Published in final edited form as:

J Mol Biol. 2008 April 18; 378(1): 227–242.

Crystal Structures of the Response Regulator DosR from *Mycobacterium tuberculosis* Suggest a Helix Rearrangement Mechanism for Phosphorylation Activation

Goragot Wisedchaisri^{1,2,3}, Meiting Wu^{1,2,3}, David R. Sherman⁴, and Wim G. J. Hol^{1,2,3*}

¹Department of Biochemistry, University of Washington, Seattle, Washington 98195.

²Biomolecular Structure Center, University of Washington, Seattle, Washington 98195.

³Biomolecular Structure and Design (BMSD) Graduate Program, University of Washington, Seattle, Washington 98195.

⁴Department of Pathobiology, University of Washington, Seattle, Washington 98195.

Abstract

The response regulator DosR is essential for promoting long-term survival of *Mycobacterium tuberculosis* under low oxygen conditions in a dormant state and may be responsible for latent tuberculosis in one third of the world's population. Here we report crystal structures of full-length unphosphorylated DosR at 2.2 Å and its C-terminal DNA-binding domain at 1.7 Å resolution. The full-length DosR structure reveals several features never seen before in other response regulators. The N-terminal domain of the full-length DosR structure has an unexpected (β α)₄ topology instead of the canonical (β α)₅ fold observed in other response regulators. The linker region adopts a unique conformation which contains two helices forming a four-helix bundle with two helices from another subunit, resulting in dimer formation. The C-terminal domain in the full-length DosR structure displays a novel location of helix α 10 which provides Gln199 to interact with the catalytic Asp54 residue of the N-terminal domain. In contrast, the structure of the DosR C-terminal domain alone displays a remarkable unstructured conformation for helix α 10 residues different from the well-defined helical conformations in all other known structures, indicating considerable flexibility within the C-terminal domain. Our structures suggest a mode of DosR activation by phosphorylation *via* a helix rearrangement mechanism.

INTRODUCTION

With eight million new cases reported annually,¹ tuberculosis remains a major global infectious disease responsible for approximately two million deaths per year worldwide.¹ It is estimated that one third of the world's population is latently infected with its causative agent, *Mycobacterium tuberculosis*.² The ability to persist in infected individuals without producing any symptoms for extended periods of time makes *M. tuberculosis* a remarkably successful pathogen. Latent tuberculosis is a major public health concern because it acts as a reservoir of *M. tuberculosis* that can remain undetected for decades before reemerging as active disease. Current treatments for tuberculosis are highly effective only when the bacteria are actively

*Correspondence should be addressed to Wim G. J. Hol, Tel: (206) 685-7044, Fax: (206) 685-7002, E-mail: wghol@u.washington.edu.

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.

growing. Tuberculosis infections require prolonged drug therapy, presumably due to the persistence of the bacteria that make them refractory to the drug treatment.

Two-component regulatory systems in bacteria are a major class of signal transduction proteins involved in adaptation to environmental changes.³ Typical two-component regulatory systems contain (i) a membrane-bound sensor kinase that plays a crucial role in sensing environmental stimuli, and (ii) a cytosolic response regulator, usually consisting of an N-terminal domain, a linker, and a C-terminal domain. This pair of proteins functions as a molecular switch that controls diverse adaptive environmental responses (for reviews see ^{3–5}). In the presence of the appropriate stimulus, a sensor kinase catalyzes an ATP-dependent autophosphorylation of a specific His residue in the cytoplasmic domain. The signaling cascade proceeds next by a transfer of the phosphoryl group from the histidine of the sensor kinase to a conserved aspartate in the N-terminal domain of the response regulator. Phosphorylation of the response regulator switches on its C-terminal domain to perform its associated function such as DNA-binding or enzymatic activity. Multi-domain response regulators containing a C-terminal DNA-binding domain function in most cases as transcription factors by binding to promoter regions of specific genes to activate gene expression. This family of response regulators can be classified into three major subfamilies: the NarL,⁶ the OmpR,^{7,8} and the NtrC^{9,10} subfamilies, based on their C-terminal domain sequences.^{3–5} So far, unphosphorylated full-length structures have only been reported for the NarL (represented by NarL^{6,11} and StyR¹²) and OmpR (represented by DrrD¹³ and DrrB¹⁴) subfamilies. No full-length structures of multi-domain response regulators in the active phosphorylated state have been reported to date.

A two-component regulatory system *dosS-dosR* (also called *devS-devR*¹⁵) is among the *M. tuberculosis* genes induced by hypoxia or NO exposure,^{16,17} two conditions frequently associated with the onset and maintenance of latent tuberculosis.^{18,19} Sherman *et al.*¹⁶ proposed that *dosS* and *dosR* form a two-component signaling system involved in the adaptation of bacilli to low oxygen conditions within the host. Nearly all *M. tuberculosis* genes rapidly up-regulated in response to low oxygen and NO levels require DosR for their induction.^{17,20} The sensor kinase DosS (whose gene is upstream of the *dosR* gene) and an orphan sensor kinase DosT (located elsewhere) phosphorylate their conserved histidine (His395 for DosS and His392 for DosT) and subsequently transfer the phosphoryl group to Asp54 of DosR using Mg²⁺ as a cofactor.^{21–23} Phosphorylation of the catalytic Asp54 presumably leads to a conformational change which enhances the binding affinity of DosR to its cognate DNA sequence and activates transcription of regulated genes.²¹

By sequence homology, the N-terminal domain of DosR belongs to a larger CheY family of regulatory domains with a (β α)₅ fold.²⁴ The activation mechanism in this family of regulatory domain has been studied extensively through comparisons of available native and phosphorylation-activated regulatory domain structures of FixJ,^{25,26} NtrC,^{27,28} and Spo0A,^{29,30} and BeF₃[–]-activated structures of CheY,^{31–33} DctD,^{34,35} NtrC,^{28,36} Spo0F,^{37–40} CheY2,⁴¹ ArcA,⁴² PhoB⁴³ and NtrC1.^{9,44} Six conserved residues have been reported to be involved in the phosphorylation activation mechanism: three aspartates (equivalent to Asp8, Asp9, and Asp54 in DosR), one Thr/Ser (Thr82), one Tyr/Phe (Tyr101), and one Lys (Lys104)⁴⁵ (for residue numbers of selected proteins, see Figure 1). All these proteins share a general mechanism of “coupled rearrangement” of the Thr and Tyr residues after modification of the third Asp occurs (by a phosphoryl group or its analog BeF₃[–]).^{4,36,45} Upon activation, the Thr residue changes conformation to form a hydrogen bond with an oxygen atom of a phosphoaspartate (or a fluoride atom in the BeF₃[–]-activated proteins). This causes the Tyr residue to move from a solvent-exposed outward position to a buried inward position which induces conformational changes of nearby residues. Different response regulators have different magnitudes of conformational change³⁸ and the outcome is different from case to case. For example, a phosphorylation-induced conformational change (i) triggers domain

separation in FixJ⁴⁶ and NarL,⁴⁷ (ii) causes dimerization of FixJ,^{25,48} StyR,^{12,49} and PhoB,⁴³ (iii) results in oligomerization of NtrC,⁵⁰ and (iv) allows binding to different protein interaction partners with different binding affinities in the case of CheY.^{51–54} Therefore, while having similar global architectures, two-component response regulators act with great diversity at the functional level.

Our previous crystal structures of the DosR C-terminal domain (DosR^C) and of its complex with DNA identify DosR as a member of the NarL subfamily of response regulators on the basis of the structural topology of the C-terminal domain.⁵⁵ A DosR^C dimer is the functional unit for DNA binding where residues in the C-terminal helix $\alpha 10$ form a dimerization interface.⁵⁵ Two such DNA-binding dimers assemble as a tetramer via an interface involving helices $\alpha 7$ and $\alpha 8$ of the C-terminal domain.⁵⁵ However, due to the absence of the N-terminal regulatory domain, a molecular mechanism for phosphorylation activation of DosR remained unclear. Here we report a crystal structure of full-length unphosphorylated DosR at 2.2 Å resolution. Our structure reveals a novel topology for both the N-terminal domain and the linker region, and a unique conformation for the C-terminal domain. In addition, we present a 1.7 Å resolution structure of the C-terminal domain fragment (DosR^C) with yet another conformation of the residues forming helix $\alpha 10$ which is different from our previous structures of DosR^C⁵⁵ and, to the best of our knowledge, all related structures.^{6,11,12,56–62} Our structures reported here provide insight on how DosR conformational changes might be induced by phosphorylation through a helix rearrangement mechanism.

RESULTS

Overall structure of full-length DosR

The structure of full-length DosR was solved by the SeMet MAD technique and refined using native data to 2.2 Å resolution (Table 1). The asymmetric unit contains two DosR molecules forming a dimer related by a non-crystallographic two-fold axis. The electron density is well defined for the first 210 residues with C-terminal tail residues 211–217 being disordered (Figure 1). The N-terminal domain of DosR (formed by residues 1 to 97) is connected to the C-terminal domain (residues 150–210) by a two-helix linker (Figure 2). Quite remarkably, the structure of each of these three elements is different from other members of the NarL subfamily to which DosR belongs.

The N-terminal regulatory domain

The N-terminal domain of DosR contains an α/β fold with a $(\beta\alpha)_4$ arrangement and is topologically different from the canonical regulatory domain fold with a $(\beta\alpha)_5$ arrangement.²⁴ The canonical secondary structure elements $\alpha 4$ and $\beta 5$ appear to be absent in the DosR structure (Figure 2). Surprisingly, the DosR helix $\alpha 4$ is positioned where the canonical helix $\alpha 5$ is located in structures of other members of the subfamily (as discussed below).

The catalytic Asp54 residue, involved in the phosphorylation reaction, is located in the $\beta 3$ - $\alpha 3$ loop surrounded by the conserved Asp8 and Asp9 residues from the $\beta 1$ - $\alpha 1$ loop from the same subunit (Figure 3). The O ^{$\delta 1$} atom of Asp54 makes a hydrogen bond to a water molecule. This water molecule also forms two hydrogen bonds with the O ^{$\delta 1$} and O ^{$\delta 2$} atoms of Asp9. Interestingly, the O ^{$\delta 2$} atom of Asp54 makes two hydrogen bonds to the O $\epsilon 1$ and N $\epsilon 2$ atoms of Gln199 located in the beginning of the C-terminal domain $\alpha 10$ helix. This unique feature will also be discussed further below.

The linker region

The linker formed by residues 98 to 149, between the N- and the C-terminal domains contains helices $\alpha 5$ (residues 101 to 121) and $\alpha 6$ (residues 127 to 143). These two helices form a four-

helix bundle dimerization interface with their counterparts from another subunit of the dimer (Figure 2). According to a *B*-factor plot (not shown), helix $\alpha 6$ has the highest average *B*-factor in the structure ($78 \text{ \AA}^2$ for residues 125–144 vs. $44 \text{ \AA}^2$ for the remainder of protein), suggesting flexibility in this part of the molecule. The conserved Tyr101 and Lys104 residues of DosR are located at the beginning of helix $\alpha 5$ and their side chains point into the solvent. This arrangement is entirely different from what is seen in the structures of other NarL subfamily members.^{6,11,12}

The C-terminal DNA-binding domain

In the full-length DosR structure, the C-terminal domain folds back from the linker region to interact with the N-terminal domain of the same subunit (Figure 2). This domain has an unusual conformation where helix $\alpha 10$, which starts at Gln199, moves away from the core formed by helices $\alpha 7$ to $\alpha 9$ of the C-terminal domain to pack against helix $\alpha 1$ of the N-terminal domain (Figure 2). The O ^{ϵ 1} and N ^{ϵ 2} atoms of Gln199 from helix $\alpha 10$ form hydrogen bonds to the O ^{δ 2} atom of Asp54 of the N-terminal domain (Figure 3), shielding Asp54 and surrounding residues from solvent and stabilizing the protein in the inactive closed conformation.

Domain Interfaces within a DosR Subunit

The interactions between the N-terminal domain, the linker region, and the C-terminal domain of DosR are extensive. The largest interdomain interface occurs between the N-terminal domain plus the linker region and the C-terminal domain, burying $\sim 2,500 \text{ \AA}^2$ solvent accessible surface. This interface has a gap volume index (gap volume per interface solvent accessible surface) of $1.4 \text{ \AA}$ which is a low value indicating an excellent complementarity of the interaction.^{63,64} The N-terminal domain alone and the C-terminal domain bury $1,900 \text{ \AA}^2$ solvent accessible surface, with the same gap volume index. Interestingly, most of the interface between the N- and the C-terminal domains is due to contacts between the N-terminal domain and the $\alpha 10$ helix region (residues 194–210) which bury $1,400 \text{ \AA}^2$ solvent accessible interface with a very small gap volume index of $0.8 \text{ \AA}$, indicating a very tight interaction given that the average gap volume index is $2.5 \pm 1.0 \text{ \AA}$ for heterocomplexes.⁶⁴ The major contacts between the N-terminal domain and the $\alpha 10$ helix region are hydrogen bonds between Asp54 O ^{δ 2} and Gln199 O ^{ϵ 1} and N ^{ϵ 2} atoms; Leu81 carbonyl oxygen and Arg197 N ^{ϵ} and N ^{η 2} atoms; Thr82 O ^{γ 1} and Met194 carbonyl oxygen; Thr82 carbonyl oxygen and Arg196 main chain N; Asp87 O ^{δ 1} and Arg197 N ^{η 1}; and Asp87 O ^{δ 2} and Glu206 O ^{ϵ 2} (Figure 3). The rest of the C-terminal domain (helices $\alpha 7$ to $\alpha 9$) and the N-terminal domain bury only $500 \text{ \AA}^2$ solvent accessible interface with a gap volume index of $4.9 \text{ \AA}$, indicating minimal interactions.

Inspection of all hydrogen bonds present in all interdomain interfaces within the subunit reveals that only the hydrogen bonds between Asp54 O ^{δ 2} and Gln199 O ^{ϵ 1}, and between Asp87 O ^{δ 2} and Glu206 O ^{ϵ 2} atoms are made possible by protonation of side chain carboxylate groups at pH 3.0 of our crystallization mother liquor. Only a small Gln199 side chain rotation is required, and possible, to maintain the Gln199-Asp54 contact via the N ^{ϵ 2} of Gln199 at neutral pH. Similarly, the loss of a proton in the Asp87 to Glu206 hydrogen bond in our current structure might be followed by small rotations about the C ^{β} -C ^{γ} bond of Asp87 and the C ^{β} -C ^{γ} and C ^{γ} -C ^{δ} bonds of Glu206 so that these residues are not repulsed by the negative charges on the side chain carboxylate groups at neutral pH. Therefore, it is likely that at neutral pH, the $\alpha 10$ helix can also adopt the conformation observed in our structure. Because no contact involving protonated carboxylate groups occurs elsewhere within the same subunit, we believe that our structure may be physiologically relevant and crystallization conditions at low pH only helped stabilize the dynamically flexible helix $\alpha 10$ into a well-defined conformation which assists crystal growth.

Dimer Interface of Full-length DosR

The two subunits in the asymmetric unit form a dimer related by a non-crystallographic two-fold axis (Figure 2b). The dimerization interface buries 2,900 Å² of solvent accessible surface of which 33% is polar and its gap volume index is 2.4 Å (compared to an average of 2.2 ± 0.9 Å for homodimers).⁶⁴ The two-helix linker (helices $\alpha 5$ and $\alpha 6$) from both subunits form a four-helix bundle and bury 1,800 Å² of the solvent accessible surface, i.e. 62% of the total dimerization interface. The second part of the dimerization interface comes from interactions between the N-terminal domain helices $\alpha 2$ and $\alpha 3$ of one subunit and the C-terminal domain helices $\alpha 7$ and $\alpha 8$ of the other subunit. A sulfate molecule is found buried in the dimer interface mediating hydrogen bonds between its oxygen atoms and side chains of Arg113 and of a protonated Asp117, and mediating longer-range electrostatic interactions with the guanidinium group of Arg67 at a distance of ~ 3.6 Å.

Structure of the DosR C-terminal domain in crystal form II

The DosR^C structure in a new crystal form II (DosR^C-II), different from our crystal from I (DosR^C-I) reported previously,⁵⁵ was obtained at pH 6.5, and solved by molecular replacement and refined to 1.7 Å resolution (Table 1). Interestingly, molecular replacement using all four helices of the DosR^C-I structure was not successful, but using only helices $\alpha 7$ to $\alpha 9$ gave a clear solution. In the DosR^C-II crystals, two DosR^C molecules form a dimer in the asymmetric unit (Figure 4a). The dimerization interface buries 1,300 Å² solvent accessible surface with a gap volume index of 2.7 Å. In this structure, helices $\alpha 7$, $\alpha 8$, and $\alpha 9$ are arranged in the same manner as observed previously in the structures of the DosR^C-I and DosR^C-DNA complexes⁵⁵ (Figure 4b) and in the full-length structure reported here (Figure 2). Most surprisingly, however, helix $\alpha 10$ is not observed in this structure with residues Arg196 to Gln199 adopting a coiled coil structure and residues beyond Gln199 being disordered. The latter probably explains the differences in success with three and four helix molecular replacement probes.

Comparison of the DosR C-terminal domain structures

The full-length DosR and the new DosR^C-II structures presented here reveal two novel features with helix $\alpha 10$ separated from the rest of the C-terminal domain in both cases. In our previous structures of DosR^C-I⁵⁵ and the DosR^C-DNA complex,⁵⁵ helix $\alpha 10$ is formed by residues Arg196 to Arg209 and packs in the canonical conformation against the hydrophobic core of the rest of the domain ($\alpha 7$ – $\alpha 9$ helices). In the two new structures described here, residues Leu150 to Leu192, comprising helices $\alpha 7$, $\alpha 8$, and $\alpha 9$, superimpose well in all four available structures (DosR^C-I, DosR^C-DNA, DosR^C-II, and full-length DosR) with mutual r.m.s. deviations of less than 1 Å (Figure 5). However, helix $\alpha 10$ behaves totally different. In the current full-length structure, helix $\alpha 10$ starts at Gln199 (instead of Arg196 in the previous structures) and rotates about 113° from the canonical conformation to a new position enabling the helix to interact with the N-terminal domain. In the new DosR^C-II structure, residues Arg196 to Gln199 become a coiled coil structure and residues after Gln199 are disordered. It is possible that residues Gln199 to Arg209 may still adopt a helical conformation in the DosR^C-II crystal but due to high flexibility, this region is invisible in the electron density map. The orientation of this coiled coil structure is accompanied by a 180° kink at Gly193 compared to the DosR^C-I and the DNA complex structures (Figure 5). The three different conformations of the residues forming helix $\alpha 10$ in different crystal structures suggest major flexibility within the C-terminal domain. This flexibility is most likely of great importance for DosR function as discussed below.

Comparison with homologous structures

DosR is a member of the NarL subfamily by sequence homology of the C-terminal domain (Figure 1). Two other members of this subfamily with known full-length structures are NarL^{6,11} and StyR,¹² both in the unphosphorylated state.

The N-terminal domain—Full length DosR adopts an unusual $(\beta\alpha)_4$ fold with the canonical helix $\alpha 4$ and strand $\beta 5$ absent in the N-terminal domain compared to the $(\beta\alpha)_5$ fold of homologous structures. Superposition of the DosR N-terminal domain to its closest sequence homologue NarL¹¹ yields an r.m.s. deviation of 1.7 Å for 82 superimposable C α atoms, and 1.2 Å for the superposition of 77 C α atoms with StyR (Figure 6). Their equivalent secondary structure elements comprise $\beta 1$ to $\beta 3$ plus $\alpha 5$ in NarL and StyR and $\beta 1$ to $\beta 3$ plus $\alpha 4$ in DosR. Most unexpectedly, helix $\alpha 4$ of DosR is shifted and positioned in a place where the canonical helix $\alpha 5$ is located in NarL and StyR. In another surprise, a sequence segment Tyr101 to Gly121 of DosR, corresponding to the $\beta 5$ - $\alpha 5$ of the canonical regulatory domain becomes helix $\alpha 5$ of the linker region.

When considering conserved residues involved in phosphorylation activation, it appears that Asp8 and Asp9 in the $\beta 1$ - $\alpha 1$ loop and Asp54 in the $\beta 3$ - $\alpha 3$ loop are within 1 Å from equivalent residues in other structures. However, the key Thr82 residue located at the end of strand $\beta 4$ is shifted 3–4 Å away compared to the position of the equivalent Thr residue in other structures as the $\beta 4$ - $\alpha 4$ loop has to connect to helix $\alpha 4$ in its new position (Figure 6b and 6c). Tyr101, conserved in the primary sequence alignment (Figure 1a) and known to be involved in the activation mechanism, is positioned in the linker region with its C α atom 12 Å away from C α of Thr82. The C α of Tyr101 is shifted by 5–6 Å compared to the equivalent Tyr106 in NarL and Phe102 in StyR. A conserved Lys104, which makes a salt bridge with a phosphoryl oxygen of phosphoaspartate or BeF₃[−] analog in activated response regulator structures, is located in the $\alpha 5$ linker helix in our structure 21 Å away from the position of the equivalent Lys in NarL and StyR. This has implications for the activation mechanism as will be discussed below.

The linker region—The interdomain linker of the full-length DosR structure shows no resemblance to those of NarL and StyR.^{6,11,12} Quite surprisingly, DosR uses the amino acid sequence for the canonical $\beta 5$ - $\alpha 5$ secondary structure elements to form a “new” helix $\alpha 5$ (Figure 1 and 2) which clearly is part of the linker connecting the N-terminal and the C-terminal domains. Together with helix $\alpha 6$, the two-helix linker forms a four-helix bundle with the other subunit in the dimer (Figure 2). In contrast, in the unphosphorylated NarL structure, helix $\alpha 6$ is a linker allowing both domains to make contacts in a compact “closed” conformation (Figure 6a).^{6,11} In the case of StyR, the $\alpha 5$ helix adopts its canonical position but residues 108 to 141 form a long helical extension of helix $\alpha 5$ which separates the C-terminal domain from the N-terminal domain by 48 Å (blue helical segment in Figure 6a).¹² (Although specific residues involved in activation adopt an “active-like” conformation, the reported structure of StyR is likely to be an inactive monomer from an oligomerization point of view.¹²) In the case of DosR, however, the linker positions the C-terminal domain near the phosphorylation site to make extensive contacts with the N-terminal domain. Moreover, the DosR linker provides a dimerization interface for the unphosphorylated protein in contrast to both StyR and NarL whose structures are both monomeric (Figure 2 and Figure 6a).^{6,11,12}

The C-terminal domain—Except for helix $\alpha 10$, the C-terminal domain in the full-length DosR and the DosR^{C-II} structures are similar to the canonical conformation observed before in all related structures.^{6,11,12,55–62} In contrast, the entire C-terminal domains of the full-length NarL and StyR structures are both in the canonical conformation (Figure 6a).^{6,11,12} In all three full-length proteins, the C-terminal domain does not form the $\alpha 10$ interface dimer as

seen in the previous structure of DosR^C-I and its DNA complex⁵⁵ and the NarL^C-DNA complex.⁵⁹

DISCUSSION

How phosphorylation activates response regulators is one of the central questions in two-component regulatory system research. Response regulators are believed to be in equilibrium between inactive and active conformations, with phosphorylation acting as a molecular switch to shift this equilibrium towards the active form.^{45,65–67} In all activated structures of the N-terminal regulatory domain, phosphorylation (or BeF₃[−] activation) of the critical Asp residue causes a movement of a conserved Lys in the β5-α5 loop to form a salt bridge to the phosphoryl group. Phosphorylation also causes a conformational change of a conserved Thr residue in strand β4 to form a hydrogen bond with a phosphoryl oxygen atom. This allows the sidechain of a conserved Tyr residue in β5 to reorient from a solvent-exposed outward position to fill a pocket vacated by the Thr residue in a buried inward position with concomitant conformational changes of surrounding residues.⁴⁵

The puzzle remaining after our structure determination is that, in DosR, Tyr101 and Lys104 (i) deviate considerably from the positions seen in other structures; (ii) the sidechains in their new positions are essentially totally solvent exposed, while (iii) these residues are conserved nevertheless. The main reason for all these differences is a major rearrangement of secondary structure element such that our full length DosR structure lacks the canonical α4 and β5 secondary structure elements in the N-terminal domain in which the conserved Tyr101 and Lys104 are located. These residues are now found in helix α5 of the linker region 15–20 Å away from the catalytic Asp54.

The DosR activation mechanism may involve a phosphorylation-dependent equilibrium switching between the inactive conformation seen in our full-length structure and the active canonical conformation where the C-terminal domain forms a dimer at the α10 interface.⁵⁵ Our full-length DosR structure indicates that in the inactive conformation, the C-terminal domain is bound to the N-terminal domain where Gln199 of helix α10 may makes contacts with and buries Asp54 residues from accessibility to solvent. Interestingly, a similar interaction has been observed for Spo0F where the catalytic Asp54 residue is blocked by a salt bridge interaction of an Arg side chain from a neighboring molecule.⁶⁸ The extensive interactions between the N-terminal domain and helix α10 in the DosR structure could be assisted by hydrogen bonds from protonated carboxylate groups at low crystallization pH. It is likely that at neutral pH, helix α10 is dynamically in equilibrium between different conformations, of which at least two are seen in our structures reported here. This would allow Asp54 to be solvent exposed part of the time and to become phosphorylated in the presence of sensor His kinases or acetyl phosphate.

A proposed helix rearrangement activation mechanism

On the basis of the structures reported here and previously,⁵⁵ we propose the arrangement of helices α4, α5, and α10 from the conformations seen in our structure to new conformations in a phosphorylated active state to occur *via* a helix rearrangement mechanism (Figure 7). The unphosphorylated DosR dimer in an inactive/closed conformation (**1** in Figure 7) as seen in the crystal structure of full-length DosR, is in equilibrium with the monomer (**2**) in solution. At neutral pH, helix α10 adopts multiple conformations including one with no contacts with the N-terminal domain (**3**), allowing Asp54 to be solvent exposed part of the time. Phosphorylation of Asp54 is likely to promote a cascade of events which may be as follows. First, the new phosphoryl group will prevent helix α10 from returning to bind the N-terminal domain due to steric hindrance between the phosphoaspartate and Gln199 sidechains. Secondly, phosphorylation presumably induces a conformational change in the N-terminal

domain such that interactions with the C-terminal domain (as seen in our structure) become unfavorable. This change includes a rearrangement of the $(\beta\alpha)_4$ domain and of the $\alpha 5$ - $\alpha 6$ linker seen in our structure to become the canonical $(\beta\alpha)_5$ fold. In this case, helix $\alpha 4$ moves to resume its canonical position and helix $\alpha 5$ becomes the canonical $\beta 5$ - $\alpha 5$ unit. Such a conformational change will cause a collision between helix $\alpha 4$ and the C-terminal domain. Thirdly, the conserved Thr82 sidechain may abandon its interactions with the C-terminal domain (i.e. with residues Met194 and Arg196) in order to reorient and form a hydrogen bond with phosphoaspartate. Alternatively, the adoption of the canonical $(\beta\alpha)_5$ fold could potentially take place before phosphorylation of Asp54 occurs. In agreement with the fact that the interdomain interface in our structure occurs extensively only through the helix $\alpha 10$ region, while the rest of the interface is minimal, a separation of the two domains after phosphorylation of the N-terminal domain is likely to occur (4). This allows the C-terminal domain to arrange helix $\alpha 10$ to its canonical position, leading to an active/open conformation (5) and form a dimer with the $\alpha 10$ helix interface (6), resulting in a DosR dimer-DNA complex in the presence of cognate DNA (7).

In the absence of a phosphorylated DosR structure, this scenario remains to be validated. Nevertheless, once the C-terminal domain is released from the N-terminal domain and the $\alpha 10$ helix is reorganized to the canonical conformation, DosR becomes active and binds to DNA²¹ as a dimer as observed in the structures of the DosR^C-DNA complex.⁵⁵

MATERIAL AND METHODS

Protein expression and purification

The *E. coli* expression system for full-length DosR was constructed as described previously.²¹ The selenomethionine-substituted (SeMet) DosR with a C-terminal His-tag was expressed in M9 minimal media supplemented with selenomethionine. The growth temperature was reduced from 37°C to 30°C when the OD₆₀₀ reached 0.6 and the culture was induced with 1 mM IPTG for 4 hours. The cells were harvested by centrifugation at 5,000 rpm for 20 min at 4 °C and stored at -80 °C. Cell pellets containing SeMet DosR were resuspended in a buffer containing 20 mM TrisHCl buffer pH 8.0, 500 mM NaCl, and 10% glycerol, supplemented with 1 mM PMSF, Complete EDTA-free protease inhibitor cocktail (Roche), and benzonase nuclease, and lysed by sonication. To the lysate, protamine sulfate was added and subsequently incubated on ice for 30 minutes. The mixture was centrifuged at 20,000×g for 30 minutes and the supernatant was clarified by filtration and loaded onto a Ni-NTA affinity chromatography column. The non-specifically bound proteins were eluted with 30 mM imidazole in 20 mM Tris-HCl and 500 mM NaCl, 10% glycerol. Subsequently, SeMet DosR was eluted with 300 mM imidazole in the same buffer. Fractions containing SeMet DosR were concentrated and subsequently purified using a Superdex75 HR10/30 size-exclusion column (Amersham Biosciences) equilibrated in a buffer containing 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 10% glycerol, 1 mM EDTA, and 1 mM TCEP (buffer A). SeMet DosR from peak fractions was diluted to 3 mg/ml in the same buffer for crystallization. Native DosR was expressed in LB media and purified using the same procedure as the SeMet protein.

Crystallization

SeMet DosR crystals were grown at 4° C by sitting drop vapor diffusion by adding 2 µl of 10% glycerol, 1 µl of 10 mM MgCl₂, and 1 µl of precipitant containing 0.5 M Li₂SO₄ and 0.1 M citric acid pH 3.0 to 2 µl of protein (3 mg/ml) to minimize showers of microcrystals. Thin plate crystals appeared after 1 week and reached their maximum size after about 1 month. For acetyl phosphate soaking experiments, crystals were transferred from the mother liquor to a working solution containing 25 mM lithium potassium acetyl phosphate and 5 MgCl₂ in 1:1:1 ratio by volume of buffer A:50% glycerol:precipitant and were soaked in this solution at 4°C for 1 hour.

Crystals were subsequently transferred to a cryoprotectant solution containing 25 mM acetyl phosphate and 5 mM MgCl₂ in 1:1:1 ratio of buffer A:75% glycerol:precipitant and flash frozen under liquid nitrogen.

Native DosR was crystallized at 4° C by sitting drop vapor diffusion by adding 2 µl of 10% glycerol, 1 µl of 10 mM MnCl₂, and 1 µl of precipitant solution containing 0.5 M lithium sulfate and 0.1 M citric acid pH 3.0 to 2 µl of protein solution. Crystals were soaked for 10 seconds in a solution containing 5 mM MnCl₂ in 1:1:1 ratio by volume of buffer A:50% glycerol precipitant followed by a cryoprotectant solution containing 5 mM MnCl₂ in 1:1:1 ratio by volume of buffer A:75% glycerol precipitant, and subsequently flash frozen under liquid nitrogen.

DosR^C was expressed natively and purified as described previously for SeMet DosR^C.⁵⁵ The protein was crystallized with 24% w/v PEG 5000mme, 0.2 M ammonium sulfate, 10% glycerol and 0.1 M MES buffer pH 6.5. Crystals were transferred to a cryoprotectant solution containing the same concentration of precipitant plus additional 15% ethylene glycol and subsequently flash frozen under liquid nitrogen.

Data Collection and Structure Determination

DosR crystals often diffracted anisotropically and had very high mosaicity (2°–2.5°) unsuitable for data collection. Hundreds of crystals had to be screened for diffraction quality. Data sets were eventually collected from a native crystal to 2.2 Å resolution, and from an acetyl phosphate-soaked SeMet crystal. The SeMet data sets were collected at the Se peak and high-energy remote wavelengths at the Advanced Photon Source (APS) and processed separately (due to random crystal orientations on a goniometer) with HKL2000⁶⁹ to 2.9 Å for the peak wavelength data and 4.1 Å for the remote wavelength data. The peak wavelength was used for finding Se sites and phasing. The systematic absence of the data suggested space group P₂₁2₁2₁. The Matthews' coefficient is 2.6 Da/Å³ for 2 molecules in the asymmetric unit (with an anticipated 16 SeMet per asymmetric unit). While the self-rotation function revealed no non-origin peak, the self-Patterson function contains one strong peak at $v = 0.5$, suggesting a non-crystallographic translation along the *b* axis which may contribute to the systematic absence of 0 *k* 0 reflections. Hence, space group P₂₁2₁2 was also tested. The programs SHELX⁷⁰ and BnP (SnB)⁷¹ were used to locate heavy atoms. Heavy atom sites were refined by the programs SOLVE/RESOLVE⁷² or SHARP.^{73,74} All programs were able to find probable solutions for both P₂₁2₁2₁ and P₂₁2₁2 space groups. However, heavy atom sites in space group P₂₁2₁2₁ show clearly a drop of peak height after 16 sites were found and maintained high site occupancy after refinement, while in P₂₁2₁2 the peak heights decreased gradually without a clear cut. The program RESOLVE⁷² was able to find the NCS operators from these sites only in space group P₂₁2₁2₁ which incorporated a translation of 0.5 along *v* in agreement with the self-Patterson. Anomalous Fouriers using phases from RESOLVE confirmed the presence of all 16 sites. However, the map at this point was not interpretable. Molecular replacement with known structures of the DosR C-terminal domain⁵⁵ did not yield a solution either by using *F*_{obs} or experimental phased map. Molecular replacements using homologous protein structures of the N-terminal domain did not yield any solutions either.

The remote wavelength data set was then used to improve phasing power. The program SHELX⁷⁰ was used to search for 16 Se sites which were found and which appeared to be identical to those found previously using only the peak data. By subsequently using (i) the program SHARP^{73,74} to refine Se sites and calculate phases;. (ii) RESOLVE⁷² for density modification with NCS averaging; (iii) the RESOLVE_BUILD script for iterative model building and refinement, RESOLVE⁷⁵ was able to build 296 residues of which 80 residues had side chains assigned (out of the expected 450 residues in the asymmetric unit) with *R*_{work}/*R*_{free} = 43.9/52.1%.

This initial model was used as a search model for molecular replacement to the native data set collected at Stanford Synchrotron Radiation Laboratory (SSRL) and processed with HKL2000⁶⁹ in space group $P2_1$. The program MOLREP⁷⁶ yielded a clear solution with a 44.1% *R*-factor and 60% correlation coefficient. At this point, the native data was used for chain tracing. Cycles of RESOLVE_BUILD⁷⁵ and manual building using program COOT⁷⁷ and restrained refinement using the program REFMAC⁷⁸ were carried out. The refinement of the native DosR structure was completed with TLS refinement^{79,80} using REFMAC⁷⁸ after subsequent cycles of manual building and water addition. Data collection and refinement statistics are provided in Table 1. Subsequently, the native DosR model was used for molecular replacement for the MAD data set to verify the accuracy of chain trace using anomalous Fourier peaks of methionine residues in the MAD electron density map and to see if Asp54 was phosphorylated. After partial refinement, the SeMet DosR structure turned out to be almost identical to the native structure but, unfortunately, no electron density for a phosphorylated aspartate was observed.

A data set for the native DosR^C crystal form II was collected at the Advance Light Source (ALS). The data set was processed with HKL2000⁶⁹ in space group $P2_12_12_1$, while the SeMet DosR^C crystal form I, crystallized under similar conditions, belongs to space group $P1$.⁵⁵ The structure could not be solved by molecular replacement using the complete C-terminal domain of crystal form I⁵⁵ as search model. A new search model using only helices $\alpha 7$ to $\alpha 9$ was used and two solutions were found using the program MOLREP,⁷⁶ resulting in subunits A and B per asymmetric unit. The σ_A -weighted $F_{obs}-F_{calc}$ electron density map after several cycles of rigid body and restrained refinements using REFMAC⁷⁸ revealed extra electron density for the C-terminal portion of the molecule in a conformation different from what had been observed previously.⁵⁵ Automatic model building using RESOLVE_BUILD script⁷⁵ was carried out followed by manual building and water adding using the program COOT.⁷⁷ Subsequently REFMAC⁷⁸ was used for TLS refinement,^{79,80} yielding statistics as summarized in Table 1.

Protein Data Bank accession codes

The coordinates and structure factors of DosR and DosR^C have been deposited in the RCSB Protein Data Bank with accession codes 3C3W and 3C57, respectively.

ACKNOWLEDGEMENTS

We thank Konstantin Korotkov and Adrian Rice for providing the full-length DosR and DosR^C constructs, respectively. We gratefully acknowledge the use of the Advanced Photon Source (APS) beamline 19-BM (SBC-CAT) at Argonne National Laboratory, the Advanced Light Source (ALS) beamline 8.2.2 (HHMI) at Lawrence Berkeley National Laboratory, the Stanford Synchrotron Radiation Laboratory (SSRL) at Stanford Linear Accelerator Center, and their staffs for technical assistance. Use of the APS, ALS, and SSRL are supported by the U.S. Department of Energy. This work was supported by the Howard Hughes Medical Institute (W.G.J.H.) and by the National Institute of Health grants CA65656 to W.G.J.H. and AI47744 to D.R.S.

REFERENCES

1. Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, Dye C. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch. Intern. Med* 2003;163:1009–1021. [PubMed: 12742798]
2. Dye C, Scheele S, Dolin P, Parthania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *JAMA* 1999;282:677–686. [PubMed: 10517722]
3. West AH, Stock AM. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem. Sci* 2001;26:369–376. [PubMed: 11406410]
4. Robinson VL, Buckler DR, Stock AM. A tale of two components: a novel kinase and a regulatory switch. *Nat. Struct. Biol* 2000;7:626–633. [PubMed: 10932244]

5. Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. *Annu. Rev. Biochem* 2000;69:183–215. [PubMed: 10966457]
6. Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Grzeskowiak K, Gunsalus RP, Dickerson RE. Structure of the *Escherichia coli* response regulator NarL. *Biochemistry* 1996;35:11053–11061. [PubMed: 8780507]
7. Martinez-Hackert E, Stock AM. The DNA-binding domain of OmpR: crystal structures of a winged helix transcription factor. *Structure* 1997;5:109–124. [PubMed: 9016718]
8. Mizuno T, Tanaka I. Structure of the DNA-binding domain of the OmpR family of response regulators. *Mol. Microbiol* 1997;24:665–667. [PubMed: 9179858]
9. Lee SY, De La Torre A, Yan D, Kustu S, Nixon BT, Wemmer DE. Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. *Genes Dev* 2003;17:2552–2563. [PubMed: 14561776]
10. Pelton JG, Kustu S, Wemmer DE. Solution structure of the DNA-binding domain of NtrC with three alanine substitutions. *J. Mol. Biol* 1999;292:1095–1110. [PubMed: 10512705]
11. Baikalov I, Schroder I, Kaczor-Grzeskowiak M, Cascio D, Gunsalus RP, Dickerson RE. NarL dimerization? Suggestive evidence from a new crystal form. *Biochemistry* 1998;37:3665–3676. [PubMed: 9521685]
12. Milani M, Leoni L, Rampioni G, Zennaro E, Ascenzi P, Bolognesi M. An active-like structure in the unphosphorylated StyR response regulator suggests a phosphorylation- dependent allosteric activation mechanism. *Structure* 2005;13:1289–1297. [PubMed: 16154086]
13. Buckler DR, Zhou Y, Stock AM. Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from *Thermotoga maritima*. *Structure* 2002;10:153–164. [PubMed: 11839301]
14. 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;185:4186–4194. [PubMed: 12837793]
15. Dasgupta N, Kapur V, Singh KK, Das TK, Sachdeva S, Jyothisri K, Tyagi JS. Characterization of a two-component system, *devR-devS*, of *Mycobacterium tuberculosis*. *Tuber. Lung. Dis* 2000;80:141–159. [PubMed: 10970762]
16. Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Schoolnik GK. Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha-crystallin. *Proc. Natl. Acad. Sci. USA* 2001;98:7534–7539. [PubMed: 11416222]
17. Voskuil MI, Schnappinger D, Visconti KC, Harrell MI, Dolganov GM, Sherman DR, Schoolnik GK. Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J. Exp. Med* 2003;198:705–713. [PubMed: 12953092]
18. Nathan C, Shiloh MU. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc. Natl. Acad. Sci. USA* 2000;97:8841–8848. [PubMed: 10922044]
19. Wayne LG, Sohaskey CD. Nonreplicating persistence of *Mycobacterium tuberculosis*. *Annu. Rev. Microbiol* 2001;55:139–163. [PubMed: 11544352]
20. Park HD, Guinn KM, Harrell MI, Liao R, Voskuil MI, Tompa M, Schoolnik GK, Sherman DR. Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol* 2003;48:833–843. [PubMed: 12694625]
21. Roberts DM, Liao RP, Wisedchaisri G, Hol WG, Sherman DR. Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J. Biol. Chem* 2004;279:23082–23087. [PubMed: 15033981]
22. Saini DK, Malhotra V, Dey D, Pant N, Das TK, Tyagi JS. DevR-DevS is a bona fide two-component system of *Mycobacterium tuberculosis* that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. *Microbiology* 2004;150:865–875. [PubMed: 15073296]
23. Saini DK, Malhotra V, Tyagi JS. Cross talk between DevS sensor kinase homologue, Rv2027c, and DevR response regulator of *Mycobacterium tuberculosis*. *FEBS Lett* 2004;565:75–80. [PubMed: 15135056]
24. Stock AM, Mottonen JM, Stock JB, Schutt CE. Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. *Nature* 1989;337:745–749. [PubMed: 2645526]

25. Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, Rousseau P, Kahn D, Samama JP. Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure* 1999;7:1505–1515. [PubMed: 10647181]
26. Gouet P, Fabry B, Guillet V, Birck C, Mourey L, Kahn D, Samama JP. Structural transitions in the FixJ receiver domain. *Structure* 1999;7:1517–1526. [PubMed: 10647182]
27. Kern D, Volkman BF, Luginbuhl P, Nohaile MJ, Kustu S, Wemmer DE. Structure of a transiently phosphorylated switch in bacterial signal transduction. *Nature* 1999;402:894–898. [PubMed: 10622255]
28. Volkman BF, Nohaile MJ, Amy NK, Kustu S, Wemmer DE. Three-dimensional solution structure of the N-terminal receiver domain of NTRC. *Biochemistry* 1995;34:1413–1424. [PubMed: 7827089]
29. Lewis RJ, Brannigan JA, Muchova K, Barak I, Wilkinson AJ. Phosphorylated aspartate in the structure of a response regulator protein. *J. Mol. Biol* 1999;294:9–15. [PubMed: 10556024]
30. Lewis RJ, Muchova K, Brannigan JA, Barak I, Leonard G, Wilkinson AJ. Domain swapping in the sporulation response regulator Spo0A. *J. Mol. Biol* 2000;297:757–770. [PubMed: 10731426]
31. Cho HS, Lee SY, Yan D, Pan X, Parkinson JS, Kustu S, Wemmer DE, Pelton JG. NMR structure of activated CheY. *J. Mol. Biol* 2000;297:543–551. [PubMed: 10731410]
32. Lee SY, 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]
33. Stock AM, Martinez-Hackert E, Rasmussen BF, West AH, Stock JB, Ringe D, Petsko GA. Structure of the Mg(2+)-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. *Biochemistry* 1993;32:13375–13380. [PubMed: 8257674]
34. Meyer MG, Park S, Zeringue L, Staley M, McKinstry M, Kaufman RI, Zhang H, Yan D, Yennawar N, Yennawar H, Farber GK, Nixon BT. A dimeric two-component receiver domain inhibits the sigma54-dependent ATPase in DctD. *FASEB J* 2001;15:1326–1328. [PubMed: 11344129]
35. 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;16:1964–1966. [PubMed: 12368235]
36. Hastings CA, Lee SY, Cho HS, Yan D, Kustu S, Wemmer DE. High-resolution solution structure of the beryll fluoride-activated NtrC receiver domain. *Biochemistry* 2003;42:9081–9090. [PubMed: 12885241]
37. Feher VA, Zapf JW, Hoch JA, Whiteley JM, McIntosh LP, Rance M, Skelton NJ, Dahlquist FW, Cavanagh J. High-resolution NMR structure and backbone dynamics of the *Bacillus subtilis* response regulator, Spo0F: implications for phosphorylation and molecular recognition. *Biochemistry* 1997;36:10015–10025. [PubMed: 9254596]
38. Gardino AK, Volkman BF, Cho HS, Lee SY, Wemmer DE, Kern D. The NMR solution structure of BeF3–-activated Spo0F reveals the conformational switch in a phosphorelay system. *J. Mol. Biol* 2003;331:245–254. [PubMed: 12875849]
39. Madhusudan, Zapf J, Whiteley JM, Hoch JA, Xuong NH, Varughese KI. Crystal structure of a phosphatase-resistant mutant of sporulation response regulator Spo0F from *Bacillus subtilis*. *Structure* 1996;4:679–690. [PubMed: 8805550]
40. Varughese KI, Tsigelny I, Zhao H. The crystal structure of beryll fluoride Spo0F in complex with the phosphotransferase Spo0B represents a phosphotransfer pretransition state. *J Bacteriol* 2006;188:4970–4977. [PubMed: 16788205]
41. Riepl H, Scharf B, Schmitt R, Kalbitzer HR, Maurer T. Solution structures of the inactive and BeF3–-activated response regulator CheY2. *J. Mol. Biol* 2004;338:287–297. [PubMed: 15066432]
42. 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 alpha4-beta5-alpha5 face. *J. Mol. Biol* 2005;349:11–26. [PubMed: 15876365]
43. 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* 2005;13:1353–1363. [PubMed: 16154092]
44. Doucleff M, Chen B, Maris AE, Wemmer DE, Kondrashkina E, Nixon BT. Negative regulation of AAA + ATPase assembly by two component receiver domains: a transcription activation mechanism

that is conserved in mesophilic and extremely hyperthermophilic bacteria. *J. Mol. Biol.* 2005;353:242–255. [PubMed: 16169010]

45. Cho HS, Pelton JG, Yan D, Kustu S, Wemmer DE. Phosphoaspartates in bacterial signal transduction. *Curr. Opin. Struct. Biol* 2001;11:679–684. [PubMed: 11751048]
46. Birck C, Malfois M, Svergun D, Samama J. Insights into signal transduction revealed by the low resolution structure of the FixJ response regulator. *J. Mol. Biol* 2002;321:447–457. [PubMed: 12162958]
47. Zhang JH, Xiao G, Gunsalus RP, Hubbell WL. Phosphorylation triggers domain separation in the DNA binding response regulator NarL. *Biochemistry* 2003;42:2552–2559. [PubMed: 12614149]
48. Da Re S, Schumacher J, Rousseau P, Fourment J, Ebel C, Kahn D. Phosphorylation-induced dimerization of the FixJ receiver domain. *Mol. Microbiol* 1999;34:504–511. [PubMed: 10564492]
49. Leoni L, Ascenzi P, Bocedi A, Rampioni G, Castellini L, Zennaro E. Styrene-catabolism regulation in *Pseudomonas fluorescens* ST: phosphorylation of StyR induces dimerization and cooperative DNA-binding. *Biochem. Biophys. Res. Commun* 2003;303:926–931. [PubMed: 12670500]
50. Wyman C, Rombel I, North AK, Bustamante C, Kustu S. Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science* 1997;275:1658–1661. [PubMed: 9054362]
51. Blat Y, Eisenbach M. Phosphorylation-dependent binding of the chemotaxis signal molecule CheY to its phosphatase, CheZ. *Biochemistry* 1994;33:902–906. [PubMed: 8305438]
52. Li J, Swanson RV, Simon MI, Weis RM. The response regulators CheB and CheY exhibit competitive binding to the kinase CheA. *Biochemistry* 1995;34:14626–14636. [PubMed: 7578071]
53. Welch M, Oosawa K, Aizawa S, Eisenbach M. Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria. *Proc. Natl. Acad. Sci. USA* 1993;90:8787–8791. [PubMed: 8415608]
54. Welch M, Oosawa K, Aizawa SI, Eisenbach M. Effects of phosphorylation, Mg²⁺, and conformation of the chemotaxis protein CheY on its binding to the flagellar switch protein FliM. *Biochemistry* 1994;33:10470–10476. [PubMed: 8068685]
55. Wisedchaisri G, Wu M, Rice AE, Roberts DM, Sherman DR, Hol WG. Structures of *Mycobacterium tuberculosis* DosR and DosR-DNA Complex Involved in Gene Activation during Adaptation to Hypoxic Latency. *J. Mol. Biol* 2005;354:630–641. [PubMed: 16246368]
56. Ducros VM, Lewis RJ, Verma CS, Dodson EJ, Leonard G, Turkenburg JP, Murshudov GN, Wilkinson AJ, Brannigan JA. Crystal structure of GerE, the ultimate transcriptional regulator of spore formation in *Bacillus subtilis*. *J. Mol. Biol* 2001;306:759–771. [PubMed: 11243786]
57. Kurashima-Ito K, Kasai Y, Hosono K, Tamura K, Oue S, Isogai M, Ito Y, Nakamura H, Shiro Y. Solution Structure of the C-Terminal Transcriptional Activator Domain of FixJ from *Sinorhizobium meliloti* and Its Recognition of the fixK Promoter. *Biochemistry* 2005;44:14835–14844. [PubMed: 16274231]
58. Maris AE, Kaczor-Grzeskowiak M, Ma Z, Kopka ML, Gunsalus RP, Dickerson RE. Primary and Secondary Modes of DNA Recognition by the NarL Two-Component Response Regulator. *Biochemistry* 2005;44:14538–14552. [PubMed: 16262254]
59. Maris AE, Sawaya MR, Kaczor-Grzeskowiak M, Jarvis MR, Bearson SM, Kopka ML, Schroder I, Gunsalus RP, Dickerson RE. Dimerization allows DNA target site recognition by the NarL response regulator. *Nat. Struct. Biol* 2002;9:771–778. [PubMed: 12352954]
60. Pristovsek P, Sengupta K, Lohr F, Schafer B, von Trebra MW, Ruterjans H, Bernhard F. Structural analysis of the DNA-binding domain of the *Erwinia amylovora* RcsB protein and its interaction with the RcsAB box. *J. Biol. Chem* 2003;278:17752–17759. [PubMed: 12740396]
61. Vannini A, Volpari C, Gargioli C, Muraglia E, Cortese R, De Francesco R, Neddermann P, Marco SD. The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *EMBO J* 2002;21:4393–4401. [PubMed: 12198141]
62. Zhang RG, Pappas T, Brace JL, Miller PC, Oulmassov T, Molyneaux JM, Anderson JC, Bashkin JK, Winans SC, Joachimiak A. Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature* 2002;417:971–974. [PubMed: 12087407]
63. Jones S, Thornton JM. Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol* 1995;63:31–65. [PubMed: 7746868]

64. Jones S, Thornton JM. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* 1996;93:13–20. [PubMed: 8552589]
65. Ames SK, Frankema N, Kenney LJ. C-terminal DNA binding stimulates N-terminal phosphorylation of the outer membrane protein regulator OmpR from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 1999;96:11792–11797. [PubMed: 10518529]
66. Simonovic M, Volz K. A distinct meta-active conformation in the 1.1-Å resolution structure of wild-type ApoCheY. *J. Biol. Chem* 2001;276:28637–28640. [PubMed: 11410584]
67. Volkman BF, Lipson D, Wemmer DE, Kern D. Two-state allosteric behavior in a single-domain signaling protein. *Science* 2001;291:2429–2433. [PubMed: 11264542]
68. Madhusudan M, Zapf J, Hoch JA, Whiteley JM, Xuong NH, Varughese KI. A response regulatory protein with the site of phosphorylation blocked by an arginine interaction: crystal structure of Spo0F from *Bacillus subtilis*. *Biochemistry* 1997;36:12739–12745. [PubMed: 9335530]
69. Otwinowski Z, Minor W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* 1997;276:307–326.
70. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr. D* 2002;58:1772–1779. [PubMed: 12351820]
71. Howell PL, Blessing RH, Smith GD, Weeks CM. Optimizing DREAR and SnB parameters for determining Se-atom substructures. *Acta Crystallogr D* 2000;56(Pt 5):604–617. [PubMed: 10771430]
72. Terwilliger TC. Automated structure solution, density modification and model building. *Acta Crystallogr. D* 2002;58:1937–1940. [PubMed: 12393925]
73. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* 2003;59:2023–2030. [PubMed: 14573958]
74. La Fortelle E, Bricogne G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol* 1997;276:472–494.
75. Terwilliger TC. Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement. *Acta Crystallogr. D* 2003;59:1174–1182. [PubMed: 12832760]
76. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst* 1997;30:1022–1025.
77. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 2004;60:2126–2132. [PubMed: 15572765]
78. Murshudov GN, Vagin AA, Dodson EJ. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D* 1997;53:240–255. [PubMed: 15299926]
79. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D* 2001;57:122–133. [PubMed: 11134934]
80. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* 2003;374:300–321. [PubMed: 14696379]

Figure 1.

Sequence-based and structure-based sequence alignment of DosR and its homologues. The sequence alignment is color-coded red for > 80% identity and blue for > 80% homology. A schematic representation of secondary structure elements of full-length DosR is shown on top of the sequence alignment and colored magenta for the N-terminal regulatory domain, blue for the linker, and green for the C-terminal DNA binding domain. The secondary structure elements of NarL are shown on the bottom of the alignment for comparison. A * indicates conserved residues involved in phosphorylation and the coupled rearrangement mechanism for canonical regulatory domains. DosR, NarL, FixJ, and StyR are members of the NarL

subfamily. *T. maritima* and *E. coli* CheY are one domain proteins. Spo0A, PhoP, PhoB, and MicA are members of the OmpR subfamily and only their N-terminal domain sequences are used for alignment. (a) Sequence-based alignment. (b) Structure-based alignment of the $\beta 4$ to $\alpha 6$ region.

Figure 2.

Crystal structure of full-length DosR. (a) The DosR monomer. The structure is shown in magenta for the N-terminal domain, blue for the linker, and green for the C-terminal DNA binding domain. (b) The DosR dimer. Subunit A is shown as above. Subunit B is shown in gray. A sulfate molecule (shown as a stick model) is found in the dimer interface between the two DosR molecules.

Figure 3.

The DosR phosphorylation site. (a) The side chains of D8, D9, D54 (highlighted in yellow), T82 and D87 in the N-terminal domain (magenta), Y101 and K104 in the linker (blue), Q199 and E206 in the C-terminal domain (green) are shown in stick model. Y101 is ~12 Å away from T82. The D54-Q199 and D87-E206 hydrogen bonds contribute to contacts between helix $\alpha 10$ and the N-terminal domain. The latter hydrogen bonds require protonation of at least one of the carboxylates. (b) Stereo view of the phosphorylation site. The σ_A -weighted $2F_{\text{obs}} - F_{\text{calc}}$ electron density map of the phosphorylation site D54, contoured at the 1σ level is shown including neighboring residues.

Figure 4.

The crystal structures of the DosR C-terminal domain. (a) The DosR^C dimer structure from crystal form II (DosR^C-II). Subunit A is shown in cyan and subunit B in blue. The $\alpha 10$ helix is "melted" into a coiled coil, and the C-terminal tail after Gln199 is disordered. (b) Superposition of helices $\alpha 7$ to $\alpha 9$ of DosR^C-II subunit A (cyan) onto subunit A of the DosR^C dimer from the DosR^C-DNA complex crystal structure (subunit A shown in orange and subunit B in gold) reported previously.⁵⁵

Figure 5.

Superpositions of the DosR C-terminal domain in three different conformations. The C-terminal domain from the full-length DosR is shown in green, the DosR^C-DNA complex⁵⁵ (which is the same as DosR^C in crystal form I, not shown) in orange, and the DosR^C crystal form II in cyan. The structures are superimposed with r.m.s. deviations less than 1 Å for helices α7–α9. Helix α10 evidently adopts three entirely different positions or conformations. Only the conformation observed of the DosR^C-DNA complex and of DosR^C in crystal form I (orange) is able to bind DNA.⁵⁵

Figure 6.

Comparison of DosR with other NarL subfamily members. (a) Structures of full-length unphosphorylated DosR, NarL, and StyR monomers arranged in the same orientation of their N-terminal domains. Domains are color-coded with magenta for the N-terminal domain, blue for the linker, and green for the C-terminal domain. (b and c) N-terminal domain superpositions of full length DosR (magenta) onto the NarL (b) and StyR (c) structures (yellow). Loops are smoothened for clarity. DosR lacks the canonical $\beta 4$ and $\alpha 4$ segments of NarL or StyR (shown in cyan) and its $\alpha 4$ is situated roughly at the position of the canonical $\alpha 5$.

Figure 7.

A helix rearrangement activation mechanism of DosR by phosphorylation. The N-terminal regulatory domain is shown in magenta, the linker in blue, and the C-terminal DNA-binding domain in green. Helices are numbered. The critical Asp54 is depicted as a red star with a yellow core. Unphosphorylated DosR dimer in an inactive/closed conformation (1) as seen in the crystal structure of full-length DosR and is in a concentration-dependent equilibrium with monomer (2) in solution. At neutral pH, helix $\alpha 10$ adopts multiple conformations including one with no contacts with the N-terminal domain (3), allowing Asp54 to be solvent exposed. One of these conformations might be the "canonical" conformation, which facilitates phosphorylation of Asp54 which in its turn may contribute to shifting the equilibrium towards that of canonical conformation. The linker helix $\alpha 5$ becomes now the canonical $\beta 5$ and $\alpha 5$ secondary structure elements (4), resulting in an active/open conformation with the two domains separated by linker $\alpha 6$ (5). This allows the C-terminal domain to arrange helix $\alpha 10$ to its canonical position and interact with $\alpha 10$ from a second subunit to form the $\alpha 10$ helix interface, resulting in a DosR dimer (6). This dimer recognizes cognate DNA yielding DosR dimer-DNA complex (7).

Table 1

Data collection and refinement statistics

	SeMet derivative DosR		Native DosR	DosR ^{C-II}
<i>Data collection</i>				
Space group	P2 ₁ 2 ₁ 2 ₁		P2 ₁	P2 ₁ 2 ₁ 2 ₁
Wavelength (Å)	0.9791 (peak)	0.9641 (remote)		
Unit cell dimension	53.68, 95.81,	53.68, 95.83,	48.11, 98.08,	39.64, 47.40,
<i>a</i> , <i>b</i> , <i>c</i> (Å)	97.77	97.70	53.71, β=93.81	66.44
Resolution range (Å) ^a	50–2.88	50–4.1	50–2.2	50–1.7
	(2.98–2.88)	(4.25–4.10)	(2.28–2.20)	(1.76–1.70)
Total reflections	69,241	25,716	100,603	90,800
Unique reflections	11,508	4,334	22,915	14,128
<i>I</i> /σ(<i>I</i>) ^a	8.9 (3.5)	10.1 (6.3)	22.1 (3.1)	19.3 (2.2)
Completeness (%) ^a	95.6 (95.2)	99.5 (99.8)	90.5 (64.3)	98.5 (92.2)
<i>R</i> _{sym} ^a (%)	19.8 (33.9)	18.2 (29.8)	7.4 (39.7)	9.1 (59.0)
<i>MAD phasing</i>				
Resolution range (Å)	50.0–2.88			
No. methionine residues	16			
No. heavy-atom sites found	16			
FOM ^b	0.28			
Solvent flattening FOM ^b	0.59			
<i>TLS Refinement</i>				
Resolution range (Å)			50–2.2	38.6–1.7
Reflections used (working/free)			21,670/1,172	13,384/704
<i>R</i> _{work} (%)			18.6	18.8
<i>R</i> _{free} (%)			24.9	20.6
Average residual <i>B</i> -factor of all atoms (Å ²)			38.5	14.3
No. protein molecules per asymmetric unit			2	2
R.m.s. deviations from ideal geometry				
Bond length (Å)			0.010	0.012
Bond angles (deg.)			1.27	1.33
Ramachandran plot				
Most favored (%)			93.8	98.9
Additional allowed (%)			6.2	1.1
Generously allowed (%)			0	0
Disallowed (%)			0	0

^a Values in parentheses are for the highest-resolution shell.^b According to RESOLVE.⁷²