Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Substrate channels revealed in the trimeric *Lactobacillus reuteri* bacterial microcompartment shell protein PduB

Allan Pang,[a] Mingzhi Liang,[b,c] Michael B. Prentice[b,c,d] and Richard W. Pickersgill[a]*

[a]School of Biological and Chemical Sciences, Queen Mary University of London, 327 Mile End Road, London E1 4NS, England, [b]Department of Microbiology, University College Cork, Cork, Ireland, [c]Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland, and [d]Department of Pathology, University College Cork, Cork, Ireland

Correspondence e-mail: r.w.pickersgill@qmul.ac.uk

Received 9 August 2012
Accepted 14 September 2012

PDB Reference: PduB, 4fay

Lactobacillus reuteri metabolizes two similar three-carbon molecules, 1,2-propanediol and glycerol, within closed polyhedral subcellular bacterial organelles called bacterial microcompartments (metabolosomes). The outer shell of the propanediol-utilization (Pdu) metabolosome is composed of hundreds of mainly hexagonal protein complexes made from six types of protein subunits that share similar domain structures. The structure of the bacterial microcompartment protein PduB has a tandem structural repeat within the subunit and assembles into a trimer with pseudo-hexagonal symmetry. This trimeric structure forms sheets in the crystal lattice and is able to fit within a polymeric sheet of the major shell component PduA to assemble a facet of the polyhedron. There are three pores within the trimer and these are formed between the tandem repeats within the subunits. The structure shows that each of these pores contains three glycerol molecules that interact with conserved residues, strongly suggesting that these subunit pores channel glycerol substrate into the metabolosome. In addition to the observation of glycerol occupying the subunit channels, the presence of glycerol on the molecular threefold symmetry axis suggests a role in locking closed the central region.

1. Introduction

A wide variety of heterotrophic bacteria, including Enterobacteriaceae and Firmicutes, produce multiple polyhedral cellular inclusions containing enzymes when induced by specific substrates (Sriramulu *et al.*, 2008; Bobik *et al.*, 1999; Shively *et al.*, 1998). The archetypal structure of this type (a bacterial microcompartment) is the carboxysome, which is found in cyanobacteria and some chemoautotrophic bacteria (Kerfeld *et al.*, 2010) and consists of a thin protein shell enclosing the enzymes RuBisCO and carbonic anhydrase. Similar structures in heterotrophic bacteria are termed metabolosomes and are classified according to the substrate that they process (Brinsmade *et al.*, 2005; Parsons *et al.*, 2008; Heldt *et al.*, 2009). These include propanediol-utilization (Pdu), ethanolamine-utilization (Eut) and ethanol-utilization (Etu) metabolosomes. These different types of metabolosomes are presumed to share a common function, which is to bring together the enzymes and metabolites, increasing their effective concentrations (Price & Badger, 1989) and sequestering them from the bacterial cytoplasm. Physical isolation of

a metabolic pathway in this way is believed to provide various advantages for the bacterial cell, including protection from toxic aldehyde intermediates, which can cause DNA damage and growth arrest when present in the bacterial cytoplasm (Sampson & Bobik, 2008), and reduction in evaporative loss of these intermediates from the cell (Penrod & Roth, 2006).

Lactobacillus reuteri is a probiotic bacterium that is able to colonize the gastrointestinal tract of a wide variety of animals (Casas & Dobrogosz, 2000). It produces an antimicrobial agent called reuterin, a mixture of monomeric and dimeric forms of β-hydroxypropionaldehyde (3-HPA; Talarico & Dobrogosz, 1989), by glycerol cofermentation. The cobalamin-dependent diol dehydratase forming 3-HPA from glycerol is PduCDE (GupCDE), which is part of the Pdu metabolosome, which also carries out 1,2-propanediol metabolism (Morita *et al.*, 2008). 1,2-Propanediol metabolism involves multiple *pdu*-operon-encoded enzymes associated with the metabolosome or present in the cytoplasm, as shown in Fig. 1 (for simplicity, additional Pdu enzymes required for PduCDE reactivation are not shown; Sriramulu *et al.*, 2008). The general functional model for the Pdu metabolosome is based on *Salmonella*. Inside the microcompartment, 1,2-propanediol is converted by diol dehydratase (PduCDE) into propionaldehyde (Havemann & Bobik, 2003). The propionaldehyde is subsequently disproportionated into 1-propanol or propionyl-CoA by the aldehyde dehydrogenase PduQ and the CoA-transferase PduP, respectively. These two products are delivered to the cytoplasm, where propionyl-CoA is further converted into propionyl phosphate and propionate by PduL and PduW, respectively. The *L. reuteri* Pdu metabolosome differs from that of *Salmonella* by the association of the *L. reuteri* PduL enzyme with the metabolosome structure (Sriramulu *et al.*, 2008) and by the absence of PduT (a single-electron channel; Pang *et al.*, 2011; Crowley *et al.*, 2010). This suggests the possibility that propionyl phosphate may be formed inside the *L. reuteri* metabolosome. Onward respiration of propionate elsewhere in the cell does not occur in *L. reuteri* (Sriramulu *et al.*, 2008), unlike in *Salmonella* (Horswill & Escalante-Semerena, 1999; Price-Carter *et al.*, 2001).

Enclosing these metabolic reactions is a polyhedral shell formed by a protein layer. In *L. reuteri* six different proteins form the shell: PduA, PduB, PduJ, PduK, PduN and PduU (PduT is missing in the *L. reuteri pdu* operon; Sriramulu *et al.*, 2008). Sequence comparison of these shell proteins indicates that they are composed mainly of bacterial microcompartment (BMC) protein domains (InterPro domain IPR000249), with the exception of PduN, which has a Pf00319 domain. Crystal structures of PduA (Crowley *et al.*, 2010), PduT (Pang *et al.*, 2011; Crowley *et al.*, 2010) and PduU (Crowley *et al.*, 2008) revealed that shell proteins may consist of either single (PduA or PduU) or tandem (PduT) BMC domains per subunit that assemble into hexamers or trimers, respectively. Furthermore, the structures also showed that the shell proteins are assembled in such a way as to form a central pore at the centre of the hexamer. It has been suggested that these central pores may be used to transport substrates and products (Crowley *et al.*, 2008, 2010) as well as to channel electrons *via* a central 4Fe–4S cluster (Pang *et al.*, 2011; Crowley *et al.*, 2010). The central pore of the PduA hexamer has been proposed as the channel for 1,2-propanediol transport on the basis of its polar lining, but limited resolution and sixfold symmetry have not allowed the detection of 1,2-propanediol within the pore (Crowley *et al.*, 2010).

Sequence analysis of PduB indicated that the shell protein is formed by two BMC domains and appears to be structurally similar to its homologue EtuB. The EtuB trimer has three pores formed within subunits instead of a central pore formed by symmetry-related subunits (Heldt *et al.*, 2009). Several authors have suggested that the central pore of the hexamers and the subunit pores of the trimers are the substrate channels, but this has not been demonstrated. Here, we report the crystal structure of the *L. reuteri* Pdu shell protein PduB, a trimer with glycerol trapped in the subunit pores, revealing that these pores are the substrate channels.

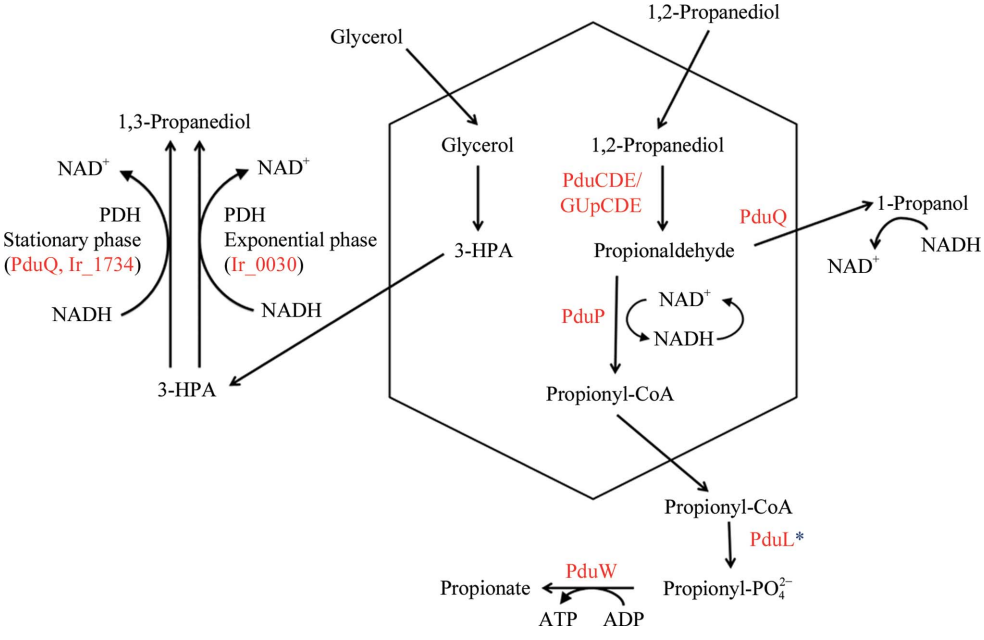


Figure 1
Schematic representation of the metabolic pathway of the *L. reuteri* propanediol- and glycerol-utilization metabolosome. Glycerol is converted to 1,3-propanediol in a two-step reaction; the first step occurs within the microcompartment and the second is mediated by PDH (propanediol dehydrogenase). 1,2-Propanediol is subsequently converted to either 1-propanol or 1-propionate, with the generation of propionaldehyde occurring inside the microcompartment. PduL may be found within the microcompartment or closely associated with the outside surface of the shell. Enzyme designations are in red.

2. Materials and methods

2.1. Production, purification and crystallization of PduB

The coding region of *pduB* (*L. reuteri* strain 20016) was cloned into pET14b to facilitate PduB overproduction in *Escherichia coli* BL21 (DE3) cells. The cells were grown in LB (Luria–Bertani) medium containing 100 mg l^{-1} ampicillin with aeration at 310 K. Upon reaching an OD_{600} of 0.6–0.8, the protein was induced with 0.4 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) and was left shaking overnight at 289 K. The cells were harvested by centrifugation (20 min, 8000 rev min^{-1}) and resuspended in binding buffer (50 m*M* Tris–HCl pH 8.0, 0.5 *M* NaCl, 10 m*M* imidazole). Cell lysis was achieved by sonication using a Sonics Vibracell Ultrasonic processor with an output of 20 W for eight 20 s bursts interspersed with 1 min of cooling. The cell lysate was applied onto a nickel-charged Sepharose column. Unbound protein was washed off with binding buffer, wash buffer I (50 m*M* Tris–HCl pH 8.0, 0.5 *M* NaCl, 50 m*M* imidazole) and wash buffer II (50 m*M* Tris–HCl pH 8.0, 0.5 *M* NaCl, 100 m*M* imidazole). Proteins were eluted with buffer consisting of 50 m*M* Tris–HCl pH 8.0, 0.5 *M* NaCl, 400 m*M* imidazole. PduB was further purified on a size-exclusion column (Superdex 200 Global 10/30) connected to an ÄKTA FPLC chromatography system. The majority of the overexpressed shell protein resided in the insoluble fraction; however, we produced a reasonable amount of protein from 7 l culture.

His-tagged PduB was concentrated to approximately 7 mg ml^{-1} in 50 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl. Initial hanging-drop vapour-equilibration crystallization trials using Hampton Research Crystal Screen and Crystal Screen 2 resulted in five conditions yielding small crystals. Crystallization conditions were further optimized using 24-well plates to produce hexagonal plate-shaped crystals (0.1 mm across). The best diffracting crystals were grown using a reservoir consisting of 0.1 *M* sodium cacodylate pH 6.5, 1.4 *M* sodium acetate with hanging drops formed from 1 µl protein solution mixed with 1 µl reservoir solution. Single crystals were harvested in litholoops, transferred through reservoir augmented with 20% glycerol as cryoprotectant and stored in liquid nitrogen prior to data collection.

2.2. Data collection and structure solution

High-quality diffraction data were collected to 1.55 Å resolution on beamline I04-1 at Diamond Light Source, Oxfordshire, England (Table 1). A number of the crystals screened exhibited severe twinning and lattice-disorder problems, as well as anisotropic diffraction. However, the crystal used to solve the PduB structure showed no sign of these problems. Data were reduced using *xdsme* (Kabsch, 2010) and *SCALA* (Evans, 2006). The crystal belonged to space group $C222_1$ and had three PduB subunits in the asymmetric unit, giving a solvent content of 36%. The structure was solved using molecular replacement with *PHENIX* (Adams *et al.*, 2010) using EtuB (PDB entry 3io0; Heldt *et al.*, 2009) as the search model. The resulting structure of PduB was rebuilt using *Coot* (Emsley *et al.*, 2010) and was refined using *REFMAC* (Murshudov *et al.*, 2011). The stereochemistry and validation statistics of the final PduB model are given in Table 1. Coordinates and structure-factor amplitudes have been deposited in the PDB as entry 4fay.

Table 1
Data-collection, processing and refinement statistics for PduB.

Values in parentheses are for the outer resolution shell. The values presented in this table were obtained using *SCALA*, *REFMAC* and *PROCHECK* from the *CCP*4 suite (Winn *et al.*, 2011).

Data collection	
Space group	$C222_1$
Unit-cell parameters (Å, °)	$a = 69.865$, $b = 120.864$, $c = 145.531$, $\alpha = \beta = \gamma = 90$
Protein molecular mass (Da)	24945.3
Molecules per asymmetric unit	3
Wavelength (Å)	0.9163
Resolution (Å)	46.51–1.55 (1.64–1.55)
No. of unique reflections	87436 (12243)
Multiplicity	4.5 (4.4)
Completeness (%)	99.0 (96.2)
R_{merge}†	0.088 (0.908)
Mean $I/\sigma(I)$	11.3 (2.0)
$R_{p.i.m.}$‡	0.045 (0.483)
R_{meas}§	0.099 (1.035)
Wilson *B* factor (Å^2)	16.0
Refinement	
Resolution (Å)	46.51–1.55
Reflections (working/test)	83762/4338
R factor/R_{free}¶	0.1762/0.2045
R.m.s.d. bonds (Å)/angles (°)	0.0198/1.7151
Ramachandran plot statistics (%)	
Residues in most favoured regions	94.35
Residues in allowed regions	5.17
Residues in disallowed regions	0.48 [Asp83]

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th observation, $\langle I(hkl)\rangle$ is the mean intensity of the reflection and the summations extend over all unique reflections (*hkl*) and all equivalents (*i*). ‡ $R_{p.i.m.}$ is a measure of the quality of the data after averaging the multiple measurements (Evans, 2006). § R_{meas} (also known as $R_{r.i.m.}$) is an improved version of the traditional R_{merge} (R_{sym}) and measures how well the different observations agree (Evans, 2006). ¶ *R* factor = $\sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} represent the observed and calculated structure factors, respectively. The *R* factor was calculated using 95% of the data, which were included in refinement, and R_{free} was calculated using 4.5% of the data, which were excluded from refinement.

2.3. Structure analysis

The protein sequence of *L. reuteri* PduB was obtained from the NCBI protein database. *ClustalW* (Larkin *et al.*, 2007) was used for sequence alignment and *PDBsum* was used to produce a schematic representation of the topology of the PduB trimer (Laskowski *et al.*, 1997). The *PISA* software (Krissinel & Henrick, 2007) was employed to analyse surface interactions, and *DaliLite* (Holm & Park, 2000) was used for pairwise structure alignment and comparison. Images of the molecular structures were generated and visualized using *PyMOL* (DeLano & Lam, 2005).

3. Results and discussion

3.1. Crystal structure of PduB

PduB is a trimeric shell protein with pseudo-hexameric symmetry, a result of each subunit having a tandem BMC repeat (Fig. 2). To date, four shell proteins, CsoS1D (Klein *et al.*, 2009), EtuB (Heldt *et al.*, 2009), EutL (Takenoya *et al.*,

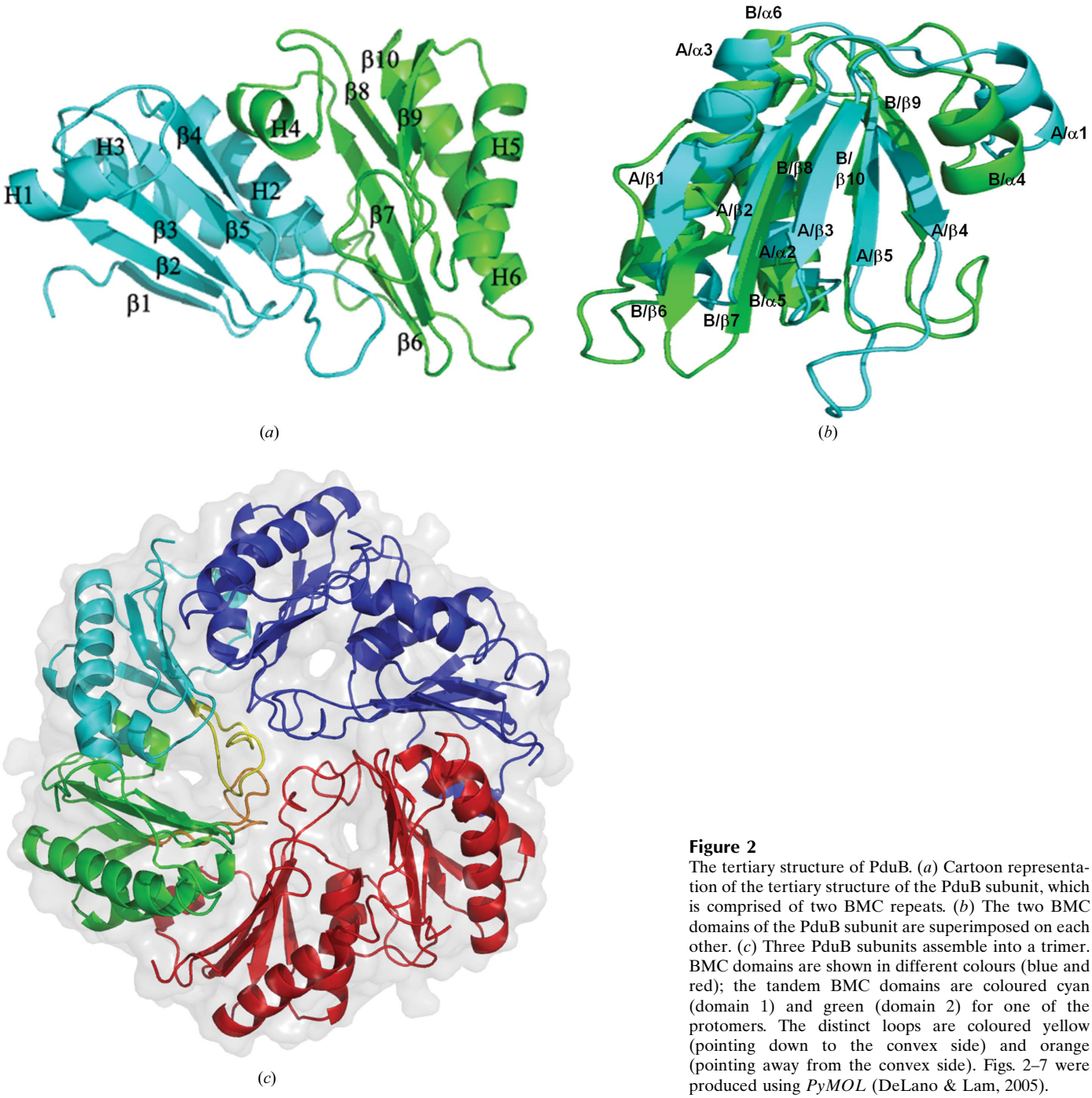


Figure 2
The tertiary structure of PduB. (*a*) Cartoon representation of the tertiary structure of the PduB subunit, which is comprised of two BMC repeats. (*b*) The two BMC domains of the PduB subunit are superimposed on each other. (*c*) Three PduB subunits assemble into a trimer. BMC domains are shown in different colours (blue and red); the tandem BMC domains are coloured cyan (domain 1) and green (domain 2) for one of the protomers. The distinct loops are coloured yellow (pointing down to the convex side) and orange (pointing away from the convex side). Figs. 2–7 were produced using *PyMOL* (DeLano & Lam, 2005).

2010; Tanaka *et al.*, 2010) and PduT (Pang *et al.*, 2011; Crowley *et al.*, 2010), are known to contain tandem BMC repeats within their subunits. PduB has clearly defined polypeptide backbone in the electron-density map for residues 11–237. The residues preceding Val11 and the C-terminal residue Lys238 are poorly defined in the electron-density map, presumably because they are flexible or statically disordered in the crystal. The side chains have clearly defined electron density, with the exception of four residues: Glu19, Arg30, Lys171 and Arg208. Glu19 and Arg30 are found in the loops of β-turn hairpins ($\beta1/\beta2$ and $\beta2/\beta3$), while Lys171 and Arg208 are on the solvent-accessible surface of α-helices $\alpha5$ and $\alpha6$. A total of 20 residues (Supplementary Table S1[1]) with two conformations were found in the PduB trimer. One residue, Asp83, falls in the disallowed region of the Ramachandran plot, but the electron-density map has clear carbonyl bulges for this region, suggesting that the main chain is strained for this residue (Supplementary Fig. S1[1]). This strained Asp83, which is

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: YT5049). Services for accessing this material are described at the back of the journal.

conserved in PduB and EtuB, may be functionally important, but this has yet to be demonstrated.

18 ligands were found in the final trimeric structure of PduB. 17 of these were glycerols, while the remaining ligand was an acetate ion. Chain *A* binds six glycerol molecules, while chain *B* binds six glycerols and an acetate ion, and four glycerols can be found in chain *C*. A further glycerol molecule is found at the centre of the trimeric structure. The differing number of ligands for each chain can be attributed to their different environments in the crystal. However, each subunit pore is occupied by three glycerol molecules binding at the same locations within the pore.

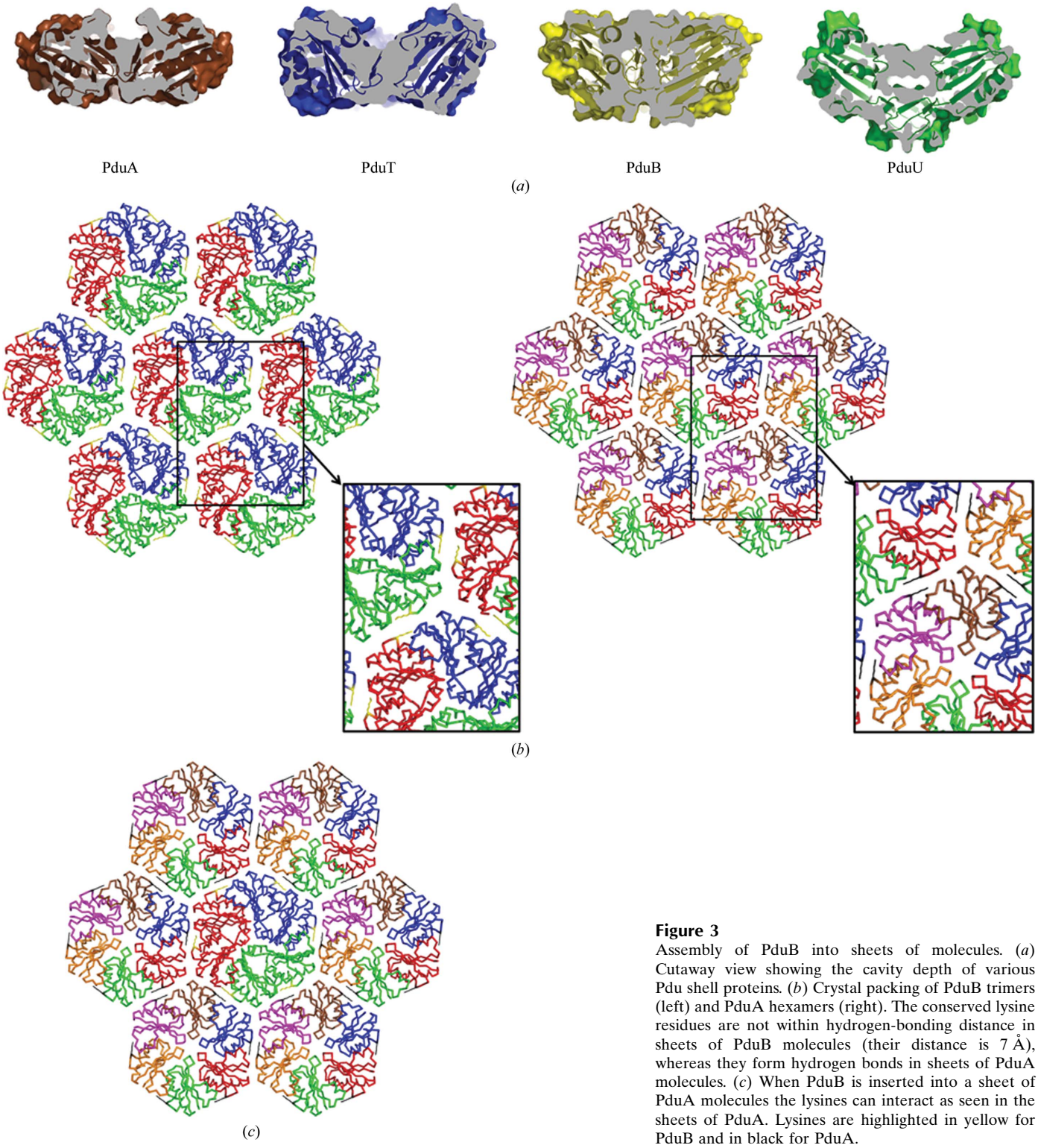


Figure 3
Assembly of PduB into sheets of molecules. (*a*) Cutaway view showing the cavity depth of various Pdu shell proteins. (*b*) Crystal packing of PduB trimers (left) and PduA hexamers (right). The conserved lysine residues are not within hydrogen-bonding distance in sheets of PduB molecules (their distance is 7 Å), whereas they form hydrogen bonds in sheets of PduA molecules. (*c*) When PduB is inserted into a sheet of PduA molecules the lysines can interact as seen in the sheets of PduA. Lysines are highlighted in yellow for PduB and in black for PduA.

3.2. The BMC domains of PduB

The first and second BMC domains of PduB are composed of 109 and 118 residues, respectively (Supplementary Fig. S2*a*). The two BMC domains share 15% sequence identity (Supplementary Fig. S2*b*) and superimpose with an r.m.s.d. of 2.7 Å with 91 aligned residues and a *Z*-score of 10 (Fig. 2*b*). The most obvious difference is in the loop regions connecting β4 and β5 and connecting β9 and β10; these loops fill the central cavity around the threefold axis (Fig. 2*c*). These loops sandwich and bind to the central glycerol molecule (discussed later).

3.3. PduB is a trimer that assembles into a sheet of molecules

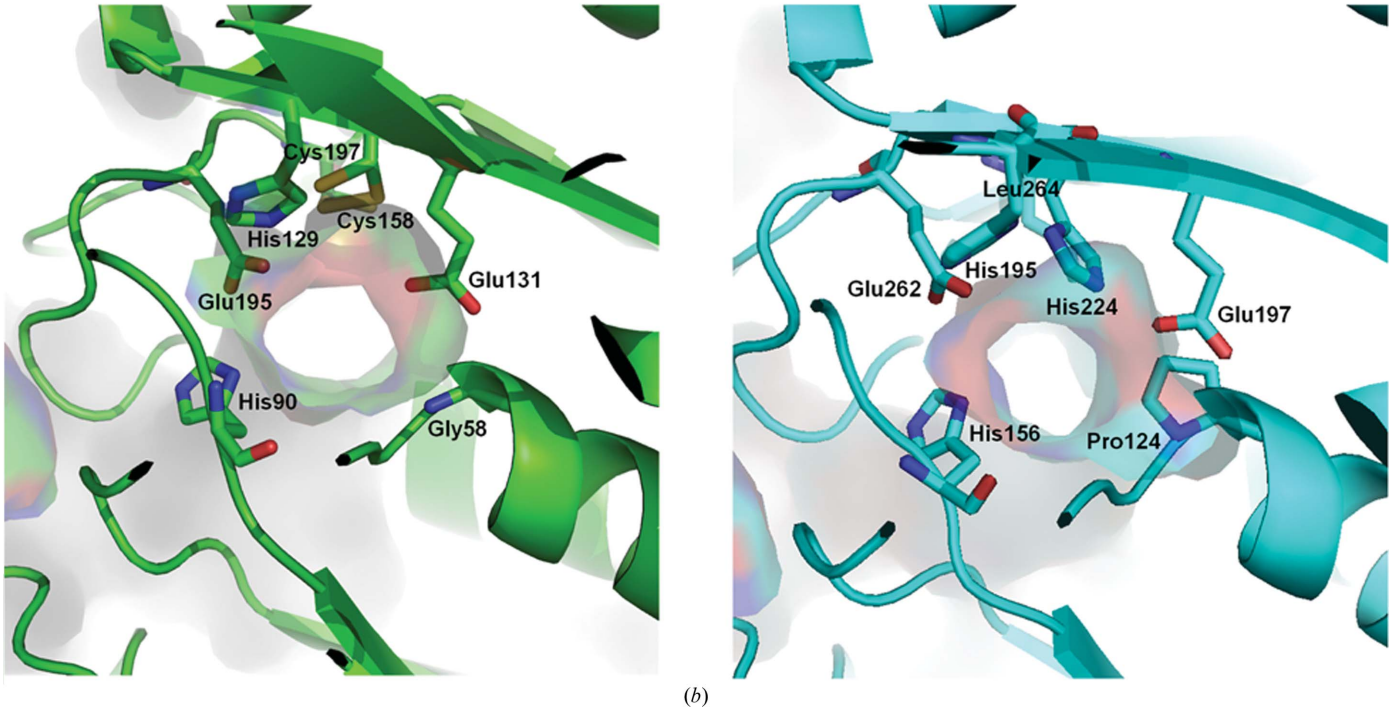
PduB assembles into a trimer with a hexameric appearance (Fig. 2*c*), a common theme for shell proteins with two BMC repeats per subunit. One face of the PduB trimer is concave and the other is convex. The concave surface lacks the distinctive nine patches of positive potential seen on the corresponding surface of EtuB (Supplementary Fig. S3). The electrostatic appearance of the convex surface of PduB is broadly similar to that of EtuB, with basic residues on the periphery and around the centre sandwiching three acidic patches (the acidic patches are more pronounced in EtuB). The conserved peripheral electrostatic appearance presumably reflects the conserved packing of trimers into sheets, while the less conserved surface features may correlate with differences in function. When the PduB trimer is viewed from the side, the bowl-shaped cavity is deeper than in PduA and PduT (Fig. 3*a*). The cavity of PduU is deeper than that of PduB owing to the presence of six tightly wound β-strands around the hexamer pore (Crowley *et al.*, 2008). The central cavities of PduA, PduT and PduU are evident, while the central cavity of PduB is not immediately apparent owing to the presence of the β9–β10 loop.

```
PduB_Salmonella     RTGAGPHIMAADEAVKATNTEVVSIELPRDTKGGAGHGSLIILGGNDVSDVKRGIEVALK 109
PduB_Citrobacter    RTGAGPHIMAADEAVKATNTEVVSIELPRDTKGGAGHGSLIILGGNDVSDVKRGIEVALK 146
PduB_Lactobacillus  RTGAGPQIMAMDEGIKATNMECIDVEWPRDTKGGGGHGCLIIIGGDDPADARQAIRVALD 113
EtuB_Clostridium    RVGAPAQMMAADEAVKGTNTEVATIELPRDTKGGAGHGIFIVLKAADVSDARRAVEIALK 179
                    *.** .::** **.:*.** *   :* *******.*** :*:: . * :*.::.:.:**.

PduB_Salmonella     ELDRTFGDVYGNEAGHIELQYTARASYALEKAFGAPIGRACGIIVGAPASVGVLMADTAL 169
PduB_Citrobacter    ELDRTFGDVYANEAGHIEMQYTARASYALEKAFGAPIGRACGVIVGAPASVGVLMADTAL 206
PduB_Lactobacillus  NLHRTFGDVYNAKAGHLELQFTARAAGAAHLGLGAVEGKAFGLICGCPSGIGVVMGDKAL 173
EtuB_Clostridium    QTDKYLGNVYLCDAGHLEVQYTARASLIFEKAFGAPSGQAFGIMHAAPAGVGMIVADTAL 239
                    : .: :*:**  .***:*:*:****:   . .:**  *:* *:: ..*:.:*:::.*.**

PduB_Salmonella     KSANVEVVAYSSPAHGT-SFSNEAILVISGDSGAVRQAVTSAREIGKTVLATLGSEPK-N 227
PduB_Citrobacter    KSANVEVVAYSSPAHGT-SFSNEAILVISGDSGAVRQAVISAREIGKTVLGTLGSEPK-N 264
PduB_Lactobacillus  KTAGVEPLNFTSPSHGT-SFSNEGCLTITGDSGAVRQAVMAGREVGLKLLSQFGEEPV-N 231
EtuB_Clostridium    KTADVKLITYGSPTNGVLSYTNEILITISGDSGAVLQSLTAARKAGLSILRSMGQDPVSM 299
                    *:*.*: : : **::*. *::**  :.*:****** *:: :.*: * .:*  :*.:*
```

(*a*)



(*b*)

Figure 4
Comparison of the subunit pores of PduB and EtuB. (*a*) Sequence comparison of EtuB and PduB generated by *ClustalW*; the pore-lining residues are underlined and the key residues shown in (*b*) are highlighted in red (Asn124 not shown). (*b*) Cutaway section looking through the subunit pore of PduB (left) and EtuB (right) showing the positions of the pore-lining residues.

The crystal structure of PduB reveals a higher order packing of trimers into uniform molecular sheets. Similar higher order packing has been observed in other Pdu shell protein crystals including PduA and PduT (Crowley *et al.*, 2010), but not PduU (Crowley *et al.*, 2008) or EtuB (Heldt *et al.*, 2009). Six conserved lysines are found within the PduB trimer that correspond to the lysines suggested to maintain the hexameric PduA within a molecular sheet. Although similar lysine residues can be found within the molecular sheets of the PduB trimer, the residues do not perfectly align as in the sheets of PduA hexamers (Fig. 3*b*) and the packing is looser. However, when the PduB trimer is placed into a sheet of PduA hexamers the trimer fits very well, with the six conserved lysine residues facing antiparallel to their corresponding adjacent lysine residues in the PduA hexamers (Fig. 3*c*). This observation is consistent with our biochemical

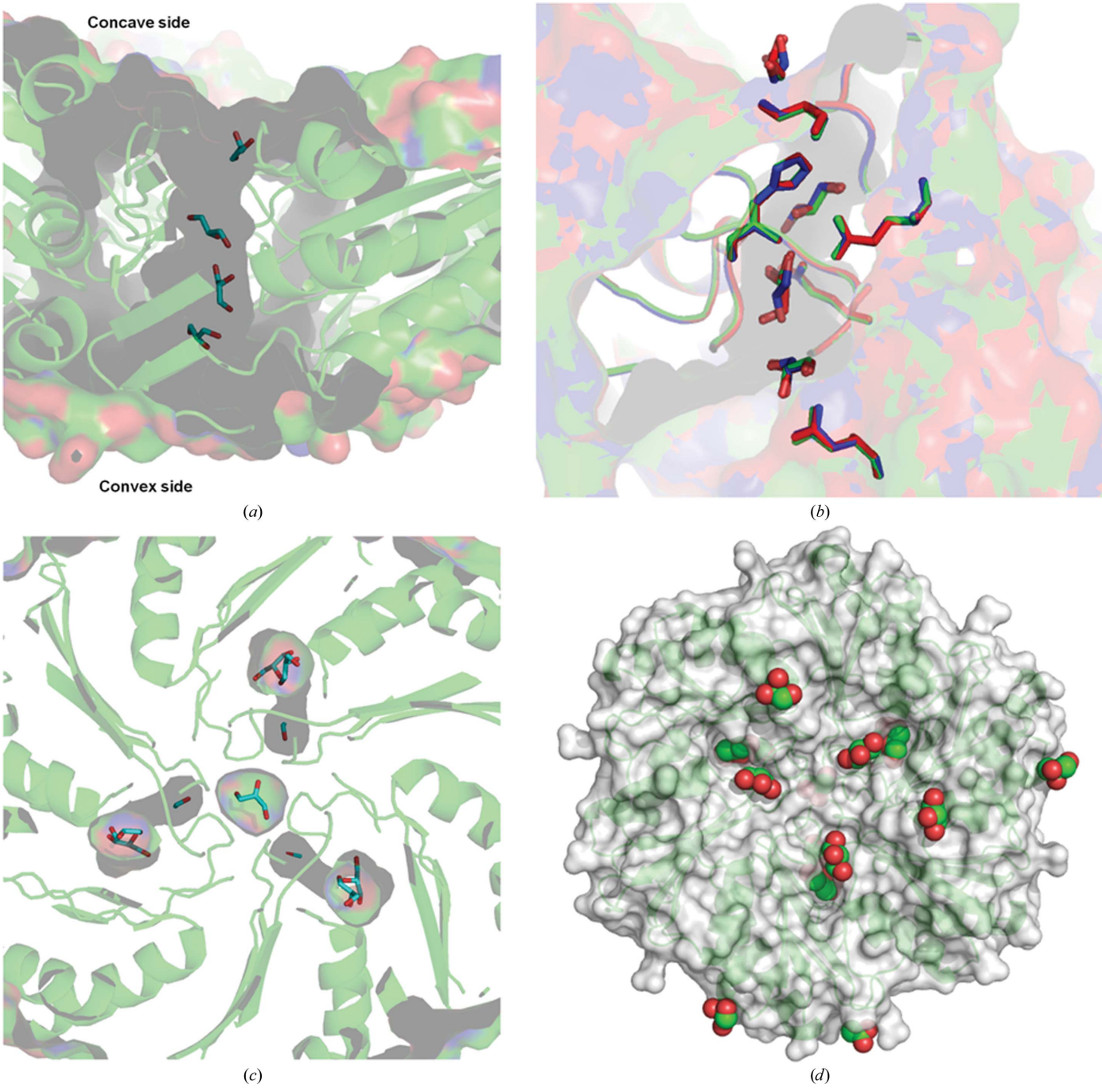
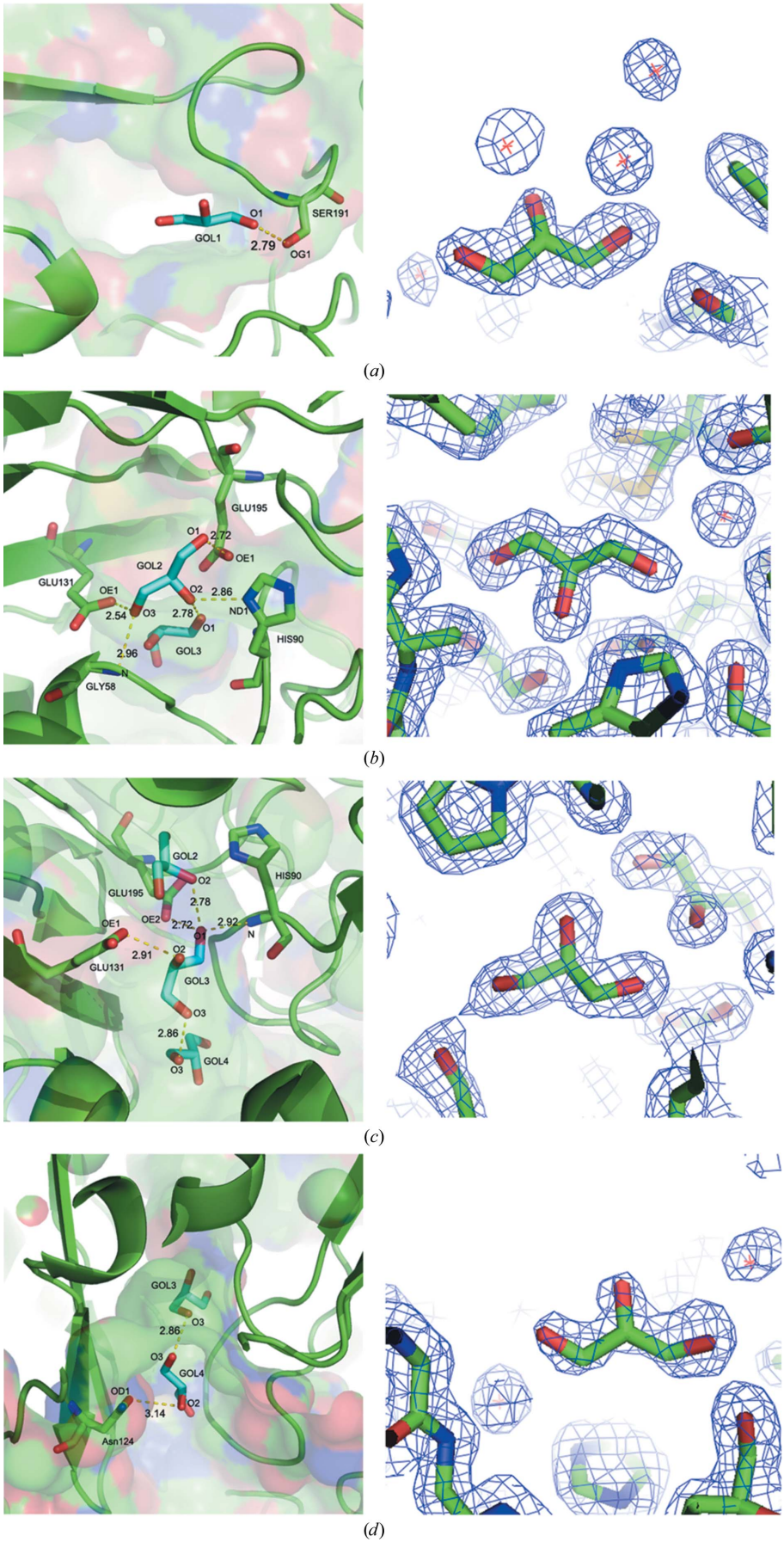


Figure 5
Ligand binding to PduB. (*a*) Four glycerols line up in the subunit pore region of PduB. (*b*) Superimposition of the three protomers, showing that the binding sites are consistent between the three protomers, which are shown in red, green and blue. Residues Gly58, His90, Asn124, Glu131, Ser191 and Glu195 are shown along with the bound glycerol molecules. (*c*) Top view looking down onto the concave side, showing glycerol molecules occupying the subunit pores and central region. (*d*) Top view looking at the concave side, showing ligand binding to this surface.



studies that suggest that PduA interacts with PduB (Parsons *et al.*, 2010), suggesting that mixed sheets of PduA and PduB are possible, although a ring-to-ring hetero-(pseudo)dodecamer cannot be ruled out.

3.4. The PduB subunit has a channel with three glycerol-binding subsites

The majority of the shell proteins solved to date have a central pore on the sixfold axis of the hexamer. In contrast, the trimeric shell proteins EtuB and EutL have pores within the subunit, a situation facilitated by the gene duplication that has given rise to the tandem repeat of the trimeric shell protein. The crystal structure of PduB reveals that this trimer has similar subunit-pore characteristics. The pores of EtuB are lined with three histidine residues (at positions 156, 195 and 224) and two glutamate residues (at positions 197 and 262) (Heldt *et al.*, 2009). Similar residues are found in PduB, with two exceptions (Fig. 4). His224 is substituted by Cys158, which forms a disulfide bridge with Cys197 (Leu264 in EtuB) and slightly widens the channel. It is not clear whether the disulfide is a feature of the *in vivo* structure or an artefact of purification. If the disulfide bond were reduced, the structure would be unlikely to change substantially as the disulfide links adjacent β-strands in the central PduB sheet, nor would the channel be likely to be narrowed. In addition to this, Pro124 (EtuB) is substituted by Gly58 (PduB), which also slightly widens the pore in PduB. Subtle widening of the channel may allow glycerol access to the channel (Fig. 4*b*).

When viewed in cross-section, the channel can be seen to be approximately 22 Å long and 7.5 Å wide. Three glycerol molecules occupy subsites within this channel. A fourth

Figure 6
Localization of glycerol in the subunit pores of PduB. (*a*) GOL1, (*b*) GOL2, (*c*) GOL3 and (*d*) GOL4 form hydrogen bonds to the residues lining the pore (left). The σ_A-weighted $2F_{obs} - F_{calc}$ Fourier synthesis contoured around the corresponding glycerol molecules at 1σ showing the quality of the electron-density map, represented as a blue chicken-wire mesh, is shown on the right.

glycerol is also seen located close to the channel opening; this glycerol appears to not be as tightly bound as the chain of three glycerol molecules occupying the channel itself (Fig. 5*a*). Starting from the concave side of the channel, the O atom (O1) of glycerol (GOL1) binds to the O atom (OG1) of the side chain of Ser191 (Fig. 6*a*). The second (GOL2) and third (GOL3) glycerols make more extensive hydrogen bonds both to each other as well as to residues lining the pore. The key residues in this interaction are Glu131 (OE1) and Glu195 (OE1 and OE2), the carboxylate O atoms of which hydrogen bond to both glycerols. The main-chain amide of Gly58 and the side-chain amide (ND1) of His90 also hydrogen bond to GOL2 (Figs. 6*b* and 6*c*). These residues are conserved across bacterial PduB sequences as well as EtuB, with the exception of Gly58, which is a proline in EtuB (Fig. 4*a*). The final glycerol (GOL4) completes the chain of three glycerol mole-

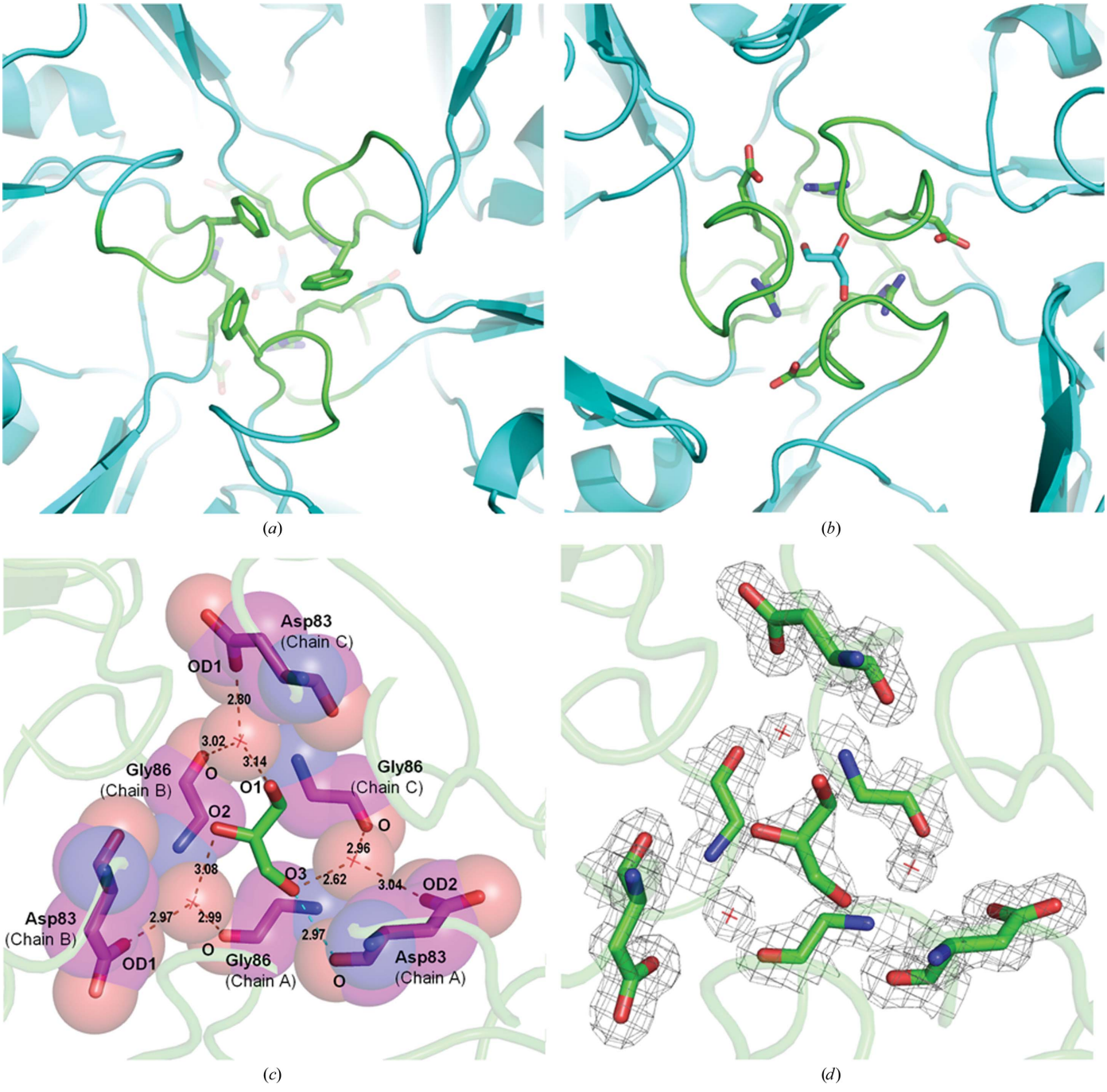


Figure 7
Threefold symmetry axis of PduB. (*a*) View of the upper loop (GTSFS) coloured green; three phenylalanine residues block this side of the central region. (*b*) View of the lower loop (RDTKGGGG) coloured green; Asp and Arg residues encircling the glycerol are shown. (*c*) A representative of one orientation of glycerol on the threefold symmetry axis of PduB surrounded by three water molecules, which in turn form hydrogen bonds to Asp83 OD1 and Gly86 O. (*d*) Electron-density map for this glycerol, surrounded by three water molecules, contoured at 1σ.

cules and appears to be more weakly associated with PduB (Fig. 6*d*). A table and a graphical representation of the interactions of the glycerols with each other and the pore-lining residues can be found in Supplementary Table S2 and Fig. 6, respectively. The binding sites for the three glycerol molecules are consistent for the three protomers (Fig. 5*b*).

The abundance of glycerol ligands binding to sites around the concave face of the shell protein PduB is suggestive of their involvement in glycerol transport (Fig. 5*d*), possibly to attract glycerol into the channel, although it should be noted the the glycerol concentration was high in this study (2.7 *M*). The diameters of the subunit pores are slightly larger than previously observed, at around 7.5 Å. The glycerol molecules trapped in the long thin channels of PduB suggest that these trimeric subunit pores are channels for substrate. The ligands of pores generally bind to a series of well defined sites within the pore. Several crystal structures of membrane-channel proteins with trapped cargo in their pore regions have been characterized; for instance, aquaglyceroporin with glycerol (Newby *et al.*, 2008), maltoporin with sucrose, trehalose and melibiose (Wang *et al.*, 1997) and a mitochondrial channel with ATP (Rostovtseva & Bezrukov, 1998). Theory provides an understanding of how binding to sites within the pore can facilitate ligand flux (Berezhkovskii & Bezrukov, 2005). The structural association of glycerol molecules with the pore allows the identification of conserved residues in PduB which are associated with binding. Based on the size and hydrogen-bonding potential of 1,2-propanediol compared with glycerol, we anticipate that the subunit channels may also be conduits for 1,2-propanediol. The pore-lining residues involved in hydrogen bonds to glycerol molecules in *L. reuteri* PduB are conserved across PduB and EtuB from different species. This is not surprising as the substrates (glycerol, 1,2-propanediol, ethanol and ethanolamine) share similar characteristics which suggests translocation using the same mechanism and closely similar pores. However, propionaldehyde and β-hydroxypropionaldehyde might not pass through the pores because of the different hydrogen-bonding potential of the aldehyde compared with a hydroxyl group. Aldehydes can accept two hydrogen bonds in a planar arrangement, while alcohols can donate one hydrogen bond and accept two with tetrahedral geometry. Both O1 and O3 of GOL2 are hydrogen-bond donors because the carboxylate groups of Glu131 and Glu195 are expected to be deprotonated (Fig. 6*b*). Similarly, O1 of GOL3 is required to donate a proton to Glu195 and the geometry at this atom is approximately tetrahedral (Fig. 6*c*). These two sites therefore appear to act as a specificity filter that allows the passage of substrates but retains reaction intermediates within the microcompartment.

3.5. Glycerol binding on the threefold symmetry axis of PduB

A further glycerol molecule is trapped within the protein on the threefold axis of the PduB structure (Fig. 5*c*). It is surrounded by three well ordered water molecules to which it makes hydrogen bonds (Fig. 7*c*). The water molecules in turn form hydrogen bonds to the O atom of the side chain of Asp83 and the O atom of the main chain of Gly86 (Fig. 7*c*). The electron density for the central glycerol is remarkably clear given the anticipated rotational averaging of the glycerol occupying this site (Fig. 7*d*); this clarity may be a consequence of the approximate threefold symmetry of the glycerol molecule (the mobility of the central glycerol is indicated in Supplementary Fig. S4). One glycerol hydroxyl makes a hydrogen bond to the carbonyl O atom of Asp83 (Fig. 7*c*). The strained conformation of Asp83 may be owing to the presence of Arg82. This glycerol-binding site has a loop above it (concave side) and a loop below it (convex side), both of which are characterized by the presence of residues with small side chains, suggesting the possibility of flexibility (breathing motions; Figs. 7*a* and 7*b*). Clearly, the loops have sufficient flexibility to allow the glycerol molecule access to this otherwise inaccessible pocket. The upper loop (189–193) has the sequence GTSFS and the lower loop (82–89) has the sequence RDTKGGGG. The three phenylalanine residues residing in the upper loop play a major role in blocking the central region of the trimer. When compared with closed EutL, a similar arrangement of three tyrosines blocks the pore of the trimer (Sagermann *et al.*, 2009; Takenoya *et al.*, 2010; Tanaka *et al.*, 2010).

The trimeric shell protein EutL has been shown to have a central channel that can be open or closed (Takenoya *et al.*, 2010; Tanaka *et al.*, 2010). The presence of the buried glycerol on the threefold axis of the PduB trimer suggests that the central loops are flexible and glycerol binding stabilizes the conformation observed in the crystal.

3.6. PduB and PduB′

PduB is known to be synthesized in two forms owing to the presence of two translation start sites on the polycistronic message (Parsons *et al.*, 2008; Havemann & Bobik, 2003). Bioinformatics suggested that the full-length PduB (*L. reuteri*) has an extra 25 N-terminal residues compared with PduB′. In this work, we crystallized the full-length PduB. Our crystal structure revealed 15 of these N-terminal residues (Supplementary Fig. S5), while the additional N-terminal ten residues (MNDFLNSTST) were not observed in the electron-density map. The observed additional residues of PduB form two antiparallel β-strands on the convex side of the trimer. The significance of the removal of these strands has yet to be established.

4. Conclusion

Pores or channels are a feature of several bacterial microcompartment shell proteins; these pores can be at the centre of the hexamer or trimer or within subunits of the trimeric shell proteins. Previous shell proteins for which structures have been determined typically have pores with a diameter ranging from 4 to 6 Å (Kerfeld *et al.*, 2005; Tanaka *et al.*, 2008, 2009; Tsai *et al.*, 2007; Yeates *et al.*, 2010). It is anticipated that these pores are functionally important in channelling substrates and metabolites. For years, there has been circumstantial evidence

such as ions observed to be trapped in the pore regions of different shell proteins (Tanaka *et al.*, 2009; Tsai *et al.*, 2007). However, a definitive demonstration of specific substrates trapped within the channels of shell proteins has been elusive.

PduB is a major shell protein of Pdu metabolosomes (Parsons *et al.*, 2010; Havemann & Bobik, 2003). We show that it has three subunit channels per trimer and is apparently capable of insertion into sheets which form the facets of the polyhedral structure. Our previous genetic and biochemical studies on the *L. reuteri* 1,2-propanediol microcompartment showed that the bacterial species is able to metabolize both glycerol and 1,2-propanediol within this organelle (Sriramulu *et al.*, 2008). The crystal structure of PduB reveals the presence of small subunit pores, as observed in the shell proteins EtuB (Heldt *et al.*, 2009) and EutL (Takenoya *et al.*, 2010; Tanaka *et al.*, 2010); more importantly, the structure shows that the pore-lining residues have affinity for glycerol molecules and could act as a channel for this substrate. The pattern of hydrogen bonds involving the central glycerol-binding sites also suggests a specificity filter to prevent aldehyde efflux. In addition to this, we found a glycerol molecule trapped in a central pocket, suggesting that it locks the central loops closed and raising the possibility of a ligand-gated channel. Our crystal structure provides the first evidence of the bacterial microcompartment pore acting as channel for natural substrate.

This work was supported by the Biotechnology and Biology Research Council (grant No. BB/H013180/1), the Higher Education Council of England (HEFCE) and Queen Mary University of London. MBP and ML were funded by the Health Research Board award HRA_POR/2011/111 and Science Foundation Ireland. We acknowledge the use of the Diamond Light Source, Oxford.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Berezhkovskii, A. M. & Bezrukov, S. M. (2005). *Biophys. J.* **88**, L17–L19.
Bobik, T. A., Havemann, G. D., Busch, R. J., Williams, D. S. & Aldrich, H. C. (1999). *J. Bacteriol.* **181**, 5967–5975.
Brinsmade, S. R., Paldon, T. & Escalante-Semerena, J. C. (2005). *J. Bacteriol.* **187**, 8039–8046.
Casas, I. A. & Dobrogosz, W. J. (2000). *Microb. Ecol. Health Dis.* **12**, 247–285.
Crowley, C. S., Cascio, D., Sawaya, M. R., Kopstein, J. S., Bobik, T. A. & Yeates, T. O. (2010). *J Biol. Chem.* **285**, 37838–37846.
Crowley, C. S., Sawaya, M. R., Bobik, T. A. & Yeates, T. O. (2008). *Structure*, **16**, 1324–1332.
DeLano, W. L. & Lam, J. W. (2005). *Abstr. Pap. Am. Chem. Soc.* **230**, u1371–u1372.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Havemann, G. D. & Bobik, T. A. (2003). *J. Bacteriol.* **185**, 5086–5095.
Heldt, D., Frank, S., Seyedarabi, A., Ladikis, D., Parsons, J. B., Warren, M. J. & Pickersgill, R. W. (2009). *Biochem. J.* **423**, 199–207.
Holm, L. & Park, J. (2000). *Bioinformatics*, **16**, 566–567.
Horswill, A. R. & Escalante-Semerena, J. C. (1999). *J. Bacteriol.* **181**, 5615–5623.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kerfeld, C. A., Heinhorst, S. & Cannon, G. C. (2010). *Annu. Rev. Microbiol.* **64**, 391–408.
Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M. & Yeates, T. O. (2005). *Science*, **309**, 936–938.
Klein, M. G., Zwart, P., Bagby, S. C., Cai, F., Chisholm, S. W., Heinhorst, S., Cannon, G. C. & Kerfeld, C. A. (2009). *J. Mol. Biol.* **392**, 319–333.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L. & Thornton, J. M. (1997). *Trends Biochem. Sci.* **22**, 488–490.
Morita, H. *et al.* (2008). *DNA Res.* **15**, 151–161.
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.* D**67**, 355–367.
Newby, Z. E., O'Connell, J., Robles-Colmenares, Y., Khademi, S., Miercke, L. J. & Stroud, R. M. (2008). *Nature Struct. Mol. Biol.* **15**, 619–625.
Pang, A., Warren, M. J. & Pickersgill, R. W. (2011). *Acta Cryst.* D**67**, 91–96.
Parsons, J. B. *et al.* (2008). *J. Biol. Chem.* **283**, 14366–14375.
Parsons, J. B., Frank, S., Bhella, D., Liang, M., Prentice, M. B., Mulvihill, D. P. & Warren, M. J. (2010). *Mol. Cell*, **38**, 305–315.
Penrod, J. T. & Roth, J. R. (2006). *J. Bacteriol.* **188**, 2865–2874.
Price, G. D. & Badger, M. R. (1989). *Plant Physiol.* **91**, 514–525.
Price-Carter, M., Tingey, J., Bobik, T. A. & Roth, J. R. (2001). *J. Bacteriol.* **183**, 2463–2475.
Rostovtseva, T. K. & Bezrukov, S. M. (1998). *Biophys. J.* **74**, 2365–2373.
Sagermann, M., Ohtaki, A. & Nikolakakis, K. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 8883–8887.
Sampson, E. M. & Bobik, T. A. (2008). *J. Bacteriol.* **190**, 2966–2971.
Shively, J. M., Bradburne, C. E., Aldrich, H. C., Bobik, T. A., Mehlman, J. L., Jin, S. & Baker, S. H. (1998). *Can. J. Bot.* **76**, 906–916.
Sriramulu, D. D., Liang, M., Hernandez-Romero, D., Raux-Deery, E., Lünsdorf, H., Parsons, J. B., Warren, M. J. & Prentice, M. B. (2008). *J. Bacteriol.* **190**, 4559–4567.
Takenoya, M., Nikolakakis, K. & Sagermann, M. (2010). *J. Bacteriol.* **192**, 6056–6063.
Talarico, T. L. & Dobrogosz, W. J. (1989). *Antimicrob. Agents Chemother.* **33**, 674–679.
Tanaka, S., Kerfeld, C. A., Sawaya, M. R., Cai, F., Heinhorst, S., Cannon, G. C. & Yeates, T. O. (2008). *Science*, **319**, 1083–1086.
Tanaka, S., Sawaya, M. R., Phillips, M. & Yeates, T. O. (2009). *Protein Sci.* **18**, 108–120.
Tanaka, S., Sawaya, M. R. & Yeates, T. O. (2010). *Science*, **327**, 81–84.
Tsai, Y., Sawaya, M. R., Cannon, G. C., Cai, F., Williams, E. B., Heinhorst, S., Kerfeld, C. A. & Yeates, T. O. (2007). *PLoS Biol.* **5**, 1345–1354.
Wang, Y.-F., Dutzler, R., Rizkallah, P. J., Rosenbusch, J. P. & Schirmer, T. (1997). *J. Mol. Biol.* **272**, 56–63.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yeates, T. O., Crowley, C. S. & Tanaka, S. (2010). *Annu. Rev. Biophys.* **39**, 185–205.