Title
Solution structure of the PhoP DNA-binding domain from Mycobacterium tuberculosis

Journal
Journal of Biomolecular NMR, 63(1)

ISSN
0925-2738

Authors
Macdonald, Ramsay
Sarkar, Dibyendu
Amer, Brendan R
et al.

Publication Date
2015-09-01

DOI
10.1007/s10858-015-9965-0

Peer reviewed

HHS Public Access

Author manuscript

J Biomol NMR. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

J Biomol NMR. 2015 September ; 63(1): 111–117. doi:10.1007/s10858-015-9965-0.

Solution Structure of the PhoP DNA-Binding Domain from *Mycobacterium tuberculosis*

Ramsay Macdonald[a,b], **Dibyendu Sarkar**[d], **Brendan R. Amer**[a,b], and **Robert T. Clubb**[a,b,c,*]

[a]Department of Chemistry and Biochemistry, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA

[b]UCLA-DOE Institute of Genomics and Proteomics, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA

[c]Molecular Biology Institute, University of California, Los Angeles, 611 Charles Young Drive East, Los Angeles, CA 90095, USA

[d]CSIR-Institute of Microbial Technology, Sector-39A, Chandigarh-160036, India

Summary

Tuberculosis caused by *Mycobacterium tuberculosis* is a leading cause of death world-wide. The PhoP protein is required for virulence and is part of the PhoPR two-component system that regulates gene expression. The NMR-derived solution structure of the PhoP C-terminal DNA-binding domain is reported. Residues 150 to 246 form a structured domain that contains a winged helix-turn-helix motif. We provide evidence that the transactivation loop postulated to contact RNA polymerase is partially disordered in solution, and that the polypeptide that connects the DNA-binding domain to the regulatory domain is unstructured.

Keywords

MTB; PhoPC; NMR; heteronuclear NOE

Biological Context

The bacterial pathogen *Mycobacterium tuberculosis* (MTB) causes tuberculosis resulting annually in ~1.5 million fatalities world-wide (WHO 2014). During the initial stage of an infection, MTB is inhaled and transported to the lungs where it is believed to infect alveolar macrophages (Russell et al. 2010). Once inside the macrophage, MTB encounters new oxidative and acidic stressors, and alters its gene expression profile to cope with this new environment (Gonzalo-Asensio et al. 2008). The MTB PhoP-PhoR (PhoPR) two-component system plays a critical role in microbial adaptation. In this system, environmental signals received by the membrane associated PhoR protein trigger its autophosphorylation. PhoR

*To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry, University of California, Los Angeles, 602 Boyer Hall, Los Angeles, CA 90095. Tel.: 310-206-2334; Fax: 310-206-4749; rclubb@mbi.ucla.edu.

Conflict of Interest: The authors declare no conflict of interest.

Ethical standards: Research does not involve human participants and/or animals.

then transfers the phosphate to the cytoplasmic PhoP response regulator, which in its phosphorylated state sequence specifically binds DNA with high affinity (Pathak et al. 2010). Several studies have shown that PhoP is required for MTB virulence (Perez et al. 2001; Walters et al. 2006; Gonzalo-Asensio et al. 2008). In particular, *phoP* mutant strains of MTB exhibit attenuated growth in murine and macrophage models of infection (Perez et al. 2001). Attenuated growth is due largely to the absence of certain complex lipids in the MTB cell envelope that protect the microbe against host defense mechanisms (Walters et al. 2006; Ryndak et al. 2008). Over 110 genes in MTB exhibit altered expression levels in the absence of PhoP (Walters et al. 2006). PhoP regulates genes involved in a variety of cellular functions, including: hypoxia response, respiratory metabolism, stress response, secretion of major T-cell antigen ESAT-6, synthesis of pathogenic lipids, and MTB persistence through transcriptional regulation of the enzyme isocitrate lyase (Gonzalo-Asensio et al. 2008). While antimicrobial therapies against MTB exist, new anti-tubercular therapeutics are needed as current approaches can be problematic to implement as they require the use of a sustained treatment regimen and new drug resistant forms of MTB have emerged (Ryndak et al. 2008).

The PhoP protein contains two autonomously folded domains, an N-terminal receiver domain that is phosphorylated by PhoR (PhoPN; spanning residues 1–138), and a C-terminal effector domain that binds DNA (PhoPC; comprising residues 150–247) (Pathak et al. 2010). It is unclear how phosphorylation of the PhoPN domain causes PhoP to bind DNA with high affinity via its PhoPC domain. Previously, we have shown that the linker connecting the PhoPN and PhoPC domains is required for phosphorylation-dependent DNA binding *in vitro*, and thus, the transmission of the phosphorylation signal between the domains (Pathak et al. 2010). To gain insight into the mechanism of DNA binding and the structural role of the interdomain linker, here we report the NMR structure and backbone dynamics of the PhoPC DNA-binding domain and several residues from the linker polypeptide that connects it to the receiver domain in the full-length protein.

Methods and results

Protein expression, purification, and NMR sample preparation

NMR was used to determine the structure of a polypeptide containing the C-terminal DNA-binding domain of PhoP (PhoPC, residues 142–247 of PhoP, Fig. 1a). The protein also contains at its N-terminus 30 amino acids that include a six residue histidine tag and residues that from the linker segment between the N- and C-terminal domains in the full length protein. Uniformly ^{13}C- and ^{15}N-labeled PhoPC was expressed in *Escherichia coli* BL21(DE3) cells grown in M9 medium supplemented with $^{15}NH_4Cl$ and [$^{13}C_6$] glucose. Cultures were grown at 37° C to an A_{600} of 0.4 before induction with isopropyl β-D-thiogalactoside to a final concentration of 1 mM. Induction proceeded at 18° C overnight before harvesting the cells by centrifugation at 5,400 × g for 20 min at 4° C. The cell pellet was then resuspended in lysis buffer consisting of: 50 mM $NaPO_4$, 500 mM NaCl, 10 mM $CaCl_2$, and 10 mM $MgCl_2$ (pH 8.0). Lysozyme was added to a final concentration of 0.1 mg/mL and the resuspended cells were then kept on ice for 30 min. EDTA and NaCl were then added to the final concentrations of 25 mM and 600 mM, respectively, along with a

protease inhibitor cocktail (Calbiochem) and phenylmethanesulfonyl fluoride (Sigma). Cells were lysed by sonication and centrifuged at 11,300 × g for 40 min at 4º C. The supernatant was then incubated with 1.0 mL of pre-equilibrated Ni^{2+} resin (Clontech) for 60 min at 4º C on a rotisserie before being transferred to a gravity column. The resin was then washed with 10 mL lysis buffer containing 1.0 M NaCl, 10 mL lysis buffer containing 15 mM imidazole, and the PhoPC protein was finally eluted by adding lysis buffer containing 300 mM imidazole. Fractions containing PhoPC were pooled and dialyzed against NMR buffer. NMR samples of PhoPC contained 1 mM U-[^{15}N,^{13}C] PhoPC dissolved in 50 mM $NaPO_4$, 300 mM NaCl, and 0.01% NaN_3 (pH 6.5). Two samples were studied by NMR that either contain 8% or 99.9% v/v of D_2O. Because of the high salt concentration in the sample buffer, a “Shaped Sample Tube” (Bruker) was employed to maximize sensitivity.

NMR spectroscopy, data collection, and chemical shift assignments

NMR experiments were performed at 298 K on Bruker Avance 500-, 600-, and 800-MHz spectrometers equipped with triple resonance cryogenic probes. NMR spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed using the programs CARA (Keller 2004) and PIPP (Garrett et al. 1991). Backbone ^{1}H, ^{13}C, and ^{15}N chemical shifts were assigned by analyzing data from the following experiments: 2D [^{1}H-^{15}N]-HSQC, 3D CBCA(CO)NH, HNCACB, HNCO, HN(CA)CO, HNCA, HNHA, HBHA(CO)NH, HBHANH, CC(CO)NH, ^{15}N-edited TOCSY-HSQC (40 ms mixing time), and ^{15}N-edited NOESY-HSQC (120 ms mixing time) (Fig. 1b). Aliphatic side-chain assignments were obtained by analyzing 2D [^{1}H-^{13}C]-HSQC (aliphatic), 3D HC(C)H-COSY, HC(C)H-TOCSY, and (H)CCH-TOCSY spectra [For a review, see (Cavanagh et al. 1995)]. Aromatic side-chain assignments were determined using 2D [^{1}H-^{13}C]-HSQC (aromatic) and 3D ^{13}C-edited NOESY-HSQC (aromatic, 150 ms mixing time) datasets. NOE distance restraints were obtained by analyzing 3D ^{15}N-edited NOESY-HSQC (120 ms mixing time), ^{13}C-edited NOESY-HSQC (aliphatic, 120 ms mixing time), and ^{13}C-edited NOESY-HSQC (aromatic, 150 ms mixing time) spectra. ^{13}C-edited NOESY experiments were optimized for either aliphatic or aromatic signal detection by positioning the ^{13}C carrier at 40.0 ppm or 80.0 ppm, respectively. Dihedral angle restraints, φ and ψ, were obtained using the program TALOS+ (Shen et al. 2009). $^3J_{HN^\alpha}$ couplings were measured using a 3D HNHA spectrum. To identify disordered regions in PhoPC, heteronuclear [^{1}H-^{15}N]-NOE values were measured (in triplicate) and analyzed using the program SPARKY (Goddard and Kneller 2006).

NMR structure determination

Simulated annealing with restrained molecular dynamics was used to calculate the structures of PhoPC using the program XPLOR-NIH (Schwieters et al. 2003). The structure was determined in an iterative manner. Initially, NOE cross-peaks in the NOESY spectra were assigned automatically using the programs ATNOS-CANDID (Herrmann et al. 2002b; Herrmann et al. 2002a) and UNIO (Herrmann 2010). To facilitate automatic assignments, the PhoPC crystal structure (pdb accession: 2PMU) was used as input for UNIO. Hydrogen atoms were added to the crystal structure using the Bax laboratory web-server (http://spin.niddk.nih.gov/bax/nmrserver/pdbutil/sa.html). The first round of UNIO calculations correctly generated the global fold of the protein. The structure was then iteratively refined

by adding new distance restraints that were manually identified in the NOESY spectra. In this procedure, additional side chain chemical shift assignments were obtained by inspecting all of the NMR spectra, and the 3D ^{15}N-edited NOESY-HSQC and 3D ^{13}C-edited NOESY-HSQC spectra were further analyzed to identify NOEs that were not originally assigned by UNIO. Importantly, in this procedure all of the NOEs assigned by UNIO were verified manually. Approximately 10 cycles of iterative refinement were performed. Fig. 3c shows a plot of the number of NOE derived distance restraints identified per residue. For the structured regions of the protein, an average of ~14 distance restraints define the conformation of each residue. As expected, fewer NOE distance restraints were obtained for residues located within surface loops and the polypeptide termini.

Towards the end of the refinement process, hydrogen bond restraints were employed that were identified from deuterium exchange data and characteristic NOE patterns present in the NOESY spectra (Wüthrich 1986). The final structures were also refined so as to agree with the $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of the protein, and $^{3}J_{HN}{}^{\alpha}$ coupling data. The range for the φ and ψ dihedral angle restraints was set to two times the error defined by the program TALOS+, or ± 30°, whichever value was larger. A total of 200 structures were calculated, of which 187 were completely compatible with the NMR data; they had no NOE, dihedral angle, or scalar coupling violations greater than 0.5 Å, 5°, or 2 Hz, respectively. Chemical shift assignments have been deposited in the Biological Magnetic Resonance Bank (BMRB 11590), and the coordinates and structure restraints have been deposited in the protein data bank (PDB 2RV8). The solution structure of PhoPC is represented by an ensemble of the 30 lowest energy conformers generated from this calculation (Fig. 2a, pdb accession 2RV8). Residues 150–205 and 211–246 in PhoPC are structured, and the root mean square deviation of their backbone and heavy atom coordinates to the mean structure is 0.48 ± 0.10 Å and 1.05 ± 0.10 Å, respectively. Table 1 contains the complete structural and restraint statistics.

Solution structure and dynamics of PhoPC

The structure of PhoPC is formed by two β-sheets that pack against a three-helix bundle (Fig. 2c). To facilitate comparisons with previously reported structures, the secondary structural elements of PhoPC are labeled based on their appearance in the full-length protein (Fig. 2c). Beginning at the N-terminus, the polypeptide adopts a four-stranded antiparallel β-sheet that is formed by residues that are contiguous in the primary sequence (strands β6-β9). The chain then forms helix α6, which packs all of the strands in the four-stranded sheet. After this two additional alpha helices (α7 and α8) occur that pack against helix α6 to form the three-helix bundle. The structure is completed by a β-hairpin constructed from strands β11 and β12, which are attached via strand β12 to the body of the protein by contacts to a short β-strand (β10) segment located between helices α6 and α7. Notably, in all the members of the ensemble, residues at the end of helix α7 form a single turn of a ^{3}10-helix (residues 203–205). The N-terminal β-sheet forms a hydrophobic interface with α6 as evidenced by NOE's between V165 (β8) and F179 (α6), L151 (β6) and I187 (α6), and F153 (β6) and V186 (α6). Hydrophobic contacts are also observed between helices α6, α7, and α8. This is supported by the observation of NOEs between the side chains of T177, T180, and L181 located in α6 and V202 in α7, between L199 (α7) and V214 (α8), as well as between the side chains of L181 (α6) and V218 (α8). A small interface joins the C-terminal

sheet to the helical bundle which is defined by NOE's between Y241 (β12) and residues V218 (α8) and I198 (α7), as well as NOE's between L233 (β11) and the side chains of L221 (α8) and I225 (α8).

Extensive chemical shift and NOE assignments form the foundation for the structure determination of PhoPC. The chemical shifts of 95%, 100% and 50% of the backbone amide (excluding proline residues), side chain methyl, and aromatic side chain atoms were assigned, respectively. For the structured regions of the protein, an average of ~14 NOE-derived distance restraints were identified per residue. These included, ~5 long range (> 4 residues apart), ~3 medium range (2 residue separation 4), and ~5 sequential NOE-derived distance restraints per residue. From the backbone chemical shifts the programs TALOS+ and CSI 2.0 predicted the presence of similar secondary structural elements in the protein (Shen et al. 2009; Hafsa and Wishart 2014). However, some of the elements predicted by the programs differ subtly in their amino acid lengths and location. For example, strands β6 and β9 are predicted by both programs to be of similar length, but they differ by one residue in their location in the primary sequence.

Heteronuclear [^{1}H-^{15}N]-NOE data provides insight into the flexibility of the polypeptide backbone. A total of 71 out of 101 backbone amide resonances are well resolved enabling their [^{1}H-^{15}N]-NOE values to be reliably determined. Overall, the DNA-binding domain of PhoPC adopts a rigid structure, as the majority of the residues spanning strands β6 to β12 exhibit [^{1}H-^{15}N]-NOE values greater than 0.6 (Fig. 3a). This finding is also compatible with order parameters predicted from the backbone chemical shift data (RCI-S^2, random coil index order parameter) as the RCI-S^2 values for residues within the domain generally range between 0.8–0.9 (Berjanskii and Wishart 2005) (Fig. 3b). Interestingly, both the [^{1}H-^{15}N]-NOE and RCI-S^2 data indicate that the loop connecting helices α7 and α8 (α7/α8 loop, residues 206–211) has elevated mobility; [^{1}H-^{15}N]-NOE and RCI-S^2 values for loop residues D206, F207, G208, and D210 are less than 0.6 (Fig. 3a) and 0.8, respectively. It is also compatible with the NMR structure, as the coordinates for residues in the α7/α8 loop are slightly more structurally disordered than the remainder of the domain (Fig. 2a). However, the coordinates of the loop are not entirely disordered because several long-range NOE distance restraints are observed between the loop and the remainder of the protein. The loops connecting α8 and β11, and the turn connecting the two strands in the C-terminal β-hairpin also exhibit slightly elevated mobility as indicated by [^{1}H-^{15}N]-NOE values less than 0.6 for K230 and V239 (Fig. 3a), along with the observation of slightly decreased RCI-S^2 values for these regions of the polypeptide (Fig. 3b). In addition to the DNA-binding domain, the PhoPC polypeptide studied here contains a portion of the interdomain linker segment that is important for phosphorylation-dependent DNA binding (Pathak et al. 2010). Three lines of evidence indicate that the linker is unstructured in PhoPC. First, the [^{1}H-^{15}N]-NOE values of residues K144 and R147 have negative [^{1}H-^{15}N]-NOE values, and residues N148 and V149 exhibit [^{1}H-^{15}N]-NOE values of only 0.18 ± 0.04 and 0.52 ± 0.03, respectively (Fig. 3a). Second, residues in the linker do not exhibit long-range NOE cross-peaks in the NOESY spectra. Finally, the RCI-S^2 values for residues spanning 142–149 have small magnitudes that progressively decrease towards the N-terminus.

Discussion and conclusions

The NMR structure of PhoPC adopts a winged-helix fold similar to members of the OmpR/PhoB subfamily of proteins (Wang et al. 2007). *In vitro* studies have demonstrated that PhoPC binds DNA with similar affinity as the full-length protein, indicating that PhoPC is the primary determinant for DNA binding (Pathak et al. 2010). The helix-turn-helix motif in PhoPC is formed from helices α7 and α8. Helix α8 likely mediates sequence specific binding to the major groove, and binding is presumably further assisted by DNA contacts from the "wing" structure formed by residues in the C-terminal β-hairpin. This binding mode is compatible with the positively charged electrostatic surface formed by residues within the C-terminal β-hairpin and α8 (Fig. 2d) (Wang et al. 2007). Most of the remainder of the protein surface is negatively charged, which could function to orient the protein on the DNA duplex (Wang et al. 2007). The notion that α8 interacts with the major groove is also strongly supported by the results of *in vitro* binding studies as mutants containing alterations in the helix are impaired in DNA binding (Das et al. 2010).

In general, the NMR structure is similar to previously reported crystal structures of PhoP. In particular, the backbone coordinates of the NMR structure have a r.m.s.d. of 1.21 and 1.25 Å to the crystal structures of PhoPC (pdb accession: 2PMU) and full-length PhoP (pdb accession: 3R0J), respectively. The largest differences in the structures occurs in the loop that connects α7 to α8 (loop 1), which has been postulated to interact with RNA polymerase to activate transcription (Wang et al. 2007) (Fig. 2b). As compared with the crystal structure of the full-length protein, the loop in the NMR structure is kinked forward, whereas in the crystal structure it is kinked backwards toward the side of helix α7 (Fig. 2b) This difference primarily originates at residues R204, Y205, D206, and G208 in the loop which have distinct φ and ψ angles in the NMR and crystal structures. However, these structural differences are likely to be due to loop motions as many of its residues have [^{1}H-^{15}N]-NOE values less than 0.6. This notion is consistent with the crystal structure of PhoPC, since interpretable electron density for the loop is available for only four of the six protein molecules within the asymmetric unit. Another less substantial structural difference is the positioning of the C-terminal β-sheet which in the crystal structures is packed slightly closer to helix α8 than in the NMR structure.

Overall, most of the residues in the NMR structure of PhoPC adopt favorable or additionally allowed backbone conformations when assessed using programs PROCHECK and MolProbity (Table 1) (Laskowski et al. 1996; Chen et al. 2010). Interestingly, in all the conformers of the ensemble residues A154 and A168 adopt a conformation in which they have a positive φ angle and are located in the disallowed and generously allowed regions, respectively. This is likely the predominant conformation of these residues in solution as they are well defined by the NOE data (Fig. 3c) and they also exhibit positive φ angles in both crystal structures of PhoP (Wang et al. 2007; Menon and Wang 2011). Residue T227 also adopts a high-energy conformation that places it in the generously allowed regions of the Ramachandran plot. However, this residue is located in loop 2 and is poorly defined by the NMR data. In comparison to the latest structure reported of PhoP (Menon and Wang 2011), the same secondary structural elements are present in the NMR structure. Strands β6, β7, β9, α8, and β11 appear at identical positions in the primary sequence. However, there are

subtle differences as the lengths of elements β8, α6, and β12 are extended by one residue in the crystal structure as compared to the NMR structure. In addition, helix α7 is two residues longer in the NMR structure as compared to the crystal structure. Finally, although strand β10 is the same length in both structures, in the NMR structure it is displaced one residue towards the C-terminus in the primary sequence.

The linker connecting the N- and C-terminal domains is important for phosphorylation-coupled DNA binding (Pathak et al. 2010). This raised the possibility that it might interact with PhoPC to modulate its structure and ability to bind DNA, and that these interactions could be dependent upon phosphorylation of PhoPN. Our data indicate that the linker is unstructured in the absence of the PhoPN domain, as its residues have low magnitude [^{1}H-^{15}N]-NOE and RCI-S^2 values, and are not defined by long-range NOEs in the NOESY spectra. This suggests that in solution there is no intrinsic propensity of the linker to interact with the PhoPC domain. This finding is compatible with the crystal structures of PhoP, as residues for the linker exhibited only partial electron density and could not be modeled in either structures of intact PhoP or PhoPC (Wang et al. 2007; Menon and Wang 2011). Studies suggest that the phosphorylation of PhoPN induces PhoP dimerization, consequently increasing DNA binding affinity in a cooperative manner (Sinha et al. 2008; He and Wang 2014). The presence of a long and flexible linker between the PhoPN and PhoPC domains could permit PhoPN dimerization, while still enabling large-scale domain rearrangements of PhoPC that are needed for it to productively engage DNA as a tandemly oriented dimer (Menon and Wang 2011). Insight into the mechanism of phosphorylation-coupled DNA binding by PhoP awaits the structure determination of the PhoP-DNA complex.

Acknowledgments

We would like to thank Albert H. Chan and Megan Sjodt for guidance throughout the structure determination process. We would like to thank Dr. Robert Peterson for assistance with NMR experiments. This work was supported by the National Institutes of Health grant AI52217 to RTC. R.M. was supported by a Cellular and Molecular Biology Training Grant (Ruth L. Kirschstein National Research Service Award GM007185). D.S. was supported by Raman Research Fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India. B.R.A. was supported by a Whitcome Predoctoral Training Grant, University of California–Los Angeles, Molecular Biology Institute. This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research program under Award Number DE-FC02-02ER63421.

References

Berjanskii MV, Wishart DS. A Simple Method To Predict Protein Flexibility Using Secondary Chemical Shifts. J Am Chem Soc. 2005; 127:14970–14971. [PubMed: 16248604]

Cavanagh, J.; Fairbrother, WJ.; Palmer, AG., III; Skelton, NJ. Protein NMR Spectroscopy: Principles and Practice. Academic Press; New York: 1995.

Chen VB, Arendall WB III, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr Sect D Biol Crystallogr. 2010; 66:12–21. [PubMed: 20057044]

Das AK, Pathak A, Sinha A, et al. A Single-Amino-Acid Substitution in the C Terminus of PhoP Determines DNA-Binding Specificity of the Virulence-Associated Response Regulator from Mycobacterium tuberculosis. J Mol Biol. 2010; 398:647–656. [PubMed: 20363229]

Delaglio F, Grzesiek S, Vuister GW, et al. NMRPipe: A multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR. 1995; 6:277–293. [PubMed: 8520220]

Delano WL. The Pymol User's Manual. 2002

Garrett DS, Powers R, Gronenborn AM, Clore GM. A Common Sense Approach to Peak Picking in Two-, Three-, and Four-Dimensional Spectra Using Automatic Computer Analysis of Contour Diagrams. J Magn Reson. 1991; 95:214–220.

Goddard TD, Kneller DG. SPARKY. 2006; 3

Gonzalo-Asensio J, Mostowy S, Harders-Westerveen J, et al. PhoP: A Missing Piece in the Intricate Puzzle of Mycobacterium tuberculosis Virulence. PLoS One. 2008; 3:e3496. [PubMed: 18946503]

Hafsa NE, Wishart DS. CSI 2.0: a significantly improved version of the Chemical Shift Index. J Biomol NMR. 2014; 60:131–146. [PubMed: 25273503]

He X, Wang S. DNA Consensus Sequence Motif for Binding Response Regulator PhoP, a Virulence Regulator of Mycobacterium tuberculosis. Biochemistry. 2014; 53:8008–8020. [PubMed: 25434965]

Herrmann T. UNIO' 10: Automated NMR Data Analysis for Protein Structure Determination and More (Version 2.0.2). 2010

Herrmann T, Güntert P, Wüthrich K. Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. J Mol Biol. 2002b; 319:209–227. [PubMed: 12051947]

Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J Biomol NMR. 2002a; 24:171–89. [PubMed: 12522306]

Keller R. The Computer Aided Resonance Assignment Tutorial. 2004

Koradi R, Billeter M, Wüthrich K. MOLMOL: A program for display and analysis of macromolecular structures. J Mol Graph. 1996; 14:51–55. [PubMed: 8744573]

Laskowski RA, Rullmann JAC, MacArthur MW, et al. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996; 8:477–486. [PubMed: 9008363]

Menon S, Wang S. Structure of the Response Regulator PhoP from Mycobacterium tuberculosis Reveals a Dimer through the Receiver Domain. Biochemistry. 2011; 50:5948–5957. [PubMed: 21634789]

Pathak A, Goyal R, Sinha A, Sarkar D. Domain Structure of Virulence-associated Response Regulator PhoP of Mycobacterium tuberculosis. J Biol Chem. 2010; 285:34309–18. [PubMed: 20814030]

Perez E, Samper S, Bordas Y, et al. An essential role for phoP in Mycobacterium tuberculosis virulence. Mol Microbiol. 2001; 41:179–187. [PubMed: 11454210]

Russell DG, VanderVen BC, Lee W, et al. Mycobacterium tuberculosis Wears What It Eats. Cell Host Microbe. 2010; 8:68–76. [PubMed: 20638643]

Ryndak M, Wang S, Smith I. PhoP, a key player in Mycobacterium tuberculosis virulence. Trends Microbiol. 2008; 16:528–34. [PubMed: 18835713]

Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. J Magn Reson. 2003; 160:65–73. [PubMed: 12565051]

Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR. 2009; 44:213–23. [PubMed: 19548092]

Sinha A, Gupta S, Bhutani S, et al. PhoP-PhoP Interaction at Adjacent PhoP Binding Sites Is Influenced by Protein Phosphorylation. J Bacteriol. 2008; 190:1317–1328. [PubMed: 18065544]

Walters SB, Dubnau E, Kolesnikova I, et al. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. Mol Microbiol. 2006; 60:312–30. [PubMed: 16573683]

Wang S, Engohang-Ndong J, Smith I. Structure of the DNA-Binding Domain of the Response Regulator PhoP from Mycobacterium tuberculosis. Biochemistry. 2007; 46:14751–14761. [PubMed: 18052041]

WHO. Global Tuberculosis Report 2014. World Health Organization; 2014.

Wüthrich, K. NMR of Proteins and Nucleic Acids. John Wiley & Sons, Inc; New York: 1986.

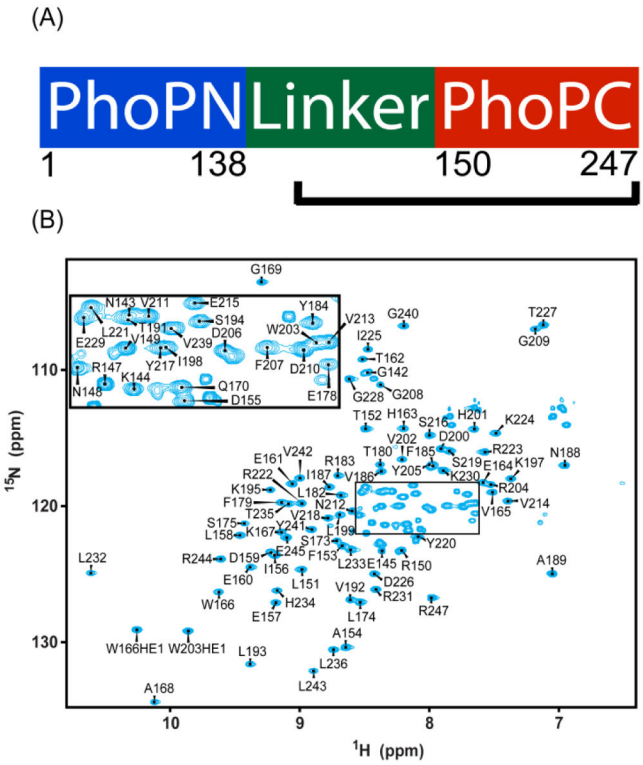

Fig. 1. NMR spectra of PhoPC

a) Schematic of full-length PhoP showing its component domains: N-terminal domain (blue), linker segment (green), and C-terminal domain (red). The bracket at the bottom illustrates the length of our construct. **b)** The [^{1}H,^{15}N]-HSQC spectrum of PhoPC (142–247) [G190 not shown].

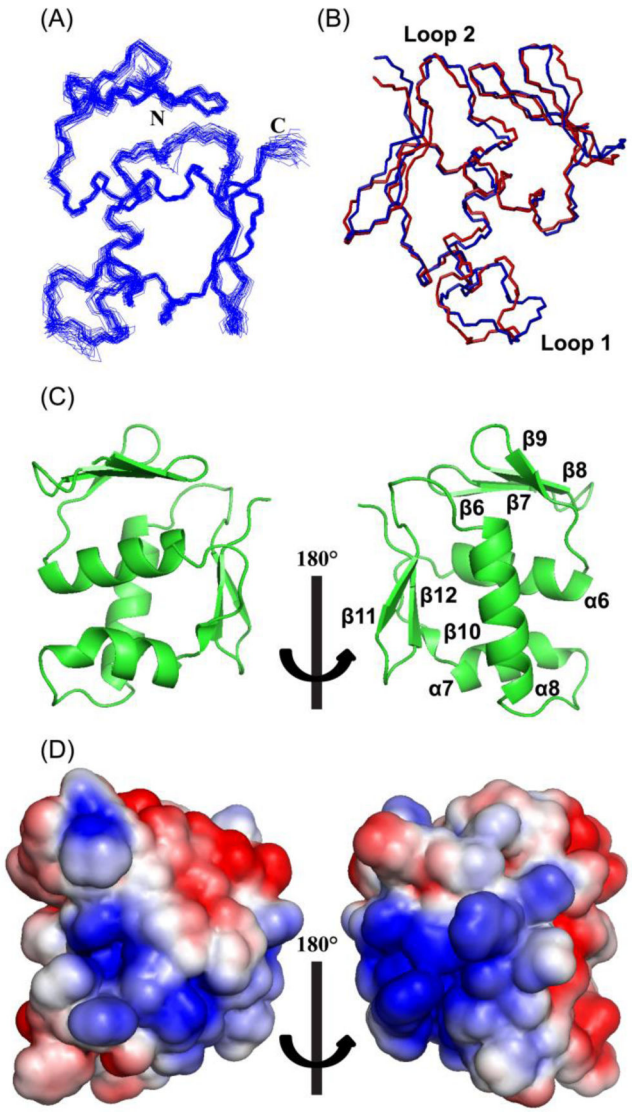

Fig. 2. NMR structure of PhoPC and structural comparisons

a) Superposition of the 30 lowest energy structures with the N- and C-termini labelled. **b)** Overlay of the lowest energy structure (blue) and the crystal structure (red) (pdb accession: 3R0J [N-terminal domain not shown]). **c)** Lowest overall energy structure with the secondary structure labelled according to the crystal structure. **d)** Electrostatic potential of the protein surface calculated in PYMOL (Delano 2002). The positively charged surface potential is shown in blue and the negatively charged surface potential is shown in red. For clarity, residues 149 to 246 are shown in the structures above and were used to calculate the electrostatic potential. Figure was generated using MOLMOL (Koradi et al. 1996) and PYMOL.

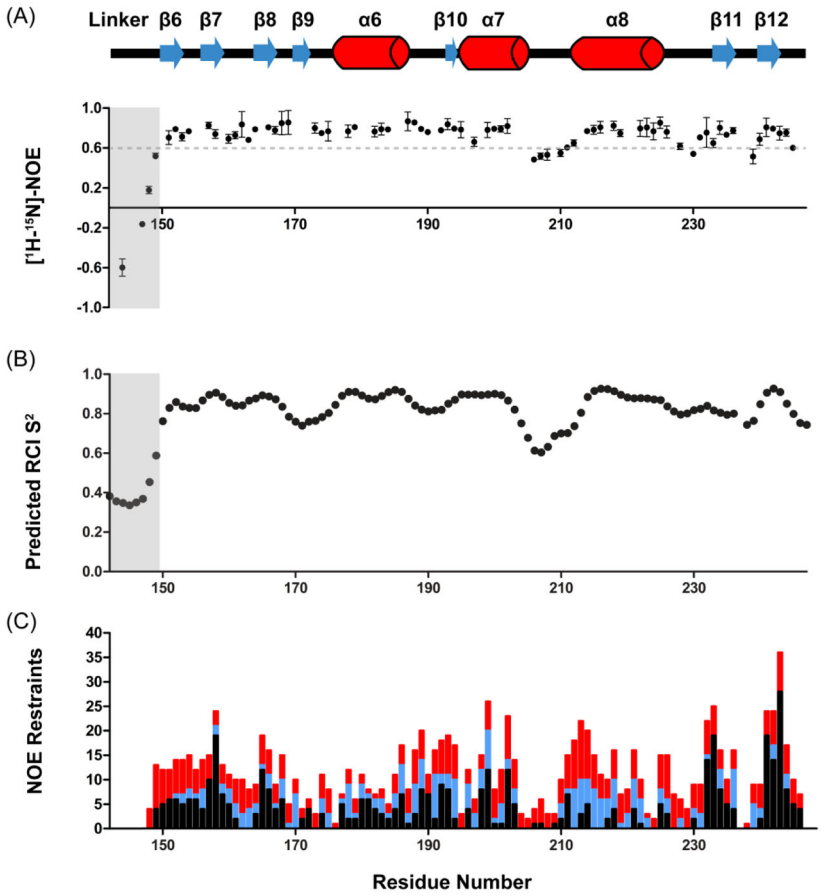

Fig. 3. NMR relaxation, chemical shift and distance restraint data

a) Plot of the heteronuclear [^{1}H-^{15}N]-NOE values versus residue number. Errors were calculated by taking the standard deviation of values taken from three separate experiments. **b)** Order parameters predicted based on backbone chemical shifts using the RCI server (RCI-S^2). Areas shaded in grey represent residues belonging to the linker. **c)** Plot showing the number of NOE-derived distance restraints as a function of residue number. Restraints are classified as long range (black,$|i - j| > 4$), medium (blue, $2\ |i - j| \geq 4$) and sequential (red, $|i - j| = 1$). Figure was generated using Graph Pad Prism (version 5.01).

TABLE 1

Statistics for the solution structure of PhoPC

	<SA>[a]
Root mean square deviations	
NOE interproton distance restraints (Å) (793)[b]	0.015 ± 0.003
Dihedral Angle Restraints (degrees)[c] (185)	0.357 ± 0.044
$^{3}J_{HN}^{\alpha}$ coupling constants (Hz) (25)	0.214 ± 0.025
Secondary ^{13}C shifts (ppm)	
$^{13}C_{\alpha}$ (ppm) (91)	1.070 ± 0.046
$^{13}C_{\beta}$ (ppm) (91)	0.986 ± 0.046
Deviations from idealized covalent geometry	
Bonds (Å)	0.0014 ± 0.0001
Angles (degrees)	0.392 ± 0.005
Impropers (degrees)	0.245 ± 0.006
PROCHECK results[d]	
Most favorable region (%)	80.4 ± 2.3
Additionally allowed region (%)	16.6 ± 2.3
Generously allowed region (%)	1.8 ± 1.0
Disallowed region (%)	1.2 ± 0.2
Coordinate precision[e]	
Protein backbone (Å)	0.48 ± 0.10
Protein heavy atoms (Å)	1.05 ± 0.10

[a] <SA> represents an ensemble of the 30 lowest energy structures calculated by simulated annealing. The number of terms for each restraint is given in parentheses. None of the structures exhibit distance violations greater than 0.5 Å, dihedral angle violations greater than 5°, or coupling constant violations greater than 2 Hz.

[b] Distance restraints: 238 sequential, 125 medium (2 ≥ residue separation ≥ 4) and 211 long range (> 4 residues apart).

[c] Experimental dihedral restraints comprised 93 φ and 92 ψ angles.

[d] PROCHECK (Laskowski et al. 1996) data include residues 150–205 and 211–246. MolProbity (Chen et al. 2010) was also used to assess the quality of the structure. For the structured regions of the protein, 96% ± 1% of the residues were in the favored or allowed regions of the Ramachandran plot.

[e] The coordinate precision is defined as the average root mean square deviation of the 30 individual simulated annealing structures and their mean coordinates. The reported values are for residues 150 to 205 and 211 to 246. The backbone value refers to the N, C_{α}, and C' atoms.