Structure of human porphobilinogen deaminase at 2.8 Å: the molecular basis of acute intermittent porphyria

Raj GILL*, Simon E. KOLSTOE*, Fiyaz MOHAMMED†[1], Abeer AL D-BASS†[2], Julie E. MOSELY†[3], Mohammed SARWAR†, Jonathan B. COOPER*, Stephen P. WOOD* and Peter M. SHOOLINGIN-JORDAN†[4]

*Centre for Amyloidosis and Acute Phase Proteins, Department of Medicine, UCL Medical School, Rowland Hill Street, London NW3 2PF, U.K., and †Biomolecular Sciences Group, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, U.K.

Mutations in the human PBGD (porphobilinogen deaminase) gene cause the inherited defect AIP (acute intermittent porphyria). In the present study we report the structure of the human uPBGD (ubiquitous PBGD) mutant, R167Q, that has been determined by X-ray crystallography and refined to 2.8 Å (1 Å = 0.1 nm) resolution ($R_{factor} = 0.26$, $R_{free} = 0.29$). The protein crystallized in space group $P2_12_12$ with two molecules in the asymmetric unit ($a = 81.0$ Å, $b = 104.4$ Å and $c = 109.7$ Å). Phases were obtained by molecular replacement using the *Escherichia coli* PBGD structure as a search model. The human enzyme is composed of three domains each of approx. 110 amino acids and possesses a dipyrromethane cofactor at the active site, which is located between domains 1 and 2. An ordered sulfate ion is hydrogen-bonded to Arg^{26} and Ser^{28} at the proposed substrate-binding site in domain 1. An insert of 29 amino acid residues, present only in mammalian PBGD enzymes, has been modelled into domain 3 where it extends helix $\alpha 2_3$ and forms a β-hairpin structure that contributes to a continuous hydrogen-bonding network spanning domains 1 and 3. The structural and functional implications of the R167Q mutation and other mutations that result in AIP are discussed.

Key words: acute intermittent porphyria, dipyrromethane, haem, human porphobilinogen deaminase, X-ray structure.

INTRODUCTION

PBGD (porphobilinogen deaminase; EC 4.3.4.8, also referred to as hydroxymethylbilane synthase, preuroporphyrinogen synthase or uroporphyrinogen I synthase) is the third enzyme in the haem biosynthesis pathway in mammals [1]. The reaction catalysed by PBGD involves the formation of preuroporphyrinogen, a linear tetrapyrrole (bilane), by the extension of an enzyme-bound DPM (dipyrromethane) cofactor that acts as a reaction primer [2,3]. This is achieved by the sequential binding, deamination and condensation of four molecules of the substrate(S) PBG (porphobilinogen) through covalently bound enzyme (E) intermediates ES, ES_2, ES_3 and ES_4 linked to the DPM cofactor. The resulting enzyme-bound hexapyrrole (ES_4) is then hydrolysed, releasing the unstable tetrapyrrole product, preuroporphyrinogen (also called 1-hydroxymethybilane) with regeneration of the enzyme with the DPM cofactor still bound at the active site.

PBGD possesses several novel features. First, the initial protein translation product is an apoenzyme that has the ability to construct its own DPM cofactor using two of the pyrrole units of preuroporphyrinogen [4]. Secondly, each of the four substrate condensation steps occurs at a single catalytic site [5] and the enzyme is therefore able to translocate the growing polypyrrole chain to vacate the substrate-binding site for the next incoming substrate. Thirdly, the enzyme can 'count' precisely to four and terminate the polymerization reaction by hydrolysis when the hexapyrrole chain (ES_4) has been assembled [3].

Comparison of the primary amino acid sequences of PBGDs from a wide range of organisms demonstrates a high degree of conservation (see Supplementary Figure S1 at http://www.BiochemJ.org/bj/420/bj420017add.htm). For example, there is 60 % similarity and 43 % identity between the deaminases from *Escherichia coli* and humans [6] (see Figure 1). PBGD from *E. coli* was the first haem biosynthesis enzyme structure to be solved, initially at 1.90 Å (1 Å = 0.1 nm) resolution [5] and subsequently refined to 1.76 Å [8]. The *E. coli* enzyme is composed of three α/β domains of approximately equal size, linked together by flexible hinge regions. Domains 1 (N-terminal) and 2 are linked by two hinge segments and each possess a similar fold based on a five-stranded, mixed β-sheet. There are relatively few direct interactions between the two domains, which form an extensive active site cleft at their interface. Both domains 1 and 2 have the same overall topology as found in the transferrins and a number of bacterial periplasmic binding proteins, which are known to adopt 'open' and 'closed' states in response to ligand binding [6]. Domain 3 (C-terminal) possesses an open-faced anti-parallel β-sheet of three strands, one face of which is covered by three α-helices. A loop from domain 3 carries an invariant cysteine that forms the covalent attachment site for the DPM cofactor. Domain 3 interacts equally with both domains 1 and 2, the packing being mediated primarily through polar interactions of loop regions.

Mutations in human PBGD cause AIP (acute intermittent porphyria), an inherited autosomal dominant disorder characterized by colicky abdominal pain and both peripheral and central

Abbreviations used: ALA, 5-aminolevulinic acid; ALAS, 5-aminolaevulinate synthase; CRIM, cross-reacting immunological material; CCD, charge-coupled device; DTT, dithiothreitol; RMSD, root mean square deviation; PBGD, porphobilinogen deaminase; uPBGD, ubiquitous PBGD.

[1] Present address: Cancer Research UK Institute for Cancer Studies, School of Cancer Sciences, University of Birmingham, Vincent Drive, Edgbaston, Birmingham B15 2TT, U.K.

[2] Present address: Department of Biochemistry, College of Sciences, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.

[3] Present address: Neurology and Gastrointestinal Centre of Excellence for Drug Discovery, GlaxoSmithKline, New Frontiers Science Park, Third Avenue, Harlow, Essex CM19 5AW, U.K.

[4] To whom correspondence should be addressed (email pmsj@soton.ac.uk).

The structure of human ubiquitous porphobilinogen deaminase has been deposited in the PDB under accession code 3EQ1.

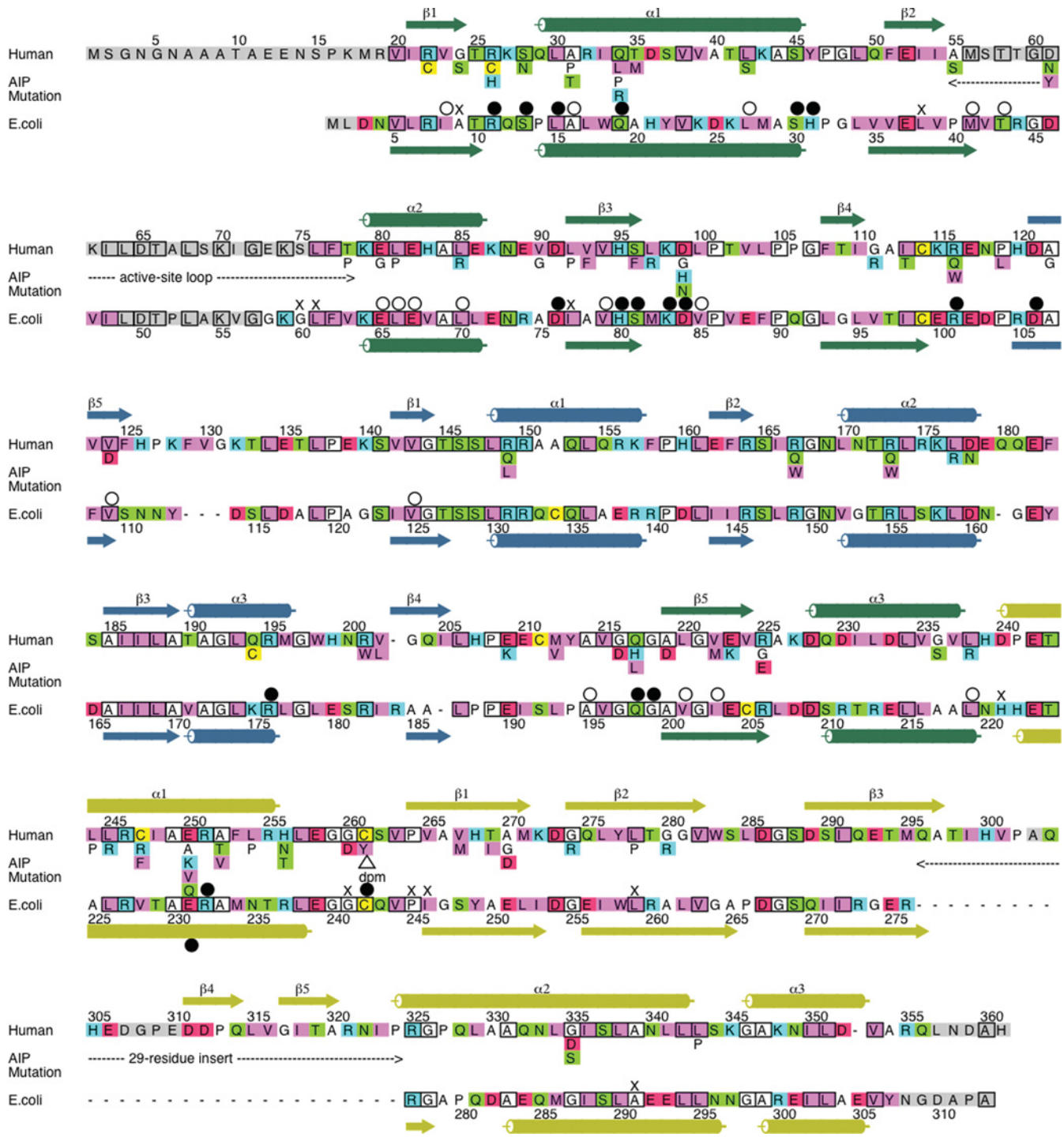


Figure 1 Structural alignment of the amino acid sequences of human uPBGD and *E. coli* PBGD

Identical residues are boxed and the secondary structural elements are coloured according to domain. N- and C-terminal residues, active site loop residues and the human domain 3 insert residues not visible in the electron density map are highlighted in grey. AIP-associated single-residue mutations are displayed beneath the human uPBGD sequence. Residues conserved across PBGDs from all species are indicated by ●, conserved substitutions are indicated by ○ and semi-conservative substitutions are indicated by ×. The sequence alignment was depicted with ALSCRIPT [7]. Amino acids are colour-coded such that red indicates acidic residues, blue indicates basic residues, green indicates neutral residues, pink indicates hydrophobic residues, yellow indicates cysteine residues and white indicates the conformationally important glycine, alanine and proline residues.

neurological symptoms accompanied by elevated levels of the haem precursors ALA (5-aminolevulinic acid) and PBG in the urine. AIP is speculated to have been the cause of the madness of King George III [9]. In humans, alternative splicing of the PBGD mRNA gives rise to two monomeric isoenzymes, uPBGD (ubiquitous PBGD) [10] that is expressed in all cells (361 residues; M_r 39620) and ePBGD (erythroid PBGD) [11] that lacks the first 17 amino acids (344 residues; M_r 38120). To date, over 250 mutations resulting in human PBGD deficiency are listed in The Human Gene Mutation Database

(http://www.hgmd.cf.ac.uk/ac/gene.php? gene=HMBS) and references therein. Over one-third of these mutations are nucleotide substitutions resulting either in single amino acid substitutions or premature chain termination. The likely impact of mutations on the structure and activity of human PBGD was originally assessed from a homology model based on the *E. coli* enzyme structure [12,13]; however, in regions of low homology and particularly in the vicinity of the 29-residue insert in the human enzyme sequence, a confident prediction of the functional consequences of mutations was not possible.

In this paper, we present the X-ray structure of the R167Q mutant of human uPBGD. This was one of the first PBGD mutations to be identified [14] and is particularly interesting because the enzyme is weakly active (8.5 %), has a reduced pH optimum and accumulates long-lived ES, ES_2, ES_3 and ES_4 enzyme–intermediate complexes that are potentially able to provide information about the mechanism of pyrrole polymerization.

MATERIALS AND METHODS

Protein expression and purification

Human uPBGD was expressed in *E. coli* strain BL21(DE3) pLysS after transformation with the pUHD2 plasmid, a derivative of pT7-7 [15]. The uPBGD insert in pUHD2 was constructed from a cDNA kindly provided by Professor Bernard Grandchamp (Hôpital Bichat, Paris, France). The R167Q mutant of human PBGD was constructed by site-directed mutagenesis and expressed and purified using a purification procedure similar to that used for the native enzyme as described below [16].

Six 800 ml flasks of 2TY containing 100 μg/ml ampicillin were inoculated with overnight cultures of transformed *E. coli* and grown at 37 °C with shaking to a D_{600} (attenuance) of 0.6. IPTG (isopropyl β-D-thiogalactoside) was then added to 1 mM and bacterial growth was continued for a further 3 h. All subsequent steps were performed at 4 °C except where noted. The cells were pelleted by centrifugation at 4500 ***g***, washed with 0.9 M NaCl and re-centrifuged. The bacterial pellet was then resuspended in 80 ml of 20 mM Tris/HCl buffer (pH 8.2) containing 5 mM DTT (dithiothreitol) and 200 μM PMSF and sonicated for 20 cycles of 30 s bursts using an MSE Soniprep 150 sonicator, interspersed with periods of cooling for 90 s. The sonicate was placed in a flask filled with N_2 gas, heated rapidly to 60 °C in a water bath for 10 min and then rapidly cooled to 4 °C by immersion in an ice/water mixture with stirring. This heat treatment step inactivated uroporphyrinogen III synthase and removed a great deal of unwanted protein. The supernatant was then applied to a Pharmacia 50 K/30 column packed with 150 ml of DEAE-Sephacel anion-exchange resin that had been equilibrated in 50 mM Tris/HCl buffer (pH 8.2) containing 5 mM DTT and 100 μM PMSF. The column was washed with 1 litre of the same buffer to remove unbound contaminating proteins. Bound proteins were eluted by the application of a linear KCl gradient (0–70 mM; 450 ml total volume at a flow rate of 0.35 ml/min). The deaminase was then eluted isocratically over approx. 100 ml at 70 mM KCl. Active fractions were pooled and concentrated to a final volume of 10 ml using an Amicon ultrafiltration cell (Millipore) fitted with a PM-10 membrane.

The deaminase, which was homogeneous as judged by SDS/PAGE (results not shown), was further concentrated to 1 ml using a Centricon YM-10 centrifugal filter (Millipore) and gel-filtered at 0.5 ml/min on a Hiload™ 16/60 Superdex G-75 (Pharmacia) preparation grade column (120 cm × 1.6 cm) equilibrated in 100 mM Tris/HCl buffer (pH 8.2) containing 5 mM DTT and 100 μM PMSF. Peak fractions with enzymatic activity were pooled and concentrated to 1 ml using a Centricon YM-10. Finally, the deaminase was buffer-exchanged into 20 mM Tris/HCl buffer (pH 8.2) containing 5 mM DTT using a Pharmacia PD10 column. The specific activities of the purified recombinant wild-type and R167Q uPBGD proteins were determined as described previously [17]. Protein concentrations were determined from the calculated molar absorption coefficient (ε) [18] (ε_{280} 0.40 for a 0.1 % solution).

Protein crystallization

Human R167Q uPBGD was crystallized using the vapour diffusion method [19]. Hanging drops were prepared by mixing equal volumes [4 μl of protein solution (20 mg/ml) in 20 mM Tris/HCl buffer (pH 8.2) containing 5 mM DTT] and reservoir solution [0.6 M ammonium sulfate, 1.2 M lithium sulfate, 5 % (v/v) ethylene glycol, 50 mM sodium citrate (pH 5.6) and 50 mM DTT]. The hanging drops were equilibrated against 1 ml of reservoir solution and placed at room temperature in the dark.

X-ray data collection and processing

Crystals of human R167Q uPBGD were transferred from mother liquor to cryoprotectant [30 % (v/v) glycerol] and flash-cooled in liquid ethane. Diffraction data were collected at the ESRF (European Synchrotron Radiation Facility, Grenoble, France) at station BM14 linked to a MAR CCD (charge-coupled device) detector at 100 K using an Oxford Cryosystems Cryostream cooler. A total of 180 0.5° oscillations were collected with an exposure time of 120 s per frame at a crystal-to-detector distance of 145 mm. Intensity data were processed with MOSFLM [20] and sorted, scaled and merged using the CCP4 (Collaborative Computational Project 4) suite of programs [21].

Structure determination and refinement

The structure of human R167Q uPBGD was solved by molecular replacement using the program MOLREP [22]. The search model consisted of the *E. coli* PBGD structure (PDB accession code 1PDA) with the DPM cofactor and water molecules omitted.

The model was refined using CNS [23] and PHENIX [24] with 21934 unique reflections in the 46.8–2.8 Å resolution range. The model was initially subjected to a round of rigid body refinement and torsion angle simulated annealing (slow cool from 5000 to 300 K in 25 K decrements) using CNS. This was followed by a number of rounds of model building in COOT [25], grouped B-factor and restrained co-ordinate refinement in the program phenix.refine [26]. Two-fold NCS (non-crystallographic symmetry) restraints, bulk solvent and anisotropy correction were used throughout the restrained refinement process, which was monitored using the R_{factor} and R_{free} statistics [27].

RESULTS

Specific activity

The specific activity at pH 8.0 of the purified R167Q mutant deaminase was 0.12 μmol of porphyrin formed/h per mg compared with the wild-type deaminase specific activity of 1.4 μmol of porphyrin formed/h per mg.

X-ray structure analysis

Thin colourless plate-like crystals grew after several weeks. The best crystal diffracted X-rays to a resolution of 2.8 Å and the

space group was determined to be orthorhombic (space group $P2_12_12$) with cell dimensions of $a = 81.0$Å, $b = 104.4$Å and $c = 109.7$Å. The calculated crystal packing parameter V_M, assuming two uPBGD molecules (84 kDa) per asymmetric unit, is 2.63 Å^3·Da^{-1}, which corresponds to a solvent content of 53 % [28]. Although the data extended beyond 2.8 Å resolution, the intensity of the diffraction pattern for this crystal form was anisotropic (being particularly weak in the direction of the k reciprocal axis); nevertheless the completeness of the data with $I/\sigma > 3$ was good.

The rotation function yielded two significant peaks at 7.5σ and 6.7σ, corresponding to the two molecules in the crystal asymmetric unit. The top rotation peak yielded a translation function peak of 13.2σ with the next peak being at 6.9σ. This solution was fixed and a second translation search was performed to determine the relative position of the second molecule, yielding a top peak of 6.0σ, with a correlation coefficient of 0.43 and an R_{factor} of 0.52. After the positioning of both molecules in the target asymmetric unit, the crystal packing of this solution was viewed using MOLPACK [29] and displayed sensible crystal contacts. After rigid body refinement and simulated annealing, the R_{factor} decreased to 0.36 ($R_{free} = 0.40$). Initial examination of the averaged electron density maps revealed prominent F_o–F_c density for the DPM cofactor and for most of the 29-residue insert in domain 3, confirming the validity of the molecular replacement solution. The model was improved by manual rebuilding, during which the uPBGD residues were fitted to the $2F_o$–F_c density. Twenty-four residues of the 29-residue insert in domain 3 and the DPM cofactor were built in during the later stages of refinement. Electron density was only visible from Val^{20} onwards at the N-terminus and residues 56–76 of the active-site loop were also disordered. This loop becomes disordered in some deaminase structures where the cofactor has become oxidized and the enzyme inactivated. Indeed, the rigid oxidized cofactor generally leads to a considerable improvement in crystal order. Nevertheless, the electron density here was of sufficient quality to build most of the human sequence into the map, to locate the cofactor and to establish the position and conformation of the 29-residue insert which is absent in the search model. The final structure was validated using PROCHECK [30] and MolProbity [31] and the co-ordinates were deposited in the PDB with the accession code 3EQ1. The final working R_{factor} was 0.26 ($R_{free} = 0.29$). The crystallographic data collection and refinement statistics are presented in Table 1.

Table 1 Data-collection and refinement statistics for human uPBGD

Values in parentheses are for the outer shell.

Parameter	Value	
Unit cell and symmetry		
Radiation source	ESRF BM14	
Radiation wavelength (Å)	1.000	
Space group	$P2_12_12$	
Unit-cell parameters		
a (Å)	81.03	
b (Å)	104.44	
c (Å)	109.73	
α (°)	90.00	
β (°)	90.00	
γ (°)	90.00	
Estimated solvent content (%)	53	
Subunits in asymmetric unit	2	
Dataset statistics		
Resolution range (Å)	46.8–2.8	(2.95–2.8)
Measured reflections	83971	(12137)
Unique reflections	21934	(3201)
R_{merge}	0.09	(0.47)
Completeness (%)	95.4	(96.4)
$\langle I/\sigma(I)\rangle$	13.2	(3.8)
Multiplicity	3.8	(3.8)
Wilson B-factor (Å^2)	43.95	
Refinement statistics		
R_{factor}	0.26	
R_{free}	0.29	
Number of reflections in free set	1711	
RMSD bond length (Å)	0.003	
RMSD bond angle (°)	0.658	
Number of protein atoms	4830	
Average B-factor protein (Å^2)	64.01	
Residues in most favoured regions (%)	94.75	
Residues in additionally allowed regions (%)	5.25	
Residues in generously allowed regions (%)	0	

Overall structure

The overall dimensions of the human uPBGD molecule are $57 \times 43 \times 32$ Å. The topologies of domain 1 (residues 20–115 and 214–238) and domain 2 (residues 116–213) are broadly similar with both consisting of a doubly wound, five-stranded mainly parallel β-sheet (one strand is antiparallel). The α-helical segments pack against each face of the β-sheets. In contrast, domain 3 (residues 239–356) comprises an open faced three-stranded antiparallel β-sheet with three α-helical segments covering one of the faces. Figure 2(A) depicts the overall three-dimensional structure of human uPBGD and a topology diagram of the structure is presented in Figure 2(B).

Electron density is clearly present for 24 residues of the 29-residue insert (electron density was absent for residues 303 and 307–310) in domain 3, which is not present in the *E. coli* enzyme and which is positioned after strand $\beta3_3$ (residues 290–298). The polypeptide chain of the human enzyme meanders away from the C-terminus of $\beta3_3$, leading to a β-hairpin with residues 312–315 ($\beta4_3$) and residues 318–321 ($\beta5_3$) making up the antiparallel strands. Main-chain hydrogen bonds (Leu^{315}NH-Ile^{318}CO and Leu^{315}CO-Ile^{318}NH) give way to a main chain–side chain interaction (Asp^{312}CO-Arg^{321}Nε) as the strands diverge. The apical residue of the hairpin, Val^{316}, hydrogen-bonds through its carbonyl oxygen with the side chain of Arg^{251}. The hairpin packs on top of $\alpha1_3$ (residues 241–256) and $\alpha2_3$ (residues 325–343) with the hydrophobic side chains Leu^{315}, Val^{316}, Ile^{318} and Ala^{320} that protrude beneath the hairpin interacting with Leu^{329}, Ile^{248} and Leu^{244} of the domain 3 helices. There are also a range of interactions between the hairpin and domain 1, in particular hydrogen bonds between the main chain atoms of the hairpin and the terminal residue of $\beta4_1$, the last strand of domain 1 (Gly^{317}CO-Ile^{110}NH and Thr^{319}NH-Ile^{110}CO). This hydrogen-bonding pattern comes close to defining an extended seven-stranded sheet linking domains 1 and 3. Beyond the hydrogen-bonded region the relative strand twist is too large to define such a sheet formally. However, there are additional interactions, notably the Gln^{314} side chain hydrogen-bonds with the backbone carbonyl of Phe^{108} and the OG atoms of Thr^{109} and Thr^{319} are hydrogen-bonded with one another, further stabilizing the interaction between domains 1 and 3. Following this excursion into the β-hairpin, helix $\alpha2_3$ (residues 325–343) extends for 19 residues (5 residues longer than the equivalent helix in the *E. coli* enzyme).

The β-hairpin is also involved in a range of packing interactions with neighbouring molecules in the crystal. The elongated loops that lead into the β-hairpin between strands $\beta4_3$ and $\beta5_3$ nestle in a solvent channel between two symmetry-related neighbours.

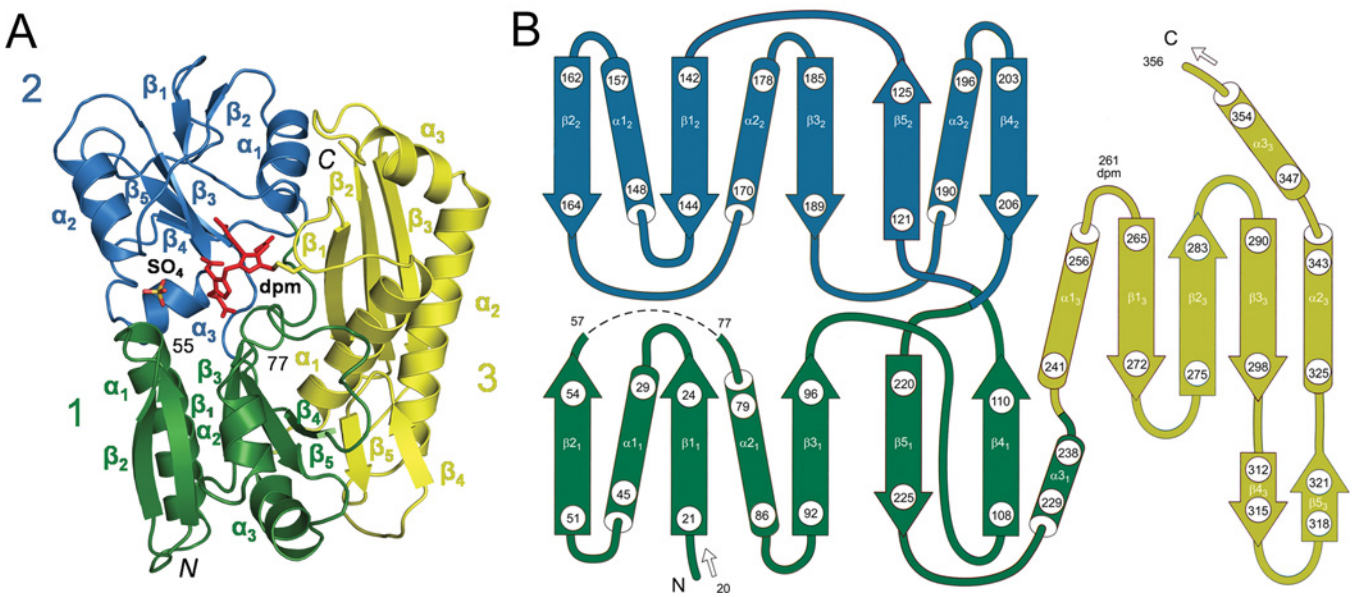


Figure 2 Tertiary and secondary structures of human uPBGD

(**A**) The structure of human uPBGD with the three domains coloured differently: domain 1 (green), domain 2 (blue) and domain 3 (yellow). The DPM cofactor (red) and sulfate ion are located at the catalytic site between domains 1 and 2. (**B**) The topology of the secondary structure of human uPBGD. Domain 1 (green) extends from residues 20 to 115 and residues 214 to 238. Domain 2 (blue) comprises residues 116–213. Domain 3 (yellow) extends from residue 239 to residue 356 and contains the cofactor-binding Cys^{261} residue. The disordered residues (56–76) and the location of the DPM cofactor are indicated.

Intermolecular contacts in the crystal

Human uPBGD crystallizes with two molecules in the asymmetric unit, related by a local 2-fold axis. Both molecules interact through a substantial interface (1472Å^2) stabilized by a series of hydrophobic contacts and a number of salt bridges and hydrogen bonds (Figure 3). The contacts arise from helix $\alpha 1_1$ (residues 29–46), the loop (residues 126–141) connecting strands $\beta 5_2$ (residues 121–125) and $\beta 1_2$ (residues 142–144), the C-terminal of helix $\alpha 3_2$ and the proceeding loop region (residues 189–196). Analysis by ProFace [32] demonstrates that the dimer interface area is only 5 % of the total solvent accessible surface area of the individual protomers, whereas the average value for a stable dimer is 16 % [33]. Furthermore, the non-polar fraction of the interface area (881 Å^2) accounts for only 14 % of the fully buried atoms in the structure, whereas the average value for a stable dimer is 65 % [34]. Indeed there is no evidence to suggest the formation of stable dimers of human uPBGD on gel filtration at neutral pH [35].

Interactions between domains

There are mainly polar contacts between the three domains in the human deaminase molecule, suggesting a flexible structure similar to that proposed for *E. coli* PBGD [5]. With the exception of the main-chain connection between domains 1 and 2, there is also a hydrophobic between Gln^{29} and Met^{196}. The 29-residue insert following the third strand in domain 3 ($\beta 3_3$), which is absent in the *E. coli* deaminase structure, is involved in packing against domain 1, the interaction being mediated primarily by hydrogen-bonding.

In addition to the mainly hydrophilic interactions at the inter-domain regions of the human structure, there are also several hydrophobic interactions. Residues at the interface between domains 1 and 3 include Leu^{97}, Leu^{244} and Cys^{247}, and between domains 2 and 3, residues Ala^{152}, Met^{212} and Leu^{285} are involved. The interface between domains 2 and 3 also contains the bulky aromatic Trp^{283} that packs against Gln^{153}.

Structural comparison between human and *E. coli* PBGDs

The three-dimensional structure of human PBGD has many features in common with the *E. coli* PBGD structure, most notably the presence of three protein domains. The main differences between the human and *E. coli* PBGD structures occur as insertions in loop regions. For example, there is a three-residue insertion occurring between strands $\beta 5_2$ and $\beta 1_2$ (residues 129–131), a single-residue insertion between helix $\alpha 2_2$ and strand $\beta 3_2$ (residue 180) and a 29-residue insertion between strand $\beta 3_3$ and $\alpha 2_3$ (residues 296–324). Of the 360 residues in human uPBGD there are only 148 sequence identities with the *E. coli* enzyme. These identities partly cluster around the active site but otherwise are distributed evenly over the protein structure. The differences are not especially concentrated in loop regions connecting secondary structure elements but often involve hydrophobic residues of the core. The close conservation of the tertiary structure is achieved by a subtle blend of compensatory substitutions arising from different parts of the sequence that come together in the fold.

Superposition of the conserved secondary structural regions of human uPBGD on to those of *E. coli* PBGD yields an RMSD (root mean square deviation) of 0.97 Å over 160 C_α atoms. Superposition of individual domains demonstrates that domain 2 of the human uPBGD structure superposes well on to the corresponding domain of *E. coli* PBGD, yielding an RMSD of 0.61 Å (over 41 C_α atoms). Domains 1 and 3 yield RMSD of 0.75 Å (over 62 C_α atoms) and 0.81 Å (over 86 C_α atoms) respectively. When corresponding domains are individually overlaid, the other two domains are also well superposed, demonstrating that there is no major perturbation of domain relationships between the bacterial and human enzymes. However,

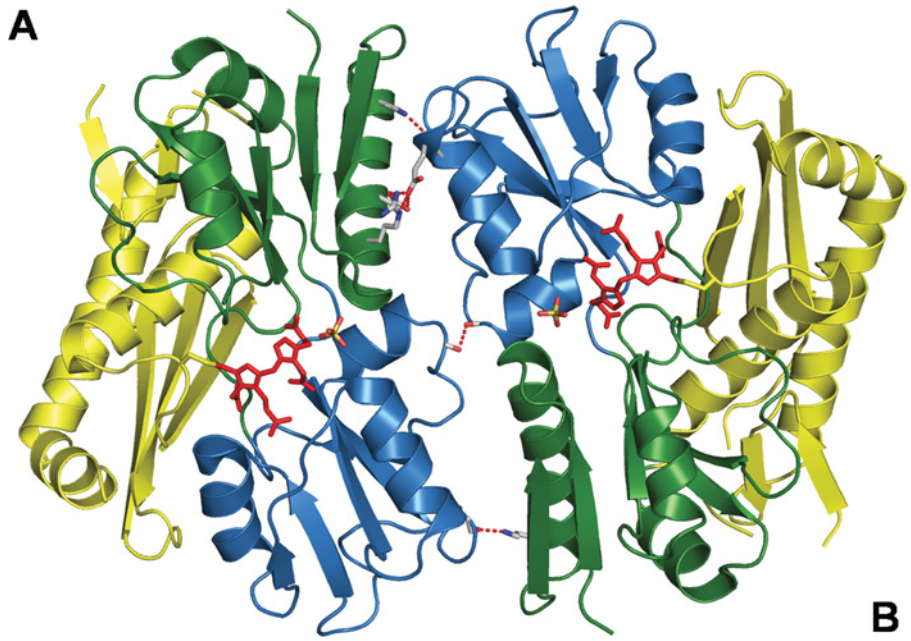


Figure 3 The two molecules of human uPBGD in the asymmetric unit

Although both molecules (labelled **A** and **B**) interact through a substantial interface stabilized by hydrophobic contacts, salt bridges and hydrogen bonds, they probably do not constitute a physiological dimer.

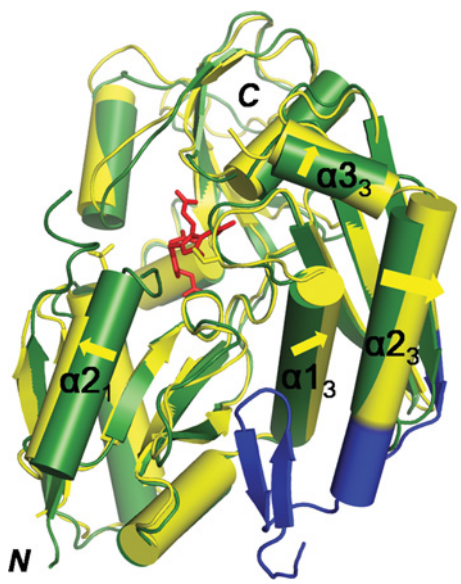


Figure 4 Although the human (yellow) and *E. coli* (green) deaminase structures superpose well, the 29-residue domain 3 insert in the human enzyme (in blue) has pushed domains 1 and 3 apart

The DPM cofactor (red) is in the oxidized conformation in the *E. coli* deaminase structure but in the reduced conformation in the human deaminase structure.

inspection of the distribution of small structural differences in the overlays demonstrates that insertion of the β-hairpin motif between $\beta 5_1$ in domain 1 and $\alpha 1_3$ and $\alpha 2_3$ in domain 3 interface has pushed these domains apart (Figure 4).

The close equivalence of the *E. coli* and human enzyme structures demonstrated by their superposition implies that the volumes of their respective hydrophobic cores are very similar in spite of the many substitutions that occur in these regions. In domain 1, Leu^{6}, Ile^{8}, Ala^{20}, Val^{36}, Leu^{38}, Ile^{77}, Leu^{93}, Ile^{203}, Cys^{205} and Leu^{216} of *E. coli* PBGD are replaced by Ile^{21}, Val^{23}, Thr^{35}, Phe^{51}, Ile^{53}, Leu^{92}, Phe^{108}, Val^{222}, Val^{224} and Val^{235} respectively in the human uPBGD structure. In domain 2, Ile^{145}, Val^{152}, Tyr^{164} and Leu^{179} in the *E. coli* deaminase structure are replaced by Phe^{163}, Leu^{170}, Phe^{183} and Trp^{198} respectively in the human deaminase structure. In all cases where there is a substantial change in side chain volume, this is accompanied by compensatory changes and adjustments in adjacent residues. Most of the residues that contribute towards the hydrophobic core in domain 3 of the human enzyme possess small aliphatic side chains that are also conserved in the *E. coli* deaminase structure, occurring between two roughly parallel α-helices ($\alpha 1_3$ and $\alpha 2_3$). There are two notable differences in this region, whereby Thr^{229} and Met^{234} in the *E. coli* PBGD structure are replaced by Ile^{248} and Phe^{253} respectively in the human uPBGD structure. Side chain substitution patterns at domain interfaces are similarly compensatory and contribute to the conservation of domain relationships.

The active site of human uPBGD and the impact of the Gln^{167} mutation

The active site cleft ($15 \times 13 \times 12$ Å) containing the DPM cofactor is located at the interface between domains 1 and 2 in a similar arrangement to that in the *E. coli* enzyme (Figure 5). The DPM cofactor is covalently bound to the protein via a thioether linkage with Cys^{261}, a residue in a type-I turn in the sequence Gly-Gly-Cys-Ser-Val-Pro (residues 259–264) located at the end of a loop connecting helix $\alpha 1_3$ and strand $\beta 1_3$ in domain 3. The electron density for the DPM cofactor indicates that it occupies the native reduced conformation, consistent with the colourless nature of the crystals. Most of the cofactor interactions with the protein are conserved between the human and *E. coli* enzymes. The acetate and propionate moieties of the DPM cofactor contribute most of the ionic interactions and hydrogen bonds to neighbouring

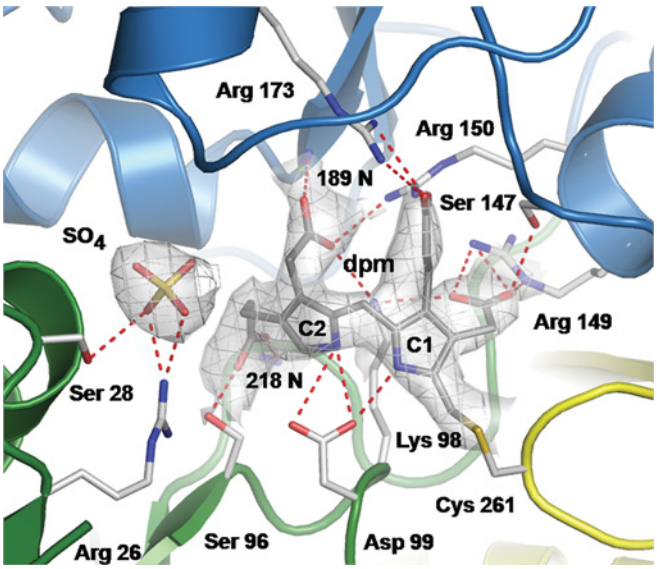


Figure 5 The active site of human uPBGD with superposed electron density (omit map) contoured at 2σ

The DPM cofactor interacts with the protein through over 25 contacts, notably with the conserved Ser^{147}, Arg^{150} and Arg^{173} residues. The catalytic Asp^{99} is hydrogen-bonded to both NH groups of the cofactor pyrrole rings and a sulfate ion is bound to Arg^{26} and Ser^{28} at the predicted substrate-binding site.

active site side chains (Figure 5). The interactions between the carboxylate groups of ring C1 and the enzyme involve several invariant arginine residues (Arg^{149}, Arg^{150} and Arg^{173}), the backbone amide nitrogen of Ser^{146} and the hydroxy and backbone carbonyl group of Ser^{147}. The DPM cofactor ring C2 is positioned towards the back of the active site cleft and interacts with Ser^{96} and the backbone amide nitrogen atoms of both Ala^{189} and Gly^{218}. The invariant Lys^{98} and Asp^{99} residues interact with both C1 and C2 rings; Lys^{98} forms salt bridges with the acetate cofactor side chains, whereas a carboxylate oxygen of the catalytic Asp^{99} residue hydrogen-bonds to both pyrrole NH atoms of the DPM cofactor.

A prominent peak of electron density was observed within hydrogen-bonding distance of Arg^{26} and Ser^{28} within the active site region and this has been refined as a sulfate ion derived from the crystal mother liquor (Figure 5). This site is occupied by the DPM ring C2 propionate in *E. coli* PBGD when the cofactor is in the oxidized linear conformation, and ring C2 occupies the proposed substrate-binding site [5]. When this site is vacant in the reduced *E. coli* enzyme, an acetate ion, again derived from the crystal mother liquor, is found at the equivalent position as the sulfate ion observed here [8]. The electron density for the mutated Gln^{167} side chain is poorly resolved in the absence of stabilizing interactions with the protein or the cofactor, as is the case for the Arg^{149} side chain in the reduced *E. coli* structure [36].

Figure 6 illustrates the remarkable stability of enzyme–intermediate complexes in the case of the R167Q mutant that hinders the turnover to form product and the regeneration of the free enzyme. These intermediates can be visualized if the samples are separated by non-denaturing gel electrophoresis at 4 °C.

DISCUSSION

Human uPBGD consists of three domains, as is the case with the prokaryotic enzyme, but is slightly larger owing to an N-terminal extension of 17 amino acids, a 29-residue insertion in domain 3 and several elongated loop regions. The enzyme crystallized

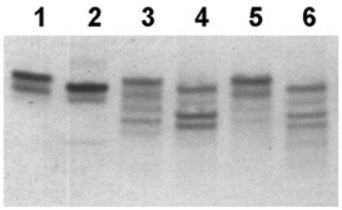


Figure 6 Non-denaturing PAGE analysis demonstrating the stability of human uPBGD enzyme–intermediate complexes generated on incubation with substrate

The enzyme was either incubated with 10 mole equivalents of PBG at 37 °C for a short period (a few seconds) in 20 mM Tris/HCl buffer (pH 8) or the incubation was prolonged until reaction of the native enzyme had proceeded to completion (1 h). All samples were cooled to 0 °C and then loaded together at 4 °C. Lane 1, native enzyme; lane 2, R167Q mutant enzyme; lane 3, native enzyme + PBG; lane 4, R167Q + PBG; lane 5, sample from lane 3 allowed to react to completion; lane 6, sample from lane 4 allowed to react to completion. Note that the purified enzyme always exhibits two protein bands on non-denaturing PAGE; the most prominent enzyme–intermediate complexes are ES_2 and ES_3.

with the DPM cofactor in the native reduced conformation with an ordered sulfate ion hydrogen-bonded to Arg^{26} and Ser^{28} at the proposed substrate-binding site.

The electron density map allowed modelling of 24 residues of the 29-residue insert in domain 3 of human uPBGD (residues 296–324) that is not present in the *E. coli* enzyme and is positioned after strand $\beta 3_3$. This insert elongates helix $\alpha 2_3$ and extends into a β-hairpin (strands $\beta 4_3$ and $\beta 5_3$) and comes close to forming an extended seven-stranded sheet linking domains 1 and 3. Because of low sequence homology in this region the insert had previously been modelled before strand $\beta 3_3$ at residue positions 289–317. Although the insert is distant from the active site, its interaction with domain 1 may modulate conformational fluctuations associated with enzyme action and perhaps promote the weak dimerization exhibited in the crystal. Although human uPBGD crystallized with two monomers in the asymmetric unit, the dimer interface is not as extensive as seen in stable (physiological) dimers of other proteins, and there is no evidence to suggest the formation of dimeric human uPBGD at neutral pH in solution. However, the possibility of a weak interaction between subunits of the human enzyme cannot be excluded, perhaps playing a role in the channelling of the unstable preuroporphyrinogen product to uroporphyrinogen III synthase or some other regulatory process.

In humans the rate-limiting reaction in haem biosynthesis is governed by ALAS (5-aminolaevulinate synthase), the first enzyme of the eight-step Shemin pathway [37]. Subsequently, PBGD has the lowest specific activity, but under normal conditions 50 % of the enzyme is adequate to sustain normal demands for haem. However, under conditions where the demand for haem is high, the induction of ALAS is triggered. Any compromise in the activity of PBGD therefore results in the accumulation of the immediate precursors, ALA and PBG, and may precipitate an attack of AIP.

The human uPBGD structure facilitates an understanding of the large number of inherited mutations that result in AIP. The original homology model of human PBGD based on the *E. coli* PBGD X-ray structure [12,13] yielded structural insights into how the many AIP-associated mutations could affect enzyme activity. Interestingly, many of the porphyria-associated mutation sites are sites that also differ between the human and *E. coli* deaminases. However, in the human PBGD structure, these isolated AIP-associated substitutions cannot sustain the active form of the enzyme.

The R167Q mutation is of particular interest since it provides insights into important aspects of the pyrrole polymerization

mechanism. The mutant enzyme is weakly active in samples taken from human blood and the recombinant human enzyme expressed in *E. coli* also exhibits an enzymatic activity of only 5–10 % of that of the native enzyme [34]. The R167Q uPBGD structure presented here demonstrates that the considerable loss of activity associated with this mutation arises from a discrete perturbation of the enzyme mechanism rather than from any major structural perturbation. Most interestingly, when incubated with PBG the R167Q mutant generates stable catalytic intermediates (ES, ES_2, ES_3 and ES_4), whereas the native enzyme turns these intermediates over rapidly. Accordingly, the R167Q mutant is CRIM (cross-reacting immunological material) positive, reflecting the accumulation of these substantially longer lived enzyme–intermediate complexes. Arg^{167} is thought to play a key role in deaminase binding to PBG by breaking the internal salt bridge between the amino and acetic acid side chains of the unbound substrate. The combination of Asp^{99} and Arg^{167} may facilitate the cleavage of the intra-molecular salt bridge to form new interactions, with the substrate amino group interacting with Asp^{99} and the substrate acetate group interacting with Arg^{167}. It has previously been observed that the equivalent Arg^{149} residue of the *E. coli* enzyme is well ordered only when interacting with the acetate side chain of a pyrrole bound at the proposed substrate-binding site [5]. The proposed role of Arg^{167} in breaking the intra-molecular salt bridge in the PBG substrate would account for the lowering of the pH optimum of human uPBGD from pH 8 to pH 6 when it is replaced by a glutamine residue, since the salt bridge is much weaker at lower pH and can break without the assistance of the enzyme. The structure presented here suggests that the mechanistic lesions associated with this mutation are only felt when the Arg^{167} side chain is required to visit the substrate site. Additionally, activity may also be reduced by the neutral Gln^{167} side chain disrupting the positively charged active site surface over which the negatively charged polypyrrole chain may 'slide' during polymerization.

PBGD mutations associated with porphyria are widely dispersed over all three domains of the enzyme structure. Over 100 of these mutations involve single base changes that result in the substitution of one amino acid residue for another (missense; summarized in Figure 1) or the formation of a truncated protein (nonsense) and it is the former group that is especially interesting for providing structure/function information for the enzyme. While the R167Q mutation discussed above exemplifies those affecting pyrrole polymerization, mutations that lead to attacks of AIP may be divided into three more broad groups according to their molecular basis, including mutations that affect protein stability and folding, cofactor assembly and the catalytic process.

Mutations of residues that are essential for protein folding and stability include those occurring in the hydrophobic core, in tight turns and other conformationally restricted areas. Their mutation may result in an abnormally folded or unstable protein that is rapidly degraded *in vivo*. Such proteins typically exhibit CRIM phenotype (CRIM negative) with antibody raised against the native protein. Because of their intrinsic instability it is not possible to study the equivalent recombinant proteins *in vitro*; however, the impact of the mutation may be assessed by inspecting its location in the human uPBGD structure. For example, Leu^{245} is a conserved buried hydrophobic side chain that interacts with a cluster of proline residues (Pro^{241}, Pro^{302} and Pro^{327}) in domain 1. Replacement of Leu^{245} by the AIP-associated Arg^{245} would bury a large positively charged side chain with resulting adverse consequences for the stability of the domain fold.

Clearly, mutations that lead to the premature introduction of a stop codon will result in a truncated protein, affecting both folding and stability. For example, the most common mutation within the PBGD gene in Swedish patients (Trp198Term) results in a severely truncated, probably inactive, and easily degraded protein, with patients exhibiting a CRIM negative phenotype [38]. The incidence of attacks of AIP in patients carrying this lesion is one of the highest, considering the low penetrance (8 %) of the disease [39]. Two truncating AIP-associated mutations also target the domain 3 insert (Gln296Term and Gln314Term) and both of these would be expected to destabilize domain 3 and interfere with the correct folding of the enzyme during biosynthesis.

Other mutations do not adversely affect apoenzyme formation but do prevent the binding, reaction and assembly of the cofactor precursor preuroporphyrinogen (1-hydroxymethylbilane). In the absence of the stabilizing influence of the DPM cofactor, the intrinsically unstable apoenzyme is denatured and broken down intracellularly. The importance of the DPM cofactor not only for enzyme activity but also for enhancing the stability of the protein is exemplified by the dramatic finding that the apoenzyme is rapidly denatured at 60 °C, whereas the holoenzyme loses little activity over 15 min at this temperature [40]. A close inspection of the structure in the vicinity of the cofactor demonstrates that it makes some 25 direct contacts with the protein. Phenotypes in this class are therefore normally CRIM negative and include, for example, the well-characterized 'base 500 mutations' that target Arg^{149} (R149Q and R149L), a residue intimately involved in cofactor assembly. Expression of recombinant R149Q human uPBGD in *E. coli* yields apoprotein with less than 5 % of wild-type activity in crude extracts and which is heat labile and completely inactive after purification [35]. A negative reaction with Ehrlich's reagent [17] confirmed that the purified mutant is devoid of the DPM cofactor, and furthermore, attempts to incorporate preuroporphyrinogen *in vitro* were unsuccessful [34].

Finally, mutations that affect catalytic residues or substrate-binding residues result in a correctly folded but inactive enzyme. Mutations at the substrate-binding site include R26H and R173Q. Mutagenesis studies at the equivalent positions in the *E. coli* enzyme (R11H and R155H [41] or R11L and R155L [42]) demonstrate that these mutations prevent the enzyme from binding the substrate. Both of these human mutant PBGDs appear to be correctly folded since they exhibit a CRIM positive phenotype and contain the DPM cofactor. However, the disruption of the PBG-binding site clearly prevents ES complex formation and both mutants are essentially inactive. Mutations involving the catalytic Asp^{99} residue are of particular interest since they trap the deaminase in the ES_2 form. Although reaction of D99G human uPBGD apoenzyme with the tetrapyrrole cofactor precursor, preuroporphyrinogen, appears to be unaffected, the resulting ES_2 complex is unable to catalyse the addition of further substrate to complete the catalytic cycle and generate the product and the DPM cofactor [43]. The related D84A mutation in *E. coli* PBGD also appears to exist as an inactive ES_2 complex [44]. Both findings support the hypothesis that the cofactor arises not from two single molecules of substrate but from a preformed tetrapyrrole [4].

All of these mutants of human PBGD have been identified on the basis that their *in vivo* activity is low enough to cause symptoms of AIP. When the CRIM status confirms that normal amounts of enzyme are produced, these mutations provide an informative repertoire of lesions in protein function that contribute towards a deeper understanding of the enzyme's mechanism of action.

ACKNOWLEDGEMENTS

We are grateful to Professor Joël Janin (Laboratoire d'Enzymologie et Biochemie Structurales, Centre National De La Researche Scientifique, Paris, France) and Dr Ranjit Bahadur (Department of Biochemistry, Bose Institute, Kolkata, India) for performing the dimer interface calculations. We thank Mr Barry Lockyer and Miss Vashti Maharaj for assisting with the preparation of the Figures.

FUNDING

The Wellcome Trust provided grants for the X-ray and MS equipment and the Biotechnology and Biological Sciences Research Council provided support for M.S. and J.M. A.A.D-B. thanks the Royal Embassy of Saudi Arabia for the award of a scholarship.

REFERENCES

1 Shoolingin-Jordan, P. M. and Cheung, K.-M. (2000) Biosynthesis of heme. In Comprehensive Natural Products Chemistry, Volume 4 (Barton, D. H. R. and and Nakanishi, K., eds.), pp. 61–107, Pergamon/Elsevier, New York
2 Jordan, P. M. and Warren, M. J. (1987) Evidence for a dipyrromethane cofactor at the catalytic site of *E. coli* porphobilinogen deaminase. FEBS Lett. **225**, 87–92
3 Warren, M. J. and Jordan, P. M. (1988) Investigation into the nature of substrate binding to the dipyrromethane cofactor of *E. coli* porphobilinogen deaminase. Biochemistry **27**, 9020–9030
4 Shoolingin-Jordan, P. M., Warren, M. J. and Awan, S. J. (1996) Discovery that the assembly of the dipyrromethane cofactor of porphobilinogen deaminase holoenzyme proceeds initially by the reaction of preuroporphyrinogen with the apoenzyme. Biochem. J. **316**, 373–376
5 Louie, G. V., Brownlie, P. D., Lambert, R., Cooper, J. B., Blundell, T. L., Wood, S. P., Warren, M. J., Woodcock, S. C. and Jordan, P. M. (1992) Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. Nature **359**, 33–39
6 Louie, G. V. (1993) Porphobilinogen deaminase and its structural similarity to the bidomain binding proteins. Curr. Opin. Struct. Biol. **3**, 401–408
7 Barton, G. J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. **6**, 37–40
8 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. and Jordan, P. M. (1996) The three-dimensional structure of *Escherichia coli* porphobilinogen deaminase at 1.76 Å resolution. Protein Struct. Funct. Genet. **25**, 48–78
9 Röhl, J. C. G., Warren, M. J. and Hunt, D. (1998) Purple Secret: Genes, 'Madness' and the Royal Houses of Europe, Bantam Press, London
10 Grandchamp, B., De Verneuil, H., Beaumont, C., Chretien, S., Walter, O. and Nordmann, Y. (1987) Tissue-specific expression of porphobilinogen deaminase: two isoenzymes from a single gene. Eur. J. Biochem. **162**, 105–110
11 Raich, N., Romeo, P. H., Dubart, A., Beaupain, D., Cohen-Solal, M. and Goossens, M. (1986) Molecular cloning and complete primary sequence of human erythrocyte porphobilinogen deaminase. Nucleic Acids Res. **14**, 5955–5968
12 Brownlie, P. D., Lambert, R., Louie, G. V., Jordan, P. M., Blundell, T. L., Warren, M. J., Cooper, J. B. and Wood, S. P. (1994) The three-dimensional structures of mutants of porphobilinogen deaminase: toward an understanding of the structural basis of acute intermittent porphyria. Protein Sci. **3**, 1644–1650
13 Wood, S. P., Lambert, R. and Jordan, P. M. (1995) Molecular basis of acute intermittent porphyria. Mol. Med. Today **1**, 232–239
14 Delfau, M. H., Picat, C., de Rooij, F. W., Hamer, K., Bogard, M., Wilson, J. H., Deybach, J. C., Nordmann, Y. and Grandchamp, B. (1990) Two different point G to A mutations in exon 10 of the porphobilinogen deaminase gene are responsible for acute intermittent porphyria. J. Clin. Invest. **86**, 1511–1516
15 Tabor, S. and Richardson, C. C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. U.S.A. **82**, 1074–1078
16 Jordan, P. M., Thomas, S. D. and Warren, M. J. (1988) Purification, crystallization and properties of porphobilinogen deaminase from a recombinant strain of *Escherichia coli* K12. Biochem. J. **254**, 427–435
17 Shoolingin-Jordan, P. M., Warren, M. J. and Awan, S. J. (1997) Dipyrromethane cofactor assembly; formation of apoenzyme and preparation of holoenzyme. Methods Enzymol. **281**, 317–327
18 Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. **182**, 319–326
19 Macpherson, A. (1982) Preparation and Analysis of Protein Crystals, John Wiley and Sons, New York
20 Leslie, A. G. W. (1992) Joint CCP4 and ESF–EACMB Newsletter on Protein Crystallography No, 26, Daresbury Laboratory, Warrington
21 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. **50**, 760–763
22 Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. **30**, 1022–1025
23 Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S. et al. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. Sect. D Biol. Crystallogr. **54**, 905–921
24 Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. Sect. D Biol. Crystallogr. **58**, 1948–1954
25 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. **60**, 2126–2132
26 Afonine, P. V., Grosse-Kunstleve, R. W. and Adams, P. D. (2005) The Phenix refinement framework. CCP4 Newsletter **42**, contribution 8
27 Brünger, A. T. (1992) Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature **355**, 472–475
28 Mathews, B. W. (1968) Solvent content of protein crystals. J. Mol. Biol. **33**, 491–497
29 Wang, D. W., Driessen, H. P. and Tickle, I. J. (1991) MOLPACK: molecular graphics for studying the packing of protein molecules in the crystallographic unit cell. J. Mol. Graphics **9**, 50–52
30 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291
31 Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S. and Richardson, D. C. (2003) Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. Protein Struct. Funct. Genet. **50**, 437–450
32 Saha, R. P., Bahadur, R. P., Pal, A., Mandal, S. and Chakrabarti, P. (2006) ProFace: a server for the analysis of the physicochemical features of protein–protein interfaces. BMC Struct. Biol. **6**, 11
33 Bahadur, R. P., Chakrabarti, P., Rodier, F. and Janin, J. (2003) Dissecting subunit interfaces in homodimeric proteins. Protein Struct. Funct. Genet. **53**, 708–719
34 Bahadur, R. P., Chakrabarti, P., Rodier, F. and Janin, J. (2004) A dissection of specific and non-specific protein–protein interfaces. J. Mol. Biol. **336**, 943–955
35 Al-Dbass, A. (2001) Structural basis of acute intermittent porphyria and the relationship between mutations in human porphobilinogen deaminase and enzyme activity, Ph.D. Thesis, University of Southampton, Southampton, U.K.
36 Hädener, A., Matzinger, P. K., Battersby, A. R., McSweeney, S., Thompson, A. W., Hammersley, A. P., Harrop, S. J., Cassetta, A., Deacon, A., Hunter, W. N. et al. (1999) Determination of the structure of selenomethionine-labelled hydroxymethylbilane synthase in its active form by multi-wavelength anomalous dispersion. Acta Crystallogr. Sect. D Biol. Crystallogr. **55**, 631–643
37 Thunell, S. (2000) Porphyrins, porphyrin metabolism and porphyrias I. Update. Scand. J. Clin. Lab. Invest. **60**, 509–540
38 Lee, J.-S. and Anvert, M. (1991) Identification of the most common mutation within the porphobilinogen deaminase gene in Swedish patients with acute intermittent porphyria. Proc. Natl. Acad. Sci. U.S.A. **88**, 10912–10915
39 Andersson, C., Floderus, Y., Wikber, A. and Lithner, F. (2000) The W198X and R173W mutations in the porphobilinogen deaminase gene in acute intermittent porphyria have higher clinical penetrance than R167W: a population-based study. Scand. J. Clin. Lab. Invest. **60**, 643–648
40 Awan, S. J., Siligardi, G., Shoolingin-Jordan, P. M. and Warren, M. J. (1997) Reconstitution of the *holo*enzyme form of *Escherichia coli* porphobilinogen deaminase from apoenzyme with porphobilinogen and preuroporphyinogen: a study using circular dichroism spectroscopy. Biochemistry **36**, 9273–9282
41 Jordan, P. and Woodcock, S. C. (1991) Mutagenesis of arginine residues in the catalytic cleft of *Escherichia coli* porphobilinogen deaminase that affects dipyrromethane cofactor assembly and tetrapyrrole chain initiation and elongation. Biochem. J. **280**, 445–449
42 Lander, M., Pitt, A. R., Alefounder, P. R., Bardy, D., Abell, C. and Battersby, A. R. (1991) Studies on the mechanism of hydroxymethylbilane synthase concerning the role of arginine residues in substrate binding. Biochem. J. **275**, 447–452
43 Shoolingin-Jordan, P. M., Al-Dbass, A., McNeill, L. A., Sarwar, M. and Butler, D. (2003) Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. Biochem. Soc. Trans. **31**, 731–735
44 Woodcock, S. C. and Jordan, P. M. (1994) Evidence for participation of aspartate-84 as a catalytic group at the active site of porphobilinogen deaminase obtained by site-directed mutagenesis of the *hemC* gene from *Escherichia coli*. Biochemistry **33**, 2688–2695

Received 13 October 2008/10 February 2009; accepted 11 February 2009
Published as BJ Immediate Publication 11 February 2009, doi:10.1042/BJ20082077

Biochem. J. (2009) **420**, 17–25 (Printed in Great Britain) doi:10.1042/BJ20082077

online data

SUPPLEMENTARY ONLINE DATA

Structure of human porphobilinogen deaminase at 2.8 Å: the molecular basis of acute intermittent porphyria

Raj GILL*, Simon E. KOLSTOE*, Fiyaz MOHAMMED†[1], Abeer AL D-BASS†[2], Julie E. MOSELY†[3], Mohammed SARWAR†, Jonathan B. COOPER*, Stephen P. WOOD* and Peter M. SHOOLINGIN-JORDAN†[4]

*Centre for Amyloidosis and Acute Phase Proteins, Department of Medicine, UCL Medical School, Rowland Hill Street, London NW3 2PF, U.K., and †Biomolecular Sciences Group, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, U.K.

Please see the following pages for Figure S1

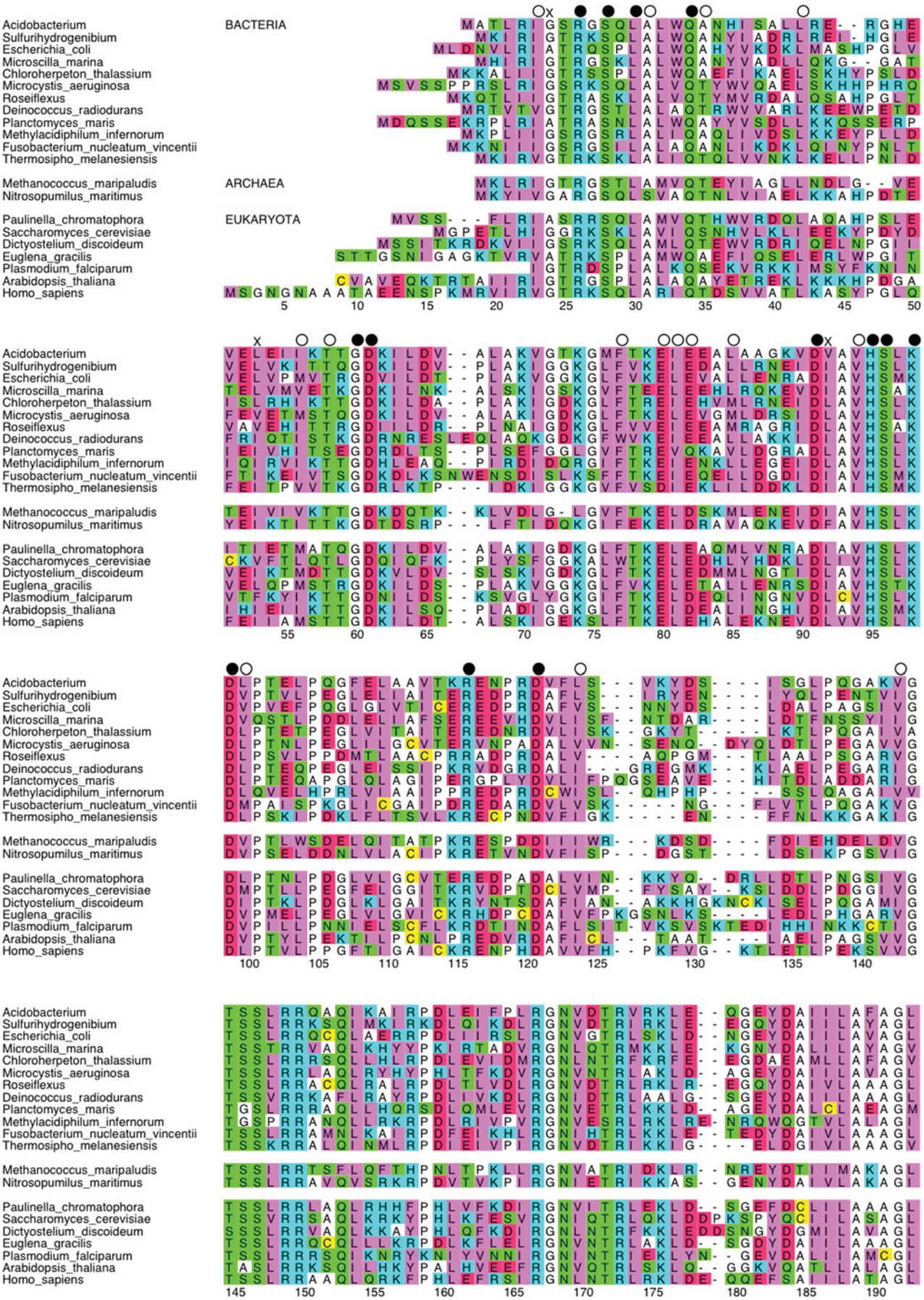
[1] Present address: Cancer Research UK Institute for Cancer Studies, School of Cancer Sciences, University of Birmingham, Vincent Drive, Edgbaston, Birmingham B15 2TT, U.K.

[2] Present address: Department of Biochemistry, College of Sciences, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.

[3] Present address: Neurology and Gastrointestinal Centre of Excellence for Drug Discovery, GlaxoSmithKline, New Frontiers Science Park, Third Avenue, Harlow, Essex CM19 5AW, U.K.

[4] To whom correspondence should be addressed (email pmsj@soton.ac.uk).

The structure of human ubiquitous porphobilinogen deaminase has been deposited in the PDB under accession code 3EQ1.



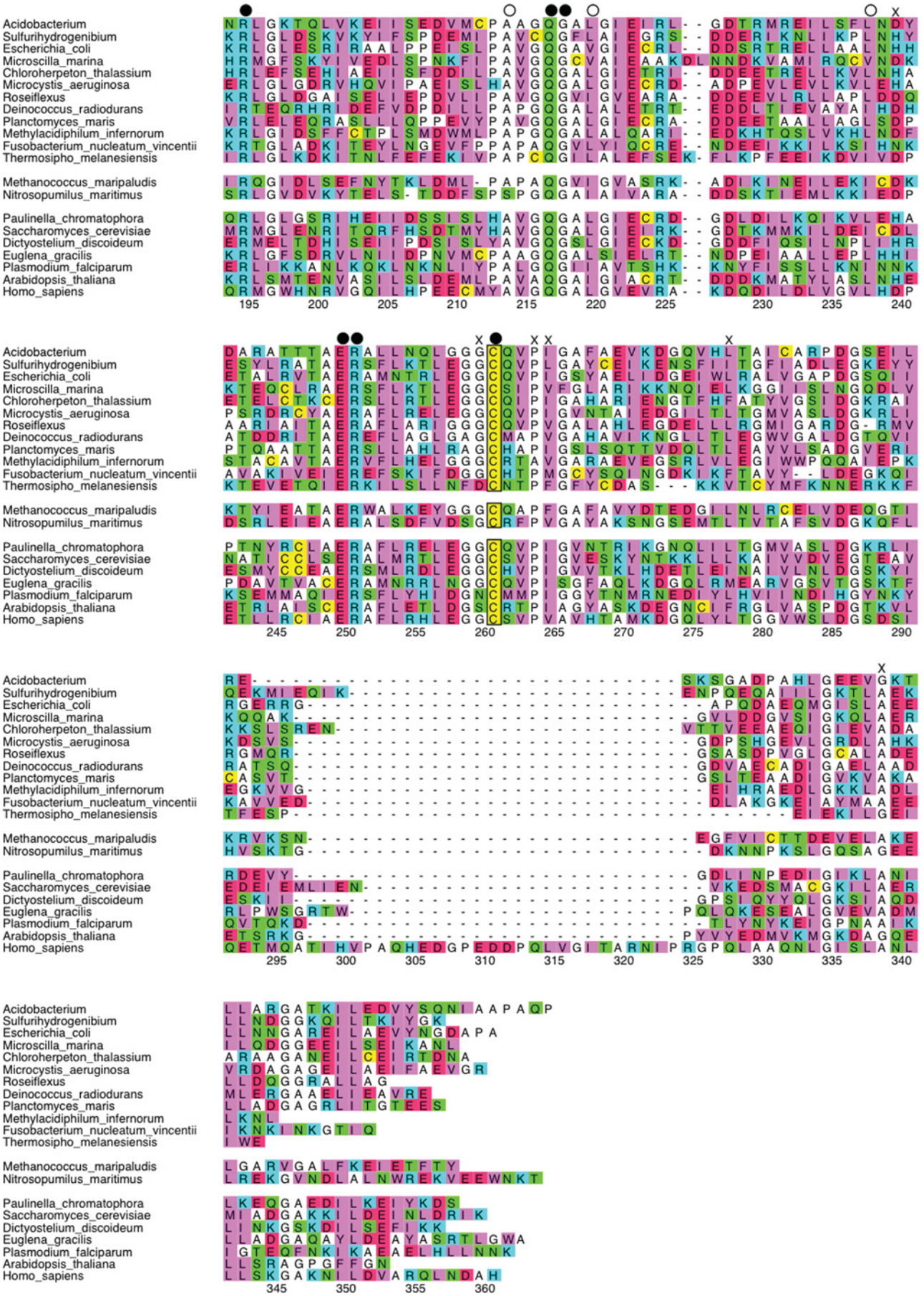


Figure S1 PBGD sequence alignments for representative bacteria, archaea and eukarya species

Representative PBGD sequences from the bacterial, archaeal and eukaryotic taxa deposited in the UniProt Knowledgebase (UniProtKB; http://www.uniprot.org) were aligned with CLUSTAL2W [1] and depicted with ALSCRIPT [2]. Residues conserved across all species are indicated by ●, conserved substitutions are indicated by ○ and semi-conservative substitutions are indicated by ×. Amino acids are colour-coded such that red indicates acidic residues, blue basic, green neutral, pink hydrophobic, yellow cysteine and white indicates the conformationally important glycine, alanine and proline residues. The representative species and sequence codes are as follows. Bacterial: Q1INI6 Acidobacteria/*Acidoacterium strain*, B2V8Q3 Aquificales/*Sulfurihydrogenibium*, P06983 Proteobacteria/*Escherichia coli* K12, A1ZNC8 Bacteroidetes/*Microscilla marina*, B3QWI1 Chlorobiacea/*Chloroherpeton thalassium*, A8YAY9 Cyanobacteria/*Microcystis aeruginosa*, A5UU37 Chloroflexi/*Roseiflexus* sp. RS-1, Q9RRY0 Deinococci/*Deinococcus radiodurans*, A6C1S2 Plantomycetes/*Plantomyces maris* DSM 8797, B3DZQ6 Verrucomicrobia/*Methylacidiphilum infernorum* V4, Q7P7E7 Fusobacteria/*Fusobacterium nucleatum*, A6LKX0 *Thermosipho melanesiensis*/*Thermosipho melanesiensis*. Archaeal: A6VFG4 Euryarchaeota/*Methanococcus maripaludis* (strain C7/ATCC BAA-1331), A9A1A1 Thermo/*Nitrosopumilus maritimus* (strain SCM1). Eukaryotic: B1X5Q3 *Rhizaria/Paulinella chromatophora*, P28789 Ascomyta/*Saccharomyces cerevisiae,* Q54P93 *Mycetozoa*/*Dictyostelium discoideum*, P13446 *Euglenozoa/Euglena gracilis*, A8CBH8 *Apicomplexa/Plasmodium falciparum*, Q43316 *Viridiplantae/Arabidopsis thaliana*, P08397 *Coelomata*/*Homo sapiens.*

REFERENCES

1 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics **23**, 2947–2948

2 Barton, G. J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. **6**, 37–40

Received 13 October 2008/10 February 2009; accepted 11 February 2009
Published as BJ Immediate Publication 11 February 2009, doi:10.1042/BJ20082077