

Crystal Structure of the Response Regulator 02 Receiver Domain, the Essential YycF Two-Component System of *Streptococcus pneumoniae* in both Complexed and Native States†

Colin J. Bent,¹ Neil W. Isaacs,¹ Timothy J. Mitchell,^{2*} and Alan Riboldi-Tunnicliffe²

Department of Chemistry¹ and Division of Infection and Immunity,² University of Glasgow,
Glasgow G12 8QQ, United Kingdom

Received 25 August 2003/Accepted 12 January 2004

A variety of bacterial cellular responses to environmental signals are mediated by two-component signal transduction systems comprising a membrane-associated histidine protein kinase and a cytoplasmic response regulator (RR), which interpret specific stimuli and produce a measured physiological response. In RR activation, transient phosphorylation of a highly conserved aspartic acid residue drives the conformation changes needed for full activation of the protein. Sequence homology reveals that RR02 from *Streptococcus pneumoniae* belongs to the OmpR subfamily of RRs. The structures of the receiver domains from four members of this family, DrrB and DrrD from *Thermotoga maritima*, PhoB from *Escherichia coli*, and PhoP from *Bacillus subtilis*, have been elucidated. These domains are globally very similar in that they are composed of a doubly wound $\alpha_5\beta_5$; however, they differ remarkably in the fine detail of the β_4 - α_4 and α_4 regions. The structures presented here reveal a further difference of the geometry in this region. RR02 has been shown to be the essential RR in the gram-positive bacterium *S. pneumoniae* (R. Lange, C. Wagner, A. de Saizieu, N. Flint, J. Molnos, M. Stieger, P. Caspers, M. Kamber, W. Keck, and K. E. Amrein, *Gene* 237:223–234, 1999; J. P. Throup, K. K. Koretke, A. P. Bryant, K. A. Ingraham, A. F. Chalker, Y. Ge, A. Marra, N. G. Wallis, J. R. Brown, D. J. Holmes, M. Rosenberg, and M. K. Burnham, *Mol. Microbiol.* 35:566–576, 2000). RR02 functions as part of a phosphotransfer system that ultimately controls the levels of competence within the bacteria. Here we report the native structure of the receiver domain of RR02 from serotype 4 *S. pneumoniae* (as well as acetate- and phosphate-bound forms) at different pH levels. Two native structures at 2.3 Å, phased by single-wavelength anomalous diffraction (xenon SAD), and 1.85 Å and a third structure at pH 5.9 revealed the presence of a phosphate ion outside the active site. The fourth structure revealed the presence of an acetate molecule in the active site.

Two-component signal transduction systems constitute a family of signaling proteins found in many bacterial species and in some higher organisms (1, 14, 41, 42, 53, 56). These systems are generally comprised of two distinct protein components, the histidine protein kinase (HPK) and its cognate response regulator (RR). In these systems the histidine kinase is autophosphorylated on a conserved histidine residue, and the RR is activated by phosphorylation on the conserved aspartic residue in an Mg^{2+} -dependent transphosphorylation reaction from the phosphohistidine of the cognate kinase. This second component of the system is usually a transcription factor consisting of two domains: a DNA-binding or effector domain and a receiver (RR) domain (that acts as a regulator of the effector domain). Activation of the RR protein results in a downstream effect, which may involve protein-protein interactions and/or effect gene regulation at the transcription level. RR proteins can be classified by similarities in their effector domains (50). The majority of RRs function as transcription factors and have DNA-binding effector domains. They are

subdivided into three groups termed the OmpR/PhoB, NarL/FixJ, and NtrC/DctD subfamilies. The RR studied here (RR02rec) belongs to the OmpR/PhoB subfamily.

RR02 (34) (also called MicA [16], VicR [54], and 492RR [51]) is a 26-kDa protein which has been shown to be essential for bacterial growth (as knockout mutations in this gene are lethal), thus making it a prime target for new antibacterial agent design (34, 51). Within the two-component signal transduction system HPK/RR02, the HPK protein contains a PAS domain, a protein-protein interaction region classically involved in redox reactions, and is involved in protecting the cell against oxidative stress by repression of competence under oxygen-limiting environments (16). The induction of competence is controlled by a peptide, which is sensed by a further HPK-RR couplet (Com D/E) (15).

The high degree of conservation among two-component systems has made them an attractive target for novel classes of antimicrobial agents (3). Although two-component systems constitute one of the largest known families of transcriptional regulators in bacteria, few have been shown to be essential for cell viability. Among these, DivJ/DivK (23, 40) and CckA/CtrA (25, 43) have been shown to control cell cycle regulation, cell division, and DNA replication in *Caulobacter crescentus* (57). The MtrB/MtrA two-component system has been shown to be essential in *Mycobacterium tuberculosis*, but its function remains unknown (58). A further essential system is the YycG/

* Corresponding author. Mailing address: University of Glasgow, Division of Infection and Immunity, Joseph Black Building, University Ave., Glasgow G12 8QQ, United Kingdom. Phone: 44 141 3303749. Fax: 44 141 3303727. E-mail: t.mitchell@bio.gla.ac.uk.

† Supplemental material for this article may be found at <http://jlb.asm.org/>.

YycF system, a member of the EnvZ/OmpR two-component system family. It is the most highly conserved two-component system in gram-positive bacteria which have low-level GC content, with YycF RRs having 70% (on average) amino acid sequence similarity. This system has been found in all low-level-GC-content gram-positive bacteria whose sequences are known and has been reported to be essential for cell growth in *Bacillus subtilis*, *Staphylococcus aureus*, *Streptococcus pyogenes*, *Streptococcus pneumoniae*, and *Listeria monocytogenes* (18, 19, 20, 31, 34, 38, 51). The reason that these proteins are essential and the signal to which YycG/YycF responds (as well as the nature of the YycF regulon) remain unknown for this family. In *S. pneumoniae*, only the RR (YycF) was essential and not the histidine kinase (YycG) (34, 51). Two-component systems have been identified as potential antibacterial targets, as they play key roles in controlling cellular processes (4). As these systems are not found in humans (making them selective and unique bacterial targets), although two-component systems have been studied from a variety of organisms there has been little work at a biochemical level on an essential system from a pathogenic organism.

We have determined the first crystal structure of the receiver domain from the YycF RR family. Four independent structures of the receiver domain of RR02 (*RR02rec*) are presented here, two bound to acetate or phosphate (P_i) and two native forms solved at different pH levels. At the gross level, these structures closely resemble those of other RR receiver domains, showing the characteristic α_5 - β_5 doubly wound motif (49). Crystals of *RR02rec* with different ligands were obtained by the addition of ligands to the crystallization medium at differing pH levels. Coordination of ligands around the active site causes the β_3 - α_3 loop to adopt a defined conformation; in the native structures presented here, this loop is seen to be mobile and therefore no electron density is visible.

MATERIALS AND METHODS

Chemicals. All chemicals were purchased at Analar grade or higher from Sigma (unless stated otherwise), and all solutions were made fresh prior to crystallization trials.

Cloning, expression, and purification of *RR02rec*. With genomic pneumococcal DNA (TIGR4 strain) used as a template, the *RR02rec* gene was isolated (5) using the PCR. Oligonucleotide primers were designed to introduce NcoI and XhoI restriction sites to aid in cloning with the initiation sequence (ATG) within the NcoI site. The *RR02rec* gene (corresponding to residues 1 to 120) was cloned into pET33b vector (Novagen) in such a way as to exclude either the N- or C-terminal polyhistidine tag coding sequence.

Expression constructs were transformed into the BL21(DE3) *Escherichia coli* host strain (Stratagene) with a heat shock technique. Several batches of Luria-Bertani-enriched growth medium were inoculated with a sample of an overnight culture (47). The cultures were shaken at 200 rpm at 310 K until an optical density at 600 nm of ~0.5 was reached. IPTG (isopropyl- β -D-thiogalactopyranoside) (Melford) was then added to achieve a final concentration of 1 mM to initiate overexpression of *RR02rec*. Cells were harvested at $5,000 \times g$ for 15 min at 277 K and resuspended in 10 ml buffer per liter of culture (25 mM Tris [pH 7.5], 5 mM benzamidine). Cells were lysed on ice by sonication, and DNase I was added to achieve a final concentration of 20 μ g/ml. Any insoluble debris present were separated from the soluble protein by centrifugation at $10,000 \times g$ for 30 min at 277 K.

The soluble fraction was then loaded onto a 1-ml Hi-Trap Q ion exchange column (Amersham-Pharmacia) previously equilibrated with 25 mM Tris (pH 7.5) and eluted using a NaCl gradient (0 to 1 M). The *RR02rec* fractions identified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis were pooled and concentrated prior to gel filtration using both Superdex-75 and Superdex-200 (Amersham-Pharmacia) (equilibrated with a buffer containing 25

TABLE 1. Crystallisation conditions and crystal data of *RR02* receiver domain forms

| Parameter | Result for crystal: | | | |
|--|--------------------------------|--------------------------------|-------------------------|-------------------------|
| | INXT | INXP | INXW | INXO |
| Buffer | 0.1 M phosphate/citrate pH 4.0 | 0.1 M phosphate/citrate pH 4.5 | 0.1 M Acetate pH 5.1 | 0.1 M Tris pH 7.0 |
| Precipitant | 40% (vol/vol) PEG-300 | 40% (vol/vol) PEG-300 | 40% (vol/vol) PEG-300 | 40% (vol/vol) PEG-300 |
| Overall pH | 5.7 | 5.9 | 6.9 | 7.1 |
| Crystal system | C-centered orthorhombic | C-centered orthorhombic | C-centered orthorhombic | C-centered orthorhombic |
| Space group | C222 ₁ | C222 ₁ | C222 ₁ | C222 ₁ |
| V_r ($\text{\AA}^3 \text{ Da}^{-1}$) | 2.63 | 2.61 | 2.59 | 2.62 |
| Solvent content (%) | 53.23 | 52.96 | 52.45 | 53.09 |
| Molecules per A.U. | 1 | 1 | 1 | 1 |
| Unit cell parameters | | | | |
| a | 79.69 \AA | 79.39 \AA | 78.35 \AA | 80.31 \AA |
| b | 92.83 \AA | 92.69 \AA | 92.62 \AA | 92.51 \AA |
| c | 36.79 \AA | 36.77 \AA | 36.89 \AA | 36.52 \AA |
| α | 90° | 90° | 90° | 90° |
| β | 90° | 90° | 90° | 90° |
| γ | 90° | 90° | 90° | 90° |
| Upper diffraction limit | 2.34 | 1.70 | 1.82 | 1.38 |

mM Tris and 300 mM NaCl [pH 7.5]), which resulted in 20 mg of pure RR02rec (as determined by silver-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis and UV-visible-spectrum spectroscopy).

Crystallization of RR02rec. The purification and crystallization of this protein have been described previously (5). Briefly, RR02rec (in a solution containing 25 mM Tris-HCl and 300 mM NaCl [pH 7.5]) was concentrated to 10 mg/ml (as determined at 280 nm from the theoretical extinction coefficient 0.188) with a Centricon 10 apparatus (Millipore). Crystals were obtained at 20°C by the sitting-drop vapor diffusion method. Initial crystallization conditions were identified with a CryoI and CryoII sparse matrix (Molecular Dimensions). After optimization, crystals of an appropriate size were grown in 7 to 10 days by mixing 3.4 µl of protein with 0.6 µl of the reservoir solution, which contained 30% (vol/vol) polyethylene glycol 300 (PEG-300) and 0.1 M Tris-HCl buffer (pH 7.0). In the case of the phosphate-containing crystals the Tri-HCl buffer was exchanged for 0.1 M phosphate-citrate buffer in a pH range of 4.0 to 4.7 (0.5 M Na₂HPO₄ and 0.5 M citric acid solutions were mixed to produce the required pH). The acetate complex was obtained by changing the buffer to 0.1 M acetate buffer in a pH range of 6.0 to 6.5 (1 M acetic acid was mixed with 1 M sodium acetate to produce the required pH). All the crystals belong to the C-centered orthorhombic space group C22₁. Table 1 contains full details of all crystallization conditions and unit cell parameters. The pH quoted for crystal growth is the overall pH of the reservoir solution prior to mixing with the protein solution.

Data collection. Crystals were cryocooled in a stream of nitrogen gas at 100 K, with the low-molecular-weight PEG from the crystallization medium acting as the cryoprotectant. The data sets were collected at 100 K on the synchrotron beam lines 14.1 and 14.2 (SRS, Daresbury, United Kingdom), XRD-1 (Elettra, Trieste, Italy), and BM-14 (ESRF, Grenoble, France). Data processing and scaling were performed using MOSFLM (A. G. W. Leslie, Abstr. Proc. Daresbury Study Weekend, SERC, Daresbury Laboratory, Warrington, United Kingdom, p. 39–50, 1987) and SCALA (P. R. Evans, Abstr. Proc. CCP4 Study Weekend Data Collection Processing, p. 114–122, 1993), with other computations carried out using the CCP4 suite of programs (11) except where otherwise stated. Full data processing statistics and parameters are shown in Table 2.

Phase determination using SAD, model building, and refinement. Crystals grown at pH 5.7 were removed from the sitting drop through the use of a loop (Hampton Research), which was then mounted in a pressure cell and subjected to xenon at high pressure (5×10^5 to 15×10^5 Pa for 2 to 20 min) before being immediately immersed in a bath of liquid nitrogen. Diffraction data were collected from an optimized xenon-derivatized crystal (8×10^5 Pa for 10 min) at 1.8 Å, a wavelength which has been shown to be sufficient to maximize the signal for both xenon and sulfur at the XRD-1 beam line (Elettra). The xenon site was located by manual inspection of the anomalous Patterson maps.

Phasing was performed using the automated phasing program SHARP (13). Four anomalous sites (which were refined as a single xenon site) and three sulfur sites (three methionine residues in RR02rec) were found. Figures of merit (FOM) for the single-wavelength anomalous diffraction (SAD)-derived phases were calculated to be 0.62 (FOM_{acent}) and 0.35 (FOM_{centric}) for data between 30.8 and 2.34 Å. The resulting electron density map was improved (using SOLOMON) (2) with several cycles of solvent flipping and one cycle of solvent flattening. The mask was calculated using a solvent content of 53%; the resulting maps allowed chain tracing for almost the whole molecule. A partial model (containing only Gly, Val, Ser, and Ala residues) was built using the automated methods incorporated in ARP and wARP (33). Residues were built into the remaining electron density according to an amino acid sequence procedure using QUANTA (Accelrys). The initial model was refined using the maximum likelihood method implemented in REFMAC5 (G. Murshudov, A. Vagin, and E. Dodson, Abstr. Proc. Daresbury Study Weekend, SERC, Daresbury Laboratory, Warrington, United Kingdom, p. 93–104, 1996) and isotropic temperature refinement and bulk solvent correction. Approximately 10% of the data were excluded from the refinement and used in the calculation of the free *R* factor (9). On completion of the protein model, solvent molecules were added to a density higher than 3σ in sigma A-weighted mF_o-DF_c maps (44) generated by REFMAC5.

In the initial structure, no electron density was visible for Met-1, Glu-57, and Ile-58 (part of the loop connecting 3α to 3β) and the final residue Glu-120. The final *R* and *R*_{free} values were 21.41 and 23.56%, respectively. The structures of a second native (higher resolution) and those in complex with phosphate and acetate were solved by molecular replacement using the RR02rec-Xe structure as the search model in MOLREP (52) followed by rigid body refinement and TLS-restrained refinement using REFMAC5 (Murshudov et al., Abstr. Proc. Daresbury Study Weekend), where T, L, and S describe mean square translation, libration, and their correlation of rigid body. Final statistics for all structures are shown in Table 2.

TABLE 2. Crystallographic data collection and refinement statistics

| Parameter | Result for crystal: | | | |
|---|---------------------------------------|--------------------|------------------------------------|--|
| | INXP | INXW | INXT | INXO |
| Beam line and location | 14.2, SRS, Daresbury, UK ^a | BM14, ESRF, France | XRD-1, Elettra, Trieste, Italy | 14.2, SRS, Daresbury, UK |
| Wavelength (Å) | 0.978 | 1.024 | 1.800 | 0.978 |
| Resolution (Å) | 31.47–1.82 | 28.87–1.92 | 30.86–2.34 | 40.16–1.85 |
| Oscillation (°) | 1.0 | 1.0 | 1.0 | 1.0 |
| No. of measurements | 199,669 | 345,059 | 110,950 | 260,225 |
| No. unique | 11,155 | 10,522 | 5,398 | 12,010 |
| Completeness (%) (outer shell) | 98.60 (99.3) | 99.04 (99.0) | 98.96 | 99.43 (100.0) |
| <i>R</i> _{sym} (%) | 7.2 (34.5) | 4.5 (21.1) | 5.2 (11.6) | 8.4 (23.7) |
| <i>I</i> (σ) | 7.7 (2.1) | 9.1 (3.6) | 11.6 (5.1) | 7.0 (2.4) |
| <i>R</i> _{cryst} (%) (outer shell) | 17.87 (22.4) | 18.46 (22.8) | 18.99 | 16.59 (23.6) |
| <i>R</i> _{free} (%) (outer shell) | 23.04 (28.3) | 22.38 (19.4) | 23.76 | 21.81 (31.4) |
| Mean temperature B-factor (Å ²) | 26.21 | 28.02 | 37.03 | 28.22 |
| Active site complex | Phosphate | Acetate | Native | Native |
| Residues missing due to poor density | Met-1 and Glu-120 | Met-1 and Glu-120 | Met-1, Glu-57, Ile-58, and Glu-120 | Met-1, Leu-55, Pro-56, Glu-57, and Glu-120 |
| Correlation coefficient after MR ^b | 0.673 | 0.684 | NA ^c | 0.667 |
| <i>R</i> _{cryst} after MR | 0.377 | 0.373 | NA | 0.381 |

^a UK, United Kingdom.

^b MR, molecular replacement.

^c NA, not applicable.



FIG. 1. Sequence alignment of RR02rec from *S. pneumoniae* and two other members of the YycF family from *S. aureus* and *B. subtilis*. The high degree of sequence similarity within this family can be seen, with over 60% of residues being identical. The linker region between the receiver domain and the effector domain shows little or no similarity. The secondary structure elements of RR02rec are shown above the alignment. The alternating pattern of helices and strands seen in all RR structures is shown. This figure was created using MultAlin (12) and ESPript (22).

RESULTS AND DISCUSSION

All coordinates have been deposited in the Protein Data Bank (PDB) (6) (see Table 2 for PDB code numbers).

Structure of RR02rec. The structure of RR02rec was solved by the SAD method from a single crystal of a xenon derivative of RR02rec. The electron density after solvent flattening allowed automatic chain tracing of almost the entire molecule. The RR02rec-PO₄ complex described in this paper was refined to 1.7 Å with crystallographic *R* of 0.177 and *R*_{free} of 0.230, respectively, the RR02rec acetate complex was refined to 1.92 Å, and the high-resolution native was refined to 1.85 Å with the statistics shown in Table 2. Binding of the ligands causes a dramatic change in the position and orientation of the β3-α3 loop, which are discussed below. The overall crystal structure of RR02rec is a homodimer, with each monomer adopting the conventional doubly wound five-stranded α/β motif found in many RR receiver domains to date (49) (Fig. 1). Two helices (α1 and α5) are found on one side of the centrally located, parallel β-sheet, and three helices (α2, α3, and α4) are found on the opposite side (Fig. 2).

RR02rec dimer interface. Examination of the crystal packing revealed the occurrence of a major interface between the RR02rec domains. In the case of RR02rec pH 5.7, this interface buries 995.06 Å² from the solvent as a result of dimer formation through a crystallographic twofold axis; this corresponds to 15.4% of the overall surface area of the monomer (28, 29). A crystallographic dimer is also observed in the recently described structure of PhoP (7); in the case of PhoP, the dimer utilizes two distinct faces, the α2-α3-α4 face of one molecule interacting with the α4-β5-α5 face of the second molecule. In RR02rec, the dimer is formed through interac-

tions utilizing the α4-β5-α5 face of the two molecules. Analysis of the residues involved in the dimer interface reveals two salt bridges observed between Asp-97 and Arg-111 (2.94 Å), Asp-96, and Arg-118 (3.18 Å) as well as a number of hydrogen bonds and hydrophobic interactions. At the center of the interface we see the residue Arg-111; this is the only positively charged residue within a group of negatively charged residues (between Asp-97 and Glu-107) from each monomer which shield the salt bridge between Arg-111 (α5) and Asp-97 (α4-β5 loop) from bulk solvent.

RR02rec active site. The active site of RR02rec has several conserved residues that are expected to play a role in the function of the protein (27, 45, 49, 56). Most importantly, the site of phosphorylation, Asp-52, is located in a cleft formed by loops 1 (Asp-9 and Glu-10), 3 (Phe-33 and Asn-34), and 5

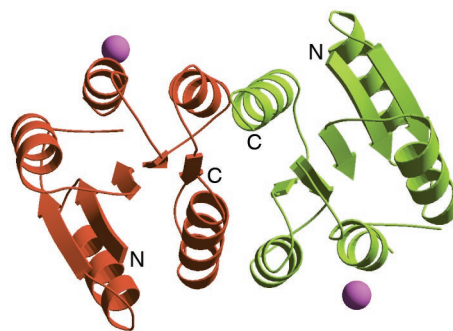


FIG. 2. Ribbon diagram of the dimer seen in the crystal structure. Figures were created using MOLSCRIPT (32), BOBSCRIPT (17), and Raster3D (39).

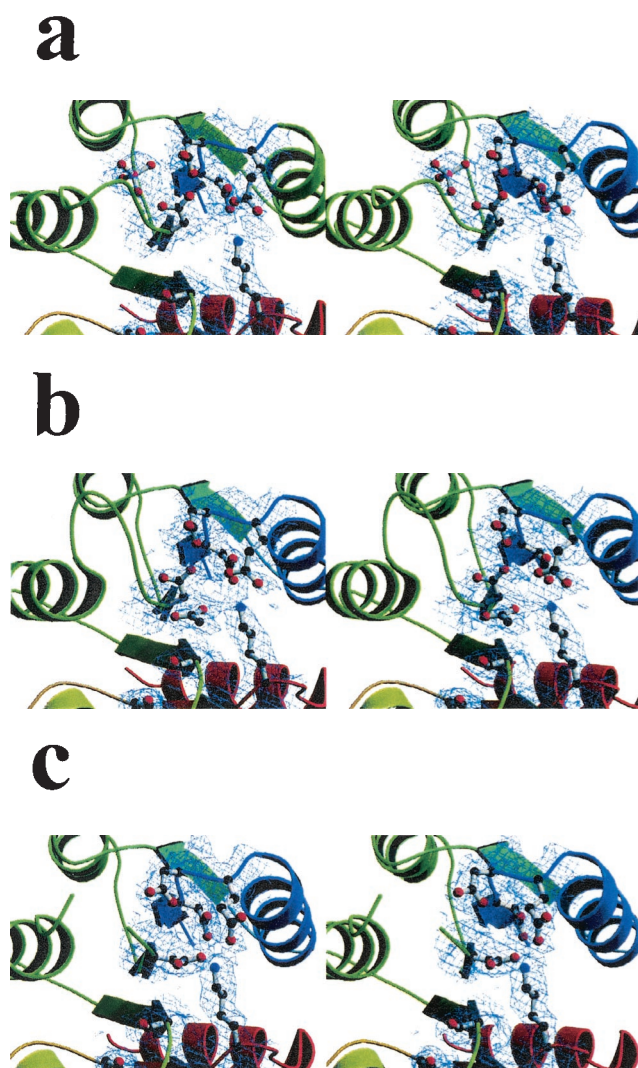


FIG. 3. Close-up stereo views of the active site of RR02rec showing several conserved residues of the RR family. (a) Position of the P_i ion outside the active site. (b) Acetate coordinated in the active site mimicking the role of a P_i ion. (c) Native RR02rec structure. Final 2FoFc maps contoured at 2 sigma. Figures were created using MOLSCRIPT (32), BOBSCRIPT (17), and Raster3D (39).

(Leu-53 to Asp-59). The carbonyl oxygens of Leu-53 and Pro-56 (loop 5) hydrogen bond to the amide nitrogens of residues Gly-60 (helix $\alpha 3$) and Gly-34 (helix $\alpha 2$). Both of these interactions help to maintain the structural conformation of the active site cleft.

In the native structure, Asp-52 establishes a characteristic hydrogen bond with a lysine residue found at the N terminus of the $\beta 5$ - $\alpha 5$ connecting loop (Lys-101) at a distance of 2.8 Å. In the complex structures this lysine residue interacts with an acidic residue found on the $\beta 1$ - $\alpha 1$ connecting loop Glu-10 (phosphate) and Asp-8 and Glu-10 (acetate).

(i) 1NXP: phosphate ion outside the active site at pH 5.9. In this structure, electron density is observed for the flexible loop region. In this case, the P_i ion is found outside the active site between $\alpha 2$ and $\alpha 3$ (Fig. 3a). Asp-52 has rotated around and is orientated differently from the orientation seen in the native

structure. The P_i ion present is too far away to be covalently bound to the side chain of Asp-52, but a strong hydrogen bond occurs between the Asp-52 side chain $O\delta^1$ and the oxygen atom P_i O4 at distance of 2.2 Å. Although this represents a short distance between the oxygen of the phosphate ion and the oxygen of Asp-52, previous structures have shown short distances of 2.3 Å (methionine adenosyltransferase 1O90) and 2.4 Å (colicin immunity protein E9 1BXI); this model has an estimated coordinate error of 0.178 Å. The side chain of Asp-52 also makes a further three hydrogen bonds with the amide nitrogens of Asp-9, Leu-53, and Met-54 (2.8, 3.0, and 2.9 Å, respectively). The P_i ion oxygen atom (O4) makes a further three interactions with the carbonyl oxygen of Asp-52, the amide nitrogen of Met-54, and the amide nitrogen of Leu-55 at respective distances of 2.66, 2.91, and 2.89 Å. The loop is also seen to make an additional three crystal contacts (not observed in the native or acetate structures). These are Pro-56 C β -Arg-106 NH2 (2.68 Å), Glu-57 OE1-Pro-102 CG (2.70 Å), and Ile-58 CD1-Val-88 CG1 (2.89 Å), which may play a role in stabilizing the loop in this conformation.

With a phosphate ion in the structure, the position of the $\beta 3$ - $\alpha 3$ loop adopts a conformation different from that seen in the native structure. This movement of the loop causes the active site cleft to be covered, closing the active site and allowing the formation of a bond between the carbonyl oxygen of Ile-58 and the amide nitrogen of Glu-62 (2.99 Å). This conformational change causes the helix $\alpha 3$ to become elongated by one residue and Asp-59 to become reoriented towards the active site and no longer to the bulk solvent (Fig. 4). With the loop in this position Ile-58 becomes the cap for helix $\alpha 3$.

(ii) 1NXW: acetate at pH 6.9. In this structure, clear density is observed within the active site that has been assigned as an acetate molecule. Two oxygens from the acetate molecule make contacts with Lys-101 NZ (2.85 Å) and water 49 (2.52 Å) (Fig. 3b). OXT of the acetate makes contacts to two main chain amides, Leu-53 (2.83 Å) and Met-54 (2.79 Å). The remaining stabilizing bond occurs between Ser-79 O γ and the OXT of the acetate (3.0 Å). In this structure electron density is observed for the $\beta 3$ - $\alpha 3$ loop (which forms a gamma turn), with the residues within the loop making several stabilizing interactions: (i) Leu-53 carbonyl oxygen is hydrogen bonded to Gly-60 amide nitrogen at a distance of 2.82 Å; (ii) Met-54 carbonyl oxygen interacts with Asp-9 $O\delta^1$ (2.75 Å); and (iii) Pro-56 carbonyl oxygen interacts with Ile-58 amide nitrogen (2.63 Å), while (iv) Glu-57 is involved in a crystal contact with Lys-23 (2.57 Å).

(iii) 1NXO: native. In this structure Asp-52 forms a salt bridge with Lys-101 $O\delta^1$ -NZ (2.82 Å) and $O\delta^2$ -NZ (2.65 Å). In this native structure, three residues (Leu-55, Pro-56, and Glu-57) within the $\beta 3$ - $\alpha 3$ loop region are disordered (Fig. 3c).

Conformational changes. In the previously determined RR structures of CheY (35), FixJ (8), and Spo0F (21), the activating mechanism has shown to involve the movement of an aromatic residue (Tyr/Phe) which rotates inwards and replaces a hydroxyl-containing residue (Ser/Thr) upon phosphorylation. This reorientation of these two key residues was observed on activation of CheY (by BeF_3) and has been termed the “active conformation.” The crystal structures presented here suggest that RR02rec is able to adopt this active form in a phosphorylation-independent manner (in that we do not observe Tyr-98

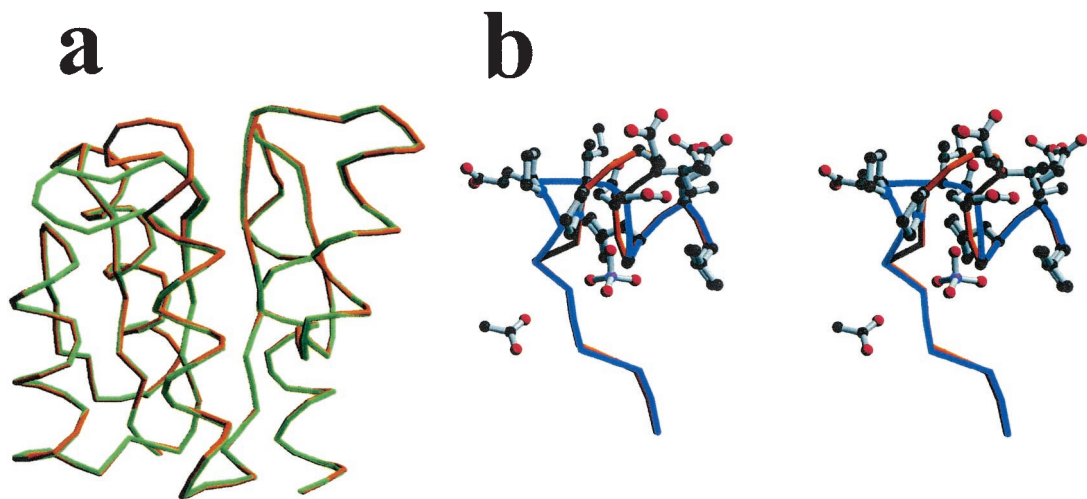


FIG. 4. (a) Overlay of two of the structures presented in this paper (acetate in yellow and P_i in green). (b) Stereo view showing the $\beta 3$ - $\alpha 3$ loop and residues 56 to 63. The orange ribbon represents the structure of acetate-coordinated RR02rec, and the purple ribbon represents the phosphate coordinated outside the active site. Figures were created using MOLSCRIPT (32), BOBSCRIPT (17), and Raster3D (39).

in a solvent-exposed conformation but only see it buried within the molecule between residues Met-77 and Lys-87, where it makes a hydrogen bond with the carbonyl oxygen of Lys-81). With the aromatic residue in this position it causes a shift in the hydroxyl group of Ser-79 and forces it to rotate towards the active site. If the aromatic residue were not in this position it would be impossible for the protein to form a dimer, as Tyr-98 would clash with Glu-107 and Arg-111. The activation of this form in a phosphorylation-independent manner was also observed in the structure of Rcp1 (24), for which it was proposed that activation was induced by the high-level ionic strength of the crystallization buffer (>3.2 M $[\text{NH}_4]_2\text{SO}_4$).

In RR02rec this may not be the case, as all structures were determined using crystals grown in concentrations of NaCl not exceeding 150 mM. We propose that the conformation of Tyr-98 and Ser-79 seen in RR02reca is concentration dependent rather than phosphorylation dependent. The effect of concentration on the oligomeric state of RR02rec has been investigated using dynamic light-scattering techniques (Table 3). The RR02-rec protein was used at various concentrations (1 to 15 mg/ml) both in phosphate and Tris-HCl buffer. The scattering of the protein solution was measured at 20°C, and at

least 10 data points were collected for each sample. The data were analyzed using DynaPro software. The results of these experiments (both with and without the presence of phosphate in the form of phosphate-citrate buffer) show that at low-level protein concentrations (less than 2 mg/ml), the only species present in solution is the monomeric form of RR02rec. When the concentration of the protein is increased to 10 mg/ml (the concentration used in crystal trials), the presence of a dimeric species is observed. Increasing the concentration still further (greater than 10 mg/ml) we see the presence of higher-order oligomers. These experiments have been confirmed by preliminary analytical ultracentrifugation experiments (data not shown). When a protein solution of 1 to 2 mg/ml was used, the technique revealed the presence of a monomeric species alone, while using solutions of 5 to 10 mg/ml we observed a dimeric species. Higher concentrations revealed an equilibrium between dimeric, tetrameric, and octameric forms. Further work is necessary to confirm these results in solution (both with and without phosphate).

Superimposition (using LSQKAB) (30) of the four RR02rec structures reveals a large shift in the positioning of the $\beta 3$ - $\alpha 3$ connecting loop (Leu-53–Asp-59) (Fig. 4a) dependent on the presence and positioning of the acetate-phosphate in the protein. In the acetate complex the loop region is in an open conformation with respect to a cleft leading into the active site formed by $\alpha 3$ and $\alpha 4$ (Fig. 4b). With no ligand present, the loop acts as a lid, closing this region. This closed conformation is also found in the structure containing a phosphate ion outside the active site in a cleft formed by $\alpha 2$ and $\alpha 3$. While none of the structures of RR02rec have a phosphate covalently bound to Asp-52, a comparison of the structure of RR02rec with acetate in the active site with the phosphate-bound structure of Spo0A (N-Spo0A~P) (36) shows that the protein adopts a similar conformation. A similar comparison of the structure of RR02rec with the structure of the phosphate found outside the active site shows the loop region in a conformation comparable with that of the unphosphorylated N-Spo0A (37).

TABLE 3. Dynamic light-scattering results using RR02rec at two different concentrations (see the supplemental material)

| RR02rec concn (mg/ml) (with or without P_i) | Diffusion coefficient | Radius (nm) | Polydispersity | Calculated mass (kDa) | Base line | SOS error |
|--|--------------------------|----------------|----------------|-----------------------------|--------------|--------------|
| 2 | 1,048 | 1.5 | 0.4 | 14 | 1.004 | 14.651 |
| 2 (+ P_i) | 1,111 | 1.4 | 0.4 | 14 | 1.002 | 12.132 |
| 10 | 806 | 2.5 | 0.4 | 30 | 1.003 | 10.532 |
| 10 (+ P_i) | 876 | 2.5 | 0.4 | 30 | 1.003 | 10.789 |

^a At the lower concentration of 2 mg/ml we see the presence of a 14-kDa species which we have assigned as the monomeric species in solution, both with and without the presence of phosphate, while at 10 mg/ml we see only the presence of a 30-kDa species, both with and without P_i , which we have assigned as the dimeric species of RR02rec.

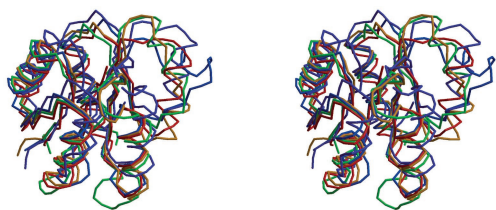


FIG. 5. Stereo diagrams of the C α traces of RR02rec. The native form (shown in red) and PhoB (in green) (PDB code 1B00) (48), DrrD (in blue) (PDB code 1KGS) (10), DrrB (in purple) (PDB code 1P2F) (46), and PhoP (in yellow) (PDB code 1MVO) (7) are shown. Superimposition and RMSDs were performed using the CCP4i program LSQKAB (30). Figures were created using MOLSCRIPT (32), BOBSCRIPT (17), and Raster3D (39).

Comparison of RR02rec with previously solved OmpR/PhoB RR receiver domain structures. The regulatory domains of PhoB, PhoP, DrrD, DrrB, and RR02 share a high degree of sequence similarity. Figure 5 shows the C α trace of the superimposition of the native protein with the receiver domains of PhoB (*E. coli*), PhoP (*B. subtilis*), DrrD (*Thermotoga maritima*), and DrrB (*T. maritima*). Superposition of the central five-strand sheet from RR02rec and the structure of DrrD indicates that although the secondary structures appear to occupy very similar positions, the overall root mean square deviation (RMSD) for all atoms is 2.9 Å whereas the area containing the α 4 helix and the β 4- α 4 loop shows most variability (with a maximum RMSD of 4.4 Å). This was also the case for DrrB (RMSD, 3.1 and 6.5 Å), PhoB (RMSD, 2.9 and 5.9 Å) and the receiver domain of PhoP (RMSD, 2.9 and 6.0 Å).

We have observed at low-level protein concentrations (less than 2 mg/ml) that RR02rec is a monomeric protein in solution. However, when the concentration is increased to more than 5 mg/ml we see the formation of dimers. We observe in the crystal structures an interface between two RR02rec monomers that causes 995 Å² to be shielded from solvent. The extent of the interacting surfaces is similar to those described for CheY and CheA (55), the homodimer formed by phosphorylated FixJN (8), and the asymmetric dimer seen in PhoP (7). The dimer formed by RR02rec is formed through twofold symmetry and involves the α 4- β 5- α 5 regions of two monomers. The formation of higher-order oligomers has been shown by dynamic light-scattering experiments. These oligomers must form by interactions other than those occurring through the α 4- β 5- α 5, although this has yet to be shown experimentally. The presence of acetate within the active site and of a phosphate ion outside the active site results in a dramatic shift of the β 3- α 3 loop, with the position of residue Pro-56 moving over 5.2 Å. This shift in the loop which occurs when acetate is bound in the active site causes the loop to move in such a way that the active site is opened, allowing access to the phosphorylatable residue Asp-52. When a P_i ion is found in the RR02rec structure we see that the accessibility of the active site is much lower. Although there is a large shift in this loop region, we see very little overall movement upon ligand binding over the structure as a whole (as was seen in the CheY D13K mutant) (26).

The structures presented here represent the first essential RR from the YycF/G family. Although we do not see the

presence of a cation within the active site (as has been seen previously within other solved RRs), this is due to the lack of Mg²⁺ or Mn²⁺ in the crystallization conditions. Further work, which will involve crystallizing the protein in the presence of cations such as Mg²⁺ or Mn²⁺ and also in the presence of a phosphate mimic (i.e., BeF₃) to produce an activated (phosphate-bound) form of this protein, is necessary to better understand this family of proteins.

ACKNOWLEDGMENTS

We thank the scientific staff of the CLRC Daresbury Laboratory (England), Sincrotrone Trieste SCpA (Italy), and the European Synchrotron Radiation Facility (France) for the provision of synchrotron-radiation beamtime and for providing excellent data collection facilities and the SHWCFG centre (University of Glasgow) for sequencing work.

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) Studentship and EU funding (grant QLRLK2-2000-00543).

REFERENCES

1. Abeyta, M., G. G. Hardy, and J. Yother. 2003. Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. *Infect. Immun.* 71:218–225.
2. Abrahams, J. P., and A. G. W. Leslie. 1996. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* 52:30–42.
3. Barrett, J. F., R. M. Goldschmidt, L. E. Lawrence, B. Folen, R. Chen, J. P. Demers, S. Johnson, R. Kanjia, J. Fernandez, J. Bernstein, L. Licata, A. Donetz, S. Huang, D. J. Hlasta, M. J. Macielag, K. Ohemeng, R. Frechette, M. B. Froese, D. H. Klaubert, J. M. Whiteley, L. Wang, and J. A. Hoch. 1998. Antibacterial agents that inhibit two-component signal transduction systems. *Proc. Natl. Acad. Sci. USA* 95:5317–5322.
4. Barrett, J. F., and J. A. Hoch. 1998. Two-component signal transduction as a target for microbial anti-infective therapy. *Antimicrob. Agents Chemother.* 42:1529–1536.
5. Bent, C. J., N. W. Isaacs, T. J. Mitchell, and A. Riboldi-Tunnicliffe. 2003. Cloning, overexpression, purification, crystallization and preliminary diffraction analysis of the receiver domain of MicA. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 59:758–760.
6. Berman, H. M., T. N. Bhat, P. E. Bourne, Z. Feng, G. Gilliland, H. Weissig, and J. Westbrook. 2000. The Protein Data Bank and the challenge of structural genomics. *Nat. Struct. Biol.* 7:957–959.
7. Birck, C., Y. Chen, F. M. Hulett, and J.-P. Samama. 2003. The crystal structure of the phosphorylation domain in PhoP reveals a functional tandem association mediated by an asymmetric interface. *J. Bacteriol.* 185:254–261.
8. Birck, C., L. Mourey, P. Gouet, B. Fabry, J. Schumacher, P. Rousseau, D. Kahn, and J. P. Samama. 1999. Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure Fold Des.* 7:1505–1515.
9. Brunger, A. T., P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Cryst.* 54D:905–921.
10. Buckler, D. R., Y. Zhou, and A. M. Stock. 2002. Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from *Thermotoga maritima*. *Structure (Cambridge)* 10:153–164.
11. Collaborative Computing Project No. 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50:760–763.
12. Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16:10881–10890.
13. de la Fortelle, E., and G. Bricogne. 1997. Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* 276:472–494.
14. Djordjevic, S., and A. M. Stock. 1998. Structural analysis of bacterial chemotaxis proteins: components of a dynamic signaling system. *J. Struct. Biol.* 124:189–200.
15. Echenique, J. R., S. Chapuy-Regaud, and M. C. Trombe. 2000. Competence regulation by oxygen in *Streptococcus pneumoniae*: involvement of ciaRH and comCDE. *Mol. Microbiol.* 36:688–696.
16. Echenique, J. R., and M. C. Trombe. 2001. Competence repression under oxygen limitation through the two-component MicAB signal-transducing system in *Streptococcus pneumoniae* and involvement of the PAS domain of MicB. *J. Bacteriol.* 183:4599–4608.

17. Esnouf, R. M. 1999. Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallogr. D* **55**: 938–940.
18. Fabret, C., and J. A. Hoch. 1998. A two-component signal transduction system essential for growth of *Bacillus subtilis*: implications for anti-infective therapy. *J. Bacteriol.* **180**:6375–6383.
19. Federle, M. J., K. S. McIver, and J. R. Scott. 1999. A response regulator that represses transcription of several virulence operons in the group A streptococcus. *J. Bacteriol.* **181**:3649–3657.
20. Fukuchi, K., Y. Kasahara, K. Asai, K. Kobayashi, S. Moriya, and N. Ogasawara. 2000. The essential two-component regulatory system encoded by *ycfF* and *ycfG* modulates expression of the *ftsAZ* operon in *Bacillus subtilis*. *Microbiology* **146**(Pt. 7):1573–1583.
21. Gardino, A. K., B. F. Volkman, H. S. Cho, S. Y. Lee, D. E. Wemmer, and D. Kern. 2003. The NMR solution structure of BeF(3)(–)-activated Spo0F reveals the conformational switch in a phosphorelay system. *J. Mol. Biol.* **331**:245–254.
22. Gouet, P., E. Courcelle, D. I. Stuart, and F. Metoz. 1999. ESPript: multiple sequence alignments in PostScript. *Bioinformatics* **15**:305–308.
23. Hecht, G. B., T. Lane, N. Ohta, J. M. Sommer, and A. Newton. 1995. An essential single domain response regulator required for normal cell division and differentiation in *Caulobacter crescentus*. *EMBO J.* **14**:3915–3924.
24. Im, Y. J., S. H. Rho, C. M. Park, S. S. Yang, J. G. Kang, J. Y. Lee, P. S. Song, and S. H. Eom. 2002. Crystal structure of a cyanobacterial phytochrome response regulator. *Protein Sci.* **11**:614–624.
25. Jacobs, C., I. J. Domian, J. R. Maddock, and L. Shapiro. 1999. Cell cycle-dependent polar localization of an essential bacterial histidine kinase that controls DNA replication and cell division. *Cell* **97**:111–120.
26. Jiang, M., R. B. Bourret, M. I. Simon, and K. Volz. 1997. Uncoupled phosphorylation and activation in bacterial chemotaxis. The 2.3 Å structure of an aspartate to lysine mutant at position 13 of CheY. *J. Biol. Chem.* **272**:11850–11855.
27. Johnson, L. N., and R. J. Lewis. 2001. Structural basis for control by phosphorylation. *Chem. Rev.* **101**:2209–2242.
28. Jones, S., and J. M. Thornton. 1996. Principles of protein-protein interaction. *Proc. Natl. Acad. Sci. USA* **93**:13–20.
29. Jones, S., and J. M. Thornton. 1995. Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.* **63**:31–165.
30. Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Cryst.* **32A**:922–923.
31. Kallipolitis, B. H., and H. Ingmer. 2001. *Listeria monocytogenes* response regulators important for stress tolerance and pathogenesis. *FEMS Microbiol. Lett.* **204**:111–115.
32. Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. App. Cryst.* **24**:946–950.
33. Lamzin, V. S., and A. Perrakis. 2000. Current state of automated crystallographic data analysis. *Nat. Struct. Biol.* **7**:978–985.
34. Lange, R., C. Wagner, A. de Saizieu, N. Flint, J. Molnos, M. Stieger, P. Caspers, M. Kamber, W. Keck, and K. E. Amrein. 1999. Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae*. *Gene* **237**:223–234.
35. Lee, S. Y., H. S. Cho, J. G. Pelton, D. Yan, E. A. Berry, and D. E. Wemmer. 2001. Crystal structure of activated CheY. Comparison with other activated receiver domains. *J. Biol. Chem.* **276**:16425–16431.
36. Lewis, R. J., J. A. Brannigan, K. Muchova, I. Barak, and A. J. Wilkinson. 1999. Phosphorylated aspartate in the structure of a response regulator protein. *J. Mol. Biol.* **294**:9–15.
37. Lewis, R. J., K. Muchova, J. A. Brannigan, I. Barak, G. Leonard, and A. J. Wilkinson. 2000. Domain swapping in the sporulation response regulator Spo0A. *J. Mol. Biol.* **297**:757–770.
38. Martin, P. K., T. L. D. Sun, D. P. Biek, and M. B. Schmid. 1999. Role in cell permeability of an essential two-component system in *Staphylococcus aureus*. *J. Bacteriol.* **181**:3666–3673.
39. Meritt, E. A., and D. J. Bacon. 1997. Raster 3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**:505–525.
40. Ohta, N., T. Lane, E. G. Ninfa, J. M. Sommer, and A. Newton. 1992. A histidine protein kinase homologue required for regulation of bacterial cell division and differentiation. *Proc. Natl. Acad. Sci. USA* **89**:10297–10301.
41. Parkinson, J. S., and E. C. Kofoid. 1992. Communication modules in bacterial signalling proteins. *Annu. Rev. Genet.* **26**:71–112.
42. Pirrung, M. C. 1999. Histidine kinases and two-component signal transduction systems. *Chem. Biol.* **6**:R167–R175.
43. Quon, K. C. 1996. Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* **84**:83–93.
44. Read, R. J. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst.* **42A**:140–149.
45. Robinson, V. L., D. R. Buckler, and A. M. Stock. 2000. A tale of two components: a novel kinase and a regulatory switch. *Nat. Struct. Biol.* **7**:626–633.
46. Robinson, V. L., T. Wu, and A. M. Stock. 2003. Structural analysis of the domain interface in DrrB, a response regulator of the OmpR/PhoB subfamily. *J. Bacteriol.* **185**:4186–4194.
47. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed., vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. Sola, M., F. X. Gomis-Ruth, L. Serrano, A. Gonzalez, and M. Coll. 1999. Three-dimensional crystal structure of the transcription factor PhoB receiver domain. *J. Mol. Biol.* **285**:675–687.
49. Stock, A. M., J. M. Mottonen, J. B. Stock, and C. E. Schutt. 1989. Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. *Nature* **337**:745–749.
50. Stock, J. B., A. J. Ninfa, and A. M. Stock. 1989. Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiol. Rev.* **53**:450–490.
51. Throup, J. P., K. K. Koretke, A. P. Bryant, K. A. Ingraham, A. F. Chalker, Y. Ge, A. Marra, N. G. Wallis, J. R. Brown, D. J. Holmes, M. Rosenberg, and M. K. Burnham. 2000. A genomic analysis of two-component signal transduction in *Streptococcus pneumoniae*. *Mol. Microbiol.* **35**:566–576.
52. Vagin, A., and A. Teplyakov. 1997. MOLREP: an automated program for molecular replacement. *J. App. Cryst.* **30**:1022–1025.
53. Varughese, K. I. 2002. Molecular recognition of bacterial phosphorelay proteins. *Curr. Opin. Microbiol.* **5**:142–148.
54. Wagner, C., A. Saizieu Ad, H. J. Schönfeld, M. Kamber, R. Lange, C. J. Thompson, and M. G. Page. 2002. Genetic analysis and functional characterization of the *Streptococcus pneumoniae* *vic* operon. *Infect. Immun.* **70**: 6121–6128.
55. Welch, M., N. Chinardet, L. Mourey, C. Birck, and J. P. Samama. 1998. Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. *Nat. Struct. Biol.* **5**:25–29.
56. West, A. H., and A. M. Stock. 2001. Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem. Sci.* **26**:369–376.
57. Wu, J., N. Ohta, and A. Newton. 1998. An essential, multicomponent signal transduction pathway required for cell cycle regulation in *Caulobacter*. *Proc. Natl. Acad. Sci. USA* **95**:1443–1448.
58. Zahrt, T. C., and V. Deretic. 2000. An essential two-component signal transduction system in *Mycobacterium tuberculosis*. *J. Bacteriol.* **182**:3832–3838.