



## Structural studies of domain movement in active-site mutants of porphobilinogen deaminase from *Bacillus megaterium*

Jingxu Guo, Peter Erskine, Alun R. Coker, Steve P. Wood and Jonathan B. Cooper

*Acta Cryst.* (2017). F73, 612–620



**IUCr Journals**  
CRYSTALLOGRAPHY JOURNALS ONLINE

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>

# Structural studies of domain movement in active-site mutants of porphobilinogen deaminase from *Bacillus megaterium*

Jingxu Guo,<sup>a</sup> Peter Erskine,<sup>a,b</sup> Alun R. Coker,<sup>a</sup> Steve P. Wood<sup>a</sup> and Jonathan B. Cooper<sup>a,b\*</sup>

Received 29 August 2017  
 Accepted 23 October 2017

Edited by F. T. Tsai, Baylor College of Medicine, Houston, USA

**Keywords:** protein crystallography; structural biology; porphobilinogen deaminase; *Bacillus megaterium*; tetrapyrrole biosynthesis.

**PDB references:** porphobilinogen deaminase, D82A mutant, 5ov4; D82E mutant, 5ov5; D82N mutant, 5ov6

**Supporting information:** this article has supporting information at journals.iucr.org/f

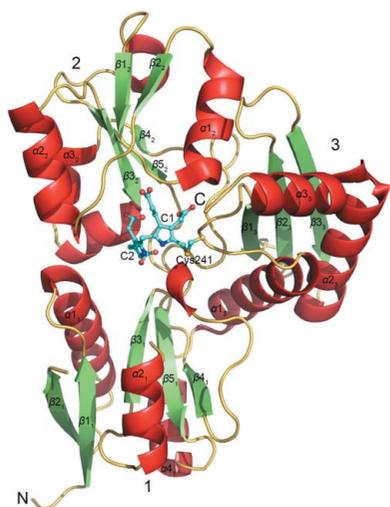
<sup>a</sup>Division of Medicine, University College London, Gower Street, London WC1E 6BT, England, and <sup>b</sup>Department of Biological Sciences, Birkbeck, University of London, Malet Street, Bloomsbury, London WC1E 7HX, England.  
 \*Correspondence e-mail: jon.cooper@ucl.ac.uk

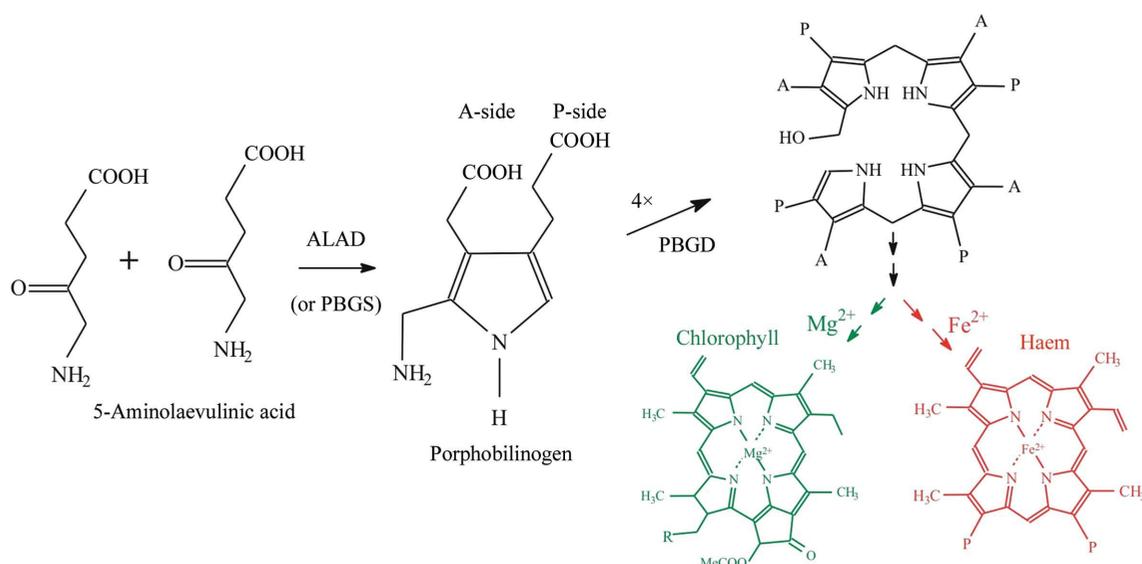
The enzyme porphobilinogen deaminase (PBGD) is one of the key enzymes in tetrapyrrole biosynthesis. It catalyses the formation of a linear tetrapyrrole from four molecules of the substrate porphobilinogen (PBG). It has a dipyrromethane cofactor (DPM) in the active site which is covalently linked to a conserved cysteine residue through a thioether bridge. The substrate molecules are linked to the cofactor in a stepwise head-to-tail manner during the reaction, which is catalysed by a conserved aspartate residue: Asp82 in the *B. megaterium* enzyme. Three mutations have been made affecting Asp82 (D82A, D82E and D82N) and their crystal structures have been determined at resolutions of 2.7, 1.8 and 1.9 Å, respectively. These structures reveal that whilst the D82E mutant possesses the DPM cofactor, in the D82N and D82A mutants the cofactor is likely to be missing, incompletely assembled or disordered. Comparison of the mutant PBGD structures with that of the wild-type enzyme shows that there are significant domain movements and suggests that the enzyme adopts ‘open’ and ‘closed’ conformations, potentially in response to substrate binding.

## 1. Introduction

Tetrapyrroles are a class of chemical compounds in which four pyrrole rings are linked together to form linear or cyclic molecules (Jordan, 1991). These natural molecules are essential for living systems since they participate in key processes such as photosynthesis and respiration. There are two pathways for tetrapyrrole biosynthesis, which are similar in all living organisms (Fig. 1). The first common intermediate in these pathways is 5-aminolaevulinic acid (ALA). Higher plants and many prokaryotes use the C5 or glutamate pathway, in which ALA is generated from glutamate (Kannangara *et al.*, 1988, 1994). Mammals and other eukaryotic organisms produce ALA from glycine and succinyl-CoA along the C4 or ‘Shemin’ pathway, which requires the enzyme ALA synthase (ALAS) (Leeper, 1985; Li *et al.*, 1989).

The next stage involves the transformation of ALA into the common precursor for tetrapyrroles, uroporphyrinogen III, in three enzymatic steps which are common to all living systems (Chadwick & Ackrill, 1994). In the first step, two molecules of ALA are condensed by the enzyme ALA dehydratase (ALAD) to form the basic pyrrole building block porphobilinogen (PBG) (Mills-Davies *et al.*, 2017; Fig. 1). In the next step, four molecules of PBG are condensed together to produce the linear tetrapyrrole preuroporphyrinogen. This reaction is catalysed by the enzyme porphobilinogen deaminase (PBGD), which is the object of this study. In the third step the linear tetrapyrrole is cyclized and rearranged.




**Figure 1**

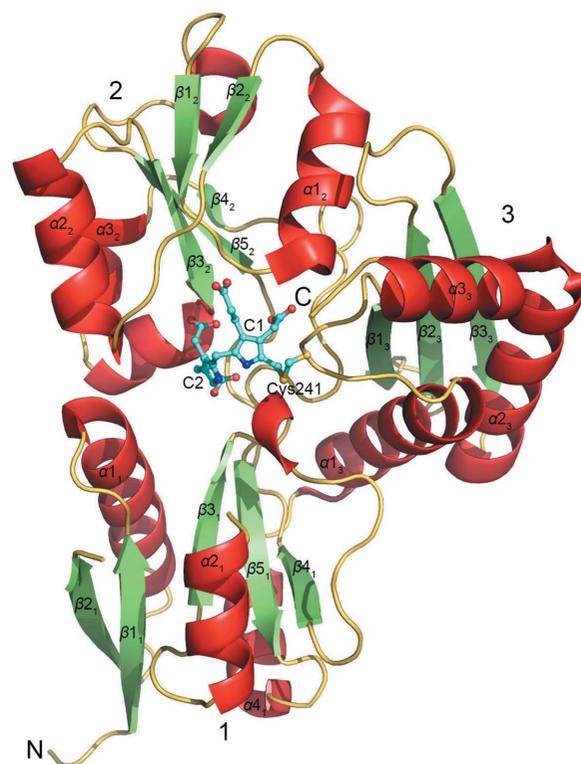
The early steps of tetrapyrrole biosynthesis showing the reactions catalysed by PBGD and the preceding enzyme ALAD or porphobilinogen synthase (PBGS).

Most haem in humans is synthesized endogenously and each tissue produces haem to satisfy its own requirements. More than 70% of the total haem in the body is synthesized in the bone marrow, where haem is incorporated into haemoglobin for erythrocyte precursors. The liver accounts for about 15% of the total haem production. There is a high demand in the liver for haem to be incorporated into mitochondrial cytochromes as well as into cytochrome P<sub>450</sub>, catalase and cytochrome b<sub>5</sub>. Haemoproteins have many roles in oxidative metabolism, including O<sub>2</sub> transport, O<sub>2</sub> sensing, oxidative stress response, oxidative phosphorylation and oxygenation reactions. They also participate in the transportation of diatomic gases, chemical catalysis, electron transfer and the sensing of diatomic gases such as nitric oxide and carbon monoxide (Rodgers, 1999).

PBGD catalyses one of the key early steps in the biosynthesis of tetrapyrroles (Warren & Smith, 2009), namely the condensation of four pyrrole moieties. Several PBGDs have been isolated from different organisms of both prokaryotic and eukaryotic origin, including *Escherichia coli* (Jordan & Warren, 1987), plants (Roberts *et al.*, 2013) and mammals (Gill *et al.*, 2009). In humans, there are two isoforms of PBGD: the erythroid-specific enzyme and the house-keeping enzyme. Both originate from a single PBGD gene on chromosome 11 and arise by alternate splicing of the primary transcript (Wood *et al.*, 1995). PBGDs have molecular weights ranging from 34 to 44 kDa and pH optima in the range 8.0–8.5. These enzymes from most organisms have great thermostability and this has been exploited during their purification (Jordan, 1991).

A number of X-ray structures of PBGD have been reported, including those from *E. coli*, human, *Arabidopsis thaliana* and *Bacillus megaterium* (Louie *et al.*, 1992; Gill *et al.*, 2009; Roberts *et al.*, 2013; Azim *et al.*, 2014; Fig. 2). *E. coli* PBGD was the first enzyme in the tetrapyrrole pathway to have its structure determined by X-ray crystallography (Louie

*et al.*, 1992). The polypeptide itself is folded into three domains (1–3) each of approximately the same size. Both domain 1 and domain 2 are composed of five-stranded mixed  $\beta$ -sheets and have a similar overall topology to the type II periplasmic binding proteins, which have been reported to adopt ‘open’


**Figure 2**

The X-ray structure of PBGD from *B. megaterium* (PDB entry 4mlv). The dipyrromethane (DPM) cofactor is covalently bound to Cys241 in the active-site cleft between domains 1 and 2 (Azim *et al.*, 2014). The domains,  $\alpha$ -helices and  $\beta$ -sheets are labelled according to the nomenclature of Louie *et al.* (1992).

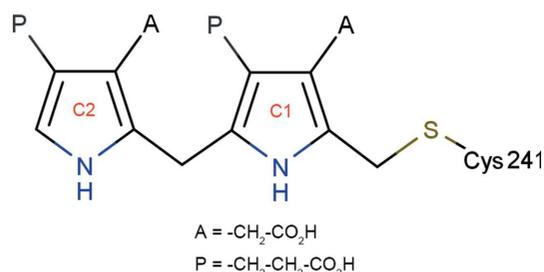
and 'closed' states in response to ligand binding (Louie, 1993; Louie *et al.*, 1992). Domain 3, which possesses an open-faced antiparallel  $\beta$ -sheet of three strands and three  $\alpha$ -helices, is folded completely differently from the other two domains.

There is a dipyrromethane (DPM) cofactor (Fig. 3) that is covalently attached to a cysteine, Cys241 in *B. megaterium*, in a loop which is located in the active-site cleft. The two pyrrole rings of the cofactor are named C1 and C2, with C1 being attached to the protein *via* a thioether link to the cysteine. The presence of the cofactor can be confirmed by treatment of PBGD with Ehrlich's reagent, which gives an absorbance peak at 565 nm and a subsequent shift to a  $\lambda_{\max}$  of 495 nm after 15 min (Jordan, 1991). The cofactor can be generated in two different ways: (i) by slow assembly from two molecules of PBG or (ii) by cleavage of the product preuroporphyrinogen, which reacts rapidly with the apoenzyme (Awan *et al.*, 1997; Shoolingin-Jordan *et al.*, 1997).

The DPM cofactor plays two different roles in the catalysis. Firstly, it acts as the primer in the active site in order to initiate the enzymatic reaction and to connect the substrate moieties to the enzyme during the elongation stage. This has been confirmed by  $^{14}\text{C}$ -labelling experiments, in which the enzyme containing a labelled cofactor was incubated with unlabelled PBG. The final product contained no labelled compounds, indicating that the DPM cofactor was not affected by the catalytic turnover and remained permanently linked to the enzyme (Warren & Jordan, 1988). Secondly, the permanent connection between the cofactor and the enzyme may help to limit the maximum number of substrate molecules that can bind and make it precisely four.

The cofactor adopts different conformations depending on its oxidation state. The C1 ring varies little between different states, while the C2 ring in the oxidized state (in which the cofactor is called dipyrromethenone) tends to occupy a space which is thought to be close to the position for the incoming substrate PBG (Louie *et al.*, 1996). In this state, both rings of the dipyrromethenone factor are found to be coplanar, whilst in the reduced dipyrromethane state the cofactor adopts a conformation in which the C2 ring occupies a more internal position in the active-site cleft (Azim *et al.*, 2014; Louie *et al.*, 1996; Roberts *et al.*, 2013).

There is also an invariant aspartic acid (Asp82 in the *B. megaterium*) enzyme that is located in a highly conserved



**Figure 3**

Structure of the dipyrromethane cofactor, which is covalently linked to Cys241 *via* a thioether linkage. Four molecules of PBG are sequentially added to the cofactor *via* the free  $\alpha$ -position of the terminal pyrrole, and the tetrapyrrole product is then released from the cofactor.

region of the PBGD sequence, VHSMKDMP, from residues 77–84 in the *B. megaterium* enzyme (Azim *et al.*, 2014). The aspartate side chain forms two hydrogen bonds to the NH groups of the DPM and is thought to catalyse the tetramerization reaction. Site-directed mutagenesis of this residue, Asp84, in *E. coli* PBGD was undertaken by Jordan and Woodcock (Jordan & Woodcock, 1991; Woodcock & Jordan, 1994), converting it to Glu, Ala and Asn. The D84E mutant had a significantly reduced activity of less than 1% of that of the wild-type (WT) enzyme, while the other two mutants, D84A and D84N, were folded but completely inactive. It was found that the cofactor of the D84E mutant was more sensitive to oxidation, and one reason for this might be that the hydrogen bond between the carboxylic group of Asp84 and the NH group of the C2 ring was weaker or lost owing to the mutation (Lambert *et al.*, 1994).

Porphyrias are a group of inherited disorders caused by enzyme defects in haem biosynthesis and the resulting accumulation of phototoxic intermediates. Acute intermittent porphyria (AIP) is an autosomal dominant metabolic disease resulting from diminished erythrocyte activity of PBGD. The disease is caused by point mutations in PBGD, which often affect the conserved pyrrole-binding arginine residues in the active site of the enzyme. In addition, other mutations destabilize the three-dimensional fold by steric and electrostatic effects (Delfau *et al.*, 1990; Jordan & Woodcock, 1991; Lander *et al.*, 1991; Scott *et al.*, 1989).

In this investigation, we have used a combination of mutagenesis of Asp82 of *B. megaterium* PBGD and X-ray crystallography to analyse the resulting cofactor-binding effects and domain movements within the protein. Our results allow quantification of the scale of domain movement within the enzyme in response to single-atom changes at the catalytic centre.

## 2. Methods

### 2.1. Mutagenesis, expression and purification

The wild-type *B. megaterium* PBGD gene in a pET-14b expression construct, as previously reported by Azim *et al.* (2013), was used with the QuikChange mutagenesis kit (Agilent Technologies, Cheshire, England) to introduce the required base changes. The following mutagenic primers were ordered from Yorkshire Bioscience (York Science Park, York, England) with the base changes shown in bold.

D82A (GAT to GCT):

Forward: 5'-GGCCGTT**C**ATAGTATGAAAG**C**TATGCCGGC-3'  
Reverse: 5'-GCCGGC**A**T**G**CTTTCATACTATGAACGGCC-3'

D82E (GAT to GAA):

Forward: 5'-GGCCGTT**C**ATAGTATGAAAG**A**ATGCCGGC-3'  
Reverse: 5'-GCCGGC**A**T**T**TCTTTCATACTATGAACGGCC-3'

D82N (GAT to AAT):

Forward: 5'-GATATGGCCGTT**C**ATAGTATGAA**A**AATATGCCGGCTG-3'  
Reverse: 5'-CAGCCGGC**A**T**A**TTTTCATACTATGAACGGCCATATC-3'

Each mutagenesis reaction was undertaken using a Techne Progene thermal cycler (Techne, Staffordshire, England)

following the protocol provided in the kit. The nicked plasmid DNA was then transformed into subcloning-efficient DH5 $\alpha$  competent cells (Invitrogen, Thermo Fisher Scientific, Dartford, England) for nick repair and DNA amplification. The plasmid DNA was then extracted and purified using an AxyPrep Plasmid Miniprep Kit (Axygen, Union City, California, USA). DNA sequencing, which was carried out at DNA Sequencing and Services (University of Dundee, Dundee, Scotland), confirmed that the base changes had been introduced successfully. Each of the mutated genes was transformed into *E. coli* Rosetta (DE3) pLysS cells (Novagen, Darmstadt, Germany) for protein expression.

Each of the mutant proteins was expressed using a heat shock in which the cells at mid-log phase were heated to 42°C for approximately 20 min prior to cooling on ice and induction with IPTG. The His-tagged proteins were purified in 50 mM Tris–HCl buffer pH 7.3 along with 100 mM NaCl, initially using a HisTrap FF column, followed by cleavage of the tag using thrombin. A HiTrap benzamidin column was then used to remove the thrombin, and the proteins were finally desalted and purified using a Superdex 75 gel-filtration column. All of the columns used were purchased from GE Healthcare (Buckinghamshire, England) and the buffers used were made according to the manufacturer's recommendations.

## 2.2. Protein crystallization

Screening for crystallization conditions for all of the PBGD mutant proteins was undertaken by the hanging-drop method using the Structure Screens 1 & 2 kit from Molecular Dimensions (Suffolk, England). 5  $\mu$ l of each mutant protein (5 mg ml<sup>-1</sup>) was mixed with 5  $\mu$ l of the corresponding well solution on a siliconized cover slip and the plates were stored at 21°C for crystallization. Showers of small yellow crystals of the D82N protein started to appear after 3 d in Structure Screen 1 condition 2 (0.2 M ammonium acetate, 0.1 M sodium acetate pH 4.6, 30% PEG 4000). Subsequent optimization revealed that better crystals of all of the mutants could be obtained reproducibly in 0.2 M ammonium acetate, 0.1 M sodium acetate pH 3.5–4.0, 22–28% PEG 4000. These crystals grew as clusters formed of thin plates, and there were no apparent morphological differences amongst crystals of the mutants.

In order to obtain protein–substrate complex structures, co-crystallization was undertaken for all PBGD mutants with the substrate PBG ranging from 140  $\mu$ M to 1.4 mM (1–10 times molar excess over the protein) using the same crystallization condition. In addition, co-crystallization of the mutants with the WT enzyme and PBG was also carried out using a WT:mutant:PBG molar ratio of 1:50:500. The idea behind this experiment was that the WT enzyme might make sufficient product, which would then bind to the inactive mutants, given the evidence that the product is actually the precursor of the cofactor (Pluscec & Bogorad, 1970; Jordan & Warren, 1987; Mauzerall & Granick, 1956).

**Table 1**

X-ray statistics for all three mutant structures.

Values in parentheses are for the outer resolution shell.

	D82A	D82E	D82N
Beamline	I03, DLS	I03, DLS	I02, DLS
Wavelength (Å)	0.9763	0.9763	0.9795
Space group	<i>P</i> <sub>2</sub> <i>1</i> <sub>2</sub> <i>1</i>	<i>P</i> <sub>2</sub> <i>1</i> <sub>2</sub> <i>1</i>	<i>P</i> <sub>2</sub> <i>1</i> <sub>2</sub> <i>1</i>
Unit-cell parameters (Å)			
<i>a</i> (Å)	49.1	49.2	49.0
<i>b</i> (Å)	62.5	62.7	62.7
<i>c</i> (Å)	91.4	91.8	91.3
Mosaic spread (°)	0.37	0.24	0.53
Resolution (Å)	29.56–2.69 (2.76–2.69)	91.83–1.81 (1.86–1.81)	36.89–1.87 (1.92–1.87)
<i>R</i> <sub>merge</sub> (%)	12.8 (48.3)	4.8 (62.9)	6.9 (54.1)
<i>R</i> <sub>meas</sub> (%)	15.1 (58.5)	5.3 (69.7)	8.2 (65.5)
CC <sub>1/2</sub> (%)	99.0 (79.7)	99.9 (76.6)	99.6 (78.8)
Completeness (%)	99.6 (97.4)	99.9 (100.0)	98.9 (96.5)
Average <i>I</i> / $\sigma$ ( <i>I</i> )	11.8 (3.2)	20.7 (2.6)	16.9 (4.0)
Multiplicity	6.4 (5.6)	6.7 (5.8)	6.2 (5.7)
No. of observed reflections	52592 (3255)	178498 (11002)	194239 (12783)
No. of unique reflections	8221 (579)	26586 (1907)	23650 (1679)
Wilson plot <i>B</i> factor (Å <sup>2</sup> )	51.8	24.0	24.0
Solvent content (%)	36.9	37.5	36.8
<i>R</i> factor (%)	20.9	19.4	21.0
<i>R</i> <sub>free</sub> (%)	23.9	23.4	25.3
R.m.s.d., bond lengths (Å)	0.002	0.007	0.003
R.m.s.d., bond angles (°)	0.479	0.971	0.515
No. of reflections			
Working set	8186	26585	23589
Test set	394	1310	1164
No. of protein atoms	2251	2266	2257
No. of ligand atoms	0	30	16
No. of water molecules	47	172	147
Mean protein <i>B</i> factor (Å <sup>2</sup> )	38.7	32.1	28.7
Ramachandran favoured (%)	97.5	98.6	99.3
Ramachandran outliers (%)	0	0	0
PDB code	5ov4	5ov5	5ov6

## 2.3. Data collection and data processing

Selected crystals were mounted in loops with approximately 30% (*v/v*) glycerol as the cryoprotectant before flash-cooling in a nitrogen-gas cryostream (Oxford Cryosystems Ltd, Oxford, England) at 100 K or immersion in liquid ethane. Crystals were then stored in pucks in liquid nitrogen prior to exposure to the X-ray beam.

X-ray data collection was performed at stations I02, I03, I04 and I04-1 at Diamond Light Source (DLS). 190° of data were collected for every crystal which yielded good diffraction, with 1 s exposure times, 1° oscillation and 20% transmission. Data processing was achieved using two different methods: either by use of the automatically processed data from *xia2* (Winter, 2010) or by manual processing. For manual data processing, *iMosflm* (Leslie, 2006; Battye *et al.*, 2011) was used to integrate the diffraction spots and produce the *reflection.mtz* file before *SCALA* (Evans, 2006) was used to scale all of the equivalent symmetry-related reflections together. For automatic data processing with *xia2*, the data were integrated and scaled with *XDS* and *XSCALE* (Kabsch, 2010).

## 2.4. Structure determination, refinement and validation

Structure analysis was undertaken by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) with the WT

enzyme structure as the search model (PDB entry 4mlv; Azim *et al.*, 2014). Following one round of refinement of the solutions, manual rebuilding and introduction of the mutant residues Ala, Glu and Asn to replace Asp82 in the WT structure was accomplished by the use of *Coot* (Emsley & Cowtan, 2004). The structures were then refined by restrained refinement with *REFMAC5* (Murshudov *et al.*, 2011; Winn *et al.*, 2001, 2003) and *phenix.refine* (Afonine *et al.*, 2012). All of the mutant structures were validated by *MolProbity* (Chen *et al.*, 2010). All of the statistics for data collection, data processing, structure determination, refinement and structure validation, together with the associated PDB codes, are given in Table 1.

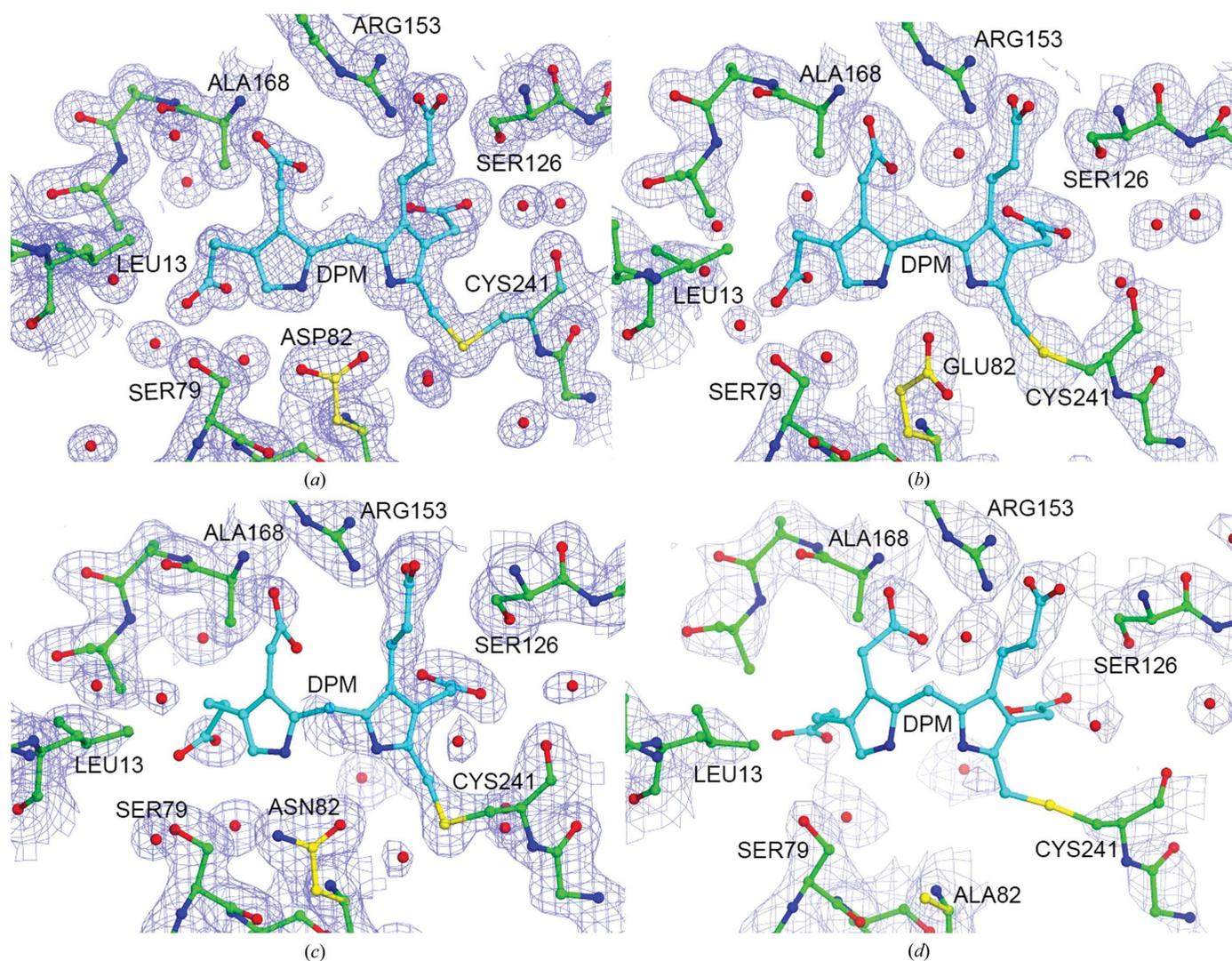
### 2.5. Biochemical characterization

Ehrlich's reagent was used to determine the status of the DPM cofactor in WT and mutant PBGD as described by Jordan and Warren (Warren & Jordan, 1988; Jordan &

Warren, 1987). The observation of an absorption peak at 565 nm which subsequently shifts to 495 nm after 15 min indicates the presence of the cofactor. Further details are given in the Supporting Information. Kinetic assays were conducted for WT and mutant PBGDs using the method described by Jordan (1991). All absorbance measurements were made using an Ultrospec 3000 UV/visible spectrophotometer (GE Healthcare, Buckinghamshire, England) and calculations were performed using *OriginPro 9.1* (OriginLab, Northampton, Massachusetts, USA).

### 2.6. Determination and classification of domain movements

Relative domain movements were analysed using the *DynDom* website (Hayward & Berendsen, 1998; Hayward & Lee, 2002). The hinge and shear classification of the domain movements was determined using the dynamic contact graphs (DCG) method described by Taylor *et al.* (2013, 2014). This



**Figure 4**  
The electron-density map for the DPM cofactor. A selection of the active-site residues surrounding the cofactor in the wild-type structure (a) and the D82E (b), D82N (c) and D82A (d) mutant structures are shown. There is almost no density for the cofactor in the D82A mutant and for the C2 ring of the cofactor in the D82N mutant, while the D82E mutant shows good density for both rings of the cofactor. The  $2F_o - F_c$  electron-density map is shown in blue contoured at 1.0 r.m.s.

method is based on the contact changes between residues from different domains as a result of the domain movements.

### 3. Results

Isomorphous crystals of all of the three mutant proteins which diffracted synchrotron radiation to medium or high resolution were obtained, with D82E diffracting the best of all with a high resolution limit of 1.81 Å. Diffraction data parameters and refinement statistics are shown in Table 1. For kinetic assays, a range of PBG concentrations from 0.05 to 2.5 mM were used and yielded a  $K_m$  of 1.38  $\mu\text{M}$  and a  $k_{\text{cat}}$  of  $5.8 \times 10^{-2} \text{ s}^{-1}$  for the wild-type enzyme. In contrast, the only mutant which possessed catalytic activity was D82E, which was found to have a  $K_m$  of 7.71  $\mu\text{M}$  and a  $k_{\text{cat}}$  of  $6.57 \times 10^{-5} \text{ s}^{-1}$ . Thus, the D82E mutant showed a significantly reduced activity, while the D82A and the D82N mutants were completely inactive.

Overall, the mutant structures superposed closely with each other and with the WT enzyme. The active sites of the WT enzyme and the D82E mutant clearly contain a well defined DPM moiety (Fig. 4). The cofactor has very good electron density for the C1 and C2 rings that confirms the ionic and hydrogen-bond interactions which the carboxylate side chains of the pyrroles make with neighbouring active-site residues. In contrast, the D82N mutant was found to only possess a single covalently bound porphobilinogen moiety which constitutes the C1 ring of the cofactor. Whilst Woodcock & Jordan (1994) claimed that mutant PBGD enzymes with lower activity could bind additional PBG moieties to form stable enzyme–substrate complexes and that these complexes could be separated by ion-exchange chromatography, none of the 16

**Table 2**

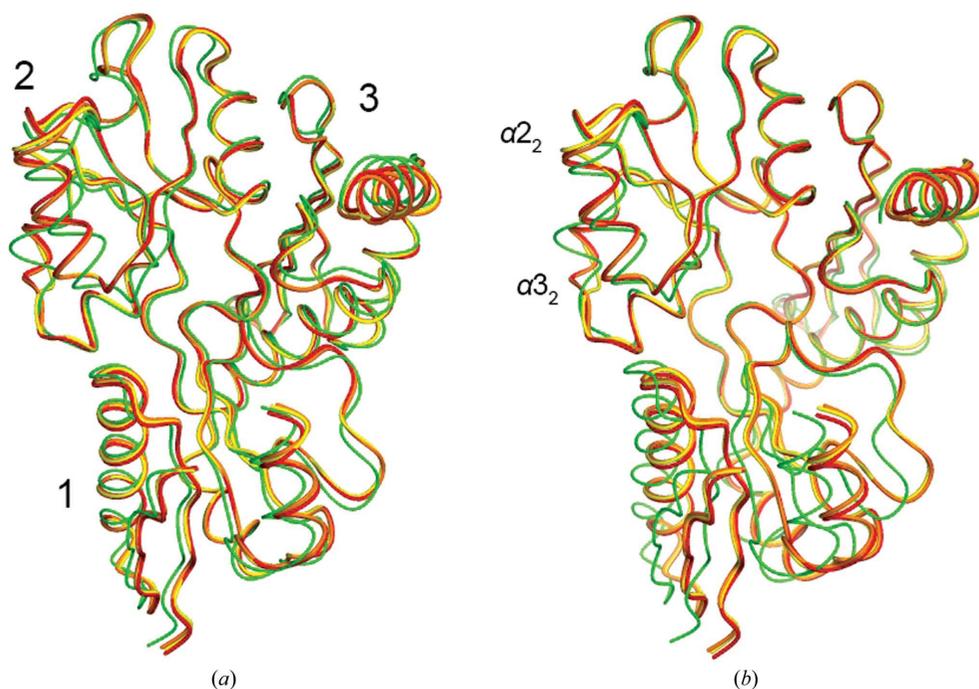
Superposition r.m.s.d. values, rotation angles and translation distances between equivalent domains of the WT and mutant PBGDs.

Mutant	R.m.s.d. (Å)				Rotation (°)	Translation (Å)
	All C $\alpha$	Domain 1	Domain 2	Domain 3		
D82A	1.12	0.43	0.87	0.49	7.4	0.2
D82E	1.04	0.37	0.86	0.49	7.2	0.1
D82N	1.07	0.42	0.89	0.44	7.2	0.2

high-resolution data sets which were collected at DLS from mutants co-crystallized with PBG demonstrated any of these putative intermediates (data not shown).

The behaviour of WT and mutant PBGD with Ehrlich's reagent is consistent with the presence of the cofactor in both the WT enzyme and the D82E mutant protein. In contrast, the absorption peaks were much lower in the D82A sample, which is consistent with the absence of the cofactor in this mutant. The intermediate behaviour of the D82N mutant is consistent with the electron-density map, which demonstrated the presence of only the C1 ring.

R.m.s.d. values of greater than 1.0 Å were observed for the mutant structures when superimposing the whole protein moiety as a rigid body with the wild-type enzyme using C $\alpha$  atoms (Fig. 5, Table 2). Superposition of domains 1 and 3 of the WT BPBGD structure individually with their counterparts in all of the mutant proteins gave r.m.s.d. values of around 0.4 Å, which demonstrates that there are no significant local differences between the structures in these domains. However, there were slightly larger differences in domain 2, with r.m.s.d. values above 0.85 Å following superposition of this domain,



**Figure 5**

Superposition of WT and mutant PBGD structures. (a) Secondary-structure superposition of the WT (green), D82A (red), D82N (yellow) and D82E (orange) PBGD structures. (b) Least-squares superposition of all of the structures by domains 2 and 3 only, indicating a clear movement of domain 1 and local movements of the  $\alpha_2$  and  $\alpha_3$  helices in domain 2.

suggesting that this domain experiences greater conformational heterogeneity than the other two domains. Analysis of the domain movements in the mutant structures relative to the WT enzyme was then undertaken using the *DynDom* server, which defines a screw axis about which the domains rotate and the associated bending residues. It was found that domains 2 and 3 tend to move together as a rigid body relative to domain 1, except for an additional local movement of the  $\alpha_2$  and  $\alpha_3$  helices in domain 2, which are at the periphery of the active-site cleft (Fig. 5). Residues 99, 100 and 198, which reside in the linker region between domain 1 and 2, are recognized by *DynDom* as bending residues, and therefore their flexibility could be key to the opening and closure of the domains during catalysis.

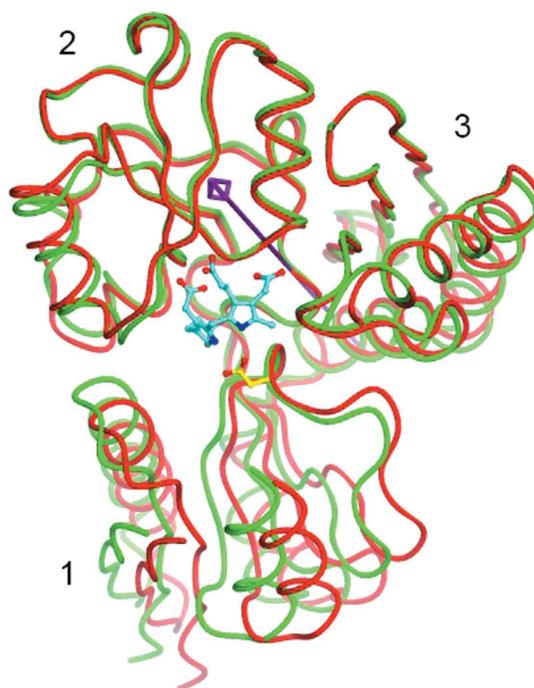
When considering domains 2 and 3 as a compound rigid body, domain 1 was found to rotate about  $7^\circ$  with negligible translation in order to superpose with the WT enzyme. Thus, the corresponding screw axis between these rigid groups in each pair of structures (WT and D82E, WT and D82N, and WT and D82A) was generated using *DynDom* (Fig. 6 shows the WT and D82A structures). *DynDom* classifies the D82A mutant as demonstrating a pure hinge motion, whilst D82E seems to have a mixed type with both hinge and shear motions. However, this is likely to be caused by the flexibility of a few side chains which are not well defined owing to the poor electron density and thus affect determination of the contact changes.

#### 4. Discussion

In superposition studies of *A. thaliana* and *E. coli* PBGD, it was found that domains 2 and 3 move in a concerted manner with respect to domain 1 (Roberts *et al.*, 2013). The same effect was also observed in this study when comparing the structures of WT and mutant *B. megaterium* enzymes. In addition, these mutant structures exhibited a marked local movement of  $\alpha$ -helices  $\alpha_2$  and  $\alpha_3$ , which further opens the active-site cleft. The differences in the relative domain orientations are emphasized most clearly when the WT and mutant structures are superposed by domains 2 and 3, as shown in Fig. 6. This is also confirmed by *DynDom*, which automatically defines domains 2 and 3 as a fixed unit and domain 1 as the moving partner. Specifically, domain 1 has a rotation of more than  $7.0^\circ$  about the inter-domain axis in all of the mutant structures compared with the WT enzyme, which leads to an opening of the active-site cleft. To date, as far as we are aware, this is the largest domain rotation that has been reported for this enzyme family. In addition there is only a small translation of 0.1–0.2 Å in the mutants, which is comparable to the errors in the atomic positions and suggests that the domains move mainly as a rotation or hinge motion (Gerstein *et al.*, 1994). Although this difference in domain orientation may partly be a consequence of small differences in crystal packing owing to the different crystallization conditions between the mutants and the wild-type enzyme, the effect is still informative on the possible mechanism of substrate entry into the active site.

Amino-acid residues 40–60 are missing in the WT and mutant PBGD structures. In the structure of the *A. thaliana* enzyme, which is 41% identical to *B. megaterium* PBGD, these residues form two  $\alpha$ -helices and loop regions which altogether cover the active site of the protein (Roberts *et al.*, 2013). The absence of this region in the *B. megaterium* enzyme could be owing to proteolysis caused by the addition of thrombin to remove the His tag or by protease impurities during purification and storage of the enzyme or crystallization (Azim *et al.*, 2014). It is reasonable to assume that this region of the enzyme is more ordered *in vivo* since it is highly conserved and forms many contacts with the DPM cofactor. Hence, we modelled the missing residues of the *B. megaterium* enzyme based on the plant PBGD structure, which is 57% identical in this region, as shown in Fig. 7. This modelling seemed to be reasonable given the high structural similarity of the enzymes (r.m.s.d. of 1.0 Å for 267 residues). If the model is correct then it is clear that the movement of domains 2 and 3 away from domain 1 (which is also enhanced by the local movement of helix  $\alpha_2$  in domain 2) could be critical to let the substrate gain access to the active site (Fig. 7). Once the substrate has gained entry, it would be quite close to both the free  $\alpha$ -position of the C2 ring of the cofactor and the catalytic residue Asp82 of the enzyme, so that condensation can occur. Further substrate moieties can then gain access to the active site and bind in a large cavity close to the catalytic aspartate such that they become connected one by one to ultimately form a hexamer.

Although the structures of the three mutants are similar to each other, they do have some slight differences. In general,



**Figure 6**  
The inter-domain screw axis. Domain 1 rotates about the axis shown in purple, which results in the 'open' (D82A, red) and 'closed' (wild type, green) conformations of the protein. The catalytic aspartate group (Asp82) is shown in yellow ball-and-stick representation with the DPM cofactor in cyan.

when they are superposed by domains 2 and 3, there are small movements in domain 1. The D82E mutant has the most closed conformation, while the D82A mutant has the most open conformation, leaving D82N as the intermediate of the three mutant structures.

Inspection of the electron density in the active site of the mutant structures shows that there is no electron density for the whole of the cofactor in the D82A mutant. In contrast, there is good density for only the C1 ring of the cofactor in the D82N mutant, but in the D82E mutant there is good electron density for both of the cofactor rings.

In the D82A and D82N mutant structures, the arginine residues (129, 130 and 153) which are close to the side chains of the C1 pyrrole ring appear to adopt native-like conformations which would allow them to interact strongly with a pyrrole (only present in D82N). In contrast, the amino acids which are close to the C2 ring in the WT structure are all moved back to some distance (>3.15 Å) in both the D82A and D82N structures, revealing for the first time the likely conformation of these residues in the apoenzyme. Compared with the native structure (PDB entry 4mlv), the side chains of residues Arg9, Arg147 and Asn149 which are close to the C2 ring of the cofactor in the native structure have moved significantly further away in the D82N and D82A structures. The side chains of these residues are well defined in the native structure and are reasonably so in the mutants, thus we are

confident about this conclusion. This finding gives an indication of the shape of the cofactor-binding pocket in the apoenzyme.

The catalytic activity measurements are consistent with the structural information that has been obtained showing that the only mutant, D82E, which has the whole cofactor bound in a well ordered manner is catalytically active, while the others are not.

### Acknowledgements

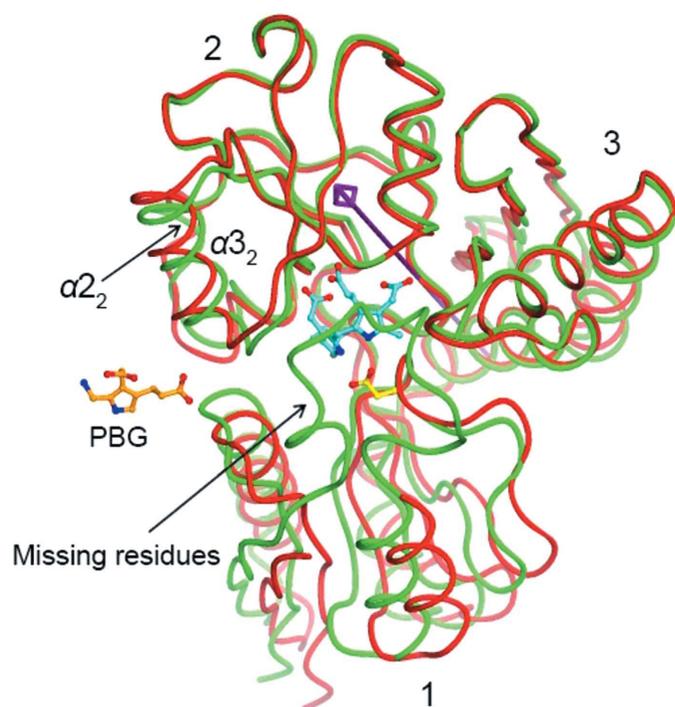
We thank Eunice Akintade for assistance with the mutagenesis.

### Funding information

We gratefully acknowledge Diamond Light Source for synchrotron X-ray beam time and travel support for data collection (award MX8922).

### References

- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352–367.
- Awan, S. J., Siligardi, G., Shoolingin-Jordan, P. M. & Warren, M. J. (1997). *Biochemistry*, **36**, 9273–9282.
- Azim, N., Deery, E., Warren, M. J., Erskine, P., Cooper, J. B., Wood, S. P. & Akhtar, M. (2013). *Acta Cryst.* **F69**, 906–908.
- Azim, N., Deery, E., Warren, M. J., Wolfenden, B. A. A., Erskine, P., Cooper, J. B., Coker, A., Wood, S. P. & Akhtar, M. (2014). *Acta Cryst.* **D70**, 744–751.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Chadwick, D. J. & Ackrill, K. (1994). *The Biosynthesis of the Tetrapyrrole Pigments*. Chichester: John Wiley & Sons.
- 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.
- Delfau, M. H., Picat, C., de Rooij, F. W., Hamer, K., Bogard, M., Wilson, J. H., Deybach, J. C., Nordmann, Y. & Grandchamp, B. (1990). *J. Clin. Invest.* **86**, 1511–1516.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Gerstein, M., Lesk, A. M. & Chothia, C. (1994). *Biochemistry*, **33**, 6739–6749.
- Gill, R., Kolstoe, S. E., Mohammed, F., Al d-Bass, A., Mosely, J. E., Sarwar, M., Cooper, J. B., Wood, S. P. & Shoolingin-Jordan, P. M. (2009). *Biochem. J.* **420**, 17–25.
- Hayward, S. & Berendsen, H. J. (1998). *Proteins*, **30**, 144–154.
- Hayward, S. & Lee, R. A. (2002). *J. Mol. Graph. Model.* **21**, 181–183.
- Jordan, P. M. (1991). *Biosynthesis of Tetrapyrroles*, edited by P. M. Jordan, pp. 1–66. Amsterdam: Elsevier. [https://doi.org/10.1016/S0167-7306\(08\)60108-8](https://doi.org/10.1016/S0167-7306(08)60108-8).
- Jordan, P. M. & Warren, M. J. (1987). *FEBS Lett.* **225**, 87–92.
- Jordan, P. M. & Woodcock, S. C. (1991). *Biochem. J.* **280**, 445–449.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 133–144.
- Kannangara, C. G., Andersen, R. V., Pontoppidan, B., Willows, R. & von Wettstein, D. (1994). *Ciba Found. Symp.* **180**, 3–20.
- Kannangara, C. G., Gough, S. P., Bruyant, P., Hooper, J. K., Kahn, A. & von Wettstein, D. (1988). *Trends Biochem. Sci.* **13**, 139–143.
- Lambert, R., Brownlie, P. D., Woodcock, S. C., Louie, G. V., Cooper, J. C., Warren, M. J., Jordan, P. M., Blundell, T. L. & Wood, S. P. (1994). *Ciba Found. Symp.* **180**, 97–104.
- Lander, M., Pitt, A. R., Alefounder, P. R., Bardy, D., Abell, C. & Battersby, A. R. (1991). *Biochem. J.* **275**, 447–452.
- Leeper, F. J. (1985). *Nat. Prod. Rep.* **2**, 561–580.



**Figure 7**  
PBGD domain movements. A superposition made using a model of the missing residues in the active-site flap of PBGD, which were built according to the structure of the *A. thaliana* enzyme. This suggests a possible route for the substrate, shown as a ball-and-stick model to the left of the protein, to gain access to the active site. The catalytic aspartate group (Asp82) is also shown in yellow ball-and-stick representation along with the DPM cofactor coloured cyan.

- Leslie, A. G. W. (2006). *Acta Cryst.* **D62**, 48–57.
- Li, J.-M., Brathwaite, O., Cosloy, S. D. & Russell, C. S. (1989). *J. Bacteriol.* **171**, 2547–2552.
- Louie, G. V. (1993). *Curr. Opin. Struct. Biol.* **3**, 401–408.
- Louie, G. V., Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Malashkevich, V. N., Hädener, A., Warren, M. J. & Shoolingin-Jordan, P. M. (1996). *Proteins*, **25**, 48–78.
- Louie, G. V., Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Warren, M. J., Woodcock, S. C. & Jordan, P. M. (1992). *Nature (London)*, **359**, 33–39.
- Mauzerall, D. & Granick, S. (1956). *J. Biol. Chem.* **219**, 435–446.
- Mills-Davies, N. *et al.* (2017). *Acta Cryst.* **D73**, 9–21.
- 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.
- Pluscec, J. & Bogorad, L. (1970). *Biochemistry*, **9**, 4736–4743.
- Roberts, A., Gill, R., Hussey, R. J., Mikolajek, H., Erskine, P. T., Cooper, J. B., Wood, S. P., Chrystal, E. J. T. & Shoolingin-Jordan, P. M. (2013). *Acta Cryst.* **D69**, 471–485.
- Rodgers, K. R. (1999). *Curr. Opin. Chem. Biol.* **3**, 158–167.
- Scott, A. I., Clemens, K. R., Stolowich, N. J., Santander, P. J., Gonzalez, M. D. & Roessner, C. A. (1989). *FEBS Lett.* **242**, 319–324.
- Shoolingin-Jordan, P. M., Warren, M. J. & Awan, S. J. (1997). *Methods Enzymol.* **281**, 317–327.
- Taylor, D., Cawley, G. & Hayward, S. (2013). *PLoS One*, **8**, e81224.
- Taylor, D., Cawley, G. & Hayward, S. (2014). *Bioinformatics*, **30**, 3189–3196.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Warren, M. J. & Jordan, P. M. (1988). *Biochemistry*, **27**, 9020–9030.
- Warren, M. J. & Smith, A. J. (2009). *Tetrapyrroles: Birth, Life and Death*. Austin: Landes Biosciences.
- Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* **D57**, 122–133.
- Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). *Methods Enzymol.* **374**, 300–321.
- Winter, G. (2010). *J. Appl. Cryst.* **43**, 186–190.
- Wood, S., Lambert, R. & Jordan, P. M. (1995). *Mol. Med. Today*, **1**, 232–239.
- Woodcock, S. C. & Jordan, P. M. (1994). *Biochemistry*, **33**, 2688–2695.