M. J. Domagalski,^{a,b,†} K. L. Tkaczuk,^{a,b,†} M. Chruszcz,^{a,b} T. Skarina,^{b,c} O. Onopriyenko,^{b,c} M. Cymborowski,^{a,b} M. Grabowski,^{a,b} A. Savchenko,^{b,c} and W. Minor^{a,b,*}

^aDepartment of Molecular Physiology and Biological Physics, University of Virginia, 1340 Jefferson Park Avenue, Jordan Hall, Charlottesville, VA 22908, USA, ^bCenter for Structural Genomics of Infectious Diseases (CSGID), USA, and ^cBanting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, ON M5G 1L6, Canada

† These authors made equal contributions.

Correspondence e-mail: wladek@iwonka.med.virginia.edu

Received 20 March 2013

Accepted 30 July 2013

PDB Reference: DhbC, 3os6

© 2013 International Union of Crystallography
All rights reserved

Structure of isochorismate synthase DhbC from *Bacillus anthracis*

The isochorismate synthase DhbC from *Bacillus anthracis* is essential for the biosynthesis of the siderophore bacillibactin by this pathogenic bacterium. The structure of the selenomethionine-substituted protein was determined to 2.4 Å resolution using single-wavelength anomalous diffraction. *B. anthracis* DhbC bears the strongest resemblance to the *Escherichia coli* isochorismate synthase EntC, which is involved in the biosynthesis of another siderophore, namely enterobactin. Both proteins adopt the characteristic fold of other chorismate-utilizing enzymes, which are involved in the biosynthesis of various products, including siderophores, menaquinone and tryptophan. The conservation of the active-site residues, as well as their spatial arrangement, suggests that these enzymes share a common Mg²⁺-dependent catalytic mechanism.

1. Introduction

Like many other microorganisms, pathogenic bacteria use siderophores, small high-affinity iron-chelating compounds, for the purpose of iron acquisition, which is an important process in microorganism development and function. In an iron-deficient environment in a host organism, siderophores give the infectious pathogen the ability to obtain this fundamental element from the host's iron-containing proteins. Without this capability, the pathogenic microorganism would be unable to grow and would gradually be defeated by the host's immune system or by iron starvation (Ratledge & Dover, 2000). *Bacillus anthracis* produces two catecholate siderophores, petrobactin (associated with the *asb* operon) and bacillibactin (associated with the *bac* operon), with the latter only being expressed under iron-limited conditions (Lee *et al.*, 2011). Five genes are involved in the bacillibactin biosynthetic pathway: *dhbABCEF* (Hotta *et al.*, 2010). The main enzymes involved in the assembly of bacillibactin are encoded by the genes of the *entA-dhbBCF* cluster. In the evolutionarily closely related species *B. subtilis*, two distinct isochorismate synthase genes, *dhbC* and *menF*, are involved in the biosynthesis of the siderophore 2,3-dihydroxybenzoate (DHB) and the respiratory-chain component menaquinone (MK), respectively. The studies of Rowland and Taber showed that DhbC can compensate for a lack of MenF, but the opposite is not true. Moreover, depletion of *dhbC* results in the absence of DHB (Rowland & Taber, 1996). Also, *dhbC* expression is strongly regulated by the iron concentration (Rowland & Taber, 1996; Baichoo *et al.*, 2002), which is not the case for *menF*. The aforementioned studies show the importance of DhbC in siderophore biosynthesis.

2. Materials and methods

Protocols and a detailed history of experiments for *B. anthracis* DhbC are publically available at http://csgid.org/csgid/space_tree/view/IDP01205.

2.1. Cloning, expression and purification

The open reading frame of *dhbC* was amplified by polymerase chain reaction from *B. anthracis* strain Ames genomic DNA using forward 5'-TACTTCCAATCCAATGCGATGAATGAATTTACG-

GCTGTAAA-3' and reverse 5'-TTATCCACTTCCAATGCTACT-TTTCATTAAGTGAAGTATC-3' primers. The gene was cloned into a pMCSG19c plasmid using ligation-independent cloning (Aslanidis & de Jong, 1990; Haun & Moss, 1992; Eschenfeldt *et al.*, 2009). This is a fusion expression vector which encodes an N-terminal maltose-binding protein (MBP) cleavable by *Tobacco vein mottling virus* (TVMV) protease and a hexahistidine tag cleavable by *Tobacco etch virus* protease (MHHHHHHSSGVDLGTEENLYFQ/SNA). This vector also carries TVMV protease, which allows *in vivo* MBP cleavage (Donnelly *et al.*, 2006). Expression levels in the first-choice vector, pMCSG7, were significantly lower and were not sufficient for protein purification. The gene was overexpressed in *Escherichia coli* BL21-CodonPlus(DE3)-RIPL cultures. Cells were grown in M9 SeMET High-Yield growth medium (Shanghai Medicilon) at 310 K to an optical density (at 600 nm) of approximately 1.2, induced with 0.4 mM isopropyl β -D-1-thiogalactopyranoside and grown overnight with shaking at 293 K. Use of selenomethionine-substituted protein was imposed by the lack of a high sequence-identity template (>30%) for molecular replacement. The harvested cells were sonicated in lysis buffer [300 mM NaCl, 50 mM HEPES pH 7.5, 5% (v/v) glycerol, 5 mM imidazole, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM benzamidine] and clarified by centrifugation; the supernatant was applied onto nickel-chelate affinity resin (Ni-NTA; Qiagen). The resin was washed with wash buffer [300 mM NaCl, 50 mM HEPES pH 7.5, 5% (v/v) glycerol, 30 mM imidazole] at 277 K and the protein was eluted using elution buffer [300 mM NaCl, 50 mM HEPES pH 7.5, 5% (v/v) glycerol, 250 mM imidazole]. The hexahistidine tag was cleaved from the protein by the addition of 1 mg recombinant His-tagged TEV protease per 15 mg of eluted protein in the presence of EDTA, TCEP and arginine (final concentrations of 1, 0.5 and 200 mM, respectively). Arginine was included in the buffer in order to suppress protein aggregation (Tsumoto *et al.*, 2004; Arakawa *et al.*, 2007). The cleavage was performed at 277 K overnight and continued during dialysis into cleavage buffer (300 mM NaCl, 50 mM HEPES pH 7.5, 0.5 mM TCEP). Cleaved protein was separated from TEV protease by running over nickel-chelating resin.

2.2. Crystallization

The crystals of selenomethionine-incorporated DhhC used for data collection were grown by the hanging-drop vapor-diffusion method. The well solution consisted of 2 M ammonium sulfate, 2% (v/v) PEG 400, 100 mM HEPES pH 7.5, 50 mM bis-tris pH 5.5. Drops were formed by mixing 2 μ l well solution and 2 μ l 16 mg ml⁻¹ protein in 300 mM NaCl, 10 mM HEPES pH 7.5, 0.5 mM TCEP. Crystals were grown at room temperature and formed after one week of incubation. Immediately after harvesting, the crystals were transferred into cryoprotectant solution consisting of 7% glycerol, 7% sucrose, 7% ethylene glycol in mother liquor, passed through Paratone oil and flash-cooled in liquid nitrogen.

2.3. Data collection and processing

Data were collected at 100 K on the 19-BM beamline of the Structural Biology Center (Rosenbaum *et al.*, 2006) at the Advanced Photon Source (Argonne National Laboratory, Argonne, Illinois, USA) controlled by HKL-3000 (Otwinowski & Minor, 1997; Minor *et al.*, 2006). Diffraction data were processed with HKL-3000 and can be obtained through the CSGID webpage (http://www.csgid.org/csgid/pages/diffraction_images). Data-collection, structure-determination and refinement statistics are summarized in Table 1.

Table 1

Data-collection, structure-determination and refinement statistics for the crystal structure of DhhC from *B. anthracis* (PDB entry 3os6).

Values in parentheses are for the highest resolution shell.

Data collection	
Wavelength (Å)	0.9793
Space group	<i>P</i> 2 ₁ 3
Unit-cell parameters (Å, °)	<i>a</i> = <i>b</i> = <i>c</i> = 201.4, $\alpha = \beta = \gamma = 90$
Resolution (Å)	50.00–2.40 (2.44–2.40)
No. of unique reflections	105217
Completeness (%)	99.40 (100)
Multiplicity	3.7 (3.7)
Mean $\langle I/\sigma(I) \rangle$	20.0 (2.2)
Molecules in asymmetric unit	4
Matthews coefficient (Å ³ Da ⁻¹)	3.78
Solvent content (%)	67.5
$R_{\text{merge}}^{\dagger}$	0.048 (0.532)
$R_{\text{meas}}^{\ddagger}$	0.054 (0.620)
$R_{\text{p.i.m.}}^{\S}$	0.026 (0.312)
Structure refinement	
R_{work}	0.171 (0.229)
R_{free}	0.212 (0.262)
No. of residues	1534
No. of protein atoms	11630
No. of water atoms	763
Average <i>B</i> factors (Å ²)	
Main chain	42
Side chains	47
Overall	44
Waters	38
Ramachandran plot ¶ , residues in (%)	
Most favored regions	97.9
Allowed regions	2.1
Disallowed regions	0.0
R.m.s. deviations from ideal values ††	
Bond lengths (Å)	0.018
Bond angles (°)	1.71
<i>MolProbity</i> ‡† statistics	
Score	1.36
Clashscore	4.39
Poor rotamers (%)	1.03

† $R_{\text{merge}} = \frac{\sum_i \sum_j |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_i \sum_j I_i(hkl)}$. ‡ $R_{\text{meas}} = \frac{\sum_i \{N(hkl) / [N(hkl) - 1]\}^{1/2} \sum_j |I_j(hkl) - \langle I(hkl) \rangle|}{\sum_i \sum_j I_j(hkl)}$. § $R_{\text{p.i.m.}} = \frac{\sum_i \{1 / [N(hkl) - 1]\}^{1/2} \sum_j |I_j(hkl) - \langle I(hkl) \rangle|}{\sum_i \sum_j I_j(hkl)}$. ¶ Emsley & Cowtan (2004). †† Engh & Huber (1991). ‡† Chen *et al.* (2010).

2.4. Structure solution and refinement

The structure of the selenomethionine-substituted protein was determined using single-wavelength anomalous diffraction (SAD) and initial models were built with HKL-3000. HKL-3000 is integrated with *SHELXC/D/E* (Sheldrick, 2008), *MLPHARE* (Otwinowski, 1991), *DM* (Cowtan & Main, 1993; Cowtan & Zhang, 1999), *ARP/wARP* (Perrakis *et al.*, 1999), *CCP4* (Winn *et al.*, 2011), *SOLVE* and *RESOLVE* (Terwilliger, 2004). The resulting model was further refined with *REFMAC5* (Murshudov *et al.*, 2011) and *Coot* (Emsley & Cowtan, 2004). *MolProbity* (Chen *et al.*, 2010) and *ADIT* (Yang *et al.*, 2004) were used for structure validation. The coordinates and experimental structure factors were deposited in the Protein Data Bank (PDB) with accession code 3os6.

2.5. Bioinformatics analyses

Sequence-homolog searches were performed with *PSI-BLAST* (Altschul *et al.*, 1997). Structural homolog searches were performed with *DALI* (Holm & Rosenström, 2010) and *HHpred* (Söding, 2005; Söding *et al.*, 2005). Structure superposition was performed with *SSM* (Krissinel & Henrick, 2004).

2.6. Spectrophotometric activity assay

An enzyme-activity assay was performed to detect isochorismate synthase activity (Fig. 1) and its dependence on Mg^{2+} ions. The assay monitors the formation of isochorismate by measuring the increase in absorbance at 278 nm (He & Toney, 2006). The kinetic assay was performed using a PHERAstar FS microplate reader at 303 K for 10 min. The 100 μl reaction mixture consisted of 50 mM HEPES pH 7.5, 300 mM NaCl, 5 mM MgCl_2 , 10 μg Dhbc, 1 mM chorismic acid. Samples were incubated for at least 10 min at room temperature prior to the addition of chorismate and measurement. The absorbance was monitored every 60 s during 10 min of reaction. Average changes in the absorbance were recorded for three repeats of the experiment. The isochorismate concentration was calculated using $\Delta\epsilon_{\text{isochorismate-chorismate}} = 10\,211\text{ M}^{-1}\text{ cm}^{-1}$.

3. Results and discussion

3.1. Overall structure

B. anthracis Dhbc crystallized in a primitive cubic space group ($P2_13$) with two dimers in the asymmetric unit, which is consistent with the gel-filtration results and PISA server analysis (Krissinel & Henrick, 2007). The structure was determined using the single-wavelength anomalous diffraction method and a selenomethionine-substituted data set (data-collection and refinement parameters are shown in Table 1). Several fragments were not modeled owing to a lack of electron density, including the N-terminal 8–9 residues (Fig. 2), the $\beta 6$ – $\beta 7$ loop and some residues in the $\beta 9$ – $\beta 10$ loop. Structural similarity searches using DALI (Holm & Rosenström, 2010) suggest that Dhbc belongs to the aminodeoxychorismate (ADC) synthase domain family according to the SCOP classification (SCOP class

d.161.1.1; Murzin *et al.*, 1995) and the PF00425 family (which contains chorismate-binding enzymes) in the Pfam classification (Bateman *et al.*, 2004). Dhbc adopts the ADC synthase-like fold containing four repeats of an α – β_2 – β motif arranged in a four-layer core structure, where the layers are $\alpha/\beta/\alpha$ with orthogonally packed β -sheets. The first β -sheet is comprised of $\beta 6$, $\beta 2$, $\beta 1$, $\beta 9$, $\beta 10$, $\beta 19$, $\beta 20$, $\beta 8$ and $\beta 7$. The second β -sheet is comprised of $\beta 3$, $\beta 5$, $\beta 4$, $\beta 17$, $\beta 18$, $\beta 11$, $\beta 12$, $\beta 16$ and $\beta 14$, and the small third sheet contains $\beta 13$, $\beta 16$ and $\beta 15$. The β -sheets are surrounded by nine α -helices. The overall structure of Dhbc is shown in Fig. 3.

3.2. Structural homologs and characteristics of Dhbc

The structures of several isochorismate-utilizing enzymes have previously been solved by X-ray crystallography (Table 2). Structurally, Dhbc strongly resembles isochorismate synthase EntC from *E. coli* (Sridharan *et al.*, 2010), as shown in Fig. 4. This enzyme is part of the *E. coli* biosynthesis pathway of the enterobactin siderophore (Sridharan *et al.*, 2010). Like *B. anthracis* and *B. subtilis*, *E. coli* also has two isochorismate synthase genes, but *entC* is not able to fully restore menaquinone deficiency in mutants with a disrupted *menF* gene (Dahm *et al.*, 1998). HHpred analysis (Söding, 2005) showed that the salicylate synthetases Irp9 from *Yersinia enterocolitica* (Kerbarh *et al.*, 2006) and MbtI from *Mycobacterium tuberculosis* (Manos-Turvey *et al.*, 2010) as well as the isochorismate synthases MenF from *E. coli* (Parsons *et al.*, 2008) and *Y. pestis* involved in the biosynthesis of menaquinone, the anthranilate synthase TrpE from *Salmonella typhimurium* (Morollo & Eck, 2001) and the 2-amino-2-desoxyisochorismate (ADIC) synthase PhzE from *Burkholderia* sp. (Li *et al.*, 2011) are also closely related structurally to Dhbc. All of these belong to the same structural family according to SCOP and

Figure 1 Bacillibactin biosynthetic pathway of *B. anthracis*. The reaction catalyzed by isochorismate synthase Dhbc is bordered by a red frame.

Table 2List of structural homologs of *B. anthracis* Dhbc identified by DALI and HHpred.

Structure superposition was performed with SSM.

Protein name	Organism	PDB code	Sequence identity (%)	R.m.s.d. (Å)	No. of superposed residues
Isochorismate synthase EntC	<i>E. coli</i>	3hwo (Sridharan <i>et al.</i> , 2010)	38	1.3	354
Salicylate synthetase Irp9	<i>Y. enterocolitica</i>	2fn0 (Kerbarh <i>et al.</i> , 2006)	25	2.0	323
Salicylate synthetase Irp9	<i>M. tuberculosis</i>	3log (Manos-Turvey <i>et al.</i> , 2010)	22	2.0	327
Isochorismate synthase MenF	<i>E. coli</i>	3bzm (Parsons <i>et al.</i> , 2008)	25	1.9	336
Isochorismate synthase MenF	<i>Y. pestis</i>	3gse (Center for Structural Genomics of Infectious Diseases, unpublished work)	22	2.2	324
Anthranilate synthase	<i>S. typhimurium</i>	1i1q (Morollo & Eck, 2001)	14	2.3	322
2-Amino-2-desoxyisochorismate (ADIC) synthase PhzE	<i>Burkholderia</i> sp.	3r75 (Li <i>et al.</i> , 2011)	16		324

they all take part in the conversion of chorismate to isochorismate (EntC and MenF), salicylate (Irp9 and MbtI), anthranilate (TrpE) or ADIC (PhzE).

3.3. Active site

The active site of *B. anthracis* Dhbc bears a strong resemblance to that of EntC from *E. coli*, as shown in Fig. 5. All catalytically

important residues are evolutionarily conserved with the exception of two: in *B. anthracis* Dhbc Leu304 is substituted by the chemically related amino acid Val305, while Phe359 is substituted by Tyr360. The latter mutation is also present in Dhbc from *B. subtilis* and in MenF from *E. coli* and *Y. pestis*; therefore, this mutation may not cause a significant change in enzyme activity. In contrast to *E. coli* EntC, neither the product of the Dhbc reaction nor a magnesium ion is bound in the active site of Dhbc. In EntC, a magnesium ion is coordinated between the side chains of Glu241 and Glu376 and the C1 carboxylate of bound isochorismate. In Dhbc the side chain of Glu241 is rotated to the outside of the active site, opening the chorismate-binding pocket. In the position corresponding to the isochorismate C1 carboxylate in EntC, a sulfate ion is coordinated by the conserved Lys381 and Ser215 and the backbone N atoms of Gly214 on one side and Gly364 on the opposite side. Comparison with EntC shows that binding of the reaction product in the active site does not affect the arrangement of the enol-pyruvyl holding residues Lys381, Ile347 and Arg348 (Lys380, Ile346 and Arg347 in EntC). The orientation of the aromatic residue pair Phe328 and Tyr360 in Dhbc differs from the orientation of the corresponding residues Phe327 and Phe359 in EntC, but is in agreement with the orientation of corresponding pairs in other chorismate-utilizing enzymes.

Experimental analyses by Sridharan *et al.* (2010) showed that Ala303 in EntC from *E. coli* (Ala304 in Dhbc from *B. anthracis*) plays an important role in positioning the peptide-bond carbonyl, enabling the formation of a proper hydrogen bond to the isochorismate C2 hydroxyl. This residue is mutated to Ser338 in *Cytophaga hutchinsonii* PabB, to Ser366 in *E. coli* PabB and to Thr348 in *Y. enterocolitica* Irp9. Mutation of the adjacent amino acid Leu304 of *E. coli* EntC to alanine resulted in loss of activity. The strict conser-

Figure 2

Typical ($2mF_o - DF_c$) electron density. A portion of the electron density covering the N-terminal residues of the crystal structure of Dhbc from *B. anthracis* is shown with a refined model and contoured at the 1.0σ level.

Figure 3

Structure of Dhbc from *B. anthracis* strain Ames (PDB entry 3os6) shown in cross-eyed stereo. The structure is colored according to the secondary-structure element order from deep blue (N-terminus) to deep red (C-terminus). The N- and C-termini, as well as the secondary-structure elements, are labeled.

vation of both of these residues in Dhbc, EntC and MenF indicates their importance in sustaining isochorismate synthase activity. Lys147 of *E. coli* (Lys142* in *B. anthracis*) is thought to act as a catalytic base by activating a nucleophilic water molecule which is also hydrogen-bonded to the C2 hydroxyl group of isochorismate (He & Toney, 2006). This residue is substituted by Glu182 in the ADC synthase from *C. hutchinsonii* (PDB entry 3h9m; New York SGX Research Center for Structural Genomics, unpublished work) and by Gln210 in that from *E. coli* (PDB entry 1k0e; Parsons *et al.*, 2002). Both of these residues possess different biochemical properties that can result in differences in activities.

Figure 4
Superposition of *B. anthracis* Dhbc (PDB entry 3os6) and *E. coli* EntC (PDB entry 3hwo; Sridharan *et al.*, 2010) isochorismate synthases. Both structures are presented in cartoon representation, with Dhbc in cyan and EntC in gray; isochorismate is shown in sphere representation and the N- and C-termini are marked with red labels.

Figure 5
Comparison of the active site of EntC from *E. coli* (a) with the proposed active site of Dhbc from *B. anthracis* (b). Isochorismate (ISC) and sulfate molecules are colored gray and yellow, respectively. Water molecules are shown as blue spheres and the magnesium ion is shown as a gray sphere. Selected hydrogen-bonding interactions are shown as dashed lines. Additional residues participating in the coordination of sulfate in the active site of Dhbc were considered and are colored gray. All O and N atoms are colored red and blue, respectively.

The F327Y mutation in EntC results in a 48-fold decrease in enzyme efficiency (Sridharan *et al.*, 2010). In both MenF and Irp9 Phe327 is substituted by Tyr, while in both Dhbc and EntC it is strongly conserved. There is a parallel-displaced π -stacking interaction between the two phenylalanines Phe327 and Phe359 in the structure of EntC; however, as can be clearly seen in Fig. 5, the position of the aromatic rings is completely different from that in Dhbc (Phe328–Tyr360). It is worth noting that the same orientation of the aromatic rings of Tyr368–Tyr399 is observed in the structure of MenF from *E. coli*. The Phe328–Tyr360 pair in Dhbc is changed to Phe362–Glu396 in ADC synthase from *C. hutchinsonii* (PDB entry 3h9m), to Trp390–Ser422 in ADC synthase from *E. coli* (PDB entry 1k0e) and to Tyr372–Gln403 in salicylate synthase from *Y. enterocolitica* (PDB entry 2fn0; Kerbarh *et al.*, 2006). In all of these cases these substitutions completely change the geometry of the interactions. There are no longer two aromatic rings that can create stacking interactions as takes place in EntC, MenF or Dhbc, which may be one of the functional determinants.

3.4. Enzymatic activity

The isochorismate synthase activity of *B. anthracis* Dhbc was tested using a spectrophotometric activity assay which monitors the formation of isochorismate by the increase in the absorbance of monochromatic light at 278 nm. For a reaction mixture containing 10 μ g enzyme, 5 mM $MgCl_2$ and 1 mM chorismic acid, we observed an average (from three repeats of the experiment) conversion rate of 33% of substrate after 10 min of reaction. For two control samples that lacked either $MgCl_2$ or Dhbc we did not detect the formation of isochorismate. The assay confirmed that Dhbc has the enzymatic function of isochorismate synthase and that its activity is dependent on the presence of Mg^{2+} ions.

4. Conclusions

In this study, we determined the crystal structure of the isochorismate synthase Dhbc from *B. anthracis* and compared it with those of other

chorismate-utilizing enzymes. The composition of its active center suggests that Dhbc uses the Mg^{2+} -dependent catalytic mechanism originally proposed by Walsh and refined by He, Kolappan and other researchers (Walsh *et al.*, 1987; He *et al.*, 2004; Kolappan *et al.*, 2007; Ziebart & Toney, 2010). In this mechanism, Lys142 is the base that is generally responsible for the activation of the water molecule for nucleophilic attack at the C2 hydroxyl group of chorismate, the backbone carbonyl of Ala304 is a hydrogen-bond acceptor for the C2 hydroxyl and Glu197 is a general acid for the loss of the C4 hydroxyl. This general mechanism is proposed here for Dhbc based on the discussed structural and experimental analyses.

This research was funded with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under contract Nos. HHSN272200700058C and HHSN272201200026C. The results shown in this report are derived from work performed at Argonne National Laboratory at the Structural Biology Center of the Advanced Photon Source. Argonne is operated by University of Chicago Argonne LLC for the US Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We thank Jing Hou for help with the design and implementation of the spectrophotometric activity assay and David Cooper, Matthew Zimmerman and Rachel Vigour for critically reading the manuscript.

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
- Arakawa, T., Tsumoto, K., Nagase, K. & Ejima, D. (2007). *Protein Expr. Purif.* **54**, 110–116.
- Aslanadis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
- Baichoo, N., Wang, T., Ye, R. & Helmann, J. D. (2002). *Mol. Microbiol.* **45**, 1613–1629.
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C. & Eddy, S. R. (2004). *Nucleic Acids Res.* **32**, D138–D141.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Cowan, K. D. & Main, P. (1993). *Acta Cryst.* **D49**, 148–157.
- Cowan, K. D. & Zhang, K. Y. J. (1999). *Prog. Biophys. Mol. Biol.* **72**, 245–270.
- Dahm, C., Müller, R., Schulte, G., Schmidt, K. & Leistner, E. (1998). *Biochim. Biophys. Acta*, **1425**, 377–386.
- Donnelly, M. I., Zhou, M., Millard, C. S., Clancy, S., Stols, L., Eschenfeldt, W. H., Collart, F. R. & Joachimiak, A. (2006). *Protein Expr. Purif.* **47**, 446–454.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Engl, R. A. & Huber, R. (1991). *Acta Cryst.* **A47**, 392–400.
- Eschenfeldt, W. H., Lucy, S., Millard, C. S., Joachimiak, A. & Mark, I. D. (2009). *Methods Mol. Biol.* **498**, 105–115.
- Haun, R. S. & Moss, J. (1992). *Gene*, **112**, 37–43.
- He, Z., Stigers Lavoie, K. D., Bartlett, P. A. & Toney, M. D. (2004). *J. Am. Chem. Soc.* **126**, 2378–2385.
- He, Z. & Toney, M. D. (2006). *Biochemistry*, **45**, 5019–5028.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Hotta, K., Kim, C.-Y., Fox, D. T. & Koppisch, A. T. (2010). *Microbiology*, **156**, 1918–1925.
- Kerbarh, O., Chirgadze, D. Y., Blundell, T. L. & Abell, C. (2006). *J. Mol. Biol.* **357**, 524–534.
- Kolappan, S., Zwahlen, J., Zhou, R., Truglio, J. J., Tonge, P. J. & Kisker, C. (2007). *Biochemistry*, **46**, 946–953.
- Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* **D60**, 2256–2268.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Lee, J. Y., Passalacqua, K. D., Hanna, P. C. & Sherman, D. H. (2011). *PLoS One*, **6**, e20777.
- Li, Q. A., Mavrodi, D. V., Thomashow, L. S., Roessle, M. & Blankenfeldt, W. (2011). *J. Biol. Chem.* **286**, 18213–18221.
- Manos-Turvey, A., Bulloch, E. M., Rutledge, P. J., Baker, E. N., Lott, J. S. & Payne, R. J. (2010). *ChemMedChem*, **5**, 1067–1079.
- Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst.* **D62**, 859–866.
- Morollo, A. A. & Eck, M. J. (2001). *Nature Struct. Biol.* **8**, 243–247.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). *J. Mol. Biol.* **247**, 536–540.
- Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Parsons, J. F., Jensen, P. Y., Pachikara, A. S., Howard, A. J., Eisenstein, E. & Ladner, J. E. (2002). *Biochemistry*, **41**, 2198–2208.
- Parsons, J. F., Shi, K. M. & Ladner, J. E. (2008). *Acta Cryst.* **D64**, 607–610.
- Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
- Ratledge, C. & Dover, L. G. (2000). *Annu. Rev. Microbiol.* **54**, 881–941.
- Rosenbaum, G. *et al.* (2006). *J. Synchrotron Rad.* **13**, 30–45.
- Rowland, B. M. & Taber, H. W. (1996). *J. Bacteriol.* **178**, 854–861.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Söding, J. (2005). *Bioinformatics*, **21**, 951–960.
- Söding, J. A., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
- Sridharan, S., Howard, N., Kerbarh, O., Błaszczyk, M., Abell, C. & Blundell, T. L. (2010). *J. Mol. Biol.* **397**, 290–300.
- Terwilliger, T. (2004). *J. Synchrotron Rad.* **11**, 49–52.
- Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S. & Arakawa, T. (2004). *Biotechnol. Prog.* **20**, 1301–1308.
- Walsh, C. T., Erion, M. D., Walts, A. E., Delany, J. J. & Berchtold, G. A. (1987). *Biochemistry*, **26**, 4734–4745.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, H., Guranovic, V., Dutta, S., Feng, Z., Berman, H. M. & Westbrook, J. D. (2004). *Acta Cryst.* **D60**, 1833–1839.
- Ziebart, K. T. & Toney, M. D. (2010). *Biochemistry*, **49**, 2851–2859.