Unique N-terminal Arm of *Mycobacterium tuberculosis* PhoP Protein Plays an Unusual Role in Its Regulatory Function

Arijit Kumar Das, Vijjamari Anil Kumar, Ritesh Rajesh Sevalkar, Roohi Bansal, and Dibyendu Sarkar

**From the CSIR-Institute of Microbial Technology, Sector 39A, Chandigarh 160036, India**

**Background:** PhoP impacts numerous aspects of *M. tuberculosis* physiology. However, the mechanism of activation of its C-terminal effector domain remains unknown.

**Results:** PhoP contains a functionally essential and unique N-terminal arm.

**Conclusion:** Conformational change of the arm-truncated variant of PhoP accounts for its loss of function.

**Significance:** The presence of the arm of PhoP helps explain the complexity of regulatory functions of the protein in *M. tuberculosis*.

*Mycobacterium tuberculosis* PhoP, a master regulator involved in complex lipid biosynthesis and expression of unknown virulence determinants, is composed of an N-terminal receiver domain and a C-terminal effector domain. The two experimentally characterized PhoP orthologs, from *Escherichia coli* and *Salmonella enterica*, display vastly different regulatory capabilities. Here, we demonstrate that the 20-residue-long N-terminal arm unique to *M. tuberculosis* PhoP plays an essential role in the expanded regulatory capabilities of this important regulator. Although the arm is not required for overall structural stability and/or phosphorylation of the PhoP N-domain, strikingly it is essential for phosphorylation-coupled transcription regulation of target genes. Consistent with this view, arm truncation of PhoP is accompanied by a conformational change of the effector domain, presenting a block in activation subsequent to phosphorylation. These results suggest that presence of the arm, unique to this regulator that shares an otherwise highly conserved domain structure with members of the protein family, contributes to the mechanism of inter-domain interactions. Thus, we propose that the N-terminal arm is an adaptable structural feature of *M. tuberculosis* PhoP, which evolved to fine-tune regulatory capabilities of the transcription factor in response to the changing physiology of the bacilli within its host.

Growing evidence in recent years suggests that *Mycobacterium tuberculosis* as an intracellular pathogen has a remarkable ability to adapt to their host environments to a large extent through signal transduction leading to switching on complex transcriptional programs (1, 2). It is now known that the major response of the bacterium to environmental changes is through classical two-component regulatory systems where the sensor histidine kinase undergoes autophosphorylation in response to an environmental change, and it subsequently transduces the signal via phosphoryl transfer onto an aspartate residue of the cognate response regulator with high specificity (3–5).

The PhoPR two-component system has drawn attention in the recent past because PhoP is required for the growth of the bacilli in both macrophages and mice (6–8). Although the reason for attenuation of *phaP* mutant remains unclear, PhoP is shown to influence numerous aspects of *M. tuberculosis* physiology, including complex lipid biosynthesis (7–10), hypoxic responses (11), secretion of important virulence factors, and specific T-cell recognition (12). Interestingly, studies aimed at understanding the mechanism of *M. tuberculosis* PhoP in controlling virulence determinants reveal that among the reasons for the attenuation of *M. tuberculosis* H37Ra is a single nucleotide polymorphism within the *phaP* gene which accounts for the following: (a) absence of polyketide-derived acyltrehaloses in H37Ra compared with H37Rv, and (b) impaired secretion of proteins that are important for virulence (12–15).

X-ray crystal structure shows that PhoP consists of an N-terminal receiver domain of (β/α)$_s$ topology and a C-terminal effector domain with a winged helix-turn-helix DNA-binding motif (Protein Data Bank code 3R01 (16)). Consistent with recent structural data of the regulator, and more recent results on regulation of lipid biosynthesis by the response regulator (10), PhoP appears to be activated by phosphorylation of the receiver domain. What remains unknown is the mechanism of how activation of the C-terminal effector domain is regulated by a covalent modification at the N-terminal receiver domain.

Previous work probing the domain structure of PhoP mapped two protease-sensitive sites of the protein. Although one of the sites was within the linker region of the protein (Arg$_{147}$/Asn$_{148}$) that links catalytic function with protein phosphorylation, the other site (Arg$_{22}$/Val$_{23}$) was located at the extreme N terminus upstream of the phosphorylation domain (17). Strikingly, trypsin cleavage of PhoP at the latter site...
releases a short N-terminal arm (comprising 20 amino acid residues), unique to this member of the response regulator family. Although linker length of PhoP has been implicated in interdomain interactions (17), the role of the N-terminal arm in regulating PhoP functionality remains unknown. Here, we show that the N-terminal arm, upstream of the phosphorylation domain, is not necessary for phosphorylation but is essential for phosphorylation-coupled transcription regulation, suggesting its role as an activator of PhoP subsequent to phosphorylation. Together, these results suggest a novel role of a unique N-terminal arm in the regulatory mechanism of PhoP-dependent gene expression.

**EXPERIMENTAL PROCEDURES**

**Bacterial Strains, Media, and Growth Conditions**—All cloning and transformations were carried out in *Escherichia coli* DH5α. For overexpression of recombinant forms of PhoP and PhoR proteins from pET15b expression plasmid, *E. coli* BL21 (DE3) served as the host. *E. coli* strains were grown in Luria Bertani (LB) broth (Difco). *Mycobacterium smegmatis* mc²155 was grown on Middlebrook 7H10 agar supplemented with oleic acid, albumin, dextrose, and catalase (Middlebrook OADC, Difco) or in Middlebrook 7H9 broth supplemented with 0.5% glycerol, 0.05% Tween 80, and 10% albumin/dextrose/catalase (Middlebrook ADC, Difco) and were transformed with test plasmids using electroporation. For *M. smegmatis* strains, hygromycin, and streptomycin were used at a concentration of 50 and 20 μg/ml, respectively. For *E. coli*, these antibiotics were used at 250 and 100 μg/ml, respectively.

**DNA Fragments and Plasmids**—For overexpression and purification of nt20-PhoP in *E. coli*, the 5'-truncated *phoP* gene extending from Arg²² to Arg²⁴⁷ was PCR-amplified from pET-phoP (18) using the primer pairs nt20-phoPstart/phoPstop (supplemental Table S1). The gel-purified PCR product was restricted with NdeI and BamHI and ligated to the NdeI/BamHI backbone fragment of pET15b (Novagen). The resulting construct was verified by DNA sequence analysis to confirm the presence of desired changes and the absence of any unintentional mutations. Plasmid pJEM-nt20-phoP, where expression of nt20-PhoP was under the control of its own promoter, was used to complement H37Ra and was constructed by two-stage overlap extension using pJEM-phoP as template and nt20-FPphoP/nt20-RPphoP as primers (supplemental Table S1). The PCR-amplified region was cloned between BamHI and KpnI sites of pJEM15, the promoter-less *E. coli*-mycobacteria shuttle vector (19). To express nt20-PhoP from pME1mL1 in *M. smegmatis* under the control of TetR, *phoP* of pME1mL1-phoP was replaced with the truncated ORF (spanning *phoP* gene encoding Arg²² to Arg²⁴⁷) using primers that introduced an NdeI site (nt20-phoPstart) at the start codon and a PstI site (MphoPstop) 3’ of the stop codon (supplemental Table S1). DNA fragments comprising the upstream regulatory region of *pks2* (pks2up) and *msl3* (msl3up) spanning −206 to +40 and −350 to +60, respectively, relative to their corresponding transcriptional start sites were PCR-amplified from the *M. tuberculosis* genomic DNA using primer pairs FFpks2up/Rfpks2up and FPmsl3up/RPmsl3up, respectively (supplemental Table S1). All of the relevant plasmids used in this study are listed in Table 1. Enzymatic manipulations of DNA were performed using reagents from New England Biolabs. Plasmid DNA isolation, purification, and recovery of DNA fragments or PCR products by gel extraction were performed using spin columns and procedures from Qiagen. PAGE-purified oligonucleotides were synthesized by Sigma.

**Protein Expression, Purification, and Phosphorylation**—Wild-type or the truncated variants of PhoP were overexpressed and purified as described (17, 18). Phosphorylation of PhoRC (C-terminal end of *M. tuberculosis* PhoR comprising residues Thr¹⁹³ to Pro⁴⁸⁵ and subsequent phosphotransfer assays to PhoP or its truncated variants were carried out as described (17, 20). To examine expression of PhoP proteins in *M. tuberculosis*, crude lysates (~5 μg of protein) were resolved by 12% SDS-PAGE and transferred onto a PVDF membrane for Western blot analysis. The blots were probed with anti-PhoP primary and HRP-conjugated anti-rabbit IgG secondary antibodies (Abexome Biosciences) and developed with Luminata Forte chemiluminescence reagent (Millipore). Phosphorylation of PhoP using AcP (~85% purity; Sigma) as the phosphodonor was examined by two-dimensional gel electrophoresis as described previously (21). The proteins were visualized by silver staining, and densitometric scans of replicates of data were used to determine efficiency of phosphorylation.

**Circular Dichroism Studies**—Circular dichroism measurements were carried out with a Jasco spectropolarimeter, model

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Description</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET-phoP</td>
<td>PhoP residues 1–247 cloned in pET15b⁺</td>
<td>18</td>
</tr>
<tr>
<td>pET-phoP/D71N</td>
<td>Asp³ mutated to Asn in pET-phoP</td>
<td>18</td>
</tr>
<tr>
<td>pET-nt20-phoP</td>
<td>PhoP residues 2–21 deleted in pET-phoP</td>
<td>This study</td>
</tr>
<tr>
<td>pJEM15</td>
<td>Promoter-less <em>E. coli</em>-mycobacteria shuttle vector</td>
<td>19</td>
</tr>
<tr>
<td>pJEM-phoP</td>
<td>Entire regulatory region along with <em>phoP</em> encoding gene (residues 1–247) cloned in pJEM15</td>
<td>This study</td>
</tr>
<tr>
<td>pJEM-phoP/D71N</td>
<td>Met⁴ to Ala⁴ mutated in <em>phoP</em> deleted in pJEM-phoP</td>
<td>This study</td>
</tr>
<tr>
<td>pSM128</td>
<td>Integrative promoter probe vector for mycobacteria</td>
<td>28</td>
</tr>
<tr>
<td>pSM- fovp</td>
<td>pkS2up-lacZ fusion in pSM128</td>
<td>10</td>
</tr>
<tr>
<td>pSM-fov</td>
<td>msl3up-lacZ fusion in pSM128</td>
<td>10</td>
</tr>
<tr>
<td>pME1mL1</td>
<td>Mycobacterial protein expression vector</td>
<td>22</td>
</tr>
<tr>
<td>pME1mL1-phoP</td>
<td>PhoP residues 1–247 cloned in pME1mL1</td>
<td>This study</td>
</tr>
<tr>
<td>pME1mL1-nt20-phoP</td>
<td>PhoP residues 2–21 deleted in pME1mL1-phoP</td>
<td>This study</td>
</tr>
<tr>
<td>pME1mL1-phoP/D71N</td>
<td>Asp³ mutated to Asn in pME1mL1-PhoP</td>
<td>10</td>
</tr>
</tbody>
</table>

* Ampr is ampicillin resistance.
* Kanr is kanamycin resistance.
* Strr is streptomycin resistance.
* Hygr is hygromycin resistance.
N-terminal Arm and Regulatory Mechanism of PhoP

J-810 (Jasco, Tokyo, Japan), using a 0.1-cm cell at 25 °C in 10 mM Tris-HCl, pH 7.5, containing 100 mM NaCl. Residue molar ellipticity (θ) was defined as 100 θ_{obs} (deg cm²)⁻¹, where θ_{obs} was observed ellipticity; l was length of the light path in centimeters, and c was residue molar concentration of each protein. Guanidine hydrochloride-induced equilibrium unfolding of PhoP and nt20-PhoP at 25 °C was monitored by changes in θ_{200} values, and normalized data from multiple experiments were combined into a single fit.

Promoter Regulation by M. tuberculosis PhoP in M. smegmatis—To express PhoP in M. smegmatis, electrocompetent cells were transformed with the indicated expression constructs expressing wild-type and truncated PhoP proteins from the P_{myc}tetO promoter under the control of the TetR repressor. Cultures of M. smegmatis strains harboring the indicated lacZ fusions along with the PhoP expression construct (or no expression construct as control) were inoculated in fresh medium (1:100) either in the absence or in the presence of 50 ng/ml anhydro-tetracycline (ATc) and grown to an OD₆₀₀≈1.5. Aliquots were withdrawn from this culture at the 24-h time point and centrifuged, and cell pellets were washed with phosphate-buffered saline. Cells were resuspended in 0.5 ml of phosphate buffer, pH 7.2, containing 75 mM NaCl, and cell suspensions were sonicated. PhoP-dependent activation of pk2 and msl3 was studied by measuring β-galactosidase activity of an aliquot of cell extracts as described previously (10). Expression of M. tuberculosis PhoP or its variants from pME1ml1 (22) was examined in ≈5 µg of crude lysates of M. smegmatis (as determined by Bradford Assay) by immunoblot analysis using anti-PhoP antibody.

Quantitative Real Time RT-PCR Analysis—Total RNA was extracted with TRIzol (Invitrogen), and RNA was treated with DNase I before analysis. Gene-specific primer pairs were synthesized by Sigma and are listed in supplemental Table S1. Both cDNA synthesis and PCRs were performed by using 200 ng of DNase-treated RNA in a reaction buffer containing 2 units of Superscript III Platinum-SYBR Green one-step quantitative RT-PCR kit (Invitrogen) with appropriate primer pairs (200 nM) according to the manufacturer’s recommendation. Quantitative real time RT-PCR was performed using a Realplex PCR detection system (Eppendorf), and the PCR data were analyzed by the comparative CT method as described previously (23). Control reactions with 2 units of Platinum TaqDNA polymerase (Invitrogen) were performed to confirm the absence of genomic DNA in all of the samples, and detection of the sigA (24) mRNA level served as an endogenous control.

EMSA—DNA-binding reactions were performed with purified proteins and 10 nM end-labeled probe for 20 min at 20 °C in EMSA buffer (50 mM Tris-HCl, pH 7.9, 50 mM NaCl, 0.2 mg/ml bovine serum albumin, 10% glycerol, 1 mM dithiothreitol, and 200 ng of sheared herring sperm DNA). Samples were resolved at 4 °C on 6% polyacrylamide gels (acylamide/bisacrylamide, 40:1.1, in 1 x Tris/EDTA buffer). Following electrophoresis and drying, the gels were visualized by autoradiography, and the radioactive bands were quantified using a phosphorimager (Fuji).

DSS Cross-linking of Proteins—Purified PhoP proteins, pre-incubated in phosphorylation mixture containing 50 mM AcP were incubated with 1 mM DSS (Sigma) in 50 mM Hepes/Na⁺, pH 7.5, 150 mM NaCl, 10 mM MgCl₂, 20% glycerol, 0.05% Tween 20 at 25 °C for 30 min as described elsewhere (25). The reaction was stopped by the addition of SDS, and samples were resolved by SDS-PAGE (12% polyacrylamide gels) followed by immunostaining using rabbit polyclonal anti-PhoP antibody (Abexome Biosciences).

Fluorescence Measurements—The bis-ANS (Molecular Probes) binding was carried out with 5 µM of PhoP or nt20-PhoP in 50 mM Tris-HCl, pH 7.90, containing 100 mM NaCl, and fluorescence measurements were carried out with a Cary Eclipse fluorescence spectrometer at 25 ± 1 °C using a sample volume of 100 µl and a 0.3-cm path length cuvette. Excitation wavelength was 387 nm, and the emission intensity was recorded at 490 nm. Bandwidths for both excitation and emission were set at 5 nm. All emission intensity values were corrected for volume changes and background fluorescence.

Limited Proteolysis of PhoP—Wild-type or PhoP variants were limit-digested by MS-grade trypsin (Sigma), and subsequent identification of proteolytic fragments was carried out by N-terminal sequencing as described previously (17).

RESULTS

PhoP Consists of a Unique 20-Residue-long N-terminal Arm—Having established the effect of phosphorylation of PhoP on its effector domain regulation (10), we sought to investigate the mechanism that allows propagation of information from the N-terminal receiver domain to the C-terminal effector domain (see Fig. 1A for domain structure of PhoP). To this end, we investigated the N-terminal domain of PhoP more closely. Previously, we have shown that the Arg²⁻Val²³ peptide bond is cleaved readily by trypsin to release the first 20 residues of the extreme N terminus (17). Also, sequence alignment of PhoP with its family members clearly shows that the region spanning Arg²⁻Arg¹³⁺ (excluding the aforementioned 20 residues) contains all of the amino acids that have been identified as being conserved in the OmpR/PhoB family of response regulators (Fig. 1B). In agreement with these results, recent structural data reveal that the PhoP residues Arg²⁻Arg¹³⁺ correspond to the (β/α)₅ topology, leaving out approximately 20 residues of the N terminus that lack sufficient electron density, suggesting disordered structure (16). Thus, we surmise that the N-terminal arm, unique to M. tuberculosis PhoP, could be a part of the unstructured region of the domain.

For in vitro studies, the recombinant N-terminal arm-truncated variant of PhoP lacking the N-terminal 20 residues spanning Lys⁴ to Arg²₂ (hereafter referred to as nt20-PhoP) was expressed and purified as described under “Experimental Procedures.” Purified nt20-PhoP, similar to full-length PhoP, eluted as a single peak from Protein Pak 300sw (Waters) size-exclusion column (supplemental Fig. S1), suggesting that the truncated protein is likely folded and remains largely monomeric in solution. This result is consistent with a recent report suggesting that E. coli-derived recombinant PhoP remains

⁵The abbreviations used are: ATc, anhydrotetracycline; DSS, disuccinimidyl suberate; AcP, acetyl phosphate; bis-ANS, 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid.
largely monomeric in solution (16). The estimated molecular mass of nt20-PhoP (27.6 kDa) as determined from the calibration curve (supplemental Fig. S1, inset) is in agreement with the molecular mass of 27,489 ± 10 and 29,495 ± 12 Da for the monomeric form of nt20-PhoP and wild-type PhoP, respectively, as determined by mass spectrometry.

On comparison of secondary structure, nt20-PhoP displayed mean residue ellipticity at 220 nm and a minimum around 208 nm, with no significant change of absolute CD intensity and the spectrum shape relative to PhoP (data not shown). Evaluation of protein stability with loss of \( \theta_{220} \) values (CD ellipticity at 220 nm) as a function of increasing guanidine hydrochloride concentration clearly showed that the truncated protein was of comparable overall stability to that of wild-type PhoP (compare free energy of unfolding of 1.8 ± 0.1 and 1.3 ± 0.1 kcal/mol for PhoP and nt20-PhoP, respectively). These results are consistent with the fact that the largely unstructured arm does not appear to contribute significantly to the stability of PhoP.

We next compared phosphorylation efficiency of a comparable amount of nt20-PhoP and wild-type PhoP using labeled PhoRC as the phospho-donor (20). Our results clearly show that, under the conditions examined, the steady-state phosphorylation reached in these reactions are largely comparable (supplemental Fig. S2A, compare lanes 1–4 (for PhoP) with lanes 6–9 (for nt20-PhoP)). Thus, we conclude that the region consisting of residues Lys\(^3\)–Arg\(^{22}\) of PhoP is not necessary for its phosphorylation function. In agreement with these results, none of the conserved residues identified to be essential for phosphorylation of PhoP (Fig. 1B and supplemental Fig. S2B) is located within the N-terminal arm.

**N-terminal Arm of PhoP Is Essential for Its in Vivo Function**—Growing evidence in recent years has shown that PhoP of *M. tuberculosis* plays a major role in complex lipid biosynthesis (7–9, 13). More recent studies from our laboratory have shown that the phosphorylated form of PhoP directly regulates expression of *pks2* and *msl3* genes (10) encoding enzymes that have been implicated in the synthesis of multiple methyl-branched fatty acids essential for biosynthesis of cell wall lipids, like sulfolipids (26), diacyltrehaloses, and polyacyltrehaloses (27). In agreement with these results, analyses of nonmucolylated extractable lipids by TLC showed that various sulfatides and diacyltrehaloses or polyacyltrehaloses were largely undetectable in the *phoP* mutant strain (7) (compared with the wild-type *M. tuberculosis* H37Rv). However, complementation of the *phoP* mutant with wild-type *phoP* showed only partial restoration of acyltrehaloses and sulfatides (7). Thus, to examine the role of the N-terminal arm of PhoP during in vivo regulation, we constructed pME1mL1-nt20-phoP, expressing the truncated PhoP from the P\(_{myc1}\)-tetO promoter under the control of the TetR repressor as described earlier (10, 22). To this end, in vivo experiments using *M. smegmatis* mc\(^{2}\)155 strains harboring pSM128 (28)-derived transcription fusion pks2up-\( \beta \)-galZ and msl3up-\( \beta \)-galZ (as described in Ref. 10) were carried out to compare \( \beta \)-galactosidase activity under conditions of inducing expression of nt20-PhoP and PhoP (Fig. 2). Note that PCR-amplified DNA fragments pks2up and msl3up, comprising nucleotides −206 to +40 and −350 to +60, with respect to their relevant translational start sites, were considered as the regulatory regions of pks2 and msl3, respectively. In sharp contrast to a 2.9 ± 0.2-fold activation of pks2 expression with inducing PhoP expression in the presence of anhydrotetracycline (ATc), there was no detectable difference in the \( \beta \)-galactosidase levels (0.95 ± 0.1-fold) in the presence or absence of inducing expression of nt20-PhoP. Likewise, in contrast to

---

**FIGURE 1. PhoP contains a unique N-terminal arm, which is not required for phosphorylation.** A, domain organization of *M. tuberculosis* PhoP showing a 20-residue-long (Lys\(^3\)–Arg\(^{22}\)) N-terminal arm, connected to the N-domain (light gray) by a trypsin-sensitive peptide bond (Arg\(^{22}\)/Val\(^{23}\)), the C-domain (dark gray), and the linker (residues 139–149) which tethers the two domains together (17). B, multiple sequence alignment of the N-terminal domain of PhoP (MtPhoP) and its family members from *Corynebacterium efficiens* (CePhoP), *Mycobacterium leprae* (MlPhoP), *Streptomyces coelicolor* (ScPhoP), *E. coli* (EcPhoB), and *Salmonella typhimurium* (StPhoP). Conserved and similar amino acids are indicated by red and blue shadings, respectively. Six highly conserved residues of the protein family important for phosphorylation and signal propagation are indicated by asterisks.
1.8 ± 0.1-fold activation of *msl3* expression by inducing PhoP expression (in the presence of ATc), we did not observe significant variation of β-galactosidase levels in the presence or absence of inducing expression of nt20-PhoP (1.1 ± 0.05-fold for *msl3*). In agreement with earlier results (10), the reporter constructs under identical experimental conditions failed to show promoter activation when expression of phosphorylation-deficient PhoP-D71N was induced. Note that a fold difference of β-galactosidase activity (ratio of enzyme activities) between induced and the uninduced samples, in all cases, never exceeded 1.05 ± 0.05-fold for strains carrying different reporter constructs but lacking a PhoP expression plasmid (Fig. 2). Corresponding insets show comparable expression of wild type and variants of *M. tuberculosis* PhoP in *M. smegmatis* in the presence or absence of ATc.
ence of 50 ng/ml ATc, thus ruling out the possibility that altered expression of the truncated PhoP might account for the failure in transcription activation. It is noteworthy that comparable in vitro stability of the E. coli-derived PhoP protein or its variants, coupled with similar in vivo expression and/or stability in M. smegmatis (Fig. 2), suggests that the failure to contribute to positive regulation of pks2, and msl3, expression (Fig. 2) is very unlikely to be due to differences in expression and/or stability of these proteins. Thus, we surmise that the N-terminal arm of M. tuberculosis PhoP is essential for its in vivo function as a transcription regulator.

To further investigate in vivo regulation of PhoP and nt20-PhoP, real time reverse transcription-PCR was used to compare pks2 and msl3 expression from M. tuberculosis H37Ra, H37Ra complemented with phoP and nt20-phoP (supplemental Fig. S3). In agreement with our previous results, pks2 and msl3 expression could be significantly activated when H37Ra was transformed with pJEM-phoP (10). However, complementation of pks2 and msl3 expression could not be achieved when H37Ra was transformed with pJEM-nt20-phoP (10). However, complementation of pks2 and msl3 expression could not be achieved when H37Ra was transformed with pJEM-nt20-phoP, carrying the nt20-phoP allele from H37Rv. Note that H37Ra expressing nt20-PhoP displayed ~4-fold lower pks2 expression (supplemental Fig. S3A) and ~11-fold lower msl3 expression (supplemental Fig. S3B), respectively, compared with H37Ra expressing wild-type PhoP. It should be noted that consistent with an earlier report (10) phosphorylation-defective PhoP-D71N, in both cases, served as an appropriate negative control. Thus, we conclude that PhoP-dependent transcription regulation of pks2 and msl3 expression is dependent on the presence of the N-terminal arm of PhoP.

DNA Binding by PhoP and Its Truncated Variant at the Target Promoters—To investigate what accounts for the failure to activate target gene expression in vivo by nt20-PhoP, we examined recruitment of the protein at the regulatory region of pks2 and msl3 by EMSA using end-labeled pks2up and msl3up DNA fragments (as described above), respectively. Interestingly, for all of the DNA substrates, purified nt20-PhoP, either in the unphosphorylated form or after preincubation in phosphorylation mix, containing acetyl phosphate (AcP) as the phospho-donor (29), was unable to generate a complex stable for gel electrophoresis (Fig. 3, A and B, lanes 3–6). However, consistent with our previous result (10), under identical conditions, full-length PhoP preincubated in phosphorylation mix containing AcP showed efficient DNA binding to both the end-labeled promoter fragments (pks2up, and msl3up, respectively) with an apparent dissociation constant of 0.15 ± 0.05 μM (Fig. 3, A and B, compare lanes 10 and 11 with lanes 5 and 6). From the EMSA data (Fig. 3), it is apparent that a 2-fold higher protein concentration clearly shows a striking difference of DNA binding function of the truncated protein compared with full-length PhoP.

In addition, from the binding data there is a clear difference of at least 7.5–10-fold pixel density against an appropriate background based on limits of detection in these assays. Together, these results suggest an effective decline of DNA binding function of the truncated mutant by at least 15–20-fold compared with PhoP. Thus, we conclude that the presence of the N-terminal arm is essential for high affinity recruitment of phospho-PhoP at the regulatory regions of its target genes.

Having shown the unexpected role of the N-terminal arm in high affinity DNA binding by the phospho-PhoP, we next investigated the role of the N-terminal arm in dimerization of phospho-PhoP. We reasoned that dimerization, which is mediated by the N-domain of many response regulators, often promotes DNA binding of these regulators (30, 31). To this end, PhoP or nt20-PhoP preincubated with AcP was subjected to cross-linking using DSS, which cross-links two amino groups that are in close proximity to each other. After analysis by SDS-PAGE, samples were visualized by immunoblotting using anti-PhoP antibody. A band that migrated corresponding to a dimer of PhoP was observed for both phospho-PhoP and phospho-nt20-PhoP (Fig. 4, compare lane 3 with lane 7). Multiple replicates of experiments clearly show that under the identical conditions examined, phospho-nt20-PhoP is as effective as the phospho-PhoP (1.4 ± 0.2-fold difference) in forming a stable cross-linked dimer. From these results, we conclude that the N-terminal arm does not appear to play a significant role in dimerization of phospho-PhoP, clearly suggesting that reduced DNA binding is not attributable to the absence of the N-terminal arm by its impact on dimerization of phospho-PhoP. It is noteworthy that the cross-linking results are in agreement with the crystal structure data of PhoP (16).

However, the absence of detectable dimer formation in gel filtration studies is likely due to the relatively lower protein concentration used, a result that is in...
N-terminal Arm and Regulatory Mechanism of PhoP

Agreement with a relatively higher $K_D$ value (high micromolar) of dimer-monomer equilibrium as suggested by previously reported analytical ultracentrifugation data (16).

Phosphorylation of nt20-PhoP by AcP—Having shown that nt20-PhoP is unable to be recruited on the target promoters, despite maintaining intact structure of both the N-terminal (phosphorylation domain) and the C-terminal (DNA binding) domain, we sought to investigate consequences of the N-terminal arm truncation on the AcP-dependent phosphorylation of the protein. To this end, both PhoP and its truncated variants were preincubated in reaction mix containing AcP as the phospho-donor. Phosphorylation of PhoP and nt20-PhoP was evaluated using two-dimensional gel electrophoresis, as addition of a phosphoryl group shifts the pI of a given protein toward acidic range (32). In the absence of AcP, recombinant PhoP and nt20-PhoP focused predominantly as a single protein species with a calculated pI of $\sim 6.9$ (Fig. 5, A and B, left panel, respectively). This is consistent with the predicted pI of 7.0 for the unphosphorylated form of PhoP (21). In contrast, PhoP and nt20-PhoP, preincubated with AcP under phosphorylation conditions, focused as two predominant protein species (Fig. 5, A and B, right panel, respectively). Notably, for these samples a PhoP species that focused at pI 6.9 and another that focused at pI 6.7 were observed. To examine whether this apparent shift in pI of PhoP following incubation with AcP is due to specific phosphorylation of PhoP, a phosphorylation-defective PhoP-D71N mutant was preincubated with AcP, and identical reactions were analyzed by two-dimensional gel electrophoresis (Fig. 5C).

In contrast to wild-type PhoP or nt20-PhoP, PhoP-D71N, both preincubated with or without AcP, focused only at pH 6.9 (Fig. 5C, compare left with right panels). Thus, incubation with AcP results in specific phosphorylation of both PhoP and nt20-PhoP. Densitometric scanning of replicates of gels clearly established that PhoP or its truncated variant is phosphorylated to $27.8 \pm 0.7$ and $33.5 \pm 0.8\%$, respectively. Together, these results are consistent with in vitro phosphorylation of PhoP and nt20-PhoP using PhoRC as the phospho-donor (supplemental Fig. S2A).

Truncation of the N-terminal Arm Causes Conformational Change of PhoP—To probe whether N-terminal arm truncation causes a conformational alteration, we next compared bis-ANS binding of PhoP and nt20-PhoP, both in the unphosphorylated and phosphorylated form (Fig. 6, A and B). The hydrophobic regions of proteins that are thought to be important for binding of this class of fluorescent probes require elements of organized structures, and in general, ANS or bis-ANS is incapable of binding to completely unfolded structures of proteins. Consistent with this notion, compared with unfolded PhoP or nt20-PhoP (when the proteins were incubated with 6 M guanidine hydrochloride at 25 °C for 8–10 h), there were substantial increases in fluorescence signal for the native form of both the wild-type as well the truncated protein, suggesting the presence of elements of folded structure (Fig. 6, A and B, compare triangles with circles). It is noteworthy that nt20-PhoP, even in the unphosphorylated form, displayed significantly higher bis-ANS binding compared with the full-length protein (Fig. 6A) suggesting that the differences in tertiary structure and not the level of phosphorylation contribute to enhanced bis-ANS binding. More importantly, a comparable concentration of phosphorylated nt20-PhoP showed significant fluorescence enhancement compared with wild-type PhoP (Fig. 6B, compare filled circles with open circles), indicating that truncation of the N-terminal arm is accompanied by alteration in the folded structure. In view of the fact that PhoP is capable of forming a dimeric structure (16), altered quaternary structure of the arm-truncated protein may also account for enhanced bis-ANS binding.

As change in folded structure often influences proteolytic susceptibility, we utilized a limited proteolysis approach to further probe conformational change of PhoP and nt20-PhoP,
N-terminal Arm and Regulatory Mechanism of PhoP

Effectors proteins involved in transmitting information from the environment to the cells often adopt multiple homeostatic mechanisms to cause physiological changes at the cellular level in response to environmental conditions. What remains unknown is which molecular determinant(s) enable these regulators to operate in multiple molecular forms. Recent work from our laboratory showed that <i>M. tuberculosis</i> PhoP is capable of recognizing its target promoters with significantly different binding affinities by at least two different mechanisms, namely phosphorylation-independent and phosphorylation-coupled. In this work, we sought to investigate how phosphorylation-coupled activation of PhoP is propagated to the C-terminal effector domain.

Although members of the response regulator family share similar structural organization, it has been appreciated that structural complexity of the two-domain response regulators (like PhoP) significantly contributes to the functional diversity of the regulators. One of the major differences that could account for the two distinctly different mechanisms of action (by OmpR and PhoB) resides in the inter-domain linker region (33). Consistent with this view, we previously showed high affinity recruitment of phospho-PhoP at the regulatory region of target genes using a linker length-dependent DNA-binding
What is particularly novel and we think mechanistically interesting is what we show here, i.e. the involvement of additional regulatory control by the N-terminal arm, a region other than the structured domains and the inter-domain linker, which enables appropriate in vivo functioning of the DNA-binding transcription factor.

Genetic and biochemical studies demonstrate that PhoP regulates biosynthesis of complex lipids, the absence of some of which in a phoP mutant is one of the major reasons for its attenuated growth in animal models (6–8). Despite the availability of high resolution structural data of the effector domain as well as full-length PhoP (16, 34), and our recent understanding at the molecular level on how phospho-PhoP regulates lipid biosynthesis (10), what remains unclear is the mechanism of how phosphorylation-coupled activation of the regulator is propagated to the effector domain. It should be noted that the PhoP C-terminal fragment (lacking the N-terminal phosphorylation domain), which includes all of the residues necessary and sufficient for sequence-specific DNA-protein interactions (34, 35), is unable to bind the upstream regulatory region of genes involved in complex lipid biosynthesis (10). In the results reported here, we identify an N-terminal arm, unique to M. tuberculosis PhoP among members of the OmpR/PhoB family, and demonstrate the importance of the arm as a key molecular determinant to phosphorylation-coupled activation of the regulator.

We present evidence showing that the N-terminal arm, which remains linked to the receiver domain by a trypsin-sensitive peptide bond (Arg22/Val23; Fig. 1A), is not necessary for phosphorylation (supplemental Fig. S2A). However, the arm is necessary for phosphorylation-coupled high affinity DNA binding by the regulator, a result of unusual significance contributing to the insights into the mechanism of action of the regulator. These results suggest critical involvement of the arm in functional coupling of the two domains necessary for transcriptional activation of target genes. The fact that truncation of the arm abrogates phosphorylation-coupled promoter recruitment (of nt20-PhoP) (Fig. 3) by changing the conformation of the effector domain (Figs. 6 and 7) without influencing the phosphorylation efficiency of PhoP (supplemental Fig. S2A and Fig. 5B) is consistent with our earlier result suggesting that phosphorylation of the N-domain alone is not sufficient for activation of target genes (17). More importantly, our subsequent results show that the arm-truncated protein is of different conformation at or around the linker (Fig. 7), a region that has been shown to influence high affinity promoter recognition by PhoP (17). What offers a new mechanistic insight is the finding that the presence of the N-terminal arm contributes to specific conformation of PhoP (Figs. 6 and 7), which, in turn, is

![Figure 7](http://www.jbc.org/content/288/40/29190/F7.large.jpg)
N-terminal Arm and Regulatory Mechanism of PhoP

distinctive aspect of PhoP structure shows that the N-terminal arm remains spatially proximal to the unstructured linker (Fig. 8). Although it is possible that the arm can interact with the region at or around the linker, clearly absence of the arm is likely to introduce a conformational change around the PhoP linker. Most likely this is what accounts for the differential bis-ANS binding (Fig. 6) and altered proteolysis profile (Fig. 7) of the arm-truncated protein compared with the full-length PhoP. We speculate that the conformational change perhaps modifies the orientation of the effector domain with respect to the corresponding receiver domain of the regulator leading to its inactivation. However, the N-terminal arm being sufficiently away from the dimeric interface (α4-β5-α5; Refs. 16, 18, 36) is unlikely to influence protein dimerization, a result confirmed by DSS cross-linking studies (Fig. 4).

The presence of a new architectural feature in addition to what is already available to other members of the family is probably not surprising, and even expected because very little sequence homology within an array of diverse target promoters are being recognized by the conserved structural motif of PhoP. Nevertheless, unique properties of PhoP attributable to the N-terminal arm suggest an additional layer of complexity that has evolved to further tune gene expression in response to the continuously changing environment of the bacteria.

Acknowledgments—We are grateful to G. Marcela Rodriguez and Issar Smith (Public Health Research Institute, UMDNJ) for the kind gift of plasmid pSM128 and Sabine Ehrt (Weil Medical College of Cornell University) for pME1mL1 expression vector. We thank Subramanian Karthikeyan for very useful discussions on PhoP structure; Sankalp Gupta and Akesh Sinha for preliminary studies with nt20-PhoP; Rajini Goyal for comments and suggestions; Neha Rana, Paramjit Kaur and Girish Sahni for N-terminal sequencing of the peptides; Sharanjeet Kaur and Purnananda Guptasarma for circular dichroism studies; Renu Sharma for technical assistance and for help with the preparation of the manuscript.

REFERENCES


Unique N-terminal Arm of *Mycobacterium tuberculosis* PhoP Protein Plays an Unusual Role in Its Regulatory Function

Arijit Kumar Das, Vijjamari Anil Kumar, Ritesh Rajesh Sevalkar, Roohi Bansal and Dibyendu Sarkar

doi: 10.1074/jbc.M113.499905 originally published online August 20, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.499905

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/08/20/M113.499905.DC1

This article cites 35 references, 17 of which can be accessed free at
http://www.jbc.org/content/288/40/29182.full.html#ref-list-1