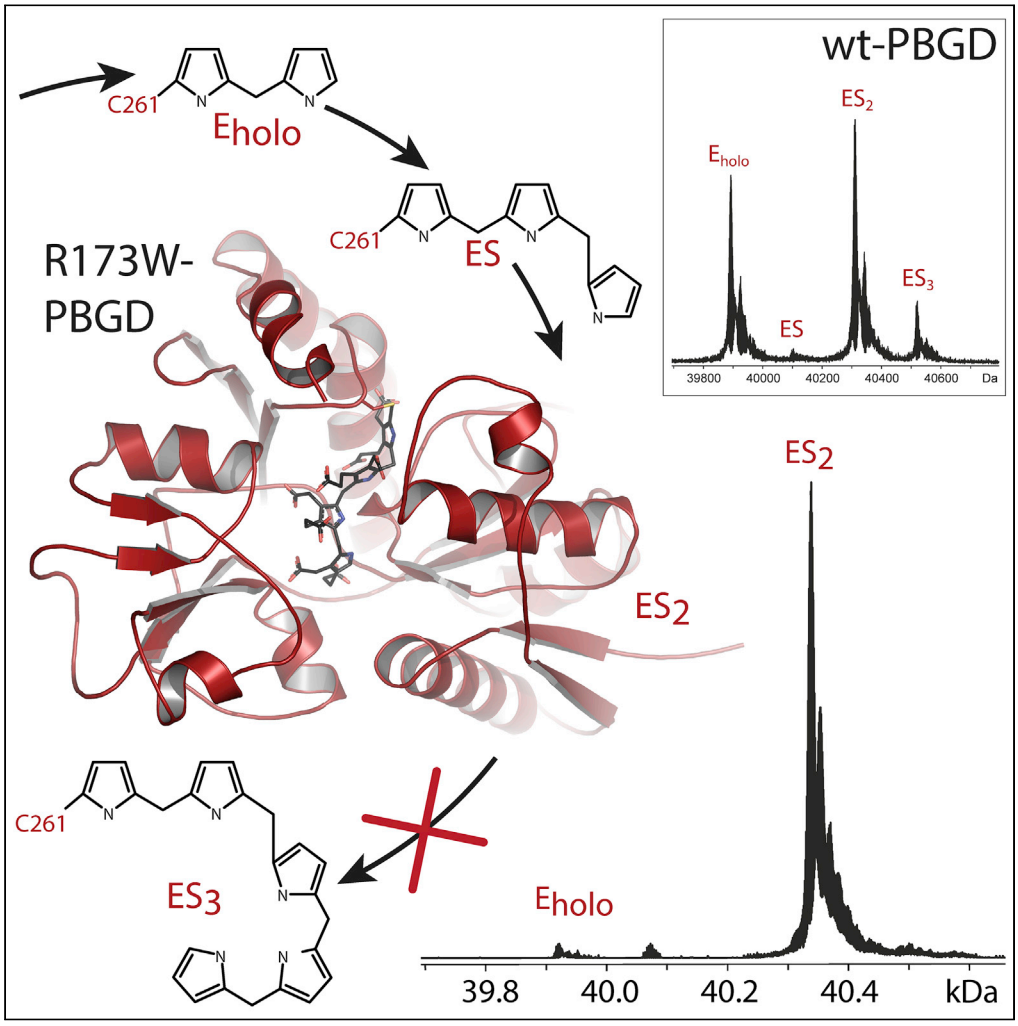
iScience

CellPress
OPEN ACCESS

Article

Characterization of porphobilinogen deaminase mutants reveals that arginine-173 is crucial for polypyrrole elongation mechanism



Helene J. Bustad, Juha P. Kallio, Mikko Laitaoja, Karen Toska, Inari Kursula, Aurora Martinez, Janne Jänis

aurora.martinez@uib.no

HIGHLIGHTS

FT-ICR MS provides molecular information of enzyme-intermediate complexes of PBGD

FT-ICR MS is optimal to elucidate the catalytic defect in AIP-associated PBGD mutants

The structure of R173W-PBGD shows a disease mutant trapped in a reaction intermediate

Arg173 is crucial for the polypyrrole elongation beyond the ES_2 intermediate state

Bustad et al., iScience *24*, 102152
March 19, 2021
https://doi.org/10.1016/j.isci.2021.102152

iScience

CellPress
OPEN ACCESS

Article

Characterization of porphobilinogen deaminase mutants reveals that arginine-173 is crucial for polypyrrole elongation mechanism

Helene J. Bustad,[1,5] Juha P. Kallio,[1,5] Mikko Laitaoja,[2,5] Karen Toska,[3] Inari Kursula,[1,4] Aurora Martinez,[1,6,*] and Janne Jänis[2]

SUMMARY

Porphobilinogen deaminase (PBGD), the third enzyme in the heme biosynthesis, catalyzes the sequential coupling of four porphobilinogen (PBG) molecules into a heme precursor. Mutations in PBGD are associated with acute intermittent porphyria (AIP), a rare metabolic disorder. We used Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) to demonstrate that wild-type PBGD and AIP-associated mutant R167W both existed as holoenzymes (E_{holo}) covalently attached to the dipyrromethane cofactor, and three intermediate complexes, ES, ES_2, and ES_3, where S represents PBG. In contrast, only ES_2 was detected in AIP-associated mutant R173W, indicating that the formation of ES_3 is inhibited. The R173W crystal structure in the ES_2-state revealed major rearrangements of the loops around the active site, compared to wild-type PBGD in the E_{holo}-state. These results contribute to elucidating the structural pathogenesis of two common AIP-associated mutations and reveal the important structural role of Arg173 in the polypyrrole elongation mechanism.

INTRODUCTION

Porphobilinogen deaminase (PBGD; EC 2.5.1.61), also known as hydroxymethylbilane synthase (HMBS), is the third enzyme in the heme biosynthetic pathway. Heme is an important biomolecule that participates in many essential functions in humans, in particular, oxygen transport in blood. PBGD catalyzes four consecutive reactions to convert porphobilinogen (PBG) into hydroxymethylbilane (HMB), a linear tetrapyrrole heme precursor. Mutations in the *HMBS* gene are associated with a genetic metabolic disorder known as acute intermittent porphyria (AIP), giving reduced heme production and severe metabolic and neurological symptoms (Bonkovsky et al., 2019). Missense mutations constitute ~32% of the more than 500 known mutations of human PBGD (hPBGD), and cause destabilization of the enzyme and/or a direct effect on catalysis (Chen et al., 2019; Scott et al., 1988). PBGD is expressed in a tissue-specific manner, with two isoforms produced by different promotor usage and alternative splicing, with the ubiquitously expressed housekeeping PBGD containing 17 extra residues in the N-terminus compared to the erythroid-specific isoform (Grandchamp et al., 1987). However, the catalytic activity of these two isoforms is similar, and the extra N-terminal region has no known function (Brons-Poulsen et al., 2005).

A wealth of structural and functional information is available for both wild-type (wt) and mutant PBGDs, including several three-dimensional structures as well as kinetic, biophysical, and computational data (Awan et al., 1997; Bung et al., 2019; Bustad et al., 2013; Hadener et al., 1999; Louie et al., 1992; Niemann et al., 1994; Pluta et al., 2018; Roberts et al., 2013; Shoolingin-Jordan et al., 1996; Song et al., 2009). The crystal structures of wt-PBGD have so far been solved for human (PDB: 3ECR, 5M7F (Pluta et al., 2018; Song et al., 2009)), *Escherichia coli* (PDB: 1PDA (Louie et al., 1992)), *Arabidopsis thaliana* (PDB: 4HTG (Roberts et al., 2013)), *Bacillus megaterium* (PDB: 4MLV (Azim et al., 2014)) and *Vibrio cholera* (PDB: 5H6O (Uchida et al., 2018)) enzymes, showing the same basic topology with three separate domains (Figure 1A). A deep cleft within the active site connects domains 1 and 2 (residues 1–114 and 120–212, respectively, in hPBGD) by several interactions that stabilize the overall structure. The holoenzyme (E_{holo}) carries a dipyrromethane (DPM) cofactor, which is covalently linked to a conserved cysteine residue (Cys261 in hPBGD) through a thioether bond (Song et al., 2009). Domain 3 (residues 241–361) contains a loop that includes the active site Cys261. DPM derives from two PBG molecules, and functions as an

[1]Department of Biomedicine, University of Bergen, Jonas Lies vei 91, 5009 Bergen, Norway

[2]Department of Chemistry, University of Eastern Finland, 80130 Joensuu, Finland

[3]Norwegian Porphyria Centre (NAPOS), Department for Medical Biochemistry and Pharmacology, Haukeland University Hospital, 5021 Bergen, Norway

[4]Faculty of Biochemistry and Molecular Medicine, University of Oulu, 90570 Oulu, Finland

[5]These authors contributed equally

[6]Lead contact

*Correspondence: aurora.martinez@uib.no

https://doi.org/10.1016/j.isci.2021.102152

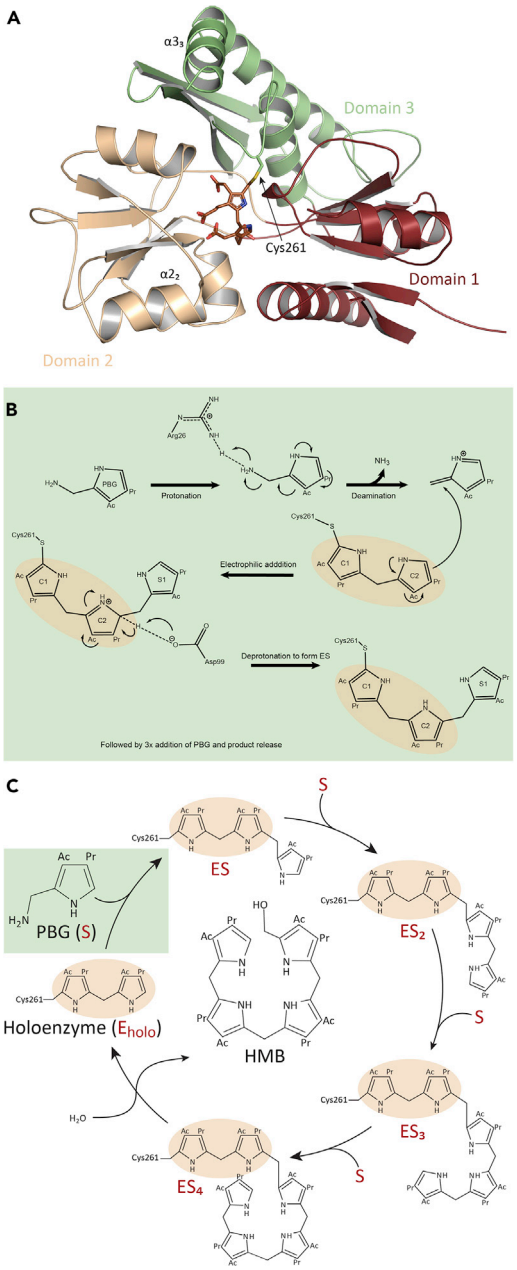


Figure 1. Crystal structure of hPBGD wt-E_{holo} and schematic representation of the polypyrrole elongation mechanism

(A) Overall representation of the crystal structure of wt-hPBGD (PDB: 7AAJ), providing the expected holoenzyme (E_{holo}) with three separate domains (domain 1 is presented in red, domain 2 in light brown, and domain 3 in green). The bound dipyrromethane (DPM) cofactor is shown in a brown stick representation.

(B) Schematic representation of the general understanding of the mechanism of a single step in the polypyrrole elongation.

(C) The polypyrrole elongation catalyzed by PBGD. PBGD with attached DPM (E_{holo}) subsequently binds four porphobilinogen (PBG) substrates (S), and generates the enzyme intermediates ES, ES_2, ES_3, and ES_4. The linear product, hydroxymethylbilane (HMB), is released by hydrolysis and cyclized by the next enzyme in the heme biosynthesis. The sidechains of the substrates, acetate (CH_2CO_2H) and propionate ($CH_2CH_2CO_2H$) are denoted Ac and Pr, respectively. Figure 1C is modified from (Jordan and Woodcock, 1991).

anchor for the sequential coupling of additional four PBG molecules by deamination (Figure 1B) (Layer et al., 2010). When the polypyrrole chain elongation is complete, HMB is released by hydrolysis of the thioether bond, leaving DPM behind, after which the cycle starts again (Figure 1C). As a part of the head-to-tail polymerization process, several covalent enzyme-substrate intermediates are formed, first described in the early 1980s (Anderson and Desnick, 1980; Berry et al., 1981; Jordan and Berry, 1981), which are usually denoted as ES, ES_2, ES_3, and ES_4, where E represents E_{holo}, and S represents the reacted PBG molecule.

The exact mechanism of the polypyrrole formation is still not completely understood. Recently, Pluta et al. published crystal structures of the wt-hPBGD E_{holo} and its ES_2 intermediate (PDB: 5M7F and 5M6R, respectively (Pluta et al., 2018)). These structures allowed the authors to propose a reaction mechanism for PBGD, highlighting the importance of the flexible loop Leu257–Val263 that gives room for the elongation from E_{holo} to ES_2. This is contrary to recent molecular dynamics (MD) simulations complemented with site-directed mutagenesis where very little structural rearrangements upon polypyrrole formation were associated with catalysis and the authors suggested that only a few specific residues, namely Asp99 and Arg26 (Figure 1B), would be responsible for the elongation process (Bung et al., 2014, 2018, 2019).

In this work, we analyzed wt-hPBGD and two recurrent AIP-associated mutants using high-resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS) and X-ray crystallography. The mutants, R167W and R173W, have different mechanistic effects, but are both associated with a severe AIP phenotype (Andersson et al., 2000; Bustad et al., 2013; Mustajoki et al., 2000). R167W represents a catalytically impaired mutant owing to high K_m for PBG and low V_{max} leading to slower polypyrrole elongation (Bustad et al., 2013; Fu et al., 2019; Solis et al., 2004). R173W, on the other hand, is catalytically deleterious with an activity of 0.6% relative to wt and indeterminable K_m and V_{max}, together with an obstructed substrate elongation (Bustad et al., 2013; Fu et al., 2019). Our results indicate that Arg173 is essential for the formation of the ES_3 intermediate, with an important structural role that is relevant for its catalytic contribution. Considering this, we provide unique molecular details on the reaction intermediates for wt-hPBGD and the AIP-associated mutants and shed light on the catalytic and pathogenic mechanisms.

RESULTS

Characterization of enzyme intermediates in the wild-type hPBGD

The recombinantly expressed and purified hPBGD typically consists of a heterogeneous mixture of enzyme-intermediates (Bustad et al., 2013; Shoolingin-Jordan et al., 2003), and isolation of each intermediate is not customarily performed. In the ESI-FT-ICR mass spectra of wt-hPBGD, all the expected enzyme intermediates, i.e., E_{holo}, ES, ES_2, and ES_3, were detected (Figures 2A and 2B). In addition, a small amount of the apoenzyme (E_{apo}) was also observed, which is surprising since E_{apo} is expected to be highly unstable and has not been observed in earlier studies of PBGD kinetic intermediates. However, ESI-FT-ICR MS is a much more sensitive detection technique than most other biophysical methods previously applied. The experimentally determined masses of the intermediates perfectly matched with the theoretical masses considering covalently linked reaction intermediates. The greatest abundance was observed for ES_2, followed by E_{holo}, ES_3, and ES complexes (Figure 2B), corroborating that ES_2 is kinetically the most stable reaction intermediate. As expected, ES_4 was not detected at all, as this intermediate is short-lived and is rapidly hydrolyzed into the linear HMB product (Warren and Jordan, 1988). This observed distribution is consistent with results from earlier studies (Bustad et al., 2013; Niemann et al., 1994; Shoolingin-Jordan et al., 2003), demonstrating the ability of ESI-FT-ICR MS to separate and directly identify different co-existing enzyme-substrate intermediates through intact protein mass analysis. In addition, the mass accuracy was high enough to directly distinguish between the reduced (DPM) and the oxidized (dipyrromethene) form of the cofactor (i.e., 2 Da mass difference in a 40 kDa protein), and our results conclusively indicated that the DPM cofactor existed exclusively in its reduced DPM form (Figure S1). There was no evidence of further cofactor oxidation to a dipyrromethanone (+16 Da) form, which has been observed in the crystal structure of *B. megaterium* PBGD and suggested to be responsible for the pink color of the enzyme in solution (Azim et al., 2014).

We also confirmed that the substrate binding exclusively occurs through the covalent linkage to Cys261. The tryptic digestion of wt-hPBGD resulted in nearly 100% sequence coverage with 58 identified specific

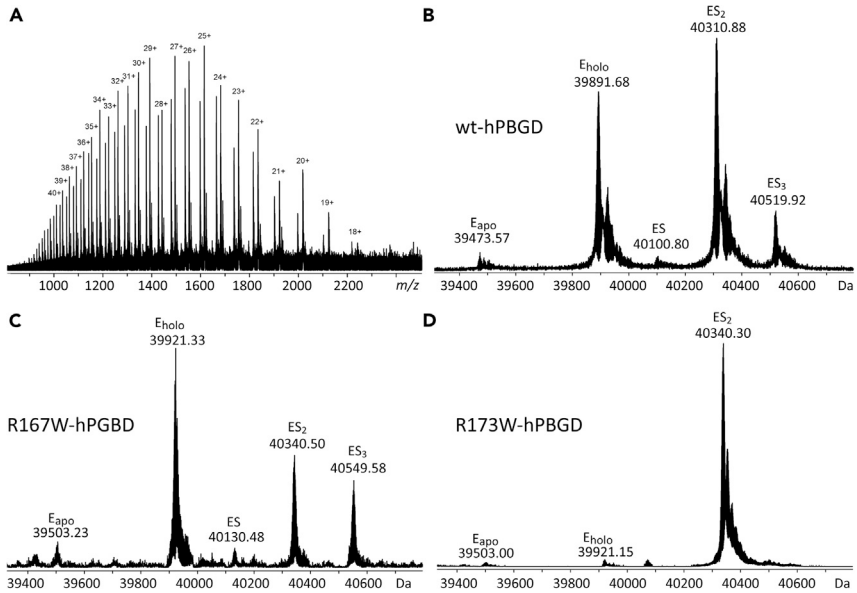


Figure 2. High-resolution mass spectrometry of hPBGD enzymes

The ESI FT-ICR mass spectra were measured at denaturing conditions with 5 μM protein.

(A) Broadband mass spectrum of wt-hPBGD with numbers denoting different protein ion-charge states. A wide charge state distribution from 18 + to 45+ is consistent with the protein being fully unfolded.

(B) Charge-deconvoluted mass spectrum showing peaks representing different enzyme-intermediates in wt-hPBGD.

(C) and (D) Charge-deconvoluted mass spectra of the hPBGD mutants R167W and R173W, respectively. The peaks representing different enzyme-intermediates are assigned.

See also Figures S1–S3, and Tables S1 and S2.

tryptic peptides (Table S1 and Figure S2). The peptide 226–272 was only identified when the tetrapyrrole (corresponding to ES_2) was included as the variable modification in the peptide fingerprinting.

Furthermore, a comparison of the ESI FT-ICR mass spectra of wt-hPBGD in denaturing (Figure 2B) and native conditions (10 mM ammonium acetate pH 6.9; Figure S1) showed similar enzyme-intermediates, thus, implying that all enzyme-substrate complexes are covalent in nature. The only difference between the spectra was that weak signals were detected at *m/z* 3800–4500 at native conditions, possibly representing a very low proportion of a noncovalent protein dimer (Figure S1).

The relative amounts of the reaction intermediates are different for the AIP-associated hPBGD mutants compared to the wild-type enzyme

Two active site hPBGD mutants, R167W and R173W, both showing catalytic dysfunction and a high association with AIP (Bustad et al., 2013), were selected for mass spectrometric analysis. For R167W, the observed enzyme intermediates by ESI FT-ICR MS in denaturing conditions were like wt-hPBGD (Figures 2B and 2C). However, a much higher relative abundance was seen for E_{holo} as compared to the other reaction intermediates, suggesting that the mutation R167W causes a perturbed binding of the first PBG molecule to the E_{holo} and decreasing the rate of HMB synthesis.

In contrast, only a single reaction intermediate was observed for R173W, with a mass corresponding to ES_2 (Figure 2D). This result is consistent with the previous native PAGE analysis of this mutant with a single protein band, and a mild conformational defect as seen by thermal circular dichroism spectroscopy and differential scanning fluorimetry (Bustad et al., 2013). This suggests that productive binding of a third substrate molecule is inhibited when Arg173 is mutated to tryptophan, leading to the accumulation of the ES_2 intermediate without turnover, in agreement with the more severe catalytic dysfunction (<1% residual activity) and a more severe AIP phenotype for the R173W than for the R167W mutant (Fu et al., 2019). Additional trypsin digestion experiments verified that the cofactor or the growing pyrrole chain were bound exclusively to Cys261 also in R173W-hPBGD (Table S2 and Figure S3).

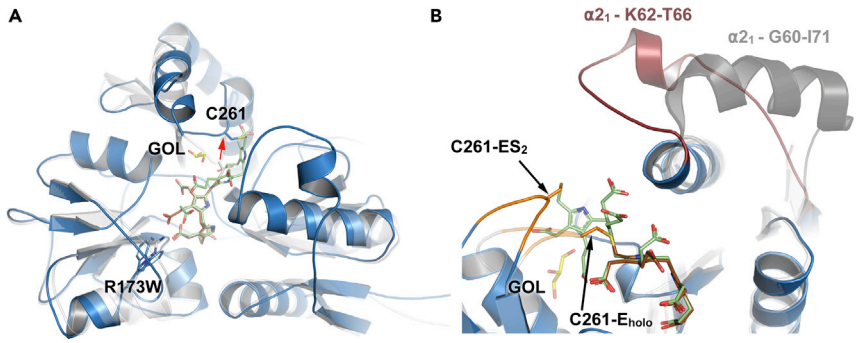


Figure 3. The crystal structure of hPBGD wt-E_{holo} and R173W-ES_2

(A) Overall cartoon representation of wt-E_{holo} (gray; PDB: 7AAJ) and mutant R173W-ES_2 (blue; PDB: 7AAK) superimposed. DPM cofactor with C1 and C2 units of wt-E_{holo} is shown in brown and elongation product of R173W-ES_2 including S1 and S2 units is shown in green. The mutated residue studied here, R173W, is shown as sticks. The cofactor-binding loop and cofactor are rearranged (red arrow) in the structure, allowing incoming substrate pyrroles (S1 and S2) substitute C1 and C2 at equal positions as in the E_{holo}.

(B) The active-site loop (residues 57–74) orientation in wt-ES_2 (PDB: 5M6R; dark gray (Pluta et al., 2018)) is compared to the loop orientation in R173W-ES_2 (red). Formed $\alpha 2_1$ helixes are labeled. Close-up also shows the movement of the cofactor-binding loop (orange; residues 257–263), and the elongation product in the R173W mutant. The position of cofactor-binding Cys261 has been indicated with labels C261-E_{holo} and C261-ES_2 for wt-E_{holo} and R173W-ES_2 structures, respectively. Glycerol (GOL) partially filling the solvent cavity under the cofactor-binding loop in the R173W-ES_2 structure is shown as sticks (yellow).

See also Figures S4 and S5.

High-resolution crystal structures provide important insights into the catalytic mechanism

To obtain further structural insight into the catalytic mechanism we crystallized wt-hPBGD and the AIP-associated mutant R173W. The three-dimensional structures were determined to 1.8 (PBD: 7AAJ) and 1.7 Å resolution (PBD: 7AAK), respectively (Table 1). The overall three-domain structure of both proteins, as well as the active site architecture of wt-hPBGD, are very similar to those of the previously published structure, with an RMSD of 0.256 Å between monomers of our wt-hPBGD and PDB: 3ECR (Figure 1A). Furthermore, both our structures contain two monomers in the asymmetric unit without indication of dimerization as also supported by the mass spectrometry results showing that except for minor dimeric forms, E_{holo} and the other enzyme-intermediates are monomeric. Our crystallization trials for the R167W mutant were unsuccessful. However, a previously published structure of the R167Q mutant (PDB: 3EQ1 (Gill et al., 2009)) also represents the E_{holo}-state only, as the wt-hPBGD.

In agreement with the ESI FT-ICR MS analyses of wt-hPBGD, a fully reduced covalently attached DPM occupies the active site cleft, showing a ~120° angle between the pyrrole rings, instead of the coplanar conformation of oxidized DPM (Azim et al., 2014). The crystal structure only represents the E_{holo}-state and is denoted as wt-E_{holo} hereafter. The binding mode (Figures 3 and 4) and interactions (Figures 5A and 6A) of DPM are the same as in the previously published structure (PDB: 3ECR). Electron density is observed for the two neighboring residues of Cys261, i.e., Gly260 and Gly259, which are also visible in another recent structure of wt-hPBGD (PDB: 5M7F) but not in the earlier structure (PDB: 3ECR). The variability of the electron density quality as well as elevated B-factors indicate a dynamic nature of the residues Leu257–Val263 that constitute the cofactor-binding loop (Figure 3B). Electron density is missing for the first 18 residues at the N-terminus and for the active-site loop residues Ser57–Lys74. Unfortunately, surface exposed residues in flexible sidechains, including Arg167, are not described in the electron density maps. This prevented us to draw further conclusions on the possible catalytic role of this residue.

The structure of the mutant R173W-hPBGD agrees with the ESI FT-ICR MS analysis, representing only ES_2 in its reduced form (denoted R173W-ES_2) and thus revealing for the first time an AIP-associated mutant trapped in a reaction intermediate state. Electron density for the N-terminus was again missing, and the main chain atoms of the active-site loop Ser57–Lys74 could be built only in subunit A (Figure 4A). Within this loop, residues Lys62–Thr66 form a short α-helix (Figure 3B) shown with individual residues in the electron density map in Figure 4A. Crystal packing prevents the loop in the other subunit, B, to adopt the same conformation as in subunit A, however, only traces of the electron density for the loop in subunit B can

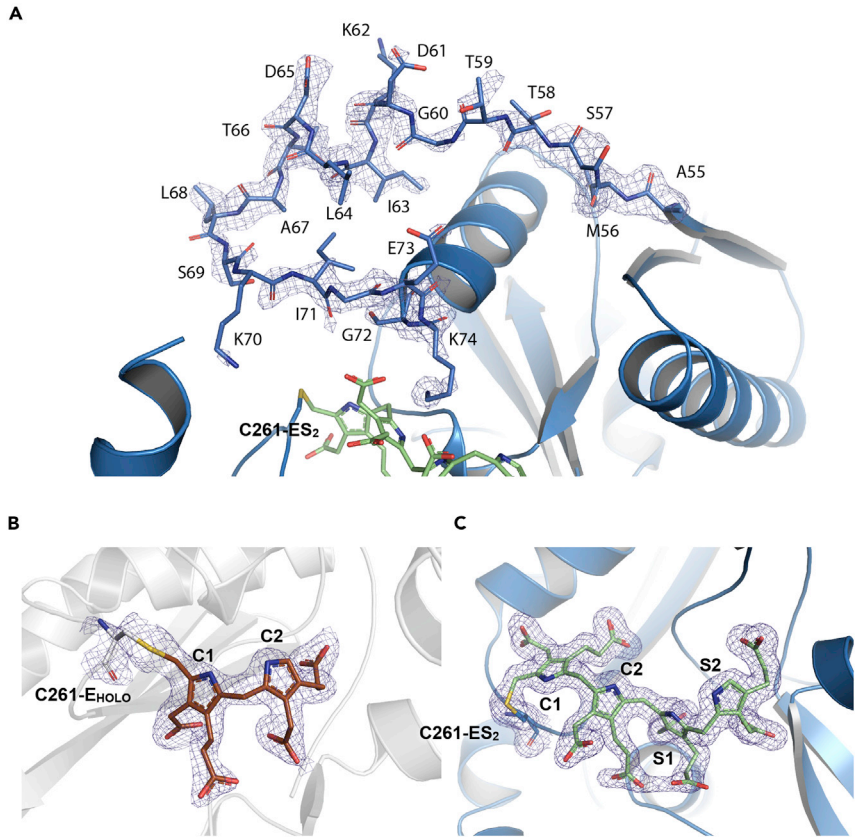


Figure 4. Electron density for the structural features

Calculated 2mFo-DFc-electron density for (A) the active-site loop in R173W-ES_2, (B) bound cofactor in wt-E_{holo} and (C) the polypyrrole chain in R173W-ES_2. Electron density is contoured with sigma level of 1.0.

be seen and was thus not built. Trp173 has different side chain conformations in the two subunits and has two alternative conformations in subunit A (Figure S4). In addition, Ser146 has alternative conformations and seems to move in concert with Trp173 (Figure S4). We also discovered rearrangement in the helix $\alpha 2_2$ (residues 170–179; Figure S4) in subunit B. This rearrangement allows Trp173 to adopt a completely different conformation than seen in subunit A. From this point forward, we will only discuss the structure of subunit A, unless otherwise stated.

The additional pyrrole rings seen in the R173W-ES_2 structure are denoted S1 and S2, in addition to the C1 and C2 rings of the original DPM cofactor (Figures 4, 5, and 6). The incorporation of S1 and S2 in R173W-ES_2 causes a major rearrangement in the cofactor-binding loop Leu257–Val263 including Cys261 to which the reaction intermediate is covalently bound. In contrast to wt-E_{holo}, the active-site loop Ser57–Lys74 in R173W-ES_2 is reoriented toward the C-terminal helix, $\alpha 3_3$, allowing the accommodation of the additional pyrrole rings in the active site cleft between the domains 1 and 2 (Figure 3). The created cavity is not filled by the relocated C1 and C2 rings, and in the structure, a glycerol molecule occupies the space between C1 and the protein core (Figure 3). Because of this rearrangement, the new pyrrole rings S1 and S2 take the original places of the C1 and C2 rings of the DPM cofactor in the wt-E_{holo} structure (Figures 3B and 5).

Structural comparison between R173W-ES_2, wt-E_{holo} and wt-ES_2

All three domains participate in the formation of the interaction network around the cofactor in E_{holo} or the four pyrrole rings in ES_2 in the active site. These interactions are described in detail in Table 2, and Figures 5 and 6. It is noticeable that the interactions change for C1 and C2 when S1 and S2 subsequently are incorporated in the R173W-ES_2 structure. Thus, C1 interacts with Thr102 and Val215 (main chain), and C2 interacts with Lys98 (main chain) and Ser262 (Figures 5 and 6). S1, which in R173W-ES_2 occupies the original position of C1 in E_{holo}, interacts with Ser147 and Arg149 through the acetate side chain. The propionate side

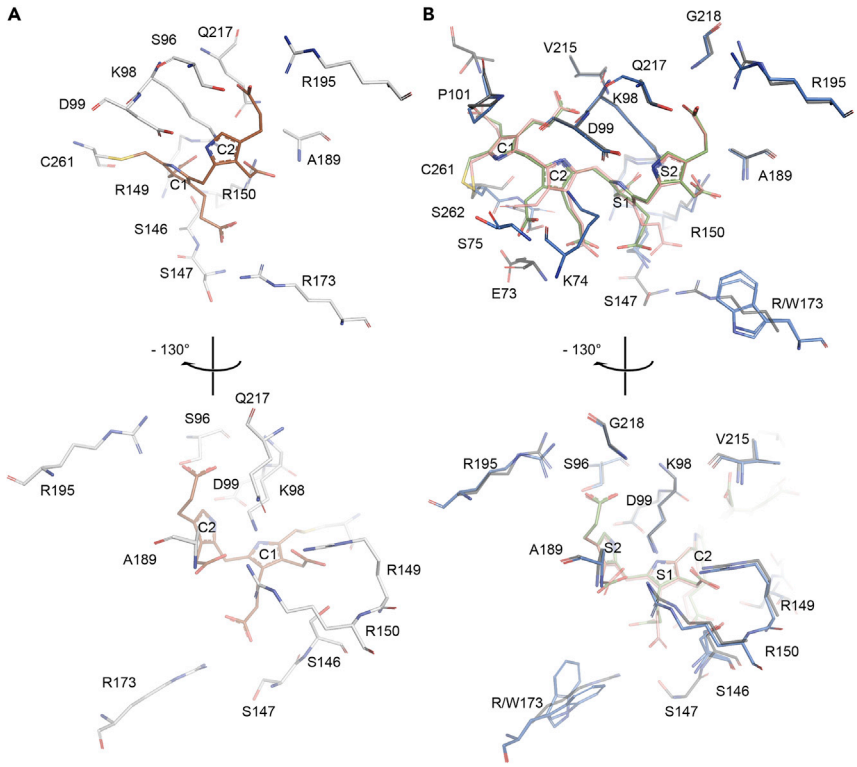


Figure 5. Stick representation of the active site

The image shows active site interactions of (A) our wt-E_{holo} (side chains gray and substrate brown), and (B) R173W-ES_2 (side chains blue and substrate green) superimposed with wt-ES_2 (PDB: 5M6R; side chains dark gray and substrate pink (Pluta et al., 2018)). Incoming PBG units in the ES_2 intermediates substitute C1 and C2 at equal positions as in the E_{holo}. Interactions are described in detail in Table 2.

chain of S1 does not make hydrogen bonds to any atoms in the R173W-ES_2 structure. S2 occupies the equivalent position of C2 in E_{holo}, where the acetate side chain interacts with Arg98, Arg150 and Ala189 and the propionate side chain creates hydrogen bonds with Ser96, Arg195, and Gly218. Lys98 forms salt bridges with both acetate side chains of rings S1 and S2. Interestingly Asp99 interacts with both pyrrole N atoms of C1 and C2 in wt-E_{holo}, whereas in the R173W-ES_2 there is only an interaction to pyrrole N in S1 (Table 2 and Figures 5 and 6).

The active-site loop Ser57–Lys74 in R173W-ES_2 has also been observed in the E_{holo} structure of PBGD from *A. thaliana* (AtPBGD; PDB: 4HTG) and in human wt-hPBGD recently crystallized in the ES_2-state (wt-ES_2; PDB: 5M6R). The loop appears to close the active site like a lid, and is mostly unstructured in AtPBGD, whereas it includes a more defined α-helix in both R173W-ES_2 (Figure 3B) and wt-ES_2 (PDB: 5M6R). In wt-ES_2 (PDB: 5M6R), the α-helix is three residues longer and the loop partially covers the active site, without direct interaction with the cofactor or substrates (Pluta et al., 2018). In R173W-ES_2 presented here, the active-site loop is oriented in a significantly different conformation, partially facing away from the surface with Lys74 establishing several interactions with C2 and S1 (Figures 5 and 6). In wt-ES_2 (PDB: 5M6R), the pyrrole rings present a similar conformation as in R173W-ES_2. However, in the mutant structure, we observe a different conformation of the propionate side chain of C1 than in wt-ES_2 (PDB: 5M6R) (Figure 5), whereas this propionate forms an electrostatic and hydrogen-bonding interaction network with Arg173 and Ser147, the mutation R173W abolishes the interaction and in the crystal structure the propionate is bent and is no longer in contact with the protein (Table 2 and Figure 5).

DISCUSSION

Despite more than 500 AIP-associated *HMBS* variants discovered to date, little is known about their structural effects. Using high-resolution ESI FT-ICR MS and X-ray crystallography, we obtained crucial structural information on two common AIP-associated mutations, R167W and R173W. Our results are an important

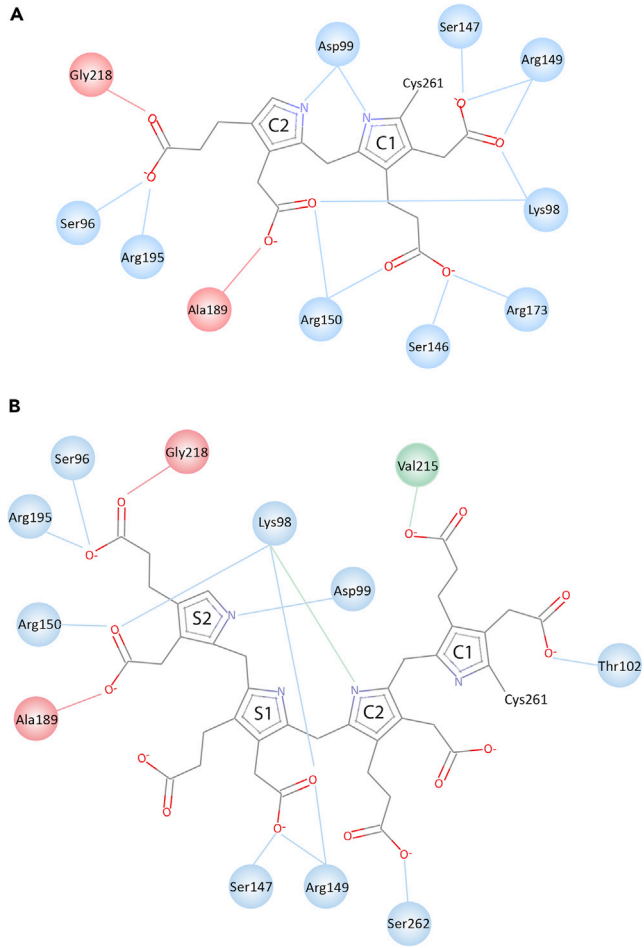


Figure 6. Schematic view of the interactions between the pyrrole rings and the hPBGD protein in the crystal structures of wt-E_{holo} and R173W-ES_2 mutant

(A) The hydrogen bond interactions of the DPM cofactor with the wt-hPBGD residues in the active site.

(B) The interactions between the pyrrole chain intermediate as seen in the crystal structure of the R173W-hPBGD mutant. Side chain H-bond interactions (blue), H-bond interactions to main chain carbonyl oxygen (green) and H-bond interactions to nitrogen (red).

step toward understanding the pathogenic molecular mechanisms leading to AIP, as well as the pyrrole-chain elongation-mechanism.

ESI FT-ICR MS allowed us to determine the distribution of the enzyme-intermediates in wt-hPBGD in a quantitative manner, which has not been possible by other methods due to their limited resolution. The ES_2 intermediate was the most abundant, whereas ES was only found in a very small amount and ES_4 was not detected at all. ES is kinetically less stable than ES_2 or ES_3, and ES_2 accumulates during the reaction, in agreement with a slow rate of the $ES_2 \rightarrow ES_3$ step (Niemann et al., 1994; Warren and Jordan, 1988). The presence of the apoenzyme in wt-hPBGD is remarkable and has not been observed before in any enzyme preparation is from prokaryote expression, as it is assumed to be unstable and less structurally compact than E_{holo} (Awan et al., 1997; Scott et al., 1989).

Several arginine residues are conserved in PBGD across species. They are involved in or even crucial for either catalysis (e.g., Arg26), structural stability (e.g., Arg251) or implicated in the cofactor binding (Arg150) (Jordan and Woodcock, 1991; Lander et al., 1991). Arg167 is one of the highly conserved residues and has been proposed to act as a gatekeeper for incoming substrates and to be important in breaking of the salt bridges in PBG prior to catalysis (Brownlie et al., 1994; Bung et al., 2018; Gill et al., 2009; Shoolingin-Jordan et al., 2003). Missense mutations of Arg167 affect both the affinity for PBG and the catalytic efficiency of the enzyme rather than instability or misfolding (Bustad et al., 2013; Gill et al., 2009). The proposed effect of Arg167 mutations has recently been

Table 1. Data collection and refinement statistics

	wt-E_{holo}	R173W-ES_2
Data collection		
PDB ID	7AAJ	7AAK
Resolution range	65.08–1.8 (1.864–1.8)	59.1–1.7 (1.761–1.7)
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Unit cell	81.2 84.6 108.9 90 90 90	81.2 86.1 107.4 90 90 90
Total reflections	528551 (52,834)	551220 (55,773)
Unique reflections	69,757 (6877)	83,105 (8199)
Multiplicity	7.6 (7.7)	6.6 (6.8)
Completeness (%)	98.9 (97.3)	99.47 (99.22)
Mean I/sigma(I)	10.24 (0.52)	15.48 (1.62)
Wilson B-factor	34.2	25.41
R-merge	0.095 (3.76)	0.068(1.56)
R-meas	0.102 (4.02)	0.074 (1.69)
CC1/2	0.99 (0.54)	1 (0.73)
Refinement		
Reflections used in refinement	69,336 (6735)	82,994 (8182)
Reflections used for R-free	3510 (280)	4022 (398)
R_{work}	24.9 (62.2)	18.3(34.8)
R_{free}	29.4 (65.3)	21.2 (35.8)
Number of non-hydrogen atoms	5239	5918
Macromolecules	5000	5267
Ligands	78	144
Solvent	161	507
Protein residues	647	669
RMSD (bonds)	0.007	0.007
RMSD (angles)	0.8	0.8
Validation		
Ramachandran favored (%)	96.8	98.2
Ramachandran allowed (%)	2.2	1.8
Ramachandran outliers (%)	0	0
Clashscore	4.7	2.8
Average B-factor	56.8	36.8
Macromolecules	56.9	36.4
Ligands	59.8	32.4
Solvent	51.2	42.2

attributed to the alteration of the binding site leading to both decreased pyrrole chain elongation and blocking of the HMB release (Bung et al., 2018). Our results clearly demonstrate that the elongation process is indeed perturbed in the R167W mutant. This is consistent with the ~30-fold higher K_m and reduced V_{max} relative to the wt enzyme (Bustad et al., 2013), which results in slower elongation and accumulation of the reaction intermediates (Bung et al., 2018; Gill et al., 2009; Jordan and Woodcock, 1991; Shoolingin-Jordan et al., 2003). Our results thus do not support the participation of Arg167 solely in product release, since the distribution of enzyme intermediates was rather similar to the wt enzyme, and ES_4 was not detected at all.

Arg173 is also highly conserved and is considered important for substrate docking to the active site (Louie et al., 1996). The substitution of Arg173 with tryptophan introduces a large hydrophobic amino acid,

Table 2. Interactions between protein and cofactor and/or substrate

Ring	Wt-E_{holo}	Interaction	Ring	R173W-ES_2	Wt-ES_2 (56MR)	Interaction	Mutations	
			C1	P101	P101	Pyrrole: π-stack	–	
				T102	T102	Ac: H-bond	–	
				V215	V215	Pr: H-bond (main chain)	V215E/M	(Bustad et al., 2013; Schneider-Yin et al., 2008)
			C2	$S75^A$	–	Ac: vdW		
				$K74^A$	–	Ac: H-bond	–	
				–	$F77^B$	Pyrrole: π-stack	–	
				K98	K98	N: H-bond (main chain)	K→R	(Kauppinen et al., 1995)
				D99	$D99^A$	Pyrrole: vdW	D→G/H/N	(Floderus et al., 2002; Kauppinen and von und zu Fraunberg, 2002)
				S262	S262	Pr: H-bond	–	
					$S262^B$	Ac: H-bond	–	
C1	K98	Ac: H-bond	S1	K98	K98	Ac: H-bond	K→R	(Kauppinen et al., 1995)
	D99	N: H-bond		D99	D99	N: H-bond	D→G/H/N	(Floderus et al., 2002; Kauppinen and von und zu Fraunberg, 2002)
	S146	Pr: vdW		–	S146	Pr: vdW	–	
	S147 S147	Ac: H-bond Pr: H-bond (main chain)		S147	S147	Ac: H-bond	S→P	(Whatley et al., 2009)
	R149	Ac: Salt bridge		R149	R149	Ac: Salt bridge	R→L/P/Q	(Delfau et al., 1991; Gu et al., 1994; Kauppinen et al., 1995; Yang et al., 2008)
	R150	Pr: Salt bridge		–	–			
	R173	Pr: Salt bridge		–	R173	Pr: Salt bridge	R→G/Q/W	(Delfau et al., 1990; Kauppinen et al., 1992; Mendez et al., 2009)
C2	S96	Pr: H-bond	S2	S96	S96	Pr: H bond	S→F	(Kauppinen and von und zu Fraunberg, 2002)
	K98	Ac: H-bond		K98	K98	Ac: H-bond	K→R	(Kauppinen et al., 1995)
	D99	N: H-bond		–	$D99^A$	N: H-bond	D→G/H/N	(Floderus et al., 2002; Kauppinen and von und zu Fraunberg, 2002)
	R150	Pr: Salt bridge		R150	R150	Ac: Salt bridge	–	
	A189	Ac: H-bond (main chain)		A189	A189	Ac: H bond (main chain)	–	
	R195	Pr: Salt bridge		R195	R195	Pr: salt bridge	R→C/H	(Kauppinen et al., 1995; Whatley et al., 2009)
	Q217	Pr: vdW		–	$Q217^B$	Pr: vdW	Q→H/L/R	(Kuo et al., 2011; Puy et al., 1997; Schneider-Yin et al., 2000, 2006)
				G218	G218	Pr: H bond (main chain)	G→R	(Yang et al., 2008)

$^{A/B}$Refers to the subunits of the asymmetric unit in the crystal structure.

predicted to hinder the cofactor and/or substrate interaction severely (Jordan and Woodcock, 1991). Both ESI FT-ICR MS and X-ray crystallography show that this mutant accumulates the ES_2 intermediate with a greatly reduced turnover. Although Arg173 forms hydrogen bonds with the propionate side chains of C1

in the E_{holo} and of S1 in the ES_2 state (Table 2 and Figure 6) (Pluta et al., 2018; Song et al., 2009), the tryptophan substitution does not hinder the binding of the cofactor and the reaction of the first two PBG substrates. Hence, Arg173 is important for docking the third PBG substrate to the growing pyrrole chain.

Wt-hPBGD is very thermostable with a T_m of ~74°C (Bustad et al., 2013, 2020). As seen from our wt-E_{holo} crystal structure and the reported wt-ES_2 (PDB: 5M6R), this can be attributed to the strong hydrogen-bonding network, in which the cofactor in E_{holo} or the pyrrole chain in ES_2 engage Arg173 and interacting residues in domain 2, including the cofactor-binding loop Leu257–Val263 in domain 1, with Cys261. In wt-hPBGD, Arg173 might interact directly with S2 and the entering S3 through salt bridges with the acetate and propionate side chains, substituting the interaction of Arg150 with S2 upon ES_3 formation. The Arg-to-Trp mutation results in a disruption of these interactions, partially explaining the large reduction of thermal stability (Bustad et al., 2013). The conformational change seen in subunit B of the R173W-ES_2 structure, where the bulky tryptophan residue turns away from the active site forcing a change in the domain structure, probably also affects the loss in thermal stability (Figure S4). Together with its important catalytic function, the role of Arg173 as a stabilizer in the active-site hydrogen-bonding network correlates with the loss of the stability and activity upon Arg-to-Trp mutation and thus a severe AIP outcome.

A striking structural feature of R173W-ES_2 compared to wt-E_{holo} is that the cofactor-binding loop rearranges to make more space for the two incoming PBG substrates. Furthermore, S1 and S2 in ES_2 occupy *exactly* the same positions as C1 and C2 of the cofactor in E_{holo} with nearly identical interactions (Table 2 and Figure 6). Despite the alteration of the interactions involving Arg173, the loop conformation and the pyrrole ring locations in our R173W-ES_2 structure correspond to the structure of wt-ES_2 (5M6R; Figure 4B), strongly indicating a correct conformation of the intermediate in the mutant. Thus, Arg173 participates in the interaction network involved in the cofactor (C1) and the substrate (S1) binding in the E_{holo} and ES_2 states, respectively (Table 2 and Figures 5 and 6) (Pluta et al., 2018), but is not important to define the proper conformation of the pyrrole chain. These data corroborate the findings from ESI FT-ICR MS, indicating that Arg173 has a crucial function in the third elongation step from ES_2 to ES_3. The catalytic defect is a major consequence of the R173W mutation, despite having mostly been associated with a conformationally unstable protein and concomitantly reduced activity (Bustad et al., 2013; Mustajoki and Desnick, 1985). Its role has, however, not been clearly elucidated in the previous investigations (Bung et al., 2018; Pluta et al., 2018), but its relevance in the interaction network in AIP pathology has been discussed recently (Fu et al., 2019). Nevertheless, our results imply that Arg173 is essential for orienting the intermediate to a specific conformation, in either correctly positioning the pyrrole chain or docking an incoming substrate, allowing the elongation from ES_2 to ES_3.

Recent *in silico* investigation of the interaction network during different intermediate states indicates dynamic movement of the active-site loop, as well as specific interactions between the loop and cofactor during the elongation (Chakrabarty et al., 2020). However, the authors do not discuss the rearrangement of the cofactor-binding loop and how the interaction network is affected by this. The active-site loop (Ser57–Lys74) in R173W-ES_2 adopts a different conformation than in the other crystal structures with electron density describing these residues, i.e., AtPBGD (PDB: 4HTG) or wt-ES_2 (PDB: 5M6R). In contrast to wt-ES_2 (PDB: 5M6R), only a short helical turn (Lys62–Thr66) is present in R173W-ES_2, resulting in a more open conformation including only one clear interaction between the loop and the substrate pyrroles, and Lys74 is hydrogen bonded to the acetate side chain of the C2 pyrrole ring.

Based on the crystal structure of the wt-ES_2 (PDB: 5M6R), Pluta et al. proposed a mechanism for the reaction progression relying on the further movement on the cofactor-binding loop in the formation of ES_3 and beyond (Pluta et al., 2018). However, this movement might cause steric issues and major rearrangement of the $\alpha 3_3$ helix as well as a disturbance of the network of hydrophobic interactions around this helix (Figure S5). Pluta et al. also proposed that Arg26 and Asp99 are the only residues responsible for the pyrrole ring condensations as well as for the release of the product, consistent with the recent computational work by Bung et al. (Bung et al., 2014). The effect caused by the R173W mutantion fits with this mechanism; however, we demonstrate by using this AIP associate mutant rather than the wt enzyme that the substrate elongation from ES_2 to ES_3 is crucially dependent on Arg173. Although the Arg-to-Trp substitution does not seem to largely affect the structure of hPBGD in the ES_2-state, it affects the enzyme stability and the polypyrrole elongation beyond ES_2, most probably due to the disruption of the Arg173-centered interactions in domains 1 and 2 (Jordan and Woodcock, 1991; Lander et al., 1991).

Understanding the details for the exact elongation mechanism remains unsolved. MD simulations propose a mechanism relying on the protonation of incoming PBGs by the Arg26, and electrophilic addition and deprotonation in concert with Asp99 (Bung et al., 2018, 2019). However, these studies have not considered the movement of the cofactor-binding loop upon the elongation; instead, they are relying on direct helicoidal elongation with only the active-site loop moving. This causes a steric problem for Arg26 and Asp99 where these residues no longer are correctly positioned for the full catalysis cycle. The available crystal structures do not provide detailed information on the dynamic movement of the active-site loop beyond the ES_2 state, but the wt-ES_2 (PDB: 5M6R) and R173W-ES_2 structures clearly show the movement of the cofactor-binding loop.

In conclusion, we show for the first time the direct effect of a mutation associated with AIP on the structure and function of PBGD, revealing the importance of the interaction network around Arg173. Using X-ray crystallography, we trapped the disease-causing mutant R173W in a reaction intermediate and combined with ESI-FT-ICR MS we pinpoint the crucial responsibility of Arg173 in the catalytic mechanism in stabilizing the structure and ensuring proper interaction with the entering substrate to form ES_3. Furthermore, our work highlights the strength of ESI-FT-ICR MS as a high-resolution technique that quantifies co-existing ES_n intermediates, representing an effective procedure to elucidate the catalytic effect of the AIP mutations. Hence, the distribution of the intermediates for the R167W and R173W mutants agrees with the severity of the respective associated AIP phenotype and provides important insights into the catalytic mechanism of PBGD. We propose that as high-resolution MS allows the direct analysis of intermediate distribution in PBGD, it could also be effective for drug screening of e.g., pharmacological chaperones (Bustad et al., 2020), aiming at the stabilization of PBGD and unstable AIP-associated mutants and/or the specific correction of altered distributions of enzyme intermediates.

Limitations of the study

This study characterized only two AIP-associated mutants. Investigating additional mutants will increase our knowledge on genotype-phenotype relationships and will aid to elucidate the polypyrrole elongation mechanism. The crystallization of the wt-PBGD was performed from a mixture of intermediates, and the heterogeneity of the sample, as well as the flexible parts of the enzyme, such as the N-terminus and the loop including residues 57–74, most likely affect the quality of the data. In the future, this could possibly be overcome by crystallizing isolated intermediates. In addition, a crystal structure of PBGD in the ES_3-state and/or ES_4-state is essential for understanding the complete mechanism.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Aurora Martinez (aurora.martinez@uib.no).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The accession numbers for the protein crystal structure reported in this paper are PDB: 7AAJ, 7AAK. Original data have been deposited to RCSB Protein Data Bank at (https://www.rcsb.org).

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2021.102152.

ACKNOWLEDGMENTS

We thank Dr. Marta Vorland, Prof. Sverre Sandberg, and Prof. Aasne K. Aarsand for lab facilities, discussions, and collaboration. This work was supported by grants from the Research Council of Norway (Program Biotek2021, project 285295 to JK and AM, and Program Toppforsk, project to HJB and AM), the Western Norway Regional Health Authority (to JK, IK, and AM (project 912246 to AM)), the Norwegian Porphyria Center (to KT) and European Union's Horizon 2020 Research and Innovation Program (grant agreement

no. 731077). The FT-ICR MS facility is supported by Biocenter Kuopio, Biocenter Finland (FINStruct) and European Regional Development Fund (grant A70135). For accessing to synchrotron facilities, we would also like to acknowledge beamlines P13 and P14 operated by EMBL Hamburg at the PETRA III storage ring and to thank Johanna Hakanpää for the assistance in using the beamline. The use of the infrastructure at the core facility BiSS at the University of Bergen is gratefully acknowledged. Prof. Juha Rouvinen (University of Eastern Finland) is thanked for critical reading of the manuscript and valuable comments.

AUTHOR CONTRIBUTIONS

Conceptualization: A.M., H.J.B., and J.J.; funding acquisition: A.M. and J.J.; investigation: H.J.B., J.P.K., and M.L.; project administration: A.M. and J.J.; resources: A.M., H.J.B, K.T., J.J, J.P.K. and I.K.; supervision: A.M., J.J., and I.K.; validation: A.M., H.J.B., J.K.P., and J.J.; visualization: H.J.B, J.P.K., and J.J.; writing – original draft: H.J.B.; writing – review & editing: H.J.B., J.P.K., A.M., J.J., K.T., and I.K.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: September 11, 2020
Revised: December 3, 2020
Accepted: February 2, 2021
Published: March 19, 2021

REFERENCES

Anderson, P.M., and Desnick, R.J. (1980). Purification and properties of uroporphyrinogen I synthase from human erythrocytes. Identification of stable enzyme-substrate intermediates. J. Biol. Chem. *255*, 1993–1999.

Andersson, C., Floderus, Y., Wikberg, 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.

Awan, S.J., Siligardi, G., Shoolingin-Jordan, P.M., and Warren, M.J. (1997). Reconstitution of the holoenzyme form of Escherichia coli porphobilinogen deaminase from apoenzyme with porphobilinogen and preuroporphyrinogen: a study using circular dichroism spectroscopy. Biochemistry *36*, 9273–9282.

Azim, N., Deery, E., Warren, M.J., Wolfenden, B.A., Erskine, P., Cooper, J.B., Coker, A., Wood, S.P., and Akhtar, M. (2014). Structural evidence for the partially oxidized dipyrromethene and dipyrromethanone forms of the cofactor of porphobilinogen deaminase: structures of the Bacillus megaterium enzyme at near-atomic resolution. Acta Crystallographica. Section D Biol. Crystallogr. *70*, 744–751.

Berry, A., Jordan, P.M., and Seehra, J.S. (1981). The isolation and characterization of catalytically competent porphobilinogen deaminase-intermediate complexes. FEBS Lett. *129*, 220–224.

Bonkovsky, H.L., Dixon, N., and Rudnick, S. (2019). Pathogenesis and clinical features of the acute hepatic porphyrias (AHPs). Mol. Genet. Metab. *128*, 213–218.

Brons-Poulsen, J., Christiansen, L., Petersen, N.E., Horder, M., and Kristiansen, K. (2005). Characterization of two isoalleles and three mutations in both isoforms of purified recombinant human porphobilinogen deaminase. Scand. J. Clin. Lab Invest. *65*, 93–105.

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.

Bung, N., Pradhan, M., Srinivasan, H., and Bulusu, G. (2014). Structural insights into E. coli porphobilinogen deaminase during synthesis and exit of 1-hydroxymethylbilane. PLoS Comput. Biol. *10*, e1003484.

Bung, N., Roy, A., Chen, B., Das, D., Pradhan, M., Yasuda, M., New, M.I., Desnick, R.J., and Bulusu, G. (2018). Human hydroxymethylbilane synthase: molecular dynamics of the pyrrole chain elongation identifies step-specific residues that cause AIP. Proc. Natl. Acad. Sci. U S A *115*, E4071–E4080.

Bung, N., Roy, A., Priyakumar, U.D., and Bulusu, G. (2019). Computational modeling of the catalytic mechanism of hydroxymethylbilane synthase. Phys. Chem. Chem. Phys. *21*, 7932–7940.

Bustad, H.J., Toska, K., Schmitt, C., Vorland, M., Skjaerven, L., Kallio, J.P., Simonin, S., Letteron, P., Underhaug, J., Sandberg, S., et al. (2020). A pharmacological chaperone therapy for acute intermittent porphyria. Mol. Ther. *28*, 677–689.

Bustad, H.J., Vorland, M., Ronneseth, E., Sandberg, S., Martinez, A., and Toska, K. (2013). Conformational stability and activity analysis of two hydroxymethylbilane synthase mutants, K132N and V215E, with different phenotypic association with acute intermittent porphyria. Biosci. Rep. *33*, e00056.

Chakrabarty, B., Das, D., Bung, N., Roy, A., and Bulusu, G. (2020). Network analysis of hydroxymethylbilane synthase dynamics. J. Mol. Graphics Model. *99*, 107641.

Chen, B., Solis-Villa, C., Erwin, A.L., Balwani, M., Nazarenko, I., Phillips, J.D., Desnick, R.J., and Yasuda, M. (2019). Identification and characterization of 40 novel hydroxymethylbilane synthase mutations that cause acute intermittent porphyria. J. Inherit. Metab. Dis. *42*, 186–194.

Delfau, M.H., Picat, C., De Rooij, F., Voortman, G., Deybach, J.C., Nordmann, Y., and Grandchamp, B. (1991). Molecular heterogeneity of acute intermittent porphyria: identification of four additional mutations resulting in the CRIM-negative subtype of the disease. Am. J. Hum. Genet. *49*, 421–428.

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.

Floderus, Y., Shoolingin-Jordan, P.M., and Harper, P. (2002). Acute intermittent porphyria in Sweden. Molecular, functional and clinical consequences of some new mutations found in the porphobilinogen deaminase gene. Clin. Genet. *62*, 288–297.

Fu, Y., Jia, J., Yue, L., Yang, R., Guo, Y., Ni, X., and Shi, T. (2019). Systematically analyzing the pathogenic variations for acute intermittent porphyria. Front. Pharmacol. *10*, 1018.

Gill, R., Kolstoe, S.E., Mohammed, F., Al, D.B.A., Mosely, J.E., Sarwar, M., Cooper, J.B., Wood, S.P., and Shoolingin-Jordan, P.M. (2009). Structure of human porphobilinogen deaminase at 2.8 A: the molecular basis of acute intermittent porphyria. Biochem. J. *420*, 17–25.

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./FEBS *162*, 105–110.

Gu, X.F., de Rooij, F., Voortman, G., Te Velde, K., Deybach, J.C., Nordmann, Y., and Grandchamp, B. (1994). Detection of eleven mutations causing acute intermittent porphyria using denaturing gradient gel electrophoresis. Hum. Genet. *93*, 47–52.

Hadener, 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 seleno-methionine-labelled hydroxymethylbilane synthase in its active form by multi-wavelength anomalous dispersion. Acta Crystallographica. Section D Biol. Crystallogr. *55*, 631–643.

Jordan, P.M., and Berry, A. (1981). Mechanism of action of porphobilinogen deaminase. The participation of stable enzyme substrate covalent intermediates between porphobilinogen and the porphobilinogen deaminase from Rhodopseudomonas spheroides. Biochem. J. *195*, 177–181.

Jordan, P.M., 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* (Pt 2), 445–449.

Kauppinen, R., Mustajoki, S., Pihlaja, H., Peltonen, L., and Mustajoki, P. (1995). Acute intermittent porphyria in Finland: 19 mutations in the porphobilinogen deaminase gene. Hum. Mol. Genet. *4*, 215–222.

Kauppinen, R., Peltonen, L., Pihlaja, H., and Mustajoki, P. (1992). CRIM-positive mutations of acute intermittent porphyria in Finland. Hum. Mutat. *1*, 392–396.

Kauppinen, R., and von und zu Fraunberg, M. (2002). Molecular and biochemical studies of acute intermittent porphyria in 196 patients and their families. Clin. Chem. *48*, 1891–1900.

Kuo, H.C., Huang, C.C., Chu, C.C., Lee, M.J., Chuang, W.L., Wu, C.L., Wu, T., Ning, H.C., and Liu, C.Y. (2011). Neurological complications of acute intermittent porphyria. Eur. Neurol. *66*, 247–252.

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* (Pt 2), 447–452.

Layer, G., Reichelt, J., Jahn, D., and Heinz, D.W. (2010). Structure and function of enzymes in heme biosynthesis. Protein Sci. *19*, 1137–1161.

Louie, G.V., Brownlie, P.D., Lambert, R., Cooper, J.B., Blundell, T.L., Wood, S.P., Malashkevich, V.N., Hadener, A., Warren, M.J., and Shoolingin-Jordan, P.M. (1996). The three-dimensional structure of Escherichia coli porphobilinogen deaminase at 1.76-A resolution. 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., and Jordan, P.M. (1992). Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. Nature *359*, 33–39.

Mendez, M., Moran-Jimenez, M.J., Gomez-Abecia, S., Garcia-Bravo, M., Garrido-Astray, M.C., Fontanellas, A., Poblete-Gutierrez, P., Frank, J., and Enriquez de Salamanca, R. (2009). Identification and characterization of HMBS gene mutations in Spanish patients with acute intermittent porphyria. Cell. Mol. Biol. (Noisy-le-grand) *55*, 55–63.

Mustajoki, P., and Desnick, R.J. (1985). Genetic heterogeneity in acute intermittent porphyria: characterisation and frequency of porphobilinogen deaminase mutations in Finland. Br. Med. J. *291*, 505–509.

Mustajoki, S., Laine, M., Lahtela, M., Mustajoki, P., Peltonen, L., and Kauppinen, R. (2000). Acute intermittent porphyria: expression of mutant and wild-type porphobilinogen deaminase in COS-1 cells. Mol. Med. *6*, 670–679.

Niemann, A.C., Matzinger, P.K., and Hadener, A. (1994). A kinetic-analysis of the reaction catalyzed by (Hydroxymethyl)Bilane synthase. Helvetica Chim. Acta *77*, 1791–1809.

Pluta, P., Roversi, P., Bernardo-Seisdedos, G., Rojas, A.L., Cooper, J.B., Gu, S., Pickersgill, R.W., and Millet, O. (2018). Structural basis of pyrrole polymerization in human porphobilinogen deaminase. Bba-Gen Subjects *1862*, 1948–1955.

Puy, H., Deybach, J.C., Lamoril, J., Robreau, A.M., Da Silva, V., Gouya, L., Grandchamp, B., and Nordmann, Y. (1997). Molecular epidemiology and diagnosis of PBG deaminase gene defects in acute intermittent porphyria. Am. J. Hum. Genet. *60*, 1373–1383.

Roberts, A., Gill, R., Hussey, R.J., Mikolajek, H., Erskine, P.T., Cooper, J.B., Wood, S.P., Chrystal, E.J., and Shoolingin-Jordan, P.M. (2013). Insights into the mechanism of pyrrole polymerization catalysed by porphobilinogen deaminase: high-resolution X-ray studies of the Arabidopsis thaliana enzyme. Acta Crystallographica. Section D Biol. Crystallogr. *69*, 471–485.

Schneider-Yin, X., Bogard, C., Rufenacht, U.B., Puy, H., Nordmann, Y., Minder, E.I., and Deybach, J. (2000). Identification of a prevalent nonsense mutation (W283X) and two novel mutations in the porphobilinogen deaminase gene of Swiss patients with acute intermittent porphyria. Hum. Hered. *50*, 247–250.

Schneider-Yin, X., Szlendak, U., Lipniacka, A.I., Minder, E.I., and Gregor, A. (2006). Nine novel mutations in the hydroxymethylbilane synthase gene of Polish patients with acute intermittent porphyria. Clin. Genet. *69*, 284–286.

Schneider-Yin, X., Ulbrichova, D., Mamet, R., Martasek, P., Marohnic, C.C., Goren, A., Minder, E.I., and Schoenfeld, N. (2008). Characterization of two missense variants in the hydroxymethylbilane synthase gene in the Israeli population, which differ in their associations with acute intermittent porphyria. Mol. Genet. Metab. *94*, 343–346.

Scott, A.I., Clemens, K.R., Stolowich, N.J., Santander, P.J., Gonzalez, M.D., and Roessner, C.A. (1989). Reconstitution of apo-porphobilinogen deaminase: structural changes induced by cofactor binding. FEBS Lett. *242*, 319–324.

Scott, A.I., Roessner, C.A., Stolowich, N.J., Karuso, P., Williams, H.J., Grant, S.K., Gonzalez, M.D., and Hoshino, T. (1988). Site-directed mutagenesis and high-resolution NMR spectroscopy of the active site of porphobilinogen deaminase. Biochemistry *27*, 7984–7990.

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.

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* (Pt 2), 373–376.

Solis, C., Martinez-Bermejo, A., Naidich, T.P., Kaufmann, W.E., Astrin, K.H., Bishop, D.F., and Desnick, R.J. (2004). Acute intermittent porphyria: studies of the severe homozygous dominant disease provides insights into the neurologic attacks in acute porphyrias. Arch. Neurol. *61*, 1764–1770.

Song, G., Li, Y., Cheng, C., Zhao, Y., Gao, A., Zhang, R., Joachimiak, A., Shaw, N., and Liu, Z.J. (2009). Structural insight into acute intermittent porphyria. FASEB J. *23*, 396–404.

Uchida, T., Funamizu, T., Chen, M., Tanaka, Y., and Ishimori, K. (2018). Heme binding to porphobilinogen deaminase from Vibrio cholerae decelerates the formation of 1-hydroxymethylbilane. ACS Chem. Biol. *13*, 750–760.

Warren, M.J., and Jordan, P.M. (1988). Investigation into the nature of substrate binding to the dipyrromethane cofactor of Escherichia coli porphobilinogen deaminase. Biochemistry *27*, 9020–9030.

Whatley, S.D., Mason, N.G., Woolf, J.R., Newcombe, R.G., Elder, G.H., and Badminton, M.N. (2009). Diagnostic strategies for autosomal dominant acute porphyrias: retrospective analysis of 467 unrelated patients referred for mutational analysis of the HMBS, CPOX, or PPOX gene. Clin. Chem. *55*, 1406–1414.

Yang, C.C., Kuo, H.C., You, H.L., Wang, J., Huang, C.C., Liu, C.Y., Lan, M.Y., Stephenson, D.A., and Lee, M.J. (2008). HMBS mutations in Chinese patients with acute intermittent porphyria. Ann. Hum. Genet. *72*, 683–686.

iScience, Volume *24*

Supplemental information

Characterization of porphobilinogen deaminase mutants reveals that arginine-173 is crucial for polypyrrole elongation mechanism

Helene J. Bustad, Juha P. Kallio, Mikko Laitaoja, Karen Toska, Inari Kursula, Aurora Martinez, and Janne Jänis

Supplemental Information

Supplemental Figures

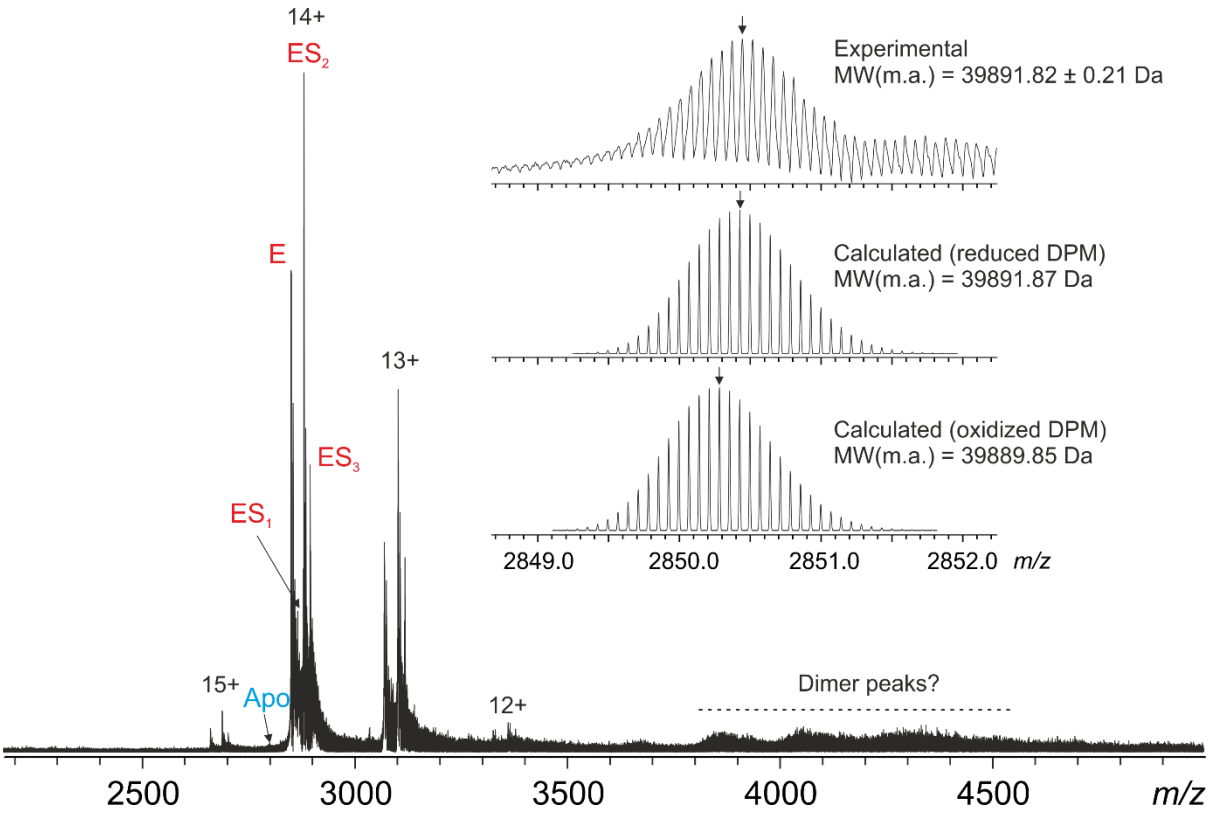


Figure S1. Native mass spectrometry of wt-hPBGD (Related to Figure 2). The ESI FT-ICR mass spectrum was acquired with10 µM protein in 10 mM ammonium acetate pH 6.9. Narrow protein ion-charge state-distribution (12+ to 15+) is consistent with a tightly folded protein conformation. Peaks representing different enzyme-intermediates are assigned for charge state 14+. The inset shows an expanded view on the charge state 14+ of the holoenzyme (E): experimental isotopic pattern (top); calculated isotopic patterns with the reduced (middle) or oxidized (bottom) dipyrromethane (DPM) cofactor present in the structure. The peaks representing the most abundant isotopic masses (MW (m.a.)) are indicated by small arrows. Faint poorly resolved signals at *m/z* 3800–4500, possibly representing non-covalent hPBGD dimer are also indicated.

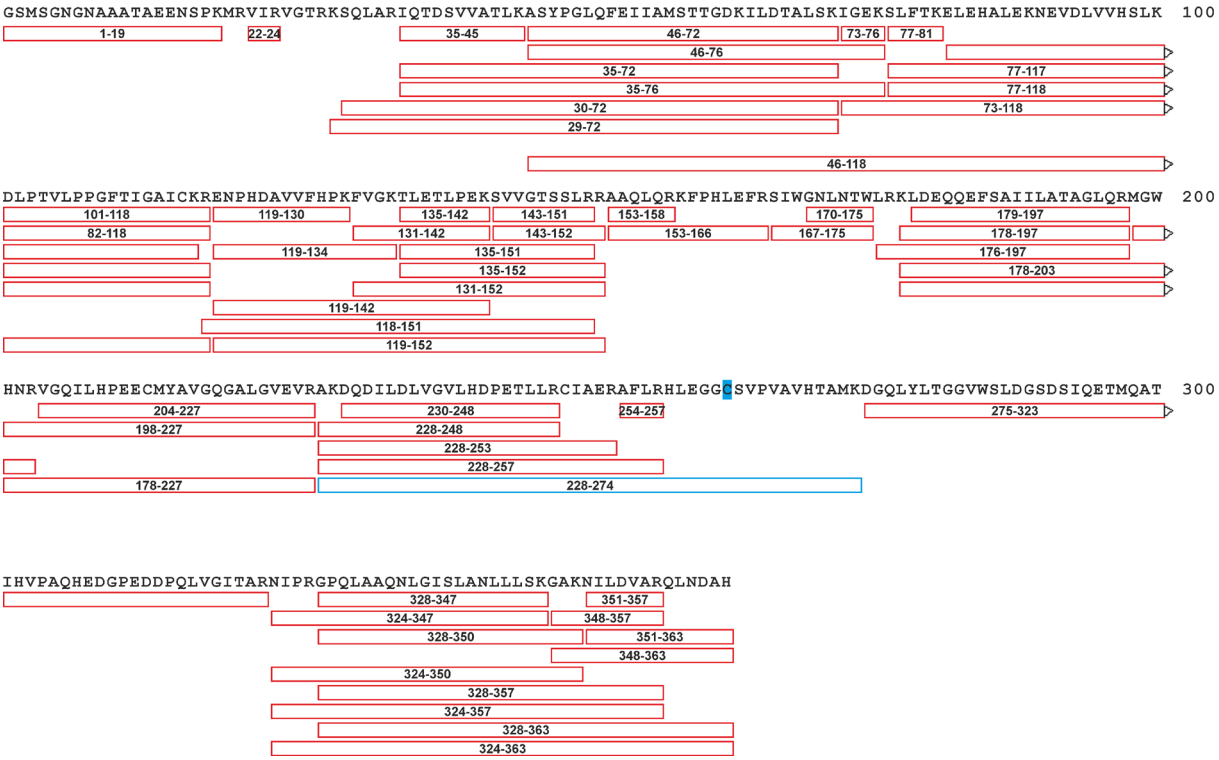


Figure S2. A tryptic peptide map for wt-hPBGD (Related to Figure 2). The numbers in the red boxes indicate corresponding residues for each observed tryptic peptide (note: the numbering here corresponds to the produced protein with two additional N-terminal residues (GS), remnant from the expression vector construct). The blue box indicates the peptide with a covalently modified Cys261 residue (tetrapyrrole = ES_2 intermediate).

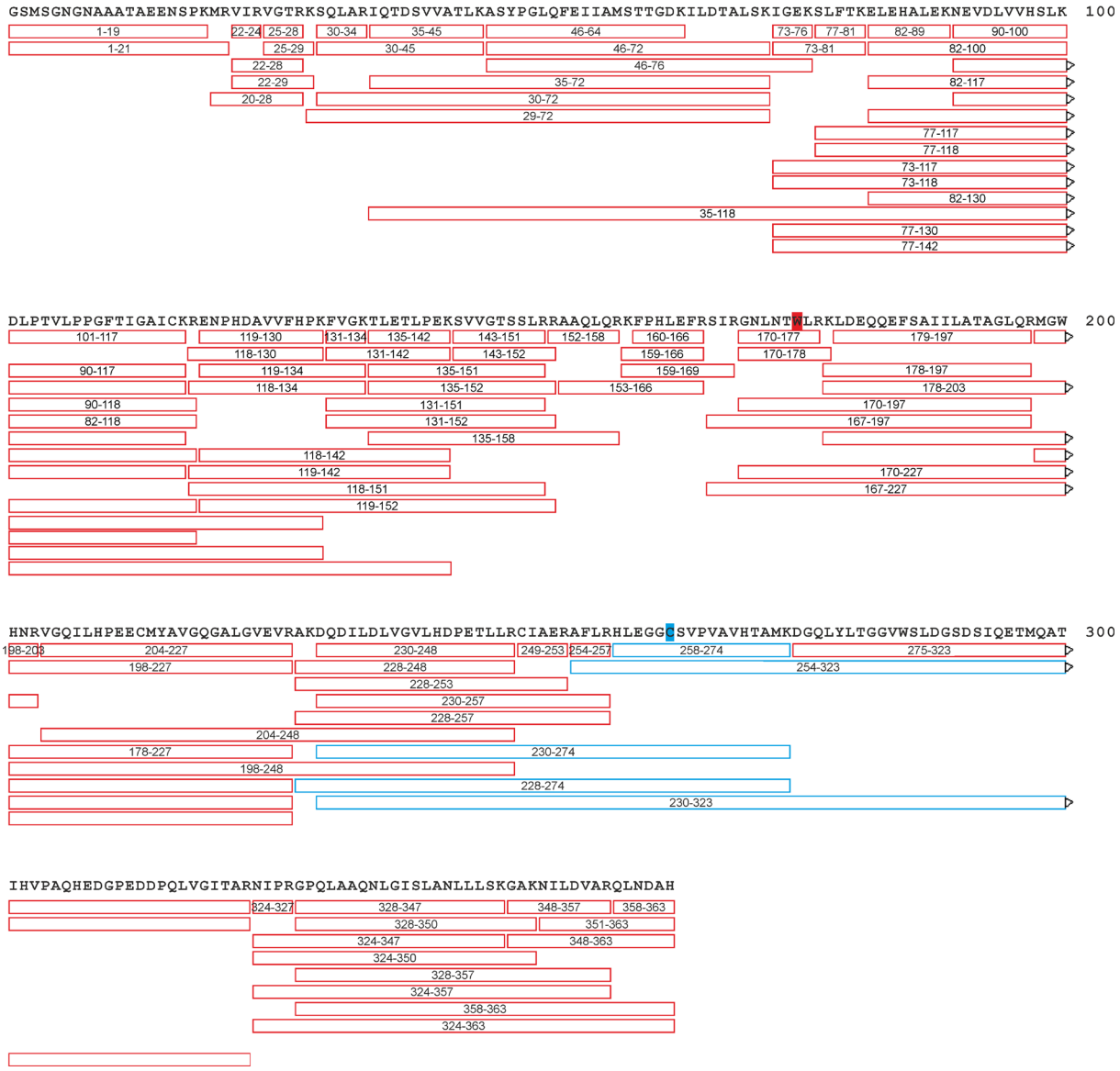


Figure S3. A tryptic peptide map for hPBGD-R173W (Related to Figure 2). The numbers in the red boxes indicate corresponding residues for each observed tryptic peptide (note: the numbering here corresponds to the produced protein with two additional N-terminal residues (GS), remnant from the expression vector construct). The blue box indicates the peptide with a covalently modified Cys261 residue (a tetrapyrrole = ES_2 intermediate).

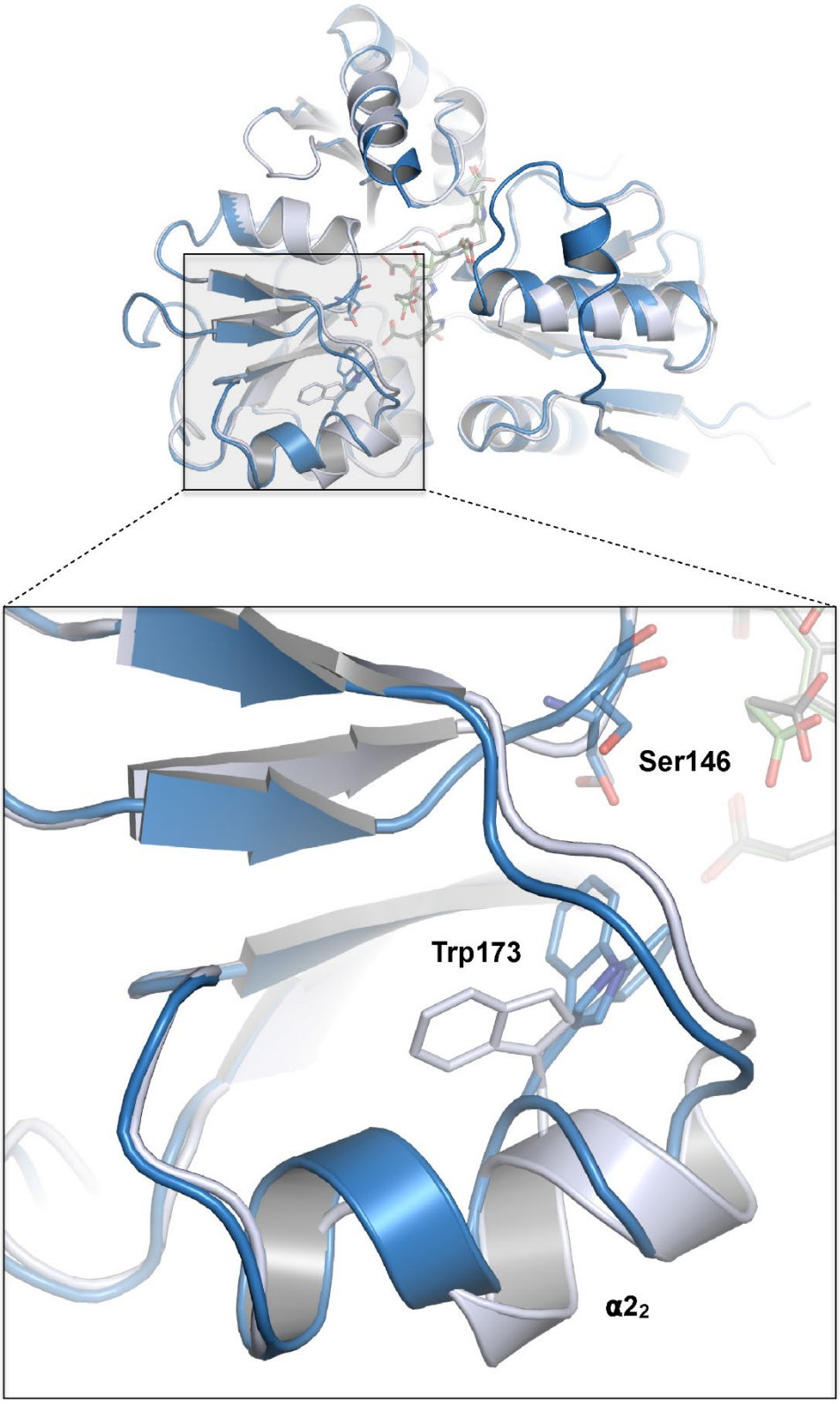


Figure S4. Comparison of the differences between the two molecules in the asymmetric unit in R173W-ES$_2$ (Related to Figure 3). Molecule A (chain A) and molecule B (chain B) are shown in blue and light blue, respectively. Inset for domain 2, with the enlargement below, shows the alternative conformations of Ser146 and Trp173 as well as the rearrangement of in the $\alpha 2_2$ (residues 170–179) in subunit B.

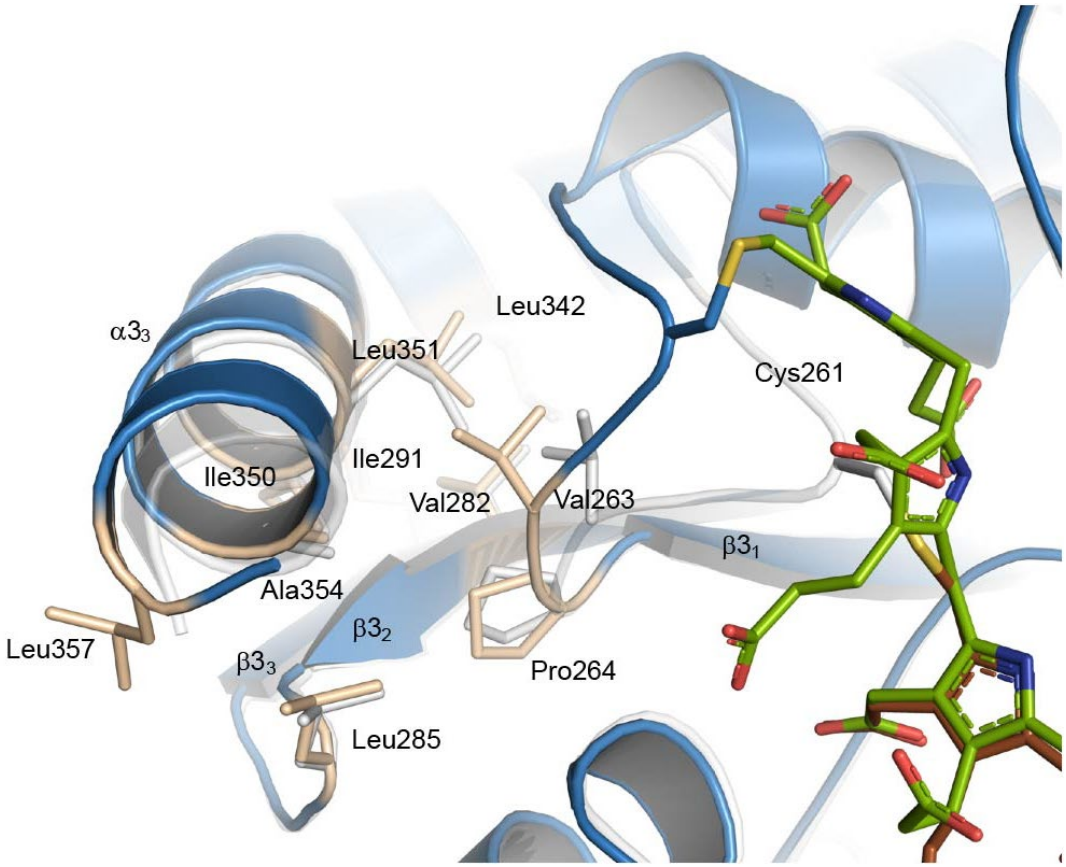


Figure S5. Cartoon representation of mutant R173W-ES_2 (blue) and wt-E_{holo} (grey) superimposed showing the hydrophobic interactions anchoring the $\alpha3_3$ in proximity of the cofactor binding loop (Related to Figure 3). Elongation product of R173W-ES_2 is shown in green and DPM cofactor with of wt-E_{holo} is shown in brown. Hydrophobic residues on R173W-ES_2 between the C-terminal helix ($\alpha3_3$) and the β-sheet ($\beta3_{1\text{-}3}$) are shown in tan. Upon the movement of cofactor-binding loop from E_{holo} to ES_2 only small change in Val263 can be detected.

Supplemental Tables

Table S1. List of tryptic peptides observed for wt-hPBGD (Related to Figure 2). The sequence in blue indicates the peptide with a covalently modified Cys261 residue (tetrapyrrole = ES_2 intermediate).

Experimental mass (Da)	Theoretical mass (Da)	Error (ppm)	Residues (start-end)		Sequence
386.26396	386.2642	0	22	24	VIR
445.2543	445.2537	-1	73	76	IGEK
505.30252	505.3013	-2	254	257	AFLR
594.33923	594.3377	-3	77	81	SLFTK
673.35252	673.3507	-3	170	175	GNLNTR
685.3891	685.3871	-3	153	158	AAQLQR
799.45691	799.4552	-2	351	357	NILDVAR
904.49974	904.4978	-2	143	151	SVVGTSSLR
929.5091	929.507	-2	135	142	TLETLPEK
1029.56994	1029.5679	-2	167	175	SIRGNLNTR
1055.61139	1055.6087	-3	348	357	GAKNILDVAR
1060.60153	1060.5989	-2	143	152	SVVGTSSLRR
1173.6635	1173.6605	-3	35	45	IQTDSVVATLK
1360.76376	1360.7602	-3	131	142	FVGKTLETLPEK
1388.68765	1388.6837	-3	119	130	ENPHDAVVFHPK
1477.76755	1477.7637	-3	351	363	NILDVARQLNDAH
1733.92213	1733.9173	-3	348	363	GAKNILDVARQLNDAH
1739.96241	1739.9584	-2	153	166	AAQLQRKFPHLEFR
1791.77445	1791.7694	-3	1	19	GSMSGNGNAAATAEENSPKMR
1815.99992	1815.9942	-3	135	151	TLETLPEKSVVGTSSLR
1819.94175	1819.937	-3	119	134	ENPHDAVVFHPKFVGK
1897.05478	1897.0495	-3	101	118	DLPTVLPPGFTIGAICKR
1972.10014	1972.0953	-2	135	152	TLETLPEKSVVGTSSLRR
2020.17372	2020.1681	-3	328	347	GPQLAAQNLGISLANLLLSK
2102.10677	2102.1008	-3	179	197	LDEQQEFSAIILATAGLQR
2160.14881	2160.1427	-3	230	248	DQDILDLVGVLHDPETLLR
2230.20143	2230.1958	-3	178	197	KLDEQQEFSAIILATAGLQR
2276.32759	2276.3216	-3	328	350	GPQLAAQNLGISLANLLLSKGAK
2359.28078	2359.2747	-3	228	248	AKDQDILDLVGVLHDPETLLR
2403.3552	2403.3486	-3	131	152	FVGKTLETLPEKSVVGTSSLRR
2499.38561	2499.3809	-2	176	197	LRKLDEQQEFSAIILATAGLQR
2500.45497	2500.4489	-2	324	347	NIPRGPQLAAQNLGISLANLLLSK
2554.2737	2554.2672	-3	204	227	VGQILHPEECMYAVGQGALGVEVR
2731.44078	2731.4334	-3	119	142	ENPHDAVVFHPKFVGKTLETLPEK
2756.61002	2756.6025	-3	324	350	NIPRGPQLAAQNLGISLANLLLSKGAK
2869.48535	2869.4783	-2	46	72	ASYPGLQFEIIAMSTTGDKILDTALSK
2931.55694	2931.5488	-3	228	253	AKDQDILDLVGVLHDPETLLRCIAER

3011.54833	3011.54	-3	178	203	KLDEQQEFSAIILATAGLQRMGWHNR
3057.77481	3057.7663	-3	328	357	GPQLAAQNLGISLANLLLSKGAKNILDVAR
3296.73031	3296.7214	-3	46	76	ASYPGLQFEIIAMSTTGDKILDTALSKIGEK
3335.6183	3335.6114	-2	198	227	MGWHNRVGQILHPEECMYAVGQGALGVEVR
3418.84888	3418.8395	-3	228	257	AKDQDILDLVGVLHDPETLLRCIAERAFLR
3538.05541	3538.0471	-2	324	357	NIPRGPQLAAQNLGISLANLLLSKGAKNILDVAR
3736.08447	3736.0748	-3	328	363	GPQLAAQNLGISLANLLLSKGAKNILDVARQLNDAH
3774.03047	3774.0217	-2	119	152	RENPHDAVVFHPKFVGKTLETLPEKSVVGTSSLR
4025.13911	4025.1283	-3	35	72	IQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSK
4080.21877	4080.2082	-3	82	118	ELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
4216.36769	4216.3557	-3	324	363	NIPRGPQLAAQNLGISLANLLLSKGAKNILDVARQLNDAH
4452.38149	4452.3713	-2	35	76	IQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSKIGEK
4500.44618	4500.4342	-3	77	117	SLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICK
4580.45307	4580.4412	-3	30	72	SQLARIQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSK
4656.54721	4656.5353	-3	77	118	SLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
4708.54925	4708.5361	-3	29	72	KSQLARIQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSK
5083.78831	5083.7784	-2	73	118	IGEKSLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
5246.4875	5246.4739	-3	275	323	DGQLYLTGGVWSLDGSDSIQETMQATIHVPAQHEDGPEDDPQLVGITAR
5547.80978	5547.7966	-2	178	227	KLDEQQEFSAIILATAGLQRMGWHNRVGQILHPEECMYAVGQGALGVEVR
5971.97033	5971.9587	-2	228	274	AKDQDILDLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVAVHTAMK (ES_2)
7935.26651	7935.2462	-3	46	118	ASYPGLQFEIIAMSTTGDKILDTALSKIGEKSLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR

Table S2. List of tryptic peptides observed for R173W-hPBGD (Related to Figure 2).
The sequences in blue indicate the peptide with a covalently modified Cys261 residue (tetrapyrrole = ES_2 intermediate).

Experimental mass (Da)	Theoretical mass (Da)	Error (ppm)	Residues (start-end)		Sequence
386.2641	386.2642	0	22	24	VIR
431.2502	431.2492	2	25	28	VGTR
445.2550	445.2537	3	73	76	IGEK
449.2653	449.2638	3	131	134	FVGK
498.2932	498.2914	4	324	327	NIPR
505.3034	505.3013	4	254	257	AFLR
559.3457	559.3442	3	25	29	VGTRK
573.3254	573.3235	3	30	34	SQLAR
590.2862	590.2846	3	249	253	CIAER
594.3396	594.3377	3	77	81	SLFTK
696.3211	696.3191	3	358	363	QLNDAH
799.3582	799.3548	4	198	203	MGWHNR
799.5046	799.5028	2	22	28	VIRVGTR
841.4910	841.4882	3	152	158	RAAQLQR
904.5002	904.4978	3	143	151	SVVGTSSLR
927.6016	927.5978	4	22	29	VIRVGTRK
929.5067	929.5070	0	135	142	TLETLPEK
944.4900	944.4868	3	160	166	FPHLEFR
967.5006	967.4975	3	82	89	ELEHALEK
972.5154	972.5141	1	170	177	GNLNTWLR
1021.5833	1021.5808	2	73	81	IGEKSLFTK
1055.6128	1055.6087	4	348	357	GAKNILDVAR
1060.6028	1060.5989	4	143	152	SVVGTSSLRR
1072.5847	1072.5818	3	159	166	KFPHLEFR
1086.6480	1086.6444	3	20	28	MRVIRVGTR
1100.6130	1100.6091	4	170	178	GNLNTWLRK
1173.6598	1173.6605	-1	35	45	IQTDSVVATLK
1251.6853	1251.6823	2	90	100	NEVDLVVHSLK
1360.7645	1360.7602	3	131	142	FVGKTLETLPEK
1388.6879	1388.6837	3	119	130	ENPHDAVVFHPK
1428.8047	1428.7990	4	159	169	KFPHLEFRSIR
1477.7664	1477.7637	2	351	363	NILDVARQLNDAH
1544.7885	1544.7848	2	118	130	RENPHDAVVFHPK
1728.9760	1728.9734	2	30	45	SQLARIQTDSVVATLK
1733.9201	1733.9173	2	348	363	GAKNILDVARQLNDAH
1739.9624	1739.9584	2	153	166	AAQLQRKFPHLEFR
1740.9509	1740.9484	1	101	117	DLPTVLPPGFTIGAICK
1791.7720	1791.7694	1	1	19	GSMSGNGNAAATAEENSPK
1815.9974	1815.9942	2	135	151	TLETLPEKSVVGTSSLR

1819.9433	1819.9370	3	119	134	ENPHDAVVFHPKFVGK
1972.0989	1972.0953	2	135	152	TLETLPEKSVVGTSSLRR
1976.0431	1976.0381	3	118	134	RENPHDAVVFHPKFVGK
2020.1705	2020.1681	1	328	347	GPQLAAQNLGISLANLLLSK
2027.9875	2027.9874	0	46	64	ASYPGLQFEIIAMSTTGDK
2078.9126	2078.9110	1	1	21	GSMSGNGNAAATAEENSPKMR
2102.1069	2102.1008	3	179	197	LDEQQEFSAIILATAGLQR
2160.1414	2160.1427	-1	230	248	DQDILDLVGVLHDPETLLR
2201.1730	2201.1692	2	82	100	ELEHALEKNEVDLVVHSLK
2230.1973	2230.1958	1	178	197	KLDEQQEFSAIILATAGLQR
2247.2525	2247.2475	2	131	151	FVGKTLETLPEKSVVGTSSLR
2276.3270	2276.3216	2	328	350	GPQLAAQNLGISLANLLLSKGAK
2359.2785	2359.2747	2	228	248	AKDQDILDLVGVLHDPETLLR
2403.3518	2403.3486	1	131	152	FVGKTLETLPEKSVVGTSSLRR
2500.4516	2500.4489	1	324	347	NIPRGPQLAAQNLGISLANLLLSK
2554.2674	2554.2672	0	204	227	VGQILHPEECMYAVGQGALGVEVR
2571.1231	2571.1298	-3	258	274	HLEGGCSVPVAVHTAMK (ES_2)
2639.4745	2639.4719	1	135	158	TLETLPEKSVVGTSSLRRAAQLQR
2731.4361	2731.4334	1	119	142	ENPHDAVVFHPKFVGKTLETLPEK
2756.6053	2756.6025	1	324	350	NIPRGPQLAAQNLGISLANLLLSKGAK
2869.4780	2869.4783	0	46	72	ASYPGLQFEIIAMSTTGDKILDTALSK
2887.5403	2887.5345	2	118	142	RENPHDAVVFHPKFVGKTLETLPEK
2931.5500	2931.5488	0	228	253	AKDQDILDLVGVLHDPETLLRCIAER
2974.6231	2974.6202	1	90	117	NEVDLVVHSLKDLPTVLPPGFTIGAICK
3011.5415	3011.5400	0	178	203	KLDEQQEFSAIILATAGLQRMGWHNR
3057.7675	3057.7663	0	328	357	GPQLAAQNLGISLANLLLSKGAKNILDVAR
3130.7239	3130.7213	1	90	118	NEVDLVVHSLKDLPTVLPPGFTIGAICKR
3184.6994	3184.6993	0	170	197	GNLNTWLRKLDEQQEFSAIILATAGLQR
3219.7042	3219.7074	-1	230	257	DQDILDLVGVLHDPETLLRCIAERAFLR
3296.7213	3296.7214	0	46	76	ASYPGLQFEIIAMSTTGDKILDTALSKIGEK
3335.6085	3335.6114	-1	198	227	MGWHNRVGQILHPEECMYAVGQGALGVEVR
3418.8392	3418.8395	0	228	257	AKDQDILDLVGVLHDPETLLRCIAERAFLR
3538.0508	3538.0471	1	324	357	NIPRGPQLAAQNLGISLANLLLSKGAKNILDVAR
3540.9120	3540.9165	-1	167	197	SIRGNLNTWLRKLDEQQEFSAIILATAGLQR
3736.0670	3736.0748	-2	328	363	GPQLAAQNLGISLANLLLSKGAKNILDVARQLNDAH
3774.0233	3774.0217	0	118	151	RENPHDAVVFHPKFVGKTLETLPEKSVVGTSSLR
3774.0233	3774.0217	0	119	152	ENPHDAVVFHPKFVGKTLETLPEKSVVGTSSLRR
3924.1041	3924.1071	-1	82	117	ELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICK
4025.1273	4025.1283	0	35	72	IQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSK
4080.1971	4080.2082	-3	82	118	ELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
4216.3556	4216.3557	0	324	363	NIPRGPQLAAQNLGISLANLLLSKGAKNILDVARQLNDAH

4500.4310	4500.4342	-1	77	117	SLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICK
4580.4320	4580.4412	-2	30	72	SQLARIQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSK
4656.5325	4656.5353	-1	77	118	SLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
4708.5311	4708.5361	-1	29	72	KSQLARIQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSK
4895.5283	4895.5314	-1	204	248	VGQILHPEECMYAVGQGALGVEVRAKDQDILDLVGVLHDPETLLR
4927.6665	4927.6773	-2	73	117	IGEKSLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICK
5083.7747	5083.7784	-1	73	118	IGEKSLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
5246.4611	5246.4739	-2	275	323	DGQLYLTGGVWSLDGSDSIQETMQATIHVPAQHEDGPEDDPQLVGITAR (ES_2)
5450.8746	5450.8813	-1	82	130	ELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKRENPHDAVVFHPK
5547.7773	5547.7966	-3	178	227	KLDEQQEFSAIILATAGLQRMGWHNRVGQILHPEECMYAVGQGALGVEVR
5676.8489	5676.8756	-5	198	248	MGWHNRVGQILHPEECMYAVGQGALGVEVRAKDQDILDLVGVLHDPETLLR
5772.8026	5772.8266	-4	230	274	DQDILDLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVAVHTAMK (ES_2)
5971.9411	5971.9587	-3	228	274	AKDQDILDLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVAVHTAMK (ES_2)
6027.1990	6027.2085	-2	77	130	SLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKRENPHDAVVFHPK
6502.2782	6502.3002	-3	170	227	GNLNTWLRKLDEQQEFSAIILATAGLQRMGWHNRVGQILHPEECMYAVGQGALGVEVR
6858.4893	6858.5174	-4	167	227	SIRGNLNTWLRKLDEQQEFSAIILATAGLQRMGWHNRVGQILHPEECMYAVGQGALGVEVR
7369.9376	7369.9581	-3	77	142	SLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKRENPHDAVVFHPKFVGKTLETLPEK
8286.8513	8286.8838	-4	254	323	AFLRHLEGGCSVPVAVHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIHVPAQHEDGPEDDPQLVGITAR
9090.8595	9090.8961	-4	35	118	IQTDSVVATLKASYPGLQFEIIAMSTTGDKILDTALSKIGEKSLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKR
11001.3389	11001.2900	4	230	323	DQDILDLVGVLHDPETLLRCIAERAFLRHLEGGCSVPVAVHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIHVPAQHEDGPEDDPQLVGITAR (ES_2)

Transparent Methods

Expression and purification of PBGD proteins

Recombinant human wild-type PBGD and the mutants R167W and R173W were expressed in *Escherichia coli* BL21 (DE3)pLysS (Agilent technologies, Santa Clara, CA, USA) as glutathione S-transferase fusion proteins with a thrombin cleavage site (Bustad et al., 2013). The plasmid for protein expression was generously provided by Professor Pavel Martasek and Dana Ulbrichova as a pGEX4T-1-expression vector (Ulbrichova et al., 2006). Overnight 50 ml pre-cultures supplemented with 100 µM ampicillin and 34 µg/ml chloramphenicol were diluted into 950 ml Luria Bertani broth supplemented with the same antibiotics and 2 g/l glucose and grown at 37 °C in a shaking incubator at 200 rpm until OD_{600} ~0.8. Protein expression was induced by adding 1 mM IPTG (isopropyl thio-*β*-D-galactoside) and grown overnight at 28 °C, 200 rpm. Cells were harvested by centrifugation at 4000 *g* and 4 °C for 15 min at, and stored at −80 ºC. Lysis was performed by sonication on ice after resuspension in lysis buffer (PBS; 140 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4 and 1.8 mM KH2PO4; 1 mM EDTA, 0.5 mM PMSF, 1 mM benzamidine and cOmplete™ Protease Inhibitor tablet (Roche Applied Science, Penzberg, Germany)), pH 7.4) with three 45-second rounds, 20 W output and 9 sec pulses. Soluble protein was obtained by centrifugation at 14 000 *g* for 45 min at 4 °C, and loaded onto a gluthatione-sepharose 4B column (GE Healthcare, Chicago, Illinois, USA) and washed with 10 × CV of PBS containing 1 mM EDTA, followed by 2 × CV of PBS. The fusion protein was cleaved by adding thrombin (200 units/l) and incubating for 2 hrs at 4 °C on rotation, in 50 mM Tris-HCl, pH 8.0, containing 2 mM CaCl2, 1 mM DTT (dithiothreitol) and 150 mM NaCl. The protein were concentrated with Amicon Ultra centrifugal 30 kDa cut-off filters (Merck Millipore, Burlington, MA, USA), followed by size exclusion chromatography using a Superdex 200 HR16/60 column (GE Healthcare) in 20 mM HEPES with 150 mM NaCl, pH 8.2. After cleavage of the fusion protein, two additional amino acid residues (glycine and serine,

which form part of the cutting site to provide the necessary flexibility for efficient cleavage) remain attached at the protein N-terminus. The protein concentration was measured using a NanoDrop® spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) with the appropriate extinction coefficient predicted by the ExPASy ProtParam web server. The average molecular masses of apo- and holo-PBGD and its intermediates were predicted adding these residues to the ExPASy ProtParam web server based on the following wt-sequence (catalytic Cys261 underlined); mutant residues were substituted where applicable:

(GS) MSGNGNAAATAEENSPKMRVIRVGTRKSQLARIQTDSVVATLKASYPGLQFEIIAMSTTG

DKILDTALSKIGEKSLFTKELEHALEKNEVDLVVHSLKDLPTVLPPGFTIGAICKRENPH

DAVVFHPKFVGKTLETLPEKSVVGTSSLRRAAQLQRKFPHLEFRSIRGNLNTRLRKLDEQ

QEFSAIILATAGLQRMGWHNRVGQILHPEECMYAVGQGALGVEVRAKDQDILDLVGVLHD

PETLLRCIAERAFLRHLEGG<u>C</u>SVPVAVHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIH

VPAQHEDGPEDDPQLVGITARNIPRGPQLAAQNLGISLANLLLSKGAKNILDVARQLNDAH

Mass spectrometry

Prior to the MS experiments, all proteins samples were first desalted/buffer exchanged into 10 mM ammonium acetate, pH 6.9, by using Sephadex G-25 M columns (PD-10; GE Healthcare) and concentrated using Vivaspin 5K centrifugal concentrators (GE Healthcare). The final protein concentrations were estimated by absorbance at 280 nm using sequence-derived extinction coefficients.

Trypsin digestion was obtained by incubating an aliquot of the protein stock solution with TCPK-treated sequencing grade trypsin (Promega, Madison, WI, USA) at a 1:15 (w/w) enzyme-to-protein ratio at 37 °C for 4 h. The digestion was quenched by diluting the samples to the desired protein concentration with a solvent mixture of acetonitrile/water/acetic acid (49.5:49.5:1.0, v/v). The same solvent was used for the intact protein analysis in *denaturing*

solution conditions. For the analysis in *near-native* solution conditions, 10 mM ammonium acetate, pH 6.9, was used instead.

All FT-ICR MS experiments were performed on a 12-T Bruker Solarix-XR instrument (Bruker Daltonik GmbH, Bremen, Germany), equipped with an Apollo-II electrospray ionisation (ESI) source and dynamically harmonised ICR cell (Paracell). All samples were directly infused at a flow rate of 1.5 µl/min using a syringe pump. For intact mass analysis, a total of 300 time-domain transients (1 MWord each) were co-added for each spectrum and zero-filled once to obtain final 2 MWord broadband data. This provided spectral resolution of roughly 200,000 (FWHM) at *m/z* 800 providing baseline-resolved protein signals. The final mass spectra were externally calibrated with respect to the ions of ESI-L Tuning Mix (Part no. G1969-85000; Agilent technologies). The instrument was operated, and the data were acquired with the use of solarixControl 2.0 software and the data were further processed and analysed with Bruker DataAnalysis 4.2 software. Spectral deconvolution for intact protein masses (reported as the most abundant isotopic masses throughout) was accomplished with a maximum entropy deconvolution (MaxEnt), while peptide monoisotopic masses were observed by using a SNAP2 peak-picking module. Sequence analysis (mass calculations) and tryptic peptide identifications were performed with GPMAW 10.0 software (Lighthouse Data). Briefly, monoisotopic peptide masses were matched against the protein sequence within a maximum mass error of 5 ppm with the covalently attached cofactor (DPM) or the polypyrrole as the variable modification.

X-ray crystallography

Both wt-hPBGD and R173W were concentrated to desired concentrations using Amicon Ultra centrifugal filters with 30 kDa cut-off (GE Healthcare) and filtered with 0.22 µm filter prior to the screening for the crystallisation conditions. Screening was performed at two initial temperatures, +8 and +20 °C, and two concentrations, 10 and 20 mg/ml, for both proteins using

vapour diffusion method in SwissCI SD2 sitting drop plates and commercial screens; PACT premier and JSCG plus for Molecular Dimension, and Crystal Screen light and Peg/Ion from Hampton Research. Pipetting was done using Mosquito LCP crystallisation robot (TTP Labtech).

From the initial hits, the conditions were optimised for both proteins. Single crystals were obtained after 4–7 days for both proteins from similar conditions containing 20–25% polyethylene glycol 3350, 200–300 mM ammonium citrate at pH 5.1. The optimised concentrations for wt-PBGD and R173W were 10 and 14 mg/ml, respectively. The crystals for the wt-PBGD were more prone to grow in stacks, whereas R173W crystals were mostly single crystals.

For the data collection, crystals were cryoprotected by soaking for a few seconds in solution having 20% glycerol in addition to crystallisation solution prior to flash freezing. The MX data for the wt-PBGD and R173W was collected at the beamlines P13 and P14, respectively, operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). The data was processed with XDS (Kabsch, 2010) and scaled with AIMLESS (Evans and Murshudov, 2013). Phaser-MR (McCoy et al., 2007) from PHENIX suite (Adams et al., 2010) was used for the molecular replacement using PBGD structure 3ECR as a search model. Structures were refined using phenix.refine (Afonine et al., 2012) and validated with MolProbity (Adams et al., 2010). Data collection and refinement statistics are shown in Table 1. All 3D structure diagrams were prepared using Pymol (The PyMOL Molecular Graphics System, Version 2.0.4 Schrödinger, LLC.). The atomic coordinates and structure factor amplitudes have been submitted to the Protein Data Bank (PDB) under the accession codes 7AAJ and 7AAK.

Supplemental References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213-221.

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., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. *68*, 352-367.

Bustad, H.J., Vorland, M., Ronneseth, E., Sandberg, S., Martinez, A., and Toska, K. (2013). Conformational stability and activity analysis of two hydroxymethylbilane synthase mutants, K132N and V215E, with different phenotypic association with acute intermittent porphyria. Biosci. Rep. *33*, e00056.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. *69*, 1204-1214.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658-674.

Ulbrichova, D., Flachsova, E., Hrdinka, M., Saligova, J., Bazar, J., Raman, C.S., and Martasek, P. (2006). De novo mutation found in the porphobilinogen deaminase gene in Slovak acute intermittent porphyria patient: Molecular biochemical study. Physiological Research *55*, 145-154.