Published in final edited form as:

Biochem Biophys Res Commun. 2010 July 30; 398(3): 342–349. doi:10.1016/j.bbrc.2010.06.027.

Role of a PAS sensor domain in the *Mycobacterium tuberculosis* transcription regulator Rv1364c

Ravi Kumar Jaiswal, G. Manjeera, and B. Gopal*

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Abstract

The *Mycobacterium tuberculosis* transcriptional regulator Rv1364c regulates the activity of the stress response σ factor σ^F . This multi-domain protein has several components: a signaling PAS domain and an effector segment comprising of a phosphatase, a kinase and an anti-anti- σ factor domain. Based on Small Angle X-Ray Scattering (SAXS) data, Rv1364c was recently shown to be a homo-dimer and adopt an elongated conformation in solution. The PAS domain could not be modeled into the structural envelope due to poor sequence similarity with known PAS proteins. The crystal structure of the PAS domain described here provides a structural basis for the dimerization of Rv1364c. It thus appears likely that the PAS domain regulates the anti- σ activity of Rv1364c by oligomerization. A structural comparison with other characterized PAS domains reveal several sequence and conformational features that could facilitate ligand binding- a feature which suggests that the function of Rv1364c could potentially be governed by specific cellular signals or metabolic cues.

Keywords

PAS; environmental sensor; transcriptional regulator; dimerization; Rv1364c

Introduction

Mycobacterium tuberculosis bacilli survive for extended periods in the host due to an effective molecular mechanism that relates environmental conditions with changes in the transcription profile of this organism. Sigma (σ) factors, in particular σ^F , have been shown to be essential for virulence and tuberculosis infection [1]. σ^F was recently shown to be regulated by a signaling cascade that is similar to the network that governs the general stress response regulator σ^B in *Bacillus subtilis* [2]. These two systems differ in one key aspect- several regulatory components are merged in *M. tuberculosis* Rv1364c, with the result that a single protein incorporates several functional features [3]. This multi-domain protein Rv1364c has an RsbU-like domain with a phosphatase fused to a signaling PAS module, an RsbW-like domain (RsbT) that is a Ser/Thr kinase and an RsbS-like (STAS) domain at the C-terminus (Figure 1A). Rv1364c is a homo-dimer in solution in both unphosphorylated and phosphorylated forms; a difference in the radius of gyration (R_G) between these two forms suggests that changes in inter-domain arrangements are likely upon phosphorylation [4]. Here we describe the molecular characterization of Rv1364c, with a particular focus on the role of the PAS domain.

*Corresponding Author: B. Gopal, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012 Phone : (+91) 80 2293 3219 Fax: (+91) 80 2360 0535 bgopal@mbu.iisc.ernet.in.

The PAS (Per-Arnt-Sim) domain is one of the most versatile signaling domains in nature. The annotation for the PAS domain was derived from the *Drosophila* period clock protein (PER), vertebrate aryl hydrocarbon receptor nuclear transporter (ARNT), and *Drosophila* single minded protein (SIM). PAS domains function mostly as sensor domains and can respond to a variety of environmental stimuli. The sensory mechanism of PAS domains varies substantially; known mechanisms include the ability of PAS domains to bind a variety of small molecules (Heme, FAD, FMN, hydroxycinnamic acid) [5] or changes in protein-protein interactions in large multi-domain proteins (K⁺ channel encoded by HERG) [6]. In prokaryotes, these domains sense redox potentials, oxygen concentration or luminescence levels [7]. In eukaryotes, PAS domains regulate serine threonine kinase activity or act as a transcription factor. Ligand binding is not a necessary characteristic for the function of PAS domains- the functional characteristics are often related to the oligomerization of this protein, perhaps associated with environmental cues [5]. The structure of the PAS domain reported here provides a rationale to reconcile the role of this domain with the ability of Rv1364c to dimerize and regulate σ^F . The structure also provides a basis to rationalize the sequence-conformational features of a PAS domain that are optimized to recognize specific environmental or metabolic cues.

Material and Methods

Cloning, expression and purification of Rv1364c and its constituent domains

Expression constructs of Rv1364c were designed to obtain recombinant proteins having residues 1-163, 190-397 and 1-653 corresponding to the PAS domain, phosphatase domain and full length Rv1364c. A summary of the different expression constructs examined in the course of this study is compiled in Table 1.

The plasmid encoding the PAS domain was transformed into BL21(DE3) competent cells (Novagen Inc.). These cells were grown in Luria broth with ampicillin (100 μ g/ml) till the cell density reached an O.D. of 0.5-0.6 at 600 nm. 0.3 mM IPTG was used for induction and post-induction, the temperature was lowered to 23 °C for 10-12 hrs. Cells were spun down at 4 °C and the pellet was re-suspended in buffer A (50mM Tris HCl, 300mM NaCl, pH 8.0) and sonicated on ice. The cell free lysate was incubated with Ni-NTA resin (Invitrogen, Inc.). Elution of the bound protein was achieved by a gradient of buffer B (50mM Tris HCl, 300mM NaCl, 200mM Imidazole, pH 8.0). This protein was further purified by size exclusion chromatography on a Sephacryl Hiprep 16/60 S-200 column (Amersham Biosciences) in buffer A.

Crystallization and structure determination of the PAS domain

Crystals of the PAS domain were obtained by the hanging drop vapor diffusion method in two conditions at a concentration of 12 mg/ml. Condition 1 (0.3M MgCl₂.6H₂O, and 30% w/v PEG 3350, pH 7.0) had crystals in the C2 space group that diffracted up to 2.5 Å resolution. This data-set however revealed pronounced pseudo-translational symmetry and was not pursued further. In the second condition (2.5M (NH₄)₂SO₄, 5% isopropanol, 0.1M Tris-HCl, pH8.0) the SeMet derivatized protein crystallized under oil within a month of setting up the crystallization trays. These crystals were cryo-protected with 10% glycerol and were flash frozen in liquid propane.

Diffraction data upto 2.3 Å resolution were collected at 0.9786 Å at the BM-14 beam line of the ESRF, Grenoble. The data were integrated and scaled using iMosflm [8] and SCALA [9]. Initial phases were obtained using the AutoSol wizard of Phenix [10] with the data collected at the peak wavelength used as a Single wavelength Anomalous Dispersion (SAD) dataset. Subsequently, three iterations of the AutoBuild module yielded the initial model of

the PAS domain. Restrained refinement was performed using four TLS groups as obtained from the TLSMD server [11] and the fit of the model to the electron density was evaluated using COOT [12]. The geometry of the refined model was analyzed using the Molprobit module of Phenix [10].

Phosphatase Assays

Activity measurements were performed using paranitrophenol phosphate (pNPP) as substrate. The reaction buffer contained 20 mM Tris-HCl, 2 mM MnCl₂, 5 mM DTT, pH 8.0. The substrate concentration was varied from 1mM to 70 mM of pNPP in a final volume of 200 μ L. The inorganic phosphate (inhibitor) concentration used was 0.8 mM, 1.2 mM and 2.6 mM. These reactions were initiated by adding 20 μ L of the purified enzyme followed by incubation for 30 min at 37 °C. The reaction was stopped by the addition of 200 μ L of 0.5 M EDTA. Catalysis was monitored by measuring the absorbance at 405 nm and was quantified using an extinction coefficient of 16000 M⁻¹ cm⁻¹. The kinetic constants were calculated by measuring initial velocities of the dephosphorylation reaction at varying initial substrate concentrations. The data were analyzed by nonlinear regression hyperbolic fitting to a classical Michaelis-Menten model using the Sigma-Plot software (SystatSoftware).

Thermo-stability Measurements

The PAS domain, phosphatase domain and the PAS-phosphatase fusion protein were examined for their relative stability. Thermo-stability was monitored by recording the scattering at 400 nm upon increase in temperature. These experiments were performed in 10mM HEPES buffer at pH 7.5 containing 200mM NaCl in a quartz cuvette of 1cm path length. The concentration of all three protein samples was maintained at 5 μ M each. The temperature was varied from 25-80 °C at the rate of 1 °C per minute.

Simulation to analyze the dynamics of the PAS domain

The procedure for the CONCOORD (from CONstraints to COORDinates) simulation [13] was adapted from a similar study on four other PAS domains reported earlier (Protein Data Bank codes: 1DRM, 1BYW, 1G28 and 2PHY) [14]. In this analysis, 1000 structures were generated for the PAS domains. The structures generated by CONCOORD were input to the GROMACS software [15] to calculate the root mean square fluctuation (RMSF).

Structure based Phylogenetic analysis of the Rv1364c PAS domain

For the phylogenetic analysis, known PAS domains were collated and an all-against-all pairwise comparison of these domains was performed using DALI. From the pairwise alignment, a distance matrix was computed using the Z-scores. The dendrogram was generated using Kitsch, a distance based algorithm from the PHYLIP suite of programs [16].

Results

Crystal Structure of the PAS domain

The crystal structure of the PAS domain of Rv1364c was solved at a resolution of 2.3 Å by the Singlewavelength Anomalous Dispersion (SAD) technique with a Se-Methionine labeled protein (Table 2; PDB code 3KX0). This structure contains residues 5- 135 of the PAS domain. Despite poor sequence similarity, the structure of the PAS domain of Rv1364c is similar to that of the known PAS proteins, with five β - strands oriented in an anti-parallel fashion to form a β -sheet forming the core of this domain [5]. This PAS core is flanked by three α helices- α 1 (residues 4-11) and α 2 (residues 14-23) at the N-terminus and α 3 (131-137) at the C-terminal end of the domain (Figure 1C).

The PAS domain of Rv1364c forms a dimer (Figure 1D; Figure 2C). The mode of dimerization is similar to that of *Escherichia coli* DosH [17], *Sinorhizobium meliloti* FixL [18] and *Azotobacter vinelandii* NifL [19]. Dimerization involves extensive H-bonds involving residues from the helix $\alpha 1$ of both monomers. The dimer is further stabilized by salt bridges involving Arg 18 and Arg 19 of helix $\alpha 2$ with Asp 8 of helix $\alpha 1$ and Glu 121 located in $\beta 5$. The role of the $\alpha 1$ helix in dimerization is evident from the case of the *Bradyrhizobium japonicum* FixL where the N-terminal helix is not involved in subunit contacts resulting in a monomer [20]. The dimeric arrangement of the Rv1364c PAS domain appears stronger than that of *E. coli* DosH, *S. meliloti* FixL or *A. vinelandii* NifL which show less number of H-bonds and no salt bridges. The buried surface area in the case of PAS domains of *E. coli*, *A. vinelandii* and *S. meliloti* are substantially lower than that of the PAS domain of Rv1364c (Supplementary Table1).

The PAS domain as a regulator of Rv1364c

PAS domains, in general, are noted to regulate the activity of the domains that are fused at the C-terminus. This regulation is achieved either by ligand binding as in the case of *E. coli* DosH or via direct protein-protein interactions as in the case of the heterodimer between the Hypoxia inducible factor (HIF2 α) and the aryl hydrocarbon receptor nuclear translocator (ARNT) [21]. In an effort to examine if the PAS domain regulates phosphatase activity by interacting with the phosphatase domain, Surface Plasmon Resonance (SPR) experiments were performed using purified recombinant PAS and phosphatase domains (the RsbU component) of Rv1364c. SPR experiments suggest that these two domains do not interact with each other (data not shown). The activity assays for the phosphatase domain were performed using a generic substrate, para-nitrophenol phosphate (pNPP). The catalytic parameters corresponding to the phosphatase activity of the RsbU domain are compiled in Table 3. The Vmax for the phosphatase activity is similar to that reported for Rv1364c [4]. The Km of the PAS-phosphatase fusion construct is closer to that of the full length Rv1364c. The difference in the Km value between the phosphatase domain alone and the other constructs perhaps reflects the effect of the dimeric nature of the full length protein as well as the phosphatase domain tethered to the dimeric PAS domain. As the catalytic efficiency of the phosphatase domain was rather low, we further examined the effect of inorganic phosphate, Pi, on phosphatase activity (Figure 2A and 2B). The inhibition constant (0.32 μ M) was found closer to the reported K_i for *E. coli* alkaline phosphatases that vary from 1 μ M and 0.6 μ M [22] [23]. To evaluate the structural contribution of the PAS domain to the PAS-phosphatase fusion, the stability of the individual domains was examined vis-à-vis the fusion construct. The thermostability of the domains was monitored by measuring the scattering at 400 nm as a function of temperature. The aggregation profile of the PAS domain and phosphatase domain revealed a T_{agg} of 48 °C and 54 °C whereas the PAS-phosphatase fusion protein had a T_{agg} of 52 °C. We also note that while the individual domains start to aggregate much earlier (42 °C), the fusion protein is relatively unaffected till ca 50 °C. These data suggest that the PAS domain could influence the stability of the fusion construct (Figure 3).

The role of conformational dynamics in the ability of the PAS domain to influence the activity of the adjacent domains was examined using a procedure outlined earlier [13, 14]. In their analyses, referred to as CONCOORD, Vreede *et al* suggested a link between the conformational dynamics of the PAS domain to its functional role as a receptor of different stimuli. The PAS domain of Rv1364c shows a similar flexibility profile to other PAS domains despite poor sequence similarity. It thus appears likely that the PAS domain of Rv1364c may have the same mode of signal transduction as in the other characterized PAS domains. (Supplementary figure 1).

Sequence and Conformational determinants for ligand binding

PAS domains interact with a variety of ligands and metabolites (Figure 4). For example, FMN is a cofactor that has been reported to play a crucial role in PAS domains that respond to light stimuli [24]. An analysis of PAS structures with a bound FMN cofactor revealed that the residues interacting with FMN are conserved in all characterized structures. This observation prompted us to examine these conserved residues in other PAS domains whose functions are yet to be assigned. A search using the sequence of the *Chlamydomonas reinhardtii* PAS domain revealed about 60 sequences with an identity between 30 to 70%. A multiple sequence alignment analysis revealed a motif NCRFLQG located on the $\alpha 3$ helix of the PAS domain (Figure 4A). In the case of the FAD binding PAS domains, a similar analysis (based on 2GJ3 and 3EWK) suggested that two residues from $\alpha 4$ and two hydrophobic residues (either Y or F and W) are conserved-these interact with the ring of the FAD cofactor (Figure 4B). In heme binding PAS (examined using 1D06, 1XJ3 and 1V9Z), two residues in $\alpha 4$, a histidine and a tyrosine, interact with the heme cofactor (Figure 4C). A conserved cysteine residue, on the other hand, plays a crucial role in the hydroxycinnamic acid binding to PAS domains (analysis based on 1MZU and 1NWZ) [25; 26] (Figure 4D). The sequence features corresponding to specific ligand interaction were pictorially depicted using the MEME server [27] and is shown in Supplementary Figure 2. Apart from the sequence characteristics, we also note distinct differences in the cavity volumes of the ligand binding pockets [28]. This data, perhaps of interest from a bioinformatics perspective, are compiled in supplementary Table 2.

The observation that sequence features in the PAS domain of Rv1364c and similar proteins was distinct from that of the other characterized ligand binding motifs led us to a structure-based phylogenetic analysis (Figure 1B, Supplementary Figure 3). We note that the PAS domain of Rv1364c clusters along with the sequences of the *Rhodococcus jostii* and *Haloarcula marismortui* PAS domains (structures 3FG8 and 3FC7). One of these structures (3FG8) also had a bound- ligand, 3-phosphonoxy-butanoic acid, similar to the isopropanol modeled in the structure of the Rv1364c PAS domain (Figure 1C). The ligands, in both cases, appear to be from the crystallization condition.

Discussion

The protein Rv1364c interacts with both σ^F as well as the anti- σ factor for σ^F , RsbW, in a yeast two-hybrid experiment [3]. Up-regulation of Rv1364c transcripts was also noted in *M. bovis* BCG during intra-macrophage survival and during starvation in *M. tuberculosis* [29]. The presence of a N-terminal sensory domain (PAS) followed by a phosphatase domain, a kinase or an anti- σ factor domain (RsbW) and a C-terminal anti- σ factor domain (RsbS) suggested that Rv1364c is likely to be a fused equivalent of the RsbP–RsbV–RsbW cluster of *B. subtilis*. In *B. subtilis* RsbP, a PAS domain senses energy stress in the cell and dephosphorylates RsbV which, in turn, binds RsbW and releases the σ factor σ^B to activate downstream processes [3]. The similarities between this cascade and the multi-domain organization of *M. tuberculosis* Rv1364c led to the hypothesis that Rv1364c could play a role in signal transduction followed by stress and relay energy stress signals to σ^F . σ^F governs several transcriptional responses in mycobacteria [30]. In addition to Rv1364c, σ^F is also controlled by its association with RsbW (Rv3286c), an anti σ factor, which is present upstream to σ^F in the operon [31; 32].

In Rv1364c, the anti- σ factor kinase domain has been demonstrated to influence the activity of the phosphatase domain [4]. Here we note that the PAS domain in Rv1364c alters the K_m corresponding to the phosphatase activity, a finding that is similar to *Streptococcus pneumoniae* WalK(Spn) where the deletion constructs of PAS showed diminished phosphatase activity [33]. The solution structural envelope derived from Small Angle X-Ray

Scattering (SAXS) experiments suggested that Rv1364c is an elongated dimer. The SAXS analysis further revealed a rearrangement between the domains of Rv1364c upon phosphorylation. This finding was consistent with the hypothesis that inter-domain communication was necessary to regulate the activity of this protein. Two reports on the activity of the RsbT domain differ in their inference of kinase activity- while Sachdeva *et al* [2] suggest that this domain is inactive, Greenstein and colleagues showed kinase activity [4]. Furthermore, it was noted that neither the PAS nor the STAS (RsbS) domains affect RsbU phosphatase activity [4]. The chimeric protein containing the phosphatase and RsbT (kinase domain) protein was a dimer in solution. In this study, the PAS domain of Rv1364c could not be modeled in the SAXS derived structural envelope due to its low sequence identity with other known PAS domains [4]. The dimeric arrangement of the PAS domain thus has a mechanistic interpretation for the activity of Rv1364c. It appears likely that the PAS domain localizes the phosphatase domain (which is monomeric in solution) between the other dimeric RsbT (kinase) domain (Figure 1E). An observation, perhaps of interest from a structural perspective, is that the dimeric arrangement of the PAS domain differs from other PAS domains, notably *B.subtilis* YtvA [34]. The N-terminus helix $\alpha 2$ in the PAS domain is amphipathic and packs against the hydrophobic surface of the β sheet of the opposite monomer whereas the YtvA PAS oligomerizes by hydrophobic anti-parallel β sheets packing against each other. The identification of a hydrophobic pocket in the PAS domain of Rv1364c alongside a pattern of sequence conservation suggests that this domain is likely to interact with specific ligands. Indeed, structure based phylogeny clusters the PAS domain of Rv1364c with other structures with a similar hydrophobic pocket (PDB IDs 3FC7 and 3FG8).

To summarize, the crystal structure of the PAS domain of Rv1364c reveals a tight dimer. This finding provides a basis to rationalize the results of the SAXS analysis that suggests that Rv1364c is a dimer in solution. The PAS domain of Rv1364c thus provides an anchor for the dimeric arrangement of Rv1364c allowing for structural rearrangements between the kinase and phosphatase domains during the course of the activity of this protein.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- [1]. Gebhard S, Humpel A, McLellan AD, Cook GM. The alternative sigma factor SigF of *Mycobacterium smegmatis* is required for survival of heat shock, acidic pH and oxidative stress. *Microbiology*. 2008; 154:2786–2795. [PubMed: 18757812]
- [2]. Sachdeva P, Narayan A, Misra R, Brahmachari V, Singh Y. Loss of kinase activity in *Mycobacterium tuberculosis* multidomain protein Rv1364c. *FEBS J*. 2008; 275:6295–6308. [PubMed: 19016841]
- [3]. Parida BK, Douglas T, Nino C, Dhandayuthapani S. Interactions of anti-sigma factor antagonists of *Mycobacterium tuberculosis* in the yeast two-hybrid system. *Tuberculosis*. 2005; 85:347–355. [PubMed: 16263329]
- [4]. Greenstein AE, Hammel M, Cavazos A, Alber T. Interdomain communication in the *Mycobacterium tuberculosis* environmental phosphatase Rv1364c. *J Biol Chem*. 2009; 284:29828–29835. [PubMed: 19700407]
- [5]. Moglich A, Ayers RA, Moffat K. Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure*. 2009; 17:1282–1294. [PubMed: 19836329]
- [6]. Zhulin IB, Taylor BL, Dixon R. PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox. *Trends Biochem Sci*. 1997; 22:331–333. [PubMed: 9301332]
- [7]. Taylor BL, Zhulin IB. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev*. 1999; 63:479–506. [PubMed: 10357859]

- [8]. Leslie AGW. Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography. 1992; 26
- [9]. The CCP4 suite: programs for protein crystallography. Acta Cryst. D. 1994; 50:760–763. [PubMed: 15299374]
- [10]. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. Acta Cryst. D. 2002; 58:1948–1954. [PubMed: 12393927]
- [11]. Painter J, Merritt EA. *TLSMD* web server for the generation of multi-group TLS models. J. Appl. Cryst. 2006; 39:109–111.
- [12]. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Cryst D. 2004; 60:2126–2132. [PubMed: 15572765]
- [13]. de Groot BL, van Aalten DMF, Scheek RM, Amadei A, Vriend G, Berendsen HJC. Prediction of protein conformational freedom from distance constraints. PROTEINS: Structure, Function, and Genetics. 1997; 29:240–251.
- [14]. Vreede J, van der Horst MA, Hellingwerf KJ, Crielaard W, van. Aalten DMF. PAS domains: Common structure and common flexibility. J Biol Chem. 2003; 278:18434–18439. [PubMed: 12639952]
- [15]. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model. 2001; 7:306–317.
- [16]. Felsenstein, J. PHYLIP (Phylogeny Inference Package). version 3.6. Department of Genome Sciences, University of Washington; Seattle: 2005. Distributed by the author
- [17]. Park H, Suquet C, Satterlee JD, Kang C. Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of Escherichia coli Dos heme domain (Ec DosH). Biochemistry. 2004; 43:2738–2746. [PubMed: 15005609]
- [18]. Miyatake H, Mukai M, Park SY, Adachi S, Tamura K, Nakamura H, Nakamura K, Tsuchiya T, Iizuka T, Shiro Y. Sensory mechanism of oxygen sensor FixL from Rhizobium meliloti: crystallographic, mutagenesis and resonance Raman spectroscopic studies. J Mol Biol. 2000; 301:415–431. [PubMed: 10926518]
- [19]. Key J, Hefti M, Purcell EB, Moffat K. Structure of the redox sensor domain of Azotobacter vinelandii NifL at atomic resolution: signaling, dimerization, and mechanism. Biochemistry. 2007; 46:3614–3623. [PubMed: 17319691]
- [20]. Gong W, Hao B, Mansy SS, Gonzalez G, Gilles-Gonzalez MA, Chan MK. Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. Proc Natl Acad Sci U S A. 1998; 95:15177–15182. [PubMed: 9860942]
- [21]. Scheuermann TH, Tomchick DR, Machius M, Guo Y, Bruick RK, Gardner KH. Artificial ligand binding within the HIF2alpha PAS-B domain of the HIF2 transcription factor. Proc Natl Acad Sci U S A. 2009; 106:450–455. [PubMed: 19129502]
- [22]. O'Brien PJ, Herschlag D. Functional interrelationships in the alkaline phosphatase superfamily: phosphodiesterase activity of Escherichia coli alkaline phosphatase. Biochemistry. 2001; 40:5691–5699. [PubMed: 11341834]
- [23]. McComb, RB.; Bowser, GN., Jr; Posen, S. Alkaline Phosphatase. Plenum Press; New York: 1979.
- [24]. Fedorov R, Schlichting I, Hartmann E, Domratcheva T, Fuhrmann M, Hegemann P. Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii. Biophys J. 2003; 84:2474–2482. [PubMed: 12668455]
- [25]. Rajagopal S, Moffat K. Crystal structure of a photoactive yellow protein from a sensor histidine kinase: conformational variability and signal transduction. Proc Natl Acad Sci U S A. 2003; 100:1649–1654. [PubMed: 12563032]
- [26]. Getzoff ED, Gutwin KN, Genick UK. Anticipatory active-site motions and chromophore distortion prime photoreceptor PYP for light activation. Nat Struct Biol. 2003; 10:663–668. [PubMed: 12872160]

- [27]. Timothy, CE.; Bailey, L. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology; Menlo Park, California: AAAI Press; 1994. p. 28-36.
- [28]. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* 2006; 34:116–118.
- [29]. Li MS, Waddell SJ, Monahan IM, Mangan JA, Martin SL, Everett MJ, Butcher PD. Increased transcription of a potential sigma factor regulatory gene Rv1364c in *Mycobacterium bovis* BCG while residing in macrophages indicates use of alternative promoters. *FEMS Microbiol Lett.* 2004; 233:333–339. [PubMed: 15063504]
- [30]. Chen P, Ruiz RE, Li Q, Silver RF, Bishai WR. Construction and characterization of a *Mycobacterium tuberculosis* mutant lacking the alternate sigma factor gene, sigF. *Infect Immun.* 2000; 68:5575–5580. [PubMed: 10992456]
- [31]. Beaucher J, Rodrigue S, Jacques PE, Smith I, Brzezinski R, Gaudreau L. Novel *Mycobacterium tuberculosis* anti-sigma factor antagonists control sigmaF activity by distinct mechanisms. *Mol Microbiol.* 2002; 45:1527–1540. [PubMed: 12354223]
- [32]. Malik SS, Luthra A, Srivastava SK, Ramachandran R. *Mycobacterium tuberculosis* UsfX (Rv3287c) exhibits novel nucleotide binding and hydrolysis properties. *Biochem Biophys Res Commun.* 2008; 375:465–470. [PubMed: 18722345]
- [33]. Gutu AD, Wayne KJ, Sham LT, Winkler ME. Kinetic characterization of the WalRKSpn (VicRK) two-component system of *Streptococcus pneumoniae*: dependence of WalKSpn (VicK) phosphatase activity on its PAS domain. *J Bacteriol.* 2010; 192:2346–2358. [PubMed: 20190050]
- [34]. Moglich A, Moffat K. Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *J Mol Biol.* 2007; 373:112–126. [PubMed: 17764689]
- [35]. Maeda Y, Koga H, Yamada H, Ueda T, Imoto T. Effective renaturation of reduced lysozyme by gentle removal of urea. *Protein Engineering.* 1995; 8:201–205. [PubMed: 7630889]

Figure 1. Sequence and Structural Features of the PAS domain of Rv1364c

A. The PAS domain of Rv1364c is a part of the RsbU module. This domain precedes the Ser/Thr phosphatase domain. **B.** A multiple sequence alignment of the PAS domain reveals a pattern of conserved residues in helices E and F. **C.** The crystal structure of the PAS domain reveals a conserved hydrophobic pocket in this protein. Isopropanol (from the crystallization condition) could be modeled into the electron density. **D.** The PAS dimer interface is fairly extensive, involving a substantial buried surface area. **E.** The dimeric nature of the PAS domain suggests that the phosphatase domain in Rv1364c is packed between two dimeric components- the PAS domain at the N-terminus and the anti- σ kinase domain at the C-terminus. This dimeric model is consistent with the SAXS data reported earlier [4].

Figure 2. A. Phosphatase activity of full-length Rv1364c

Please refer to the methods for details; the catalytic parameters are compiled in Table 3. **B.** Catalytic activity of the phosphatase domain of Rv1364c. **C.** The PAS domain of Rv1364c is a dimer in solution.

Figure 3. Thermo-stability of the PAS, the RsbU-like phosphatase domain and the PAS-phosphatase fusion protein construct

The relative changes in the aggregation profiles of the fusion protein construct vis-à-vis the individual domains suggest that the PAS domain influences the stability of the PAS-phosphatase fusion protein.

Figure 4. Sequence and structural determinants of ligand binding to PAS domains

Ligand binding features are evident based on a multiple sequence alignment. **A.** Ribbons representation of a PAS domain (1N9L) bound to FMN (left) and the multiple sequence alignment. **B.** A PAS domain (2GJ3) bound to FAD (left) and the related multiple sequence alignment. **C.** Ribbons representation of a PAS domain (1XJ3) bound to heme (left) and the corresponding multiple sequence alignment. **D.** Ribbons representation of a PAS domain (1MZU) bound to hydroxycinnamic acid (left) and the corresponding multiple sequence alignment (right).

Table 1
List of Expression constructs examined in this study

S.No.	Domain name	Length (Rv1364 sequence)	Expression Vector	Putative/Known functions	Soluble Protein/domain
1	PAS	1-163	pET15b	Sensor Domain	Yes
2	RsbU-like (phosphatase)	190- 397	pET15b, pGEX4T1	Phosphatase domain	No ^a
3	RsbW-like	404-549	pET15b pET22b	Kinase domain	No
4	STAS	551-638	pET15b, pET22b, pET28a	Anti-anti sigma	Yes
5	Rv1364c	1-653	pET15b	Full length protein	Yes
6	PAS+phosphatase	1-397	pET15b pGEX4T1		No ^a
7	RsbU +RsbW-like+STAS	190-638	pET15b		No
8	Phosphatase +RsbW like	190-549	pET15b		No
9	RsbW like+STAS	404-638	pET15b		No

^a , The recombinant protein was solubilized from inclusion bodies [35].

pET15b, pET22b, pET28a (Novagen, Inc)

pGEX4T1 (GE Lifesciences)

Table 2
Data, phasing and refinement statistics

I. Data		
Wavelength (Å)	0.9536	0.9786
Resolution(Å)	57.74-2.3 (2.42-2.30) ^a	86.56-2.3 (2.42-2.3)
Space group	P6 ₂ 22	
Unit cell parameter(Å)	a = b = 59.60, c = 173.32	
Total number of reflections	149942 (21601)	139551 (13396)
Number of unique reflections	8808 (1220)	8732 (1168)
Completeness (%)	100 (100)	99.5 (96.9)
Multiplicity	17 (17.7)	16 (11.5)
R _{merge} (%)	9.7 (57.3)	9.9 (59.9)
<I>/s(I)	21.2 (5.2)	23.2 (3.8)
II. Phasing		
	SAD (Se)	FOM
	86.56 - 11.90	0.32
	11.90 - 7.46	0.28
	7.46 - 5.81	0.34
	5.81 - 4.92	0.36
	4.92 - 4.34	0.38
	4.34 - 3.93	0.34
	3.93 - 3.61	0.31
	Mean	0.33
III. Refinement		
R _{cryst} (%)	23.1	
R _{free} (%)	27.6	
Rms _{bond} (Å)	0.007	
Rms _{angle} (degree)	1.07	

^aOuter resolution cell

R_{merge} = $\sum \sum |I(h) - \langle I(h) \rangle| / \sum \langle I(h) \rangle$, where $\langle I(h) \rangle$ is the mean intensity after rejections.

R_{cryst} = $\sum |F_p - F_c| / \sum |F_p|$; R_{free}, the same as R_{cryst} but calculated on 5% of data excluded from refinement. FOM is figure of merit.

Table 3
Catalytic parameters of the full length Rv1364c and truncated protein constructs

Protein	V _{max} ($\mu\text{mol}/\text{min}/\text{mg}$)	K _m (mM)	K _{cat} (Min^{-1})
Rv1364c	0.09 ± 0.003	3.61 ± 0.58	7.02 ± 0.4
PAS+phosphatase fusion protein	0.07 ± 0.002	2.54 ± 0.35	3.25 ± 0.1
RsbU phosphatase domain	0.07 ± 0.004	7.47 ± 1.22	1.58 ± 0.1