)). The side-chains that interact with Trp54 are
Met81 ($\beta$4), Thr83 ($\beta$4), Glu88 ($\alpha$4) and Val92 ($\alpha$4),
the latter contacting the other face of the Trp54 rings
with respect to as in WT (Figure 1(c)). The Cys102
side-chain adopts different rotamers, $g^{\prime}$ towards the
solvent in DAYC-A (Figure 2(c)) 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 $g^+$, toward the core between helix $\alpha$4 and $\beta$4 (Figures 1(c) and 2(b)). This allows
Thr83 O$^\gamma$ and Trp54 N$^\delta$ to make a hydrogen bond
(Figure 1(c)). The structures of CheY co-crystallized
with magnesium 32 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 33 (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). Interest-
ingly, in DAYC-A and -B, similar van-der Waals
interactions occur between Trp54 and Val92 (Figure
1(c)). Regarding helix $\alpha$4, the replacement of Tyr102
for a cysteine disrupts the van-der Waals interaction
that the tyrosine establishes with Arg93 in WT-
RD (Figure 1(c)). Therefore, in DAYC, the arginine
is exposed to the solvent, the last turn of helix $\alpha$4 is
unwinded so that this segment finishes at Val92
(Figure 1(c)). Besides, helix $\alpha$4 is largely displaced
along its axis and laterally, towards the active-site
groove, pushing out the whole L$\beta$3a3 and breaking
hydrogen bonds between strands $\beta$3 and $\beta$4.
Accordingly, residues in helix $\alpha$4 are considerably
placed. 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-
terminus of both helices $\alpha$4 closer to each other
(Figure 1(b)) 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 unaf-
recked. For example, in DAYC-A and -B, Glu88
makes, as in the WT-RD, interactions with the active
site (see below).

Structural analysis of the active sites of DADE
and DAYC RDs

The active site of PhoB shows two sub-sites which
we previously referred to as site M (for Mg$^{2+}$) and
site P (for phosphate). 36,40 Residues at site M are
involved in direct or indirect coordination to Mg$^{2+}$
(Figure 1(e) and (f)), and include the highly
conserved triad Glu9, Asp10 and Glu11 (L$\beta$1x1i) of
PhoB.36,40 At the other end of the cavity, Thr83 O$^\gamma$ and
Ala84 N ($\beta$4) are directly involved in phos-
phoryl oxygen (or BeF$_3^-$ ) binding, 36 thus delimiting
the site P (Figure 1(e)). The P and the M- sites are
bridged by the side-chain of Asp53, whose O$^\delta$ atom
coordinates a phosphate oxygen and O$^\gamma$ atom binds the
cation (Figure 1(e)). In parallel, the backbone of
L$\beta$3a3 is also partitioned, with the amide nitrogen
atoms of positions 54 and 55 binding the phosphate
moiety and the carbonyl oxygen of residue 55
coordinating the cation. Besides, Lys105 N$^\gamma$ (nearly
invariant among homologues) from L$\beta$5a5 binds a
phosphate oxygen atom and contacts the carboxy-
late group of Glu9 of site M. At the M-site, the
interactions between residue atoms and water
molecules are preserved in the WT-A, Mg-A and
-B, and BeF$_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

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 (Figure
2(a)) 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$^\gamma$
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 36,44). The orientation
observed for the Glu53 side-chain is possible due to
the mutation at position 10, from an aspartate to an
structures, W1 contacts Glu9 O. The location of performing similar interactions with surrounding is close to where the bivalent cation should be and, likewise, binds the amide at position 10. In WT-A, W1 further binds to Glu9 O, but in DADE Glu9 has been changed to alanine. Briefly, the active site of DADE shows that the negatively charged Glu53 O 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 is too distant from Lys105 N and makes the interaction that Asp53 and Lys105 establish in WT-A. However, F2 and F3 and the two water molecules contact Thr83 O, which in turn contacts Glu88 of helix α4 through two water bridges (involving W17, W57 and W58; Figure 1(e)). Therefore, in DADE, W13 is a linker between the M and P-sites. More interestingly, W13 is 0.80 Å from the position of the BeF3 3 fluorido atom F3 in BeF3-RD (Figure 1(e)) and, likewise, hydrogen bonds W58 similarly as to F3 interacts with W91. Besides, W17 locates close (1.0 Å) to where F2 is (Figure 1(e)). Therefore, the position of W13 and W17 recalls the location of BeF3 fluorines. However, F2 and F3 and the two water molecules contact the P-site border in a different manner: F2 interacts with Ala84 N but W13 does not, and F2 contacts Thr83 O but W17 does not either. Instead, W17 and Ala84 N interact with the Glu88 carboxylate, dragging L3αε4 towards the active site cavity (Arg85 Cε is displaced 2.10 Å). Furthermore, the similarity to the BeF3-RD active sites is only true for DADE-A, as in DADE-B there are no 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.

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 Å in average, see below for specific distances) and the coordination is not tetrahedral but octahedral-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, and (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 coordination of six oxygen ligands, with a theoretical cation-oxygen distance of 2.46 Å for Na and 2.07 Å for Mg. 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; Figure 2(b); Table 2) 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 Å to 2.44 Å, and therefore we suggest it is a Na+, as for Mg the distances would be too long as well as this cation was not added during purification and crystallization. Seven oxygen atoms coordinate the cation, which are provided by the following protein atoms (Figure 1(f)): the ligands in the plane are Asp10 O and Met55 O and a further solvent molecule W1 (W11); the Na metal coordination shell is therefore completed by three water molecules, which are at similar locations than in Mg-RD and BeF3-RD active sites. Thus, the sodium coordination atoms have the proper/active orientation regardless a Na 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 L3ε3a3 into a perfect β-ribbon, is pushed laterally by L3αε4 and helix α4 as described above. Due to the displacement of Met55 O (L3αε3) and the absence of Asp53, the metal moves about 1.1 Å towards the L3αε3 edge in both DAYC-A and -B molecules, if compared to all three Mg-RD protomers, or 0.80 Å regarding BeF3-RD. The hydrogen bond network originates 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, Leu82 O and Glu88 O (Figure 1(f)). This water tandem is located close to where the BeF3– molecule is in BeF3-RD but, unlike in DADE-A, the water molecules are not placed like the fluorine atoms (compare Figure 1(e) and (f)). From the Mg2– 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. 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 (L3αε3), whose carbonyl is involved in Na+ coordination, and to Lys105 N (L3αε5), which is implicated in phosphate binding in the WT protein (Figure 1(b) and (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.

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-Lβ5a5 interface, a contact also found in the apo and Mg2+-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 Mg2+-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 BeF3 (BeF3-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 Mg2+ binding, 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.50 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. BeF3-RD shows that upon activation the active site containing BeF3 and Mg2+ 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 this active PhoB RD form, we do not 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. DrrB59 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 antiparallel β-sheet of the ED. In contrast, in DrrB,59 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 antiparallel 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 phosphorylation 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 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 β5a5 and the N terminus of helix α5 make up a 2-fold related interface (α1–Lβ5a5 interface) not observed in other RDs. Structures of four different RD forms, namely WT,50,56 Mg2+-bound,50 DAYC and DADE have been crystallized in three different crystal lattices (P21_21_21, 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 2-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.35,38 As mentioned, the chemical stability prediction (employing the PISA server‡) 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 (six 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 BeF3-RD, also predicted to be stable, where the two partners are rotated 90° 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 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. This surface is larger and less planar than the α1–Lβ5α5 interface although it contains a higher percentage of polar interactions (about 50%; Table 1), suggesting again a transient interface.35 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 values are of −122.4 and −118.1 kcal/mol for isolated B and C molecules (PDB,1ZES), but upon complex formation the free energy is slightly positive. Interestingly, the PhoB representative in Bacillus subtilis, PhoPN, shows the same functional surface α4-β5-α5 although is packing in a head-to-tail manner against helices α3 and α4 and Lβ4α4 from the other protomer. None of the PhoB packing contacts in any of the structures presented until now resemble the asymmetrical interaction found in PhoP. Therefore, despite the high conservation within the family, the strategies of the different molecules for oligomerization 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 2-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 PhoBCA 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; (ii) expression of phnC-to-phnB operon for use of phosphonates as an alternative phosphorus source requires high level activation by phosphorylated PhoB protein; and (iii) growth on phosphonates is slow. Accordingly, standard allele-replacement methods were used to construct an E. coli K-12 strain with a

‡ http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver
chromosomal rhaBp-phoB fusion in which PhoB synthesis requires induction by rhamnose,64 a functional phnC-to-phnP operon,65 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.66 From this strain, spontaneous phosphorylation 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 dependence 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<sup>CΔ</sup> mutations, as verified by cloning and DNA sequencing and by transfer of the mutant phoB<sup>CΔ</sup> allele into a new host. The PhoB<sup>CΔ</sup> DADE and DAYC showed high level activation of Pho regulon expression that is quantitatively similar to levels in a wild-type (pho<sup>BR</sup>+) host under conditions of Pi starvation. Further, transcriptional activation by the PhoB<sup>CΔ</sup> DADE and DAYC proteins was shown to be unaffected by PhoR, CreC, or acetyl phosphate (data not shown). Expression of these PhoB<sup>CΔ</sup> 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 PhoB<sup>CΔ</sup> mutants. Details of the isolation and characterization of these and other PhoB<sup>B–A</sup> mutants will be described elsewhere (S.-K. K. & B. L. W., unpublished results).

**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) b小店, in which they transformed, were grown in LB medium containing 100 μg/ml of ampicillin at 37 °C with aeration to early exponential phase (A<sub>600</sub> nm ∼0.4–0.6). Cultures were cooled down to room temperature and 0.1 mM isopropyl-β-D-thiogalactopyranoside was added for induction. Recombinant protein production was carried out at ∼25 °C in order to decrease the formation of inclusion bodies. After 13 h, cells were collected by centrifugation at 4500 rpm. The cell pellet was resuspended in buffer A (50 mM NaCl, 1 mM EDTA, 1 mM DTT, Tris-HCl (pH 8)). Cells were disrupted by sonication, after which 100 mg/ml of phenylmethanesulfonyl fluoride and 25 μl of 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 but with 150 mM NaCl for DAYC or 200 mM 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 mM to 500 mM NaCl PhoB RD DAYC purity reached more than 98%, as confirmed by SDS–PAGE, with a yield of 25 μg 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 −80 °C. In order to verify the elution behaviour of both native PhoB crystals and mutants, analytical size exclusion chromatography was performed using a Superdex 75 HR10/30 column (Amersham).

**Protein crystalization, data collection and processing**

PhoB<sup>V</sup>–A DADE and DAYC crystals were obtained from hanging drops by the vapour-diffusion method at 20 °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 2 shows cell constants. There are two molecules in the asymmetric unit (a.u.; V<sub>m</sub> = 2.2 Å<sup>3</sup> Da<sup>−1</sup>, solvent content 43%). 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 6 μl reservoir solution (22.5% PEG 4K, 0.4 M sodium acetate, 0.1 M Tris HCl (pH 8)) rendered orthorhombic crystals (P<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub>) that were isomorphous to the WT PhoB RD crystals<sup>28,66</sup> (Table 2), with a dimer in the a.u. (V<sub>m</sub> = 2.0 Å<sup>3</sup> Da<sup>−1</sup>, 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 MOSFLM<sup>69</sup> and scaled, merged and reduced with SCALA<sup>70</sup>. Data collection and processing statistics are given in Table 2.

**Structure determination and refinement**

The DAYC and DADE RD were both solved by molecular replacement<sup>73</sup> with AMoRe<sup>74</sup> using the WT PhoB RD X-ray structure as searching model (Protein Data Bank (PDB) accession code 1B01<sup>60</sup>) and data between 15 Å 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<sub>F</sub>)<sup>72</sup> 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<sub>F</sub> 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<sup>75</sup> v. 1.0<sup>73</sup> and REFMAC5<sup>74</sup> using all data and employing bulk-solvent and anisotropic temperature factor correction. The
cross-validation $R_{	ext{free}}$ was monitored throughout with sufficient reflections (>500) 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 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 $\sigma_\alpha$-weighted $|2F_{\text{obs}} - F_{\text{calc}}|$ and $|F_{\text{obs}} - F_{\text{calc}}|$ 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 $2\sigma$. 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, Ser57A, Leu56A, Met120A, Cys512B, and Ile116B. The model of DADE RD includes residues 3A-123A and 2B-123B and 170 solvent molecules. Ramachandran plots generated with PROCHECK located no residues in disallowed regions in any of the structures. All structures show two $\text{cis}$-peptide bonds between residues 44 and 45 and 105 and 106. Final refinement statistics are in Table 2.

### Miscellaneous

Least-squares superimpositions were carried out with LSQMAN. In order to calculate distances between similar atoms of different monomers, only the cores (β-sheet strands) of the molecules were superimposed on the A monomer of WT-RD. Dimer superimpositions were calculated using all Cα atoms. The conservation values of residues were computed as follows: first, the 30 most similar sequences to PhoB were determined with BLAST and subjected to multiple alignment using CLUSTAL W. The resulting sequence superposition was submitted to ESPript which converted the similarity values to $\text{E}$-values. The corresponding stereochemical analysis was done by submitting the coordinates at the protein–protein interactions server, as well as the stability of the interactions was analysed with PISA.

### Protein Data Bank accession codes

The coordinates for the refined DADE and DAYC PhoB RD models have been deposited to the RCSB Protein Data Bank (1DER) with accession codes 2h9 and 2ba. 

During the deposition procedure renaming/renumbering of certain atoms/residues by the PDB proved unavoidable (see the REMARKS section of the PDB entries for equivalences).

---

§ www.biochem.ucl.ac.uk/bsm/PP/server/index.html
¶ http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html
1 www.ebi.ac.uk/msd

### Acknowledgements

We are in debt to 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 Research Fund of Konkuk University. A.G.B. was recipient of a fellowship from the Ministerio de Sanidad y Consumo (Spain), and P.I.B.P. 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 “Ramón 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.

### References


*Edited by R. Huber*

(Received 8 August 2006; received in revised form 7 November 2006; accepted 9 November 2006) Available online 14 November 2006