Molecular structure of EmbR, a response element of Ser/Thr kinase signaling in *Mycobacterium tuberculosis*

Luke J. Alderwick*, Virginie Molle†, Laurent Kremer‡, Alain J. Cozzone†, Timothy R. Dafforn*, Gurdyal S. Besra*, and Klaus Fütterer*§

*School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom; †Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5086, Institut de Biologie et Chimie des Protéines, 7 Passage du Vercors, 69367 Lyon Cedex 07, France; and ‡Laboratoire de Dynamique Moléculaire des Interactions Membranaires, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5539, Université Montpellier II, Case 107, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France

Edited by Gregory A. Petsko, Brandeis University, Waltham, MA, and approved December 20, 2005 (received for review September 6, 2005)

Ser/Thr phosphorylation has emerged as a critical regulatory mechanism in a number of bacteria, including *Mycobacterium tuberculosis*. This problematic pathogen encodes 11 eukaryotic-like Ser/Thr kinases, yet few substrates or signaling targets have been characterized. Here, we report the structure of EmbR (2.0 Å), a putative transcriptional regulator of key arabinosyltransferases (EmbC, -A, and -B), and an endogenous substrate of the Ser/Thr-kinase PknH. EmbR presents a unique domain architecture: the N-terminal winged-helix DNA-binding domain forms an extensive interface with the all-helical central bacterial transcriptional activation domain and is positioned adjacent to the regulatory C-terminal forkhead-associated (FHA) domain, which mediates binding to a Thr-phosphorylated site in PknH. The structure in complex with a phospho-peptide (1.9 Å) reveals a conserved mode of phospho-threonine recognition by the FHA domain and evidence for specific recognition of the cognate kinase. The present structures suggest hypotheses as to how EmbR might propagate the phospho-relay signal from its cognate kinase, while serving as a template for the structurally uncharacterized *Streptomyces* antibiotic regulatory protein family of transcription factors.

PknH | x-ray crystallography | arabinosyltransferase | transcriptional regulator | ethambutol

Transient protein phosphorylation is a ubiquitous mechanism of cellular regulation in eukaryotic organisms. The identification of eukaryotic-like Ser/Thr kinases in bacterial genomes has firmly established that Ser/Thr phosphorylation-dependent pathways exist in at least a subset of prokarya (1). The genome of *Mycobacterium tuberculosis* (*Mtb*) comprises 11 genes similar to eukaryotic Ser/Thr kinases (2, 3). Nine of these kinases are predicted to be membrane-embedded, with a sensor domain at the extracellular face and a cytoplasmic catalytic domain (3) and are showing strong autophosphorylation and phosphorylation of exogenous substrates (references 10–18 in ref. 4). Experimental reports have suggested functional roles in cell morphology (PknA and -B) (5, 6), stress response (PknH) (7), glucose transport (PknF) (8, 9), and regulation of cellular Glu/Gln levels (PknG) (10). However, endogenous substrates could be identified in only a few cases (6, 8, 11, 12). Signaling pathways and mechanisms are largely uncharted territory.

A recently characterized endogenous substrate of the sensor-like Ser/Thr kinase PknH (Rv1268c) is the putative transcriptional regulator EmbR (Rv1267c) (11). EmbR has originally been cloned from *Mycobacterium avium* in the context of identifying genes that confer ethambutol resistance (13). Ethambutol, a front-line antitubercular drug, targets a set of membrane-embedded arabinosyltransferases (*M. avium* EmbA and -B and *Mtb* EmbC, -A, and -B) that are involved in arabinogalactan and lipoarabinomannan biosynthesis, a critical component of the mycobacterial cell wall (13–17). In *M. avium*, the *embR* gene is located upstream of *embA* and *-B*, leading to the hypothesis that *embR* might control expression of the Emb arabinosyltransferases. EmbR is related to the *Streptomyces* antibiotic regulatory proteins (SARP) family of transcription factors (18); (see Fig. 6, which is published as supporting information on the PNAS web site). Members of this family share an N-terminal OmpR-like DNA-binding domain, followed by a bacterial transcriptional activation (BTA) domain, whose function or fold are unknown, and display a diverse domain organization at the C terminus (Fig. 1). Among the SARP-family proteins, EmbR and several orthologues are unique, in that they comprise a C-terminal forkhead-associated (FHA) domain, a module that mediates binding to threonine-phosphorylated sites in a sequence-specific manner. Often implicated in DNA damage repair and cell cycle control, the structure and function of eukaryotic FHA domains have been studied extensively (19). Recent work by Molle *et al.* (11) showed that the FHA domain is required for the interaction of EmbR with its cognate Ser/Thr kinase PknH. Truncation of this domain or point mutations of conserved residues in the phosphothreonine-binding pocket abrogated PknH-mediated phosphorylation of EmbR.

Thus, the data suggest that EmbR may control arabinosyltransferase activity in *Mtb* in a phosphorylation-dependent fashion, acting downstream of the Ser/Thr-kinase PknH. The structures of EmbR presented here provide a glimpse of the conserved core of SARP-family proteins and lead to hypotheses as to how the FHA domain might regulate DNA-binding of EmbR.

Results

Domain Architecture. The structure of *Mtb* EmbR was determined by multiwavelength anomalous dispersion (MAD) and refined to 2.0-Å resolution (see Table 1 and Fig. 7, which are published as supporting information on the PNAS web site). The asymmetric unit comprises one molecule, and the experimental electron density encompasses residues 10–382. EmbR splits into three distinct domains: the N-terminal OmpR/PhoB-like DNA-binding domain (residues 10–106), the central all-helical BTA domain (107–284) and the C-terminal FHA domain (285–382) (Fig. 2). Although the DNA-binding and FHA domains are adjacent, they form only few direct contacts, burying <40 Å^2 of

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: BTA, bacterial transcriptional activation; FHA, forkhead-associated; *Mtb*, *Myocbacterium tuberculosis*; pT, phospho-Thr; RMSD, root-mean-square deviation; SARP, *Streptomyces* antibiotic regulatory protein; TPR, tetratricopeptide repeat.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2FEZ and 2FF4).

§To whom correspondence should be addressed. E-mail: k.futterer@bham.ac.uk.

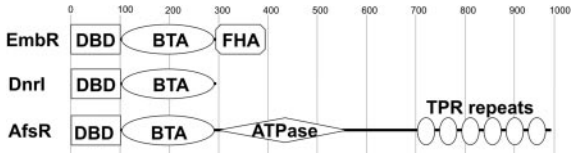

Fig. 1. Linear domain composition of SARP-family proteins. The SARP-family proteins are *S. peucetius* DnrI and *S. coelicolor* AfsR. DBD, DNA-binding domain.

solvent-accessible surface between them. In contrast, extensive interfaces are formed between the central BTA and the two terminal domains. The electrostatic surface potential (Fig. 3*A*) is highly basic around helix α3, consistent with the putative DNA-binding properties of EmbR, whereas the FHA domain is surprisingly acidic in character. A prominent L-shaped cavity, featuring a distinctly acidic surface potential, is enclosed by the FHA and the BTA domains, with dimensions of $\approx$20 Å $\times$ 15 Å $\times$ 12 Å. We have no indication that this cavity is important for function.

DNA-Binding Domain. The winged-helix DNA-binding domain of EmbR is homologous to the DNA-binding domains of the two-component system response regulators of the OmpR/PhoB family (20). However, the order of the domains in the polypeptide chain is reversed, because the DNA-binding domain is C-terminal in OmpR. A central three-helix bundle flanked by two β-sheets characterizes the fold of this domain. The closest homologue of known structure is *Escherichia coli* OmpR (PDB entry 1OPC) (20), superimposing with a root-mean-square deviation (RMSD) of 1.33 Å for 61 aligned Cα positions. Structural elements relevant for function in OmpR are conserved in EmbR, including the transactivation loop (α2–α3), which mediates interactions with RNA polymerase, the DNA recognition-helix α3, and the "wing" (β-hairpin loop β6–β7). Missing from the canonical OmpR DNA-binding domain is merely the first strand of the N-terminal β-sheet. The superposition with the PhoB DNA-binding domain–*pho* box complex [1GXP (21); RMSD, 1.7 Å; 70 Cα positions; Fig. 2*B*], provides clues for how an individual EmbR protomer might bind to DNA. In the superposition, helix α3 penetrates the major groove, whereas the wing extends toward the minor groove. PhoB/OmpR proteins insert a large side chain, located at the tip of the wing (Trp or Arg) into the minor groove (21). The corresponding position in EmbR is an alanine, flanked by Ala and Pro, implying that interactions with the minor groove may occur in a different way. Nonetheless, residues mediating nonspecific DNA contacts in the PhoB–*pho* box complex are largely conserved in EmbR, whereas sites in the DNA-binding domain of the SARP-family protein DnrI, whose mutation abrogates DNA binding, are conserved between EmbR and DnrI (Fig. 6) (22). Thus, the present structure suggests a functionally competent EmbR DNA-binding domain, which may bind to the duplex strand in a fashion similar to PhoB/OmpR. However, the superposition with the PhoB–DNA complex also suggests substantial steric clashes between protomers, if EmbR were to bind as a dimer to a *pho*-box-type seven-nucleotide direct repeat, as has been suggested for SARP-family proteins (18).

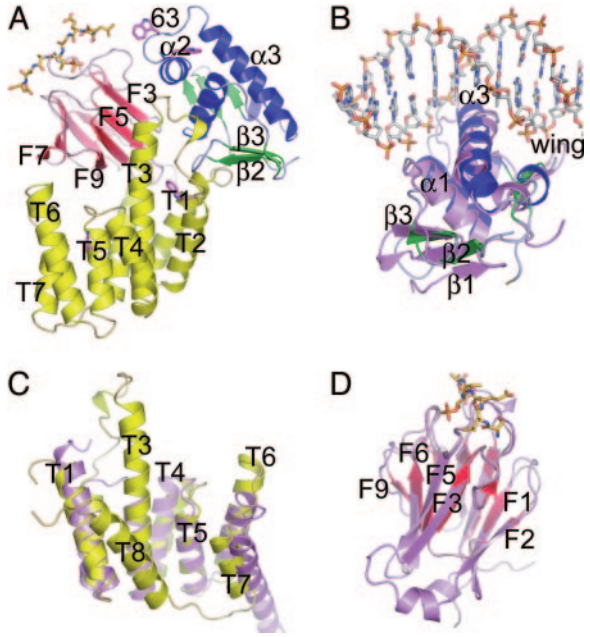

Fig. 2. Overall structure of EmbR and comparison of domains. (*A*) Ribbon diagram of phospho-peptide-bound EmbR. The DNA-binding domain is colored in blue (helices) and green (strands) and the BTA and FHA domains in yellow and red, respectively. Secondary structure elements are labeled according to the homologous domains in *B*–*D*. Trp residues are indicated as sticks in magenta, and Trp-63 is labeled. (*B*–*D*) Superposition of EmbR domains (colored as in *A*) with structural neighbors (shown in violet and determined by Dali) (24): DNA-binding (*B*), BTA (*C*), and FHA (*D*). The superimposed domains are the PhoB DNA-binding domain bound to *pho* box (1GXP) (21), the TPR domain of protein phosphatase pp5 (1A17) (23), and the FHA domain of Rad53p (1G6G) (28), respectively.

BTA Domain. Its fold previously unknown, the BTA domain resembles the tetratricopeptide repeat (TPR) domain of human protein phosphatase PP5 [1A17 (23); RMSD, 2.5 Å; 120 Cα positions] with a Z-score of 12.8 in a Dali search (www.ebi.ac.uk/dali), identifying structural neighbors in the PDB databank based on distance matrices (24) (Fig. 2*C*). TPRs are weakly conserved 34-amino acid repeats folding as pairs of antiparallel helices, with a preference for small hydrophobic residues at three positions of the repeat (25). Interestingly, the SARP-family protein AfsR is predicted to contain additional TPRs at its C terminus (Fig. 1), but it is unknown whether, together, these form a domain. The BTA domain of EmbR comprises three TPRs (helices T1–T6) plus the C-terminal helices T7 and T8. The C-terminal boundary of the domain is ambiguous. In the present structure, helix T8 forms an integral part of the BTA domain, folding back onto helices T1–T6. However, the sequence alignment of SARP-family proteins (Fig. 6) and the alignment of BTA domains in the Pfam database (PF03704) indicate that helix T8 is not conserved. Accordingly, the superposition with PP5 maps helices T1–T7 onto the seven helices in PP5 (Fig. 2*C*), whereas T8 lacks a counterpart. Yet, the orientation of helix T8 with respect to the three-TPR array is reminiscent of the peptide ligand in the structure of the adaptor protein Hop, whose TPR domains bind C-terminal motifs of the heat-shock proteins Hsp70 and Hsp90 (26). Helix T3 is strikingly elongated, protruding from the helical array and providing a major contact surface for the FHA domain (Figs. 2*C* and 3*B*). The preceding loop T2–T3 interfaces with the DNA-binding domain (Fig. 3*B*), and a flexible hairpin-like loop of six residues connects the BTA to the FHA domain.

FHA Domain. The FHA domain of EmbR displays the 11-stranded β-sandwich seen in eukaryotic FHA domains (Fig. 2*D*). The inner β-sheet (strands F3–F6 and F9) faces the DNA-binding domain and contacts helix T3 of the BTA domain (Fig. 3*B*). The outer β-sheet (strands F1, -2, -7, -8, -10, and -11) faces the solvent. The closest structural neighbor is the FHA domain of the cell-cycle-related human antigen Ki67 [1R21 (27); Z-score, 15; RMSD, 1.07 Å, 89 Cα positions]. A close structural rela-

BIOCHEMISTRY

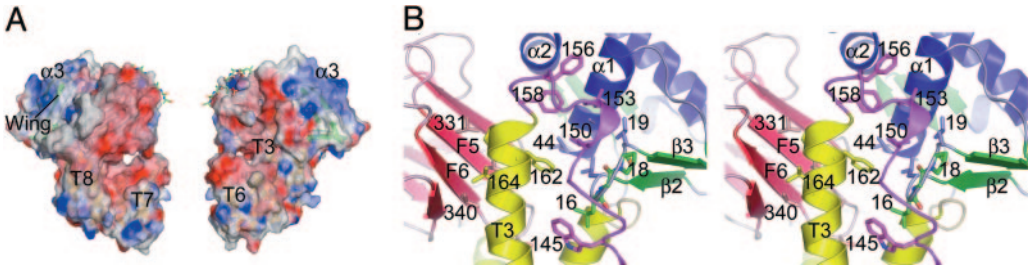

Fig. 3. Surface potential and domain interfaces. (*A*) ''Front'' and ''back'' view of the electrostatic surface of EmbR in the peptide-bound form. (*B*) Stereo view of interdomain contacts formed by the T2 to -3 loop (purple) and helix T3 (yellow). Selected side chains mentioned in the text are shown and labeled by sequence number. Ribbon colors are as in Fig. 2*A*.

tionship is also evident when superimposing EmbR with the N-terminal FHA domain of yeast Rad53p [1G6G (28), Fig. 2*D*] and that of human checkpoint kinase Chk2 [1GXC (29); RMSD, 1.1 Å; 80–86 Cα positions]. The EmbR FHA domain lacks the helical insertions displayed by some of the eukaryotic homologues, and all loop regions are short by comparison. The phosphopeptide-binding surface, formed by loops F3 to -4, F4 to -5, and F6 to -7, is separated from the DNA-binding site by ≈35 Å with respect to the phosphate position and the center of helix α3.

Domain Interfaces. The domain interfaces in EmbR bury ≈2,250 Å^2 of solvent-accessible surface. Considering that the DNA-binding and FHA domains are adjacent to each other, it is, perhaps, surprising that they share hardly any common surface (buried surface <40 Å^2). Approximately 1,400 Å^2 of solvent-accessible surface is buried between the DNA-binding and the BTA domain compared with 820 Å^2 between the FHA and BTA domains. Hydrophobic patches dominate the interfaces, in particular between the BTA and the DNA-binding domain (Fig. 3*B*). The T2- to -3 loop (residues 144–158) is well conserved across the sequence alignment of SARP-family proteins and makes contacts with residues in helices α1 and -2 and the β2 to -3 loop of the DNA-binding domain (Fig. 6). Aromatic residues (Trp-145 and Phe-156, -158, and -162) and two leucines (150 and 153) contribute a considerable fraction of contacts. For instance, the invariant Trp-145 interacts with Leu-16 of the strictly conserved Leu–Gly–Pro motif in the β2 to -3 loop, which is a variable region in OmpR-family proteins. Also highly conserved, Leu-150 makes hydrophobic contacts with residues in helix α1 (Ile-44 and Ala-40) and the flanking residues of the Leu–Gly–Pro motif in the β2 to -3 loop (Leu-15 and -19). Except for Ala-40, these sites are always hydrophobic. The hydrophobic contacts are complemented by a number of partially conserved polar or ionic contacts. Thus, the extensive interface between the DNA-binding and BTA domain appears to be governed by conserved interactions.

Helix T3 of the BTA domain provides the majority of nonbonding interactions with the inner β-sheet of the FHA domain, in particular with residues in strands F5 and -6, representing a highly conserved region in the FHA domain. Hydrophobic residues (Val-331, Val-333, Gly-336, and Val-340) and small polar residues (Thr-335 and -337) are set against small side chains in helix T3 (Pro-161, Thr-164, and Ala-165).

Phosphorylation and Phospho–Peptide Complex Structure. It has been shown that PknH-mediated Thr-phosphorylation of EmbR is abrogated when truncating the FHA domain or mutating conserved residues in the phospho-Thr binding pocket (11). Two-dimensional gel electrophoresis of EmbR after *in vitro* phosphorylation by PknH (Fig. 4*A*) indicates the presence of up to five phosphorylation states in EmbR. The dominant band (1 in Fig. 4*A*) on the autoradiograph accounts for ≈60% of the total integrated spot intensity, compared with states 2 and 3, with ≈15% each. The phosphorylation sites could not yet be assigned, because 15 of 27 Thr residues display significant solvent exposure (≥5%).

Initially, potential phospho-threonine sites on PknH to which EmbR might bind had to be inferred from data of other *Mtb* Pkn kinases, whereby phospho-Thr sites 171 and 173 in the activation loop of PknB had been reported to play a critical role for both kinase activation (30) and substrate binding (12). Of these sites, Thr-171 (Thr-170 in PknH) is conserved in all but 2 of the 11 *Mtb* Ser/Thr kinases. Hence, we chose a nine-residue peptide sequence derived from PknH residue Thr-170 (the second site, Thr-173, lines up with Leu-172 in PknH). Because EmbR contains four Trp residues, one of which (Trp-63) is located at the surface of the DNA-binding domain, ≈15 Å away from the phospho-Thr (pT)-binding pocket (Fig. 2*A*), we assessed ligand-binding by intrinsic tryptophan fluorescence. No appreciable binding of the Thr-170-phosphopeptide (DEKLpTQLGT, pT = PO_4-Thr) was detected. In line with this observation, diffraction data to 2 Å resolution of EmbR cocrystallized with a 10-fold molar excess of Thr-170-phosphopeptide revealed clearly delineated density for pT in the pT-binding pocket but none for the flanking residues (data not shown). Two further nine-residue phosphopeptides derived from Thr sites 174 and 222, chosen on the basis of a very recent tandem mass spectrometry analysis of autophosphorylated PknH (V.M., I. Zanelle-Cleon, J.-P. Robin, S. Mallejac, A.-J.C., and M. Becchi, unpublished work) similarly failed to give rise to a binding signal (data not shown). In contrast, a titration experiment using a nonspecific phosphopeptide derived from the high-affinity ligand of the FHA domain of the yeast Rad53p check-point kinase (28) (SLEVpTEADT) indicated a weakly bound complex (K_d = 185 ± 40 μM; 95% confidence interval; 105–266 mM; Fig. 4*B*). EmbR crystals grown at solution conditions essentially identical to those of the

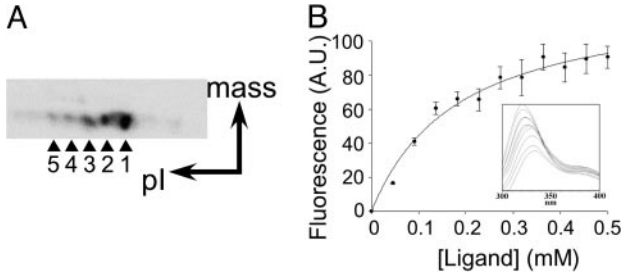

Fig. 4. Phosphorylation and phosphopeptide binding of EmbR. (*A*) A 2D-gel electrophoresis of EmbR postphosphorylation by PknH. ^{33}P-labeled EmbR is visualized by autoradiography. (*B*) Intensity of Trp fluorescence emission at 330 nm as a function of concentration of the Rad53p-ligand (SLEVpTEADT). (*Inset*) Shift of the emission peak with increasing peptide concentrations.

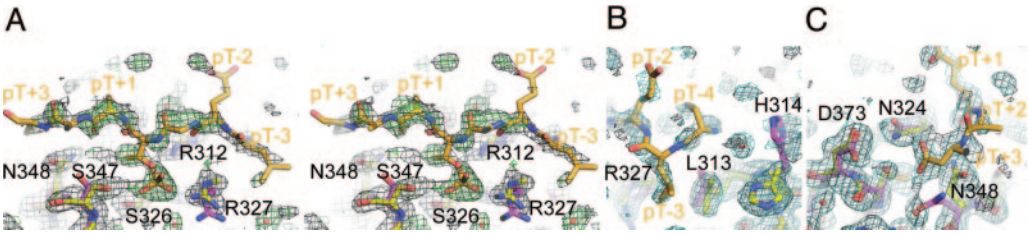

Fig. 5. Phosphopeptide complex. (*A*) View of the peptide bound to chain A (shown with carbon atoms in yellow). The σ_A-weighted 2Fo–Fc (1.9 Å, 0.8 σ, black) and Fo–Fc maps (2.5 σ, green) were calculated with model phases before building the peptide. Peptide residues (orange) are labeled relative to the pT position. (*B* and *C*) View of the peptide termini, illustrating the putative specificity-determining regions. Shown are the ligands for chains A (*B*) and B (*C*). The σ_A-weighted 2Fo–Fc map (0.8 σ) is contoured before (black) and after (cyan), including the peptide atoms in the phase calculation. Selected side chains of apo EmbR are shown in magenta.

Thr-170–peptide complex but using a 10-fold excess of Rad53p-ligand, resulted in a 1.9-Å-resolution structure of the phosphopeptide-bound complex (Table 1). Density for this ligand, albeit weak, was visible for all but one of the flanking residues of the pT (see *Materials and Methods* and Fig. 5 *A* and *B*).

Ligand-binding causes a subtle rearrangement of EmbR molecules on the crystal lattice, reducing the space group symmetry from *I*222 to *C*2, with two peptide-bound EmbR molecules per asymmetric unit (chains A and B). These complexes have been refined independently. The only change in overall conformation is a small rotation (≈6°) of the FHA domain relative to the two N-terminal domains in chain B. Because the pT-binding pockets of chains A and B face each other, a few main-chain–main-chain contacts are observed between the peptides, involving peptide residues pT−1 to pT+1. Nonetheless, the extended conformation of the ligands, straddling loops F3 to -4, F4 to -5 and F6 to -7, is entirely consistent with that seen of the high-affinity complexes in eukaryotic FHA domain–peptide complexes. The phosphate moiety is clamped between the side chains of Arg-312, Ser-326, Arg-327, and Ser-347. Ser-326 and Arg-312 are highly conserved in eukaryotic FHA domains, and Ala/Gly point mutants of either side chain show greatly diminished phosphorylation by PknH (11). Arg-327, conserved in most EmbR homologues (Fig. 6), forms a bidentate salt bridge (3.2 and 3.3 Å) with the phosphate. When superimposing the FHA domains of chains A and B, conformational differences in side chains contacting the peptide are limited to Arg-327 (RMSD, 1.1 Å-including and 0.64 Å-excluding residue 327). Still, Arg-327 forms a salt bridge to the phosphate moiety in both chains. The peptide Cα traces line up with an RMSD of 0.9 Å (six Cαs), and, as far as the peptide density is defined, no contradictory features are observed between the crystallographically distinct ligands. Compared with the apo structure, two side chains in contact with the peptide, Asn-348 and Ser-347, are rotated by ≥90° about the Cα–Cβ bond (Fig. 5). A third residue, His-314, shows a similar side-chain rotation and forms a water-mediated contact with the N terminus of the peptide. His-314 is well clear of packing interfaces, suggesting a peptide-induced conformational change.

Because a nonspecific ligand was used, specific interactions between the peptide and the FHA domain may not accurately reflect the specificity of the docking site on PknH. Yet, two features stand out. First, Leu-313 makes hydrophobic contacts with the peptide leucine pT−3 (Fig. 5*B*). This solvent-accessible leucine (exposed surface area 63 Å^2) is a variable site among *Mtb*-encoded FHA domains (see Fig. 8, which is published as supporting information on the PNAS web site) and often occupied by uncharged polar residues (Ser and Asn) in eukaryotic FHA domains. Second, the rotation of Asn-348, turning this conserved side chain into the conformation seen in the Rad53p and Chk2 FHA-domain structures, opens a deep pocket in the surface of the FHA domain. Hovering above this pocket, Asp(pT+3) forms a weak H-bond with Asn-324 and with water molecules trapped underneath. Asn-348, in turn, forms H-bonds with the backbone amide of Asp(pT+3) and the carbonyl of Glu(pT+1), contacts that are also observed in the Rad53p FHA-domain–peptide complex (28).

Discussion

Among the seven substrates of *Mtb* Ser/Thr kinases identified so far (6, 8, 11, 12, 31), EmbR is one of the few, if not the only one, for which a mechanistic hypothesis of function has been proposed (11). Compelling evidence for phosphorylation-dependent and EmbR-mediated transcriptional control of the Emb arabinosyltransferases is lacking, but data backing individual aspects of the hypothesis are beginning to emerge. For instance, when wild-type *Mtb* and a *pknH*-deletion strain were treated with sublethal concentrations of ethambutol, transcription of the *emb* operon genes was down-regulated in the knockout but not in the parental strain (32), supporting the argument that *emb* gene transcription is controlled by a PknH-dependent pathway.

EmbR shows similarities to the response regulators of the OmpR/PhoB family, but presents a unique domain architecture. The winged-helix and the central TPR domain are linked through an extensive conserved interface, firmly tethering the two domains to each other. Superimposing the two independent molecules of the peptide-bound structure with that of the ligand-free form, the two N-terminal domains line up almost perfectly (RMSD < 0.5 Å for 270 Cα positions), whereas a distinct rotation of the FHA domain can be discerned, mirrored by an RMSD of 2.1 Å (91 Cαs) for the FHA domains in the same alignment, suggesting that the winged-helix and BTA domains form an integral functional unit, whereas the coupling to the FHA domain seems less firm. Indeed, C-terminal truncation mutants of the SARP-family protein *Streptomyces coelicolor* AfsR, cutting between the BTA and ATPase domains (Fig. 1), still bind DNA (33). Conversely, C-terminal-truncation mutants of DnrI that cut into the BTA domain (note that DnrI lacks helix T8; Fig. 6) abrogate both DNA binding and induction of daunorubicin synthesis in *Streptomyces peucetius* (22). Thus, an intact conserved core (helices T1–T7) of the BTA domain seems to be required for proper function of SARP-family proteins.

Based on the homology of the DNA-binding domain to PhoB/OmpR, SARP-family proteins have been postulated to bind heptameric direct repeat sequences (18, 21). Gel mobility-shift data obtained for DnrI, indeed, indicated a 2:1 stoichiometric complex with a seven-nucleotide tandem repeat separated by four nucleotides, analogous to the DNA-binding characteristics of PhoB (21, 22). Whereas the dimensions of full-length EmbR are incompatible with the geometry of the 2:1 PhoB–DNA complex (21), stripping the FHA domain and helix T8 from the model (in effect mimicking a DnrI–DNA complex) removes the overlap between the DNA-bound protomers, suggesting that the conserved core of SARP-family proteins, in

principle, could bind to the duplex strand in a PhoB-like fashion, consistent with the DNA-binding data for DnrI (22). Moreover, in the truncated model, the concave surface formed by the three TPR repeats (helices T1–T6) presents a platform analogous to that seen in the Hop TPR1-domain–pentapeptide complex (26). As TPR domains are generally thought to mediate protein–protein interactions (25), the BTA domain may help to stabilize a DNA-bound dimer. Phosphorylation by PknH may provide a mechanism by which steric hindrance between adjacent EmbR protomers could potentially be removed. Such a concept seems plausible, considering that the SARP-family protein AfsR, which is regulated by Ser/Thr phosphorylation and comprises several domains downstream of the BTA domain shows enhanced binding to DNA in the phosphorylated state (33).

Prokaryotic FHA domains have been identified only relatively recently (34), and their presence in bacterial genomes seems to correlate with that of eukaryotic-like Ser/Thr kinases and phosphatases (35, 36). Our data show unequivocally that the FHA domain of EmbR is a bona fide pT-binding module. Peptide recognition involves the conserved residues Arg-312, Ser-326, and Asn-348. Mutating any of these sites abrogates PknH-mediated phosphorylation of EmbR (11), providing, in context with the structural data, the most direct evidence yet for FHA-domain-mediated binding of EmbR to PknH. The FHA domain mediates interaction with the cognate kinase in at least two other endogenous substrates, the putative ABC-transporter Rv1747, a substrate of PknF (8, 37), and GarA (Rv1827), a substrate of PknB (12) (endogenous substrates lacking an FHA domain have also been reported) (6). Furthermore, the FHA domains of Rv0020c and Rv1747 are *in vitro* substrates of the Ser/Thr kinases PknB, -D, -E, and -F (31). Thus, in at least a subset of cases, specific recognition at the kinase–FHA domain interface might be required to direct the phosphorylation signal to a defined pathway. Peptide library experiments have previously revealed the pT motifs recognized by eukaryotic FHA domains, indicating that residues on positions pT −4 to pT+3 contribute to specificity, with pT+1–pT+3 often dominating (19). The EmbR-phosphopeptide complex suggests selectivity for a hydrophobic residue at pT−3 and offers a binding pocket for residue pT+3, in line with recognition "rules" in eukaryotic FHA domains. Predicting the docking site on PknH based on this data is, nonetheless, difficult. Whereas the kinase-docking site of GarA was mapped to the activation loop of PknB (12), binding to either (or both) Thr-171 or Thr-173, our prediction that EmbR might bind to the corresponding site in PknH (Thr-170) was not confirmed. Neither did we detect binding to the phosphopeptide derived from Thr-174, a second, partially conserved threonine in the activation loop, recently revealed as an autophosphorylation site of PknH (V.M., I. Zanelle-Cleon, J.-P. Robin, S. Mallejac, A.-J.C., and M. Becchi, unpublished work). Thus, the crystallographic and peptide-binding data suggest that the activation loop of the *Mtb* Ser/Thr-kinases may not constitute a universal docking site for FHA-domain-containing endogenous substrates.

Materials and Methods

Cloning. The *E. coli* DH5α strain was used to propagate plasmids during cloning. Restriction enzymes, T4 DNA ligase and Deep-Vent DNA polymerase were sourced from New England Biolabs, primers from MWG Biotech, and QIAquick and QIAprep reagents (Qiagen, Valencia, CA) were used for DNA purification. The 1,167-bp region coding for *emb*R in the *Mtb* H37Rv strain was amplified by PCR using primers (restriction sites underlined) TATGGATCCATGGCTGGTAGCGCGACAGTGGAGAAGCGG (forward) and TATAAGCTTCTACGTGCCGCCATGCGTCCCCGCG (reverse). The PCR product was ligated into the BamHI and HindIII restriction sites of plasmid pET28a (Novagen), yielding the pET28a-*embR* His-6-tagged construct, which was verified by sequencing.

Protein Production and Crystallization. The plasmid pET28a-*emb*R was introduced into *E. coli* C41(DE3) or methionine-auxotroph B834(DE3) cells by transformation. Recombinant strains were used to inoculate overnight cultures (5 ml of LB and 25 μg/ml kanamycin), which were fed to bulk medium (4 × 1 liter of LB and 25 μg/ml kanamycin), grown under shaking (at 37°C) to A_{600} = 0.5, and then induced with isopropyl β-D-thiogalactoside (IPTG) (1 mM) and incubated further (12 h at 16°C). Cultures of B834(DE3) cells were first grown in LB (at 37°C) to A_{600} = 1.2, then harvested, washed, and resuspended in SeMet-supplemented M9 medium (38), followed by incubation (1 h at 16°C), induction (1 mM IPTG), and incubation at 16°C (for 12 h). The purification followed the protocol in ref. 11. EmbR-containing fractions of the Ni^{2+}-NTA column eluate were pooled and dialyzed into protein storage buffer: 40 mM Tris·HCl, pH 7.9, 0.2 M NaCl, 0.2 mM DTT, 0.2 mM EDTA, and 10% glycerol.

Crystals were grown by hanging-drop vapor diffusion over a reservoir of 8–15% (wt/vol) polyethylene glycol (PEG) 4000, 0.1 M Na-Hepes, pH 7.5, 0.2 M $MgCl_2$, and 5 mM DTT, mixing drops of 1 μl of EmbR (5 mg·ml^{-1}) with 1 μl of reservoir solution. Crystals (250 × 250 × 60 mm^3) appeared within 1–3 days. These were cryoprotected in artificial mother liquor, adding 15–17.5% (wt/vol) sucrose in steps of 5%, mounted in nylon loops, and flash-frozen in liquid nitrogen. SeMet-EmbR crystals were grown the same way. Phosphorylated peptides of >98% purity (Rad53p: Ace-SLEVpTEADT-NH_3; PknH-Thr-170: DEKLpTQLGT; -Thr-174 TQLGpTAVGT; and -Thr-222: DSAGpTLVSS) were purchased from Severn Biotech (Kidderminster, U.K.). The Rad53p- and Thr-170-peptides were dissolved in the protein storage buffer, and mixed with protein at a 10:1 molar ratio. The complexes were concentrated to ≈12 mg·ml^{-1} by ultrafiltration. Crystals of phosphopeptide-bound EmbR were obtained by vapor diffusion over a reservoir of 2.5% wt/vol PEG 8000, 0.1 M Tris·HCl, pH 8.5, and 3% ethylene glycol and were cryoprotected in 13% wt/vol sucrose and 13% ethylene glycol.

Structure Determination. X-ray diffraction data were recorded at the European Synchrotron Radiation Facility (Grenoble, France) on beamlines BM14 [SeMet multiwavelength anomalous dispersion (MAD)], ID14–3 (SeMet, single wavelength), and ID29 (peptide complex) and reduced by using the programs MOSFLM/SCALA (39, 40) or XDS (41). A Patterson-search (SOLVE) (42) using the 2.5-Å MAD data found two of four Se-positions. Phases were refined by using the program SHARP (43); a third Se site was added from difference Fourier maps, and the resulting phases [figure of merit (FOM) = 0.58, 72–73.1 Å] were subjected to density modification and phase-extension to 2.0 Å by using the program SOLOMON (40). The experimental map (Fig. 7) was traced by using the program ARP/WARP (44), generating two thirds of the backbone, which was extended in the program O (45) to comprise an undisrupted chain of 365 of 388 residues. Side chains were added, and the model was refined by using the program REFMAC5 (40). The Ramachandran plot (92%, most favored; 7%, additionally allowed region) shows two residues (Glu-61 and Cys-372) at the boundary of the disallowed region, but simulated annealing omit maps confirmed the backbone conformation. The peptide-bound structure was solved by molecular replacement by using the program PHASER (46). The apo-structure model was modified guided by σ_A-weighted Fo–Fc, 2Fo–Fc, or simulated annealing omit maps. Water molecules were added, excluding the ligand-binding region. Before building the ligands, the refinement (46–1.9 Å, REFMAC5) reached R values of ≈24.5% (test) and 22.5% (working set), with the refined 2Fo–Fc map (Fig. 5 *A–C*) displaying peptide density for residues pT−4 through pT+3 (chain A ligand) and pT−2 through pT+3 (chain B ligand). Starting with the well defined pT residues,

incremental model building led to density for additional ligand side chains. The pT residues refined to 75% occupancy (by using the program CNS) (47), which, when adopted for the whole peptide, reduced the residual density to noise level. Figs. 2, 3, and 5 were prepared by using the programs PYMOL (DeLano Scientific, San Carlos, CA), GRASP (48), DEEPVIEW (49), and POVRAY ver. 3.6 (Persistence of Vision, Williamstown, Victoria, Australia). Sequence alignments were generated by using CLUSTALW (www.ebi.ac.uk/clustalw) and formatted by using ESPRIPT (http://espript.ibcp.fr).

Phosphorylation Assay. EmbR was phosphorylated in the presence of [γ-^{32}P]ATP and PknH as described in ref. 11 and run on immobilized 7-cm pH 6.3–8.3 gradient strips on a Protein IEF Cell apparatus (Bio-Rad) for the first dimension. The second dimension was run on a 10% SDS/PAGE gel. The Coomassie-stained gel was dried onto filter paper (Whatman). Radioactivity was revealed by autoradiography and quantitatively analyzed by using a PhosphorImager (STORM), normalizing spot intensities to the total integrated intensity of the autoradiogram.

Binding Assay. Intrinsic Trp-fluorescence emission was measured between 300 and 400 nm (scan speed 100 nm/min) by using a PerkinElmer luminescence spectrometer LS 50B, exciting the sample at a wavelength of 295 nm. Spectra were recorded for each peptide aliquot (1 ml) added to 400 ml of a 5 mM solution of EmbR, until it reached a final peptide concentration of 0.5 mM (11 additions). A plot of fluorescence intensity at 330 nm vs. peptide concentration (three independent experiments) resulted in a binding isotherm that was fitted to a one-site saturation model (by using the program SIGMAPLOT, Systat).

Patrick Corsino, Elisabeth McInerney, and Melinda Carpenter contributed to crystallization efforts at different stages of the project. We thank G. Waksman for comments on the manuscript and Tony Pemberton for computer system support. Access to European Synchrotron Radiation Facility beamlines was granted under proposal MX-405, and we gratefully acknowledge support from Drs. H. Belrhali, S. Monaco, and E. Mitchell. L.J.A. is supported by a Biotechnology and Biological Sciences Research Council Studentship, T.R.D. by a Medical Research Council (MRC) Fellowship, and L.K. by a grant from the Centre National de la Recherche Scientifique (Actions Thématiques et Incitatives sur Programmes "Microbiologie Fondamentale"). G.S.B. acknowledges support as a Lister Institute Jenner Research Fellow and MRC Grants G9901077 and G9901078.

1. Kennelly, P. J. (2002) *FEMS Microbiol. Lett.* **206,** 1–8.
2. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, *et al.* (1998) *Nature* **393,** 537–544.
3. Av-Gay, Y. & Everett, M. (2000) *Trends Microbiol.* **8,** 238–244.
4. Duran, R., Villarino, A., Bellinzoni, M., Wehenkel, A., Fernandez, P., Boitel, B., Cole, S. T., Alzari, P. M. & Cervenansky, C. (2005) *Biochem. Biophys. Res. Commun.* **333,** 858–867.
5. Chopra, P., Singh, B., Singh, R., Vohra, R., Koul, A., Meena, L. S., Koduri, H., Ghildiyal, M., Deol, P., Das, T. K., *et al.* (2003) *Biochem. Biophys. Res. Commun.* **311,** 112–120.
6. Kang, C. M., Abbott, D. W., Park, S. T., Dascher, C. C., Cantley, L. C. & Husson, R. N. (2005) *Genes Dev.* **19,** 1692–1704.
7. Sharma, K., Chandra, H., Gupta, P. K., Pathak, M., Narayan, A., Meena, L. S., D'Souza, R. C., Chopra, P., Ramachandran, S. & Singh, Y. (2004) *FEMS Microbiol. Lett.* **233,** 107–113.
8. Molle, V., Soulat, D., Jault, J. M., Grangeasse, C., Cozzone, A. J. & Prost, J. F. (2004) *FEMS Microbiol. Lett.* **234,** 215–223.
9. Deol, P., Vohra, R., Saini, A. K., Singh, A., Chandra, H., Chopra, P., Das, T. K., Tyagi, A. K. & Singh, Y. (2005) *J. Bacteriol.* **187,** 3415–3420.
10. Cowley, S., Ko, M., Pick, N., Chow, R., Downing, K. J., Gordhan, B. G., Betts, J. C., Mizrahi, V., Smith, D. A., Stokes, R. W. & Av-Gay, Y. (2004) *Mol. Microbiol.* **52,** 1691–1702.
11. Molle, V., Kremer, L., Girard-Blanc, C., Besra, G. S., Cozzone, A. J. & Prost, J. F. (2003) *Biochemistry* **42,** 15300–15309.
12. Villarino, A., Duran, R., Wehenkel, A., Fernandez, P., England, P., Brodin, P., Cole, S. T., Zimny-Arndt, U., Jungblut, P. R., Cervenansky, C. & Alzari, P. M. (2005) *J. Mol. Biol.* **350,** 953–963.
13. Belanger, A. E., Besra, G. S., Ford, M. E., Mikusova, K., Belisle, J. T., Brennan, P. J. & Inamine, J. M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11919–11924.
14. Telenti, A., Philipp, W. J., Sreevatsan, S., Bernasconi, C., Stockbauer, K. E., Wieles, B., Musser, J. M. & Jacobs, W. R., Jr. (1997) *Nat. Med.* **3,** 567–570.
15. Escuyer, V. E., Lety, M. A., Torrelles, J. B., Khoo, K. H., Tang, J. B., Rithner, C. D., Frehel, C., McNeil, M. R., Brennan, P. J. & Chatterjee, D. (2001) *J. Biol. Chem.* **276,** 48854–48862.
16. Zhang, N., Torrelles, J. B., McNeil, M. R., Escuyer, V. E., Khoo, K. H., Brennan, P. J. & Chatterjee, D. (2003) *Mol. Microbiol.* **50,** 69–76.
17. Alderwick, L. J., Radmacher, E., Seidel, M., Gande, R., Hitchen, P. G., Dell, A., Sahm, H., Eggeling, L. & Besra, G. S. (2005) *J. Biol. Chem.* **280,** 32362–32371.
18. Wietzorrek, A. & Bibb, M. (1997) *Mol. Microbiol.* **25,** 1181–1184.
19. Yaffe, M. B. & Smerdon, S. J. (2004) *Annu. Rev. Biophys. Biomol. Struct.* **33,** 225–244.
20. Martinez-Hackert, E. & Stock, A. M. (1997) *Structure (Cambridge, U.K.)* **5,** 109–124.
21. Blanco, A. G., Sola, M., Gomis-Ruth, F. X. & Coll, M. (2002) *Structure (Cambridge, U.K.)* **10,** 701–713.
22. Sheldon, P. J., Busarow, S. B. & Hutchinson, C. R. (2002) *Mol. Microbiol.* **44,** 449–460.
23. Das, A. K., Cohen, P. W. & Barford, D. (1998) *EMBO J.* **17,** 1192–1199.
24. Holm, L. & Sander, C. (1998) *Nucleic Acids Res.* **26,** 316–319.
25. D'Andrea, L. D. & Regan, L. (2003) *Trends Biochem. Sci.* **28,** 655–662.
26. Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F. U. & Moarefi, I. (2000) *Cell* **101,** 199–210.
27. Li, H., Byeon, I. J., Ju, Y. & Tsai, M. D. (2004) *J. Mol. Biol.* **335,** 371–381.
28. Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J. & Yaffe, M. B. (2000) *Mol. Cell* **6,** 1169–1182.
29. Li, J., Williams, B. L., Haire, L. F., Goldberg, M., Wilker, E., Durocher, D., Yaffe, M. B., Jackson, S. P. & Smerdon, S. J. (2002) *Mol. Cell* **9,** 1045–1054.
30. Boitel, B., Ortiz-Lombardia, M., Duran, R., Pompeo, F., Cole, S. T., Cervenansky, C. & Alzari, P. M. (2003) *Mol. Microbiol.* **49,** 1493–1508.
31. Grundner, C., Gay, L. M. & Alber, T. (2005) *Protein Sci.* **14,** 1918–1921.
32. Papavinasasundaram, K. G., Chan, B., Chung, J. H., Colston, M. J., Davis, E. O. & Av-Gay, Y. (2005) *J. Bacteriol.* **187,** 5751–5760.
33. Lee, P. C., Umeyama, T. & Horinouchi, S. (2002) *Mol. Microbiol.* **43,** 1413–1430.
34. Pallen, M., Chaudhuri, R. & Khan, A. (2002) *Trends Microbiol.* **10,** 556–563.
35. Durocher, D. (2003) *Trends Microbiol.* **11,** 67–68.
36. Ponting, C. P., Aravind, L., Schultz, J., Bork, P. & Koonin, E. V. (1999) *J. Mol. Biol.* **289,** 729–745.
37. Curry, J. M., Whalan, R., Hunt, D. M., Gohil, K., Strom, M., Rickman, L., Colston, M. J., Smerdon, S. J. & Buxton, R. S. (2005) *Infect. Immun.* **73,** 4471–4477.
38. Cotton, N. P., White, S. A., Peake, S. J., McSweeney, S. & Jackson, J. B. (2001) *Structure (Cambridge, U.K.)* **9,** 165–176.
39. Leslie, A. G. W. (1992) *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**.
40. CCP4. (1994) *Acta Crystallogr. D* **50,** 760–763.
41. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26,** 795–800.
42. Terwilliger, T. C. & Berendzen, J. (1999) *Acta Crystallogr. D* **55,** 849–861.
43. De La Fortelle, E. & Bricogne, G. (1997) *Methods Enzymol.* **276,** 472–493.
44. Morris, R. J., Perrakis, A. & Lamzin, A. (2002) *Acta Crystallogr. D* **58,** 968–975.
45. Jones, T. A. & Thirup, S. (1986) *EMBO J.* **5,** 819–822.
46. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005) *Acta Crystallogr. D* **61,** 458–464.
47. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54,** 905–921.
48. Nicholls, A., Sharp, K. A. & Honig, B. (1991) *Proteins* **11,** 281–296.
49. Guex, N. & Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723.