Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the response regulator VicR DNA-binding domain

Chi-Hung Trinh, Yang Liu, Simon E. V. Phillips and Mary K. Phillips-Jones*

Astbury Centre for Structural Molecular Biology, Institute of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, England

Correspondence e-mail: m.k.phillips-jones@leeds.ac.uk

Received 15 August 2006
Accepted 18 October 2006

PDB Reference: VicR DNA-binding domain, 2hwv, r2hwvsf.

The response regulator VicR from the Gram-positive bacterium *Enterococcus faecalis* forms part of the two-component signal transduction system of the YycFG subfamily. The structure of the DNA-binding domain of VicR, $VicR_c$, has been solved and belongs to the winged helix–turn–helix family. It is very similar to the DNA-binding domains of *Escherichia coli* PhoB and OmpR, despite low sequence similarity, but differs in two important loops. The α-loop, which links the two helices of the helix–turn–helix motif, is similar to that of PhoB, where it has been implicated in contacting the σ subunit of RNA polymerase, but differs from that of OmpR. Conversely, the loop following the helix–turn–helix motif is similar to that of OmpR and differs from that of PhoB. YycF/VicR, PhoB and *Bacillus subtilis* PhoP regulators all recognize almost identical DNA sequences and although there is currently no experimental evidence linking this loop with the DNA, the structure is consistent with possible involvement in selective DNA recognition or binding.

1. Introduction

VicR is the response-regulator component of the VicKR (YycFG) two-component signal transduction system (TCS) of the infectious agent *Enterococcus faecalis* (Hancock & Perego, 2002, 2004). TCSs are the main mechanism by which bacteria sense and respond to environmental change and they are found in almost all species so far examined (Stock *et al.*, 2000). Typically, they comprise a membrane-located histidine protein kinase (HPK, responsible for environmental signal sensing) and a partner response-regulator protein (RR) that effects the appropriate response (usually activation or repression of specific genes). Extracellular stimuli are sensed by and serve to modulate the phosphorylation activities of the sensor HPK. Under appropriate conditions, the phosphoryl 'signal' is transferred to the cytoplasmic RR protein, which usually results in the activation of a downstream effector domain that elicits the specific adaptive response; for example, increased affinity for target promoter binding and hence gene transcription.

The VicKR system is a member of the YycFG/Vic subfamily that is essential for cell viability (Fabret & Hoch, 1998; Hancock & Perego, 2002; 2004; Ng & Winkler, 2004). Only a few members of this important subfamily have been identified and characterized to date, although genome database searches reveal further putative homologues in a large number of Gram-positive species, all of which exhibit high homology with known members (79–90% identity with *E. faecalis* VicR). Their essentiality makes them attractive targets for novel classes of antibacterial agents. The first member of the YycFG/Vic subfamily was reported in *Bacillus subtilis* (Fabret & Hoch, 1998). Phylogenetically, YycFG is similar to PhoRP (Fabret *et al.*, 1999) and these two systems have recently been shown to have regulatory links in this bacterium (Howell *et al.*, 2006). YycFG (Vic) homologues have subsequently been identified in pathogenic *Staphylococcus aureus* (Martin *et al.*, 1999), *Streptococcus pneumoniae* (Wagner *et al.*, 2002) and *Strep. mutans* (Senadheera *et al.*, 2005), as well as in *E. faecalis* (Hancock & Perego, 2002, 2004). In all cases, it is the RR proteins that have been shown to be essential, whilst the HPK components are

Table 1
Crystallographic summary for the structure of $VicR_c$.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters (Å)	$a = b = 36.24$, $c = 73.94$
Space group	$P4_3$
Resolution (Å)	36.2–1.9 (2.0–1.9)
No. of observed reflections	46492
No. of unique reflections	7524
Completeness (%)	99.5 (98.9)
$I/\sigma(I)$	11.2 (2.1)
R_{sym} (%)	5.3 (34.7)
Resolution range for refinement (Å)	12–1.9
R factor (%)	17.9
R_{free}† (%)	23.6
No. of protein residues	101
No. of water molecules	50
Average overall B factor (Å^2)	
Protein	18.3
Water	35.5
R.m.s.d bond lengths (Å)	0.013
R.m.s.d. bond angles (°)	1.5
Ramachandran analysis (%)	
Residues in most favoured regions	87.5
Residues in additional allowed regions	12.5

† R_{free} was calculated with 5% of the reflections set aside randomly.

not essential for all members of the YycFG/Vic TCS subfamily (Lange *et al.*, 1999; Throup *et al.*, 2000).

The target genes regulated by YycF/VicR include those critical for cell-wall biosynthesis or cell-membrane composition (Martin *et al.*, 1999; Howell *et al.*, 2003; Ng *et al.*, 2003, 2004, 2005; Dubrac & Msadek, 2004; Mohedano *et al.*, 2005), cell division (*ftsAZ*; Fukuchi *et al.*, 2000), virulence and genetic competence (Wagner *et al.*, 2002; Senadheera *et al.*, 2005; Ng *et al.*, 2005) and biofilms (Senadheera *et al.*, 2005). The DNA sequence recognized by this subfamily of 'essential' regulators is highly conserved amongst species (Dubrac & Msadek, 2004; Ng *et al.*, 2005; Senadheera *et al.*, 2005; Howell *et al.*, 2006). Despite an expanding list of target genes identified for these 'essential' RRs and the elucidation of the crystal structure of the highly conserved N-terminal receiver domain of one homologue, RR02rec of *Strep. pneumoniae* YycF (Bent *et al.*, 2004), no other structural information is available, including none at all on the effector domains of this essential group of DNA-binding proteins. Here, we present the crystal structure of the putative DNA-binding effector domain of VicR, $VicR_c$.

2. Experimental

2.1. Expression, purification and crystallization

The gene region encoding the predicted C-terminal domain of VicR (residues 134–233 inclusive) was PCR-amplified using *E. faecalis* V583 genome as template and primers VICRC-F (5′-GATCACATATGACGATTGGTGATTTAACCATTCA-3′) and VICRC-R (5′-AGCTGGATCCTAAAATTAACAGATTGAAAAAAG-3′). The product was cloned into pET14b (Novagen Inc.) as described previously for RegA (Potter *et al.*, 2002), so that the expressed $VicR_c$ protein possesses the additional N-terminal sequence $MGSS(H)_6$-SSGLVPRGSHMTIGDL . . . , where TIGDL are the first $VicR_c$ residues of the recombinant His-tagged protein.

Overexpression and purification of recombinant $VicR_c$ was carried out as described previously (Potter *et al.*, 2002), except that (i) protease-inhibitor cocktail (Sigma–Aldrich, UK) was added prior to cell breakage according to the manufacturer's instructions and (ii) nickel-affinity chromatography included an additional single wash with 100 m*M* imidazole prior to elution. $VicR_c$ was eluted with 200 m*M* sodium acetate buffer pH 4.0. Purified $VicR_c$ was confirmed by N-terminal sequencing (obtaining the expected GSSHHH), SDS–PAGE, Western blotting using HisProbe-HRP (Perbio Science UK Ltd) that detects the hexa-His tag, electrospray mass spectrometry (obtaining 14 031.54 Da, matching the expected 14 031.79 Da) and CD spectroscopy, as described previously (Potter *et al.*, 2002). Protein was purified to 95% as estimated by densitometry and concentrated to 8 mg ml^{-1} prior to crystallization. Crystals of $VicR_c$ were grown by vapour diffusion using the sitting-drop method in 1.8 *M* ammonium sulfate, 0.1 *M* sodium citrate pH 4.5, 3% dimethyl sulfoxide. Crystals were transferred to a cryoprotectant buffer consisting of the crystallization buffer with 25% glycerol prior to flash-cooling in liquid nitrogen.

2.2. Data collection, structure determination and refinement

Data were recorded at 100 K with an R-AXIS IV^{++} image-plate detector mounted on a Rigaku RU-H3R rotating-anode X-ray generator and integrated using *MOSFLM* (Leslie, 1992) to 1.9 Å resolution. Data reduction and subsequent calculations were carried out using the *CCP*4 progam suite (Collaborative Computational Project, Number 4, 1994). The crystals belong to space group $P4_3$, with unit-cell parameters $a = b = 36.24$, $c = 73.94$ Å and one molecule of $VicR_c$ per asymmetric unit (Table 1). The structure was determined by molecular replacement using the program *Phaser* (Read, 2001). Initial search models were based on the C-terminal DNA-binding domains of the response regulatory proteins OmpR from *E. coli* (Martinez-Hackert & Stock, 1997; PDB code 1opc), PrrA from *Mycobacterium tuberculosis* (Nowak *et al.*, 2006; PDB code 1ys6), DrrB from *Thermotoga maritima* (Robinson *et al.*, 2003; PDB code 1p2f) and PhoB from *E. coli* (Blanco *et al.*, 2002; PDB code 1gxq), with sequences having approximately 25, 35, 33 and 34% identity to $VicR_c$, respectively. Each of these initial models individually failed to produce a clear molecular-replacement solution, which was only found from *Phaser* when the molecular-replacement model was constructed using an ensemble of these four structures. Prior to the molecular-replacement search, the four homologous structures were superimposed using SSM superposition (Krissinel & Henrick, 2004) in *Coot* (Emsley & Cowtan, 2004) and regions showing large deviations were removed. Model building and refinement were carried out using *Coot* and *REFMAC*5 (Murshudov *et al.*, 1997), allowing the deleted regions to be rebuilt. In the later stages of refinement, TLS parameters (Winn *et al.*, 2001) based on a single-group TLS model calculated from the TLS Motion Determination server (http://skuld.bmsc.washington.edu/~tlsmd/) were refined in *REFMAC*5, leading to a final R factor and R_{free} of 17.9 and 23.6%, respectively.

3. Results and discussion

In the refined structure, the $2F_o - F_c$ electron-density maps revealed poorly ordered density at the N- and C-termini of $VicR_c$. 18 residues at the N-terminus were not visible in the density [including the six-residue His tag and five of the six residues (LVPRG) of a thrombin cleavage site originating from the plasmid vector] and the two C-terminal residues Gln233 and Glu234. The model of $VicR_c$ contains 101 residues, starting with Ser-His-Met from the plasmid vector at the N-terminus, followed by Thr135–Glu232 of $VicR_c$. The electrospray mass spectrum of the $VicR_c$ sample used for crystallization gives a molecular weight of the protein that is consistent with the expected 121 amino-acid residues.

3.1. Overview of the VicR$_c$ structure

The structure of VicR$_c$ is highly similar to the DNA-binding domains of other members of the OmpR/PhoB subfamily. It comprises an N-terminal four-stranded antiparallel β-sheet (residues 134–136, 139–141, 147–150 and 153–156), a central three α-helical bundle (α1, α2 and α3; residues 159–170, 178–185 and 194–208, respectively) and a C-terminal β-hairpin turn (residues 210–215) (Fig. 1) and belongs to the winged helix–turn–helix family of DNA-binding motifs. On DNA binding, the recognition helix α3 would be expected to occupy the major groove and to interact with bases of the consensus sequence. The loop linking α2 and α3 has been designated the α-loop in OmpR and the transactivation loop in PhoB, where it probably interacts with the σ^{70} subunit of RNA polymerase and is essential for transcription activation (Kim *et al.*, 1995; Makino *et al.*, 1996).

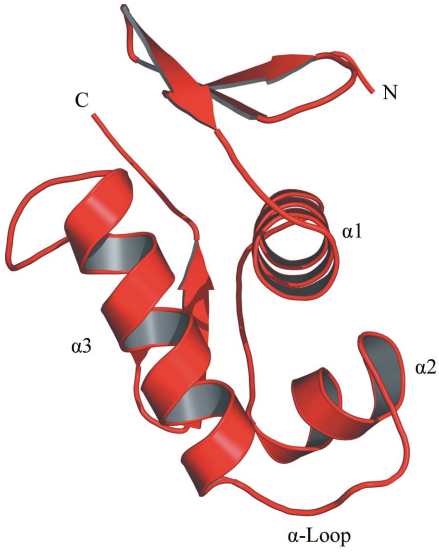

Figure 1
Ribbon diagram of VicR$_c$. Helices α2 and α3 form the helix–turn–helix motif, linked by the α-loop. All figures were created with *PyMOL* (DeLano, 2002).

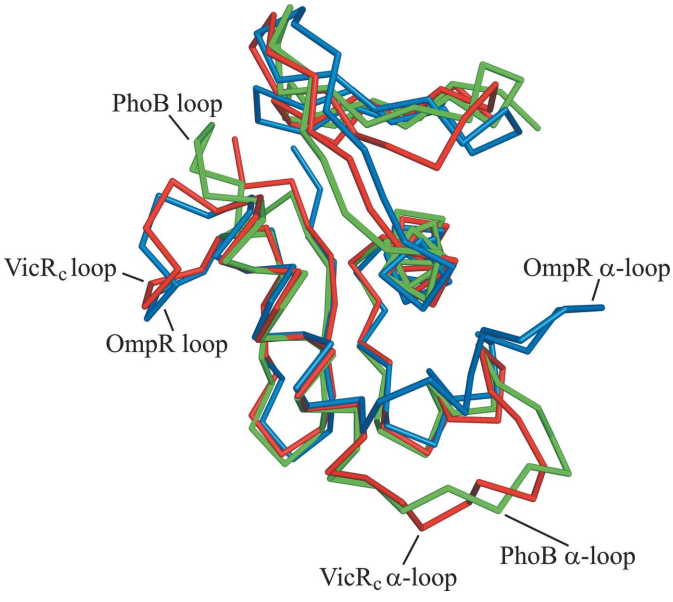

Figure 2
Structural comparison of the C$^{\alpha}$ backbones of VicR$_c$ (red), PhoB (green) and OmpR (blue). The α-loops and the loops linking α3 to the C-terminal β-hairpin are labelled.

3.2. Comparison with the OmpR/PhoB subfamily

VicR$_c$ superimposes closely on representative members of the OmpR/PhoB family, OmpR, PhoB, PrrA, DrrB and DrrD (Buckler *et al.*, 2002; PDB code 1kgs), with r.m.s.d. values of 1.5, 1.1, 1.2, 1.3 and 1.4 Å, respectively, when using C$^{\alpha}$ atoms only. Fig. 2 only shows the superposition of VicR$_c$ on OmpR and PhoB, with the other structures omitted for clarity. Although the sequence identities of VicR$_c$ with OmpR and PhoB are low (25 and 32%, respectively), the overall structures are almost identical. This is perhaps not surprising as the conserved residues play a key role in linking the secondary-structure elements responsible for DNA recognition (Blanco *et al.*, 2002). The differences arise mainly for two loop regions: the α-loop and the loop region following α3 and linking it to the C-terminal β-hairpin (Fig. 2). Both these loops were removed when the ensemble model was prepared for molecular replacement.

The positioning of the α-loop of VicR$_c$ closely matches that of the transactivation loop of PhoB, but not the equivalent α-loop of OmpR, which is positioned differently (Okamura *et al.*, 2000) and which, in common with *B. subtilis* PhoP, has a longer loop and shorter α3 than PhoB (Chen *et al.*, 2004). Both the OmpR and PhoB loops are postulated to interact with RNA polymerase for transcription activation. Evidence suggests that PhoB interacts with the σ^{70} subunit of RNA polymerase (Kim *et al.*, 1995; Makino *et al.*, 1996), whilst OmpR might associate with the α subunit (Sharif & Igo, 1993; Pratt & Silhavy, 1994; Kondo *et al.*, 1997). Therefore, our structural data here may suggest possible interactions of VicR$_c$ with the σ^{70} rather than

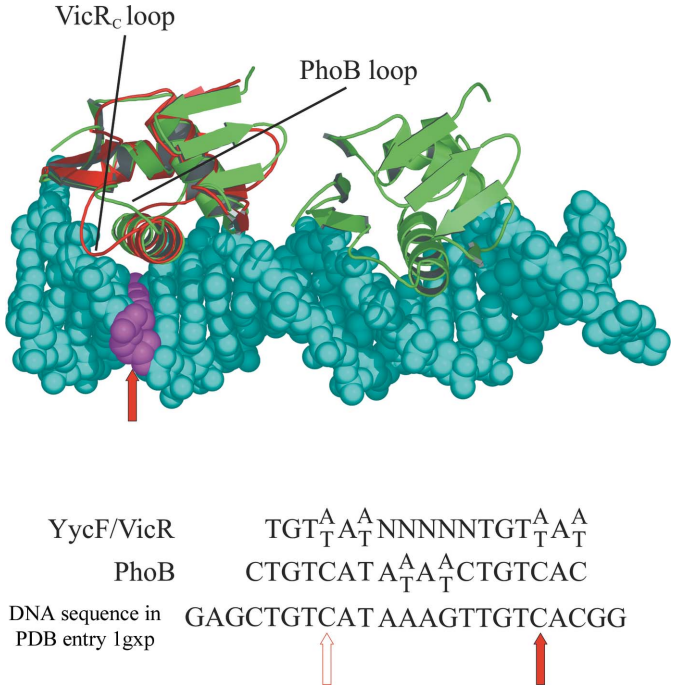

Figure 3
Ribbon representation showing two molecules of PhoB (green) in complex with DNA shown in space-filling representation (Blanco *et al.*, 2002; PDB code 1gxp). VicR$_c$ is shown in red, superimposed onto one PhoB monomer. The consensus sequences for YycF/VicR homologues and PhoB targets are also shown, with the C to A/T substitution mentioned in the text indicated by red arrows in each repeat. The filled red arrow indicates the corresponding base pair in the model, which has been coloured magenta and lies close to the recognition helix. The loop region of PhoB does not reach the DNA, while the VicR$_c$ loop could make contact with it.

the α subunit of RNA polymerase in *E. faecalis*. PhoB molecules bind DNA head-to-tail in a tandem arrangement (Okamura *et al.*, 2000; Blanco *et al.*, 2002), whilst for global OmpR binding is head-to-tail (Harlocker *et al.*, 1995) or head-to-head (Maris *et al.*, 2005). The orientation of $VicR_c$ binding to DNA has not yet been investigated. The position of the loop following $\alpha3$ in $VicR_c$ is different in PhoB, but superimposes well with that of OmpR, although its role in PhoB and OmpR is uncertain.

3.3. Modelling of the $VicR_c$–DNA complex

The specific DNA target sequence for PhoB, the *pho* box, was identified as two seven-base-pair direct repeats separated by an AT-rich region of four base pairs (Mizuno, 1997). The corresponding consensus sequence for YycF (and presumably also $VicR_c$) is almost identical, with one base change from C in PhoB to A or T in $VicR_c$ at position 5 in the repeat (with respect to the PhoB consensus; Fig. 3). This suggests that the protein–DNA interactions made by PhoB with its consensus DNA should be very similar to those in the $VicR_c$–DNA complex. Fig. 3 shows a model built by superimposing $VicR_c$ on the PhoB–DNA complex (Blanco *et al.*, 2002). In PhoB, residues Trp184 (at the C-terminal end of $\alpha2$) and Arg193, Thr194, His198 and Arg200 (from $\alpha3$) were all shown to be involved in specific DNA recognition (Okamura *et al.*, 2000; Blanco *et al.*, 2002) and all except His198 are conserved in VicR. In OmpR, Val203, Ser206 and Arg210 of $\alpha3$ have been shown to be important for target specificity (Mizuno & Tanaka, 1997). The equivalent residues in PhoB are Val197, which packs against the CG base pair at position 5, Arg200, which is nearby but not in contact, interacting with the phosphate backbone, and Lys204, which lies at the end of $\alpha3$ but does not contact the DNA. The corresponding residues in $VicR_c$ are Val199, Arg202 and Glu206. The first two are unchanged and Glu206 is too far away from the DNA to make contact, although it is close enough to influence the conformation of Arg202, which lies near the single base pair that differs between the consensus sequences. The following loop in PhoB is too short to contact the DNA, but in $VicR_c$ and OmpR it is longer (Figs. 2 and 3) and might play a role in DNA–protein interactions, although this loop has no previously reported association with DNA and whether the positions of these loops differ upon binding to DNA is not yet known.

We thank Dr Alison Ashcroft (University of Leeds) for electrospray mass spectrometry and Dr Jeff Keen (University of Leeds) for N-terminal sequencing. We are grateful to BBSRC for financial assistance (BB/D001641/1).

References

Bent, C. J., Isaacs, N. W., Mitchell, T. J. & Riboldi-Tunnicliffe, A. (2004). *J. Bacteriol.* **186**, 2872–2879.
Blanco, A. G., Sola, M., Gomis-Ruth, F. X. & Coll, M. (2002). *Structure*, **10**, 701–713.
Buckler, D. R., Zhou, Y. & Stock, A. M. (2002). *Structure*, **10**, 153–164.
Chen, Y. H., Abdel-Fattah, W. R. & Hulett, F. M. (2004). *J. Bacteriol.* **186**, 1493–1502.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA.
Dubrac, S. & Msadek, T. (2004). *J. Bacteriol.* **186**, 1175–1181.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fabret, C., Feher, V. A. & Hoch, J. A. (1999). *J. Bacteriol.* **181**, 1975–1983.
Fabret, C. & Hoch, J. A. (1998). *J. Bacteriol.* **180**, 6375–6383.
Fukuchi, K., Kasahara, Y., Asai, K., Kobayashi, K., Moriya, S. & Ogasawara, N. (2000). *Microbiology*, **146**, 1573–1583.
Hancock, L. E. & Perego, M. (2002). *J. Bacteriol.* **184**, 5819–5825.
Hancock, L. E. & Perego, M. (2004). *J. Bacteriol.* **186**, 7951–7958.
Harlocker, S. L., Bergstrom, L. & Inouye, M. (1995). *J. Biol. Chem.* **270**, 26849–26856.
Howell, A., Dubrac, S., Krogh Anderson, K., Noone, D., Fert, J., Msadek, T. & Devine, K. (2003). *Mol. Microbiol.* **49**, 1639–1655.
Howell, A., Dubrac, S., Noone, D., Varughese, K. I. & Devine, K. (2006). *Mol. Microbiol.* **59**, 1199–1215.
Kim, S.-K., Makino, K., Amemura, M., Nakata, A. & Shinagawa, H. (1995). *Mol. Gen. Genet.* **248**, 1–8.
Kondo, H., Nakagawa, A., Nishihira, J., Nishimura, Y., Mizuno, T. & Tanaka, I. (1997). *Nature Struct. Biol.* **4**, 28–31.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Lange, R., Wagner, C., de Saizieu, A., Flint, N., Molnos, J., Stieger, M., Casper, P., Kamber, M., Keck, W. & Amrein, K. E. (1999). *Gene*, **237**, 223–234.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Makino, K., Amemura, M., Kawamoto, T., Kimura, S., Shinagawa, H., Nakata, A. & Suzuki, M. (1996). *J. Mol. Biol.* **259**, 15–26.
Maris, A. E., Walthers, D., Mattison, K., Byers, N. & Kenney, L. J. (2005). *J. Mol. Biol.* **350**, 843–856.
Martin, P. K., Li, T., Sun, D., Biek, D. P. & Schmid, M. B. (1999). *J. Bacteriol.* **181**, 3666–3673.
Martinez-Hackert, E. & Stock, A. M. (1997). *Structure*, **5**, 109–124.
Mizuno, T. (1997). *DNA Res.* **4**, 161–168.
Mizuno, T. & Tanaka, I. (1997). *Mol. Microbiol.* **24**, 665–670.
Mohedano, M. L., Overweg, K., de la Fuente, A., Reuter, M., Altabe, S., Mulholland, F., de Mendoza, D., Lopez, P. & Wells, J. M. (2005). *J. Bacteriol.* **187**, 2357–2367.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Ng, W.-L., Kazmierczak, K. M. & Winkler, M. E. (2004). *Mol. Microbiol.* **53**, 1161–1175.
Ng, W.-L., Robertson, G. T., Kazmierczak, K. M., Zhao, J., Gilmour, R. & Winkler, M. E. (2003). *Mol. Microbiol.* **50**, 1647–1663.
Ng, W.-L., Tsui, H.-C. T. & Winkler, M. E. (2005). *J. Bacteriol.* **187**, 7444–7459.
Ng, W.-L. & Winkler, M. E. (2004). *Microbiology*, **150**, 3096–3098.
Nowak, E., Panjikar, S., Konarev, P., Svergun, D. I. & Tucker, P. A. (2006). *J. Biol. Chem.* **281**, 9659–9666.
Okamura, H., Hanaoka, S., Nagadoi, A., Makino, K. & Nishimura, Y. (2000). *J. Mol. Biol.* **295**, 1225–1236.
Potter, C. A., Ward, A., Laguri, C., Williamson, M. P., Henderson, P. J. F. & Phillips-Jones, M. K. (2002). *J. Mol. Biol.* **320**, 201–213.
Pratt, L. A. & Silhavy, T. J. (1994). *J. Mol. Biol.* **243**, 579–594.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Robinson, V. L., Wu, T. & Stock, A. M. (2003). *J. Bacteriol.* **185**, 4186–4194.
Senadheera, M. D., Guggenheim, B., Spatafora, G. A., Huang, Y.-C. C., Choi, J., Hung, D. C. I., Treglown, J. S., Goodman, S. D., Ellen, R. P. & Cvitkovitch, D. G. (2005). *J. Bacteriol.* **187**, 4064–4076.
Sharif, R. T. & Igo, M. M. (1993). *J. Bacteriol.* **175**, 5460–5468.
Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). *Annu. Rev. Biochem.* **69**, 183–215.
Throup, J. P., Koretke, K. K., Bryant, A. P., Ingraham, K. A., Chalker, A. F., Ge, Y., Marra, A., Wallis, N. G., Brown, J. R., Holmes, D. J., Rosenberg, M. & Burnham, M. K. R. (2000). *Mol. Microbiol.* **35**, 566–576.
Wagner, C., de Saizieu, A., Schonfeld, H.-J., Kamber, M., Lange, R., Thompson, C. J. & Page, M. G. (2002). *Infect. Immun.* **70**, 6121–6128.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.