Computational and Structural Biotechnology Journal 20 (2022) 3899–3910

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Unveiling molecular details behind improved activity at neutral to alkaline pH of an engineered DyP-type peroxidase

Patrícia T. Borges [a,1], Diogo Silva [a,1], Tomás F.D. Silva [b], Vânia Brissos [a], Marina Cañellas [c], Maria Fátima Lucas [c], Laura Masgrau [c,d], Eduardo P. Melo [e], Miguel Machuqueiro [b], Carlos Frazão [a], Lígia O. Martins [a,*]

[a] Instituto de Tecnologia Química e Biológica António Xavier, Universidade NOVA de Lisboa, Av. da República, 2780-157 Oeiras, Portugal
[b] BioISI – Biosystems & Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal
[c] Zymvol Biomodeling, Carrer Roc Boronat, 117, 08018 Barcelona, Spain
[d] Department of Chemistry, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[e] Centro de Ciências do Mar, Universidade do Algarve, 8005-139 Faro, Portugal

ARTICLE INFO

Article history:
Received 31 May 2022
Received in revised form 18 July 2022
Accepted 18 July 2022
Available online 21 July 2022

Keywords:
Biorefinery
Biocatalysis
Directed evolution
Protein stability
Structure-function relationships

ABSTRACT

DyP-type peroxidases (DyPs) are microbial enzymes that catalyze the oxidation of a wide range of substrates, including synthetic dyes, lignin-derived compounds, and metals, such as Mn^{2+} and Fe^{2+}, and have enormous biotechnological potential in biorefineries. However, many questions on the molecular basis of enzyme function and stability remain unanswered. In this work, high-resolution structures of *Pp*DyP wild-type and two engineered variants (6E10 and 29E4) generated by directed evolution were obtained. The X-ray crystal structures revealed the typical ferredoxin-like folds, with three heme access pathways, two tunnels, and one cavity, limited by three long loops including catalytic residues. Variant 6E10 displays significantly increased loops' flexibility that favors function over stability: despite the considerably higher catalytic efficiency, this variant shows poorer protein stability compared to wild-type and 29E4 variants. Constant-pH MD simulations revealed a more positively charged microenvironment near the heme pocket of variant 6E10, particularly in the neutral to alkaline pH range. This microenvironment affects enzyme activity by modulating the pK_a of essential residues in the heme vicinity and should account for variant 6E10 improved activity at pH 7–8 compared to the wild-type and 29E4 that show optimal enzymatic activity close to pH 4. Our findings shed light on the structure–function relationships of DyPs at the molecular level, including their pH-dependent conformational plasticity. These are essential for understanding and engineering the catalytic properties of DyPs for future biotechnological applications.

1. Introduction

The conversion of lignocellulosic biomass into value-added products has become a significant research area in pursuing alternative sources of fuels, chemicals, and materials. A growing scientific community is seeking sustainable lignin valorization approaches since this is the most abundant aromatic polymer on Earth and the second most abundant raw material next to cellulose [69,78]. Dye decolorizing peroxidases (DyPs) are microbial peroxidases capable of efficient oxidation of a set of structurally diverse substrates, such as synthetic anthraquinone dyes, aromatic sulfides, metals, phenolic and nonphenolic lignin units, showing desired activities for biotechnological purposes, in particular in the lignocellulose biorefinery context [15,66]. DyPs (EC 1.11.1.19) contain a ferredoxin-like fold distinct from the motifs found in the vast majority of class II peroxidases including the ligninolytic fungal lignin, versatile and manganese peroxidases, with an α-helical-based structure [29,65,68]. Furthermore, DyPs lack the highly conserved distal histidine that acts as an acid-base catalyst in the catalytic reaction of other well-characterized peroxidases. Instead, they possess a conserved distal carboxylate residue (usually aspartic acid) that, together with arginine, mediates the catalysis in DyPs. The aspartate is part of the GXXDG motif, a

* Corresponding author.
E-mail address: lmartins@itqb.unl.pt (L.O. Martins).

[1] P.T.B. and D.S. contributed equally to this work.

https://doi.org/10.1016/j.csbj.2022.07.032

well-known conserved region among DyPs, and was suggested to contribute to the optimal acidic pH of DyPs [49,63]. DyPs catalytic mechanism is proposed to resemble that of classical peroxidases [9,43,49,62,63]; the distal aspartate helps remove the proton from the proximal oxygen atom of the bound Fe(III)–O_2H_2 to generate an anionic precursor to Compound I (Cpd I) known as Compound 0 (Cpd 0; Fe(III)-O-O–H). The proton is then transferred to the distal oxygen atom of Cpd 0 to form an oxy-water complex (Fe(III)–O–OH_2), facilitating heterolytic cleavage of the O-O bond to form Cpd I ($[Fe(IV) = O]^{+\bullet}$) and release an H_2O molecule [12,63]. In classic peroxidases, the distal arginine is proposed to maintain the distal heme cavity architecture and support heterolysis of the O–O bond by electrostatically interacting with the negatively charged distal oxygen on the Fe(III)-O-O–H complex during the formation of Cpd I [50,51]. The reaction of Cpd I with one equivalent of the reducing substrate yields Compound II (Cpd II; $[Fe(IV) = O]^{+}$), which further reacts with a second equivalent of the reducing substrate producing the resting state Fe(III) peroxidase. The reduction of Cpd, I via a single 2e- reduction step was suggested in the P-class *El*DyP from *Enterobacter lignolyticus* [62]. In DyPs, the precise functional role of the distal catalytic amino acids Asp and Arg in the heterolytic cleavage of H_2O_2 is still under debate [37,68]. Three classes based on structure-based sequence alignments, V (advanced), I (intermediate), and P (primitive), are proposed for this family [81]. Even though DyPs have nearly identical active sites, they exhibit a significant variation in catalytic rates and substrate specificity. Therefore, investigating molecular details in these enzymes can provide fundamental insight into their structure–function relationships and smooth their biotechnological applications.

In recent years, we have devoted efforts to the investigation of *Bs*DyP (class I) from *Bacillus subtilis* and *Pp*DyP (class P) from *Pseudomonas putida* MET94 [55], providing mechanistic insight into the catalytic cycle and reaction intermediates [9,43,73], electrocatalytic properties and interactions modulating redox properties [42,58,59,60]. Additionally, directed evolution approaches allowed for identifying engineered variants with improved catalytic efficiency for the lignin-related phenolic 2,6-dimethoxyphenol (DMP) [6,9,54]. We successfully engineered a highly active *Pp*DyP variant (6E10), the first DyP-type peroxidase with a broad pH range and an alkaline pH optimum [9], expanding significantly the range of applications for these enzymes.

This work contains kinetic, biochemical and structural information of wild-type *P. putida Pp*DyP and of two engineered variants, 6E10 and 29E4, complemented with conformational and protonation analysis from constant-pH molecular dynamics simulations (CpHMD), providing details into the features that underlie the activity and stability of these enzymes. Further details on enzyme interaction with substrates were also studied using ensemble docking protocols. Our findings have fundamental importance in understanding the molecular features of DyPs and evaluating their biotechnological potential.

2. Material and methods

2.1. Enzyme production and purification

E. coli strain DH5α (Novagen) was used for amplification of plasmid constructs. *E. coli* Tuner (DE3, Novagen) was used to express the *ppDyP* gene previously cloned in pET21-a (+) plasmid (Novagen) (PRC-1 plasmid [55]) or its evolved variants (p6E10 [9], p29E4 [6]). In the Tuner strain, genes are under the control of the T7 promoter, and its expression is induced by isopropyl β-D-1-thiogalactopyranoside (IPTG). Luria-Bertani medium (LB) was used to grow *E. coli* strains, supplemented with 100 $\mu g \cdot mL^{-1}$ ampicillin. Recombinant enzymes were produced in 1L of LB medium in 5L-Erlenmeyers and purified as previously described [6]. Purified enzyme preparations of wild-type *Pp*DyP and variant enzymes 6E10 and 29E4 show a heme *b* content of ~1 mol per mole of protein and Reinheitszahl values between ~1 and 2 (Table S1).

2.2. Apparent steady-state kinetic analysis

Enzymatic activities of *Pp*DyP wild-type and variants were monitored using a Synergy2 microplate reader (BioTek, Vermont, USA) at 25 °C. The kinetic parameters for the reduced substrates 2,2′-azino *bis* (3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) and 2,6-dimethoxy phenol (DMP) were measured in the presence of 1.5 mM H_2O_2 for wild-type, 29E4 and 6E10. Apparent steady-state kinetic parameters (k_{cat} and K_m) were measured for ABTS (ε_{420nm} = 36,000 M^{-1} cm^{-1}) (0.1 – 3 mM) in 100 mM sodium acetate at the optimal pH of the enzymes. The kinetic parameters for DMP (ε_{468nm} = 49,600 M^{-1} cm^{-1}) (0.01–2 mM) were measured in 100 mM sodium acetate or phosphate buffer at the optimal pH of the enzymes. The concentration of purified proteins was estimated using the molar absorption coefficient of *Pp*DyP (ε_{280} = 34,850 $M^{-1} \cdot cm^{-1}$), calculated from the protein sequence using the ExPASy Bioinformatics Resource Portal (http://web.expasy.org). Kinetic data were fitted directly using the Michaelis-Menten equation or the equation for the non-linear curve that fits enzyme kinetics affected by substrate inhibition ($v = V_{max}[S]/(K_m + [S](1 + [S]/K_i))$) (Origin software).

2.3. Size exclusion chromatography

The oligomerization state of *Pp*DyP and variants was typically assessed by injecting 100 μl of purified enzyme preparations into a gel filtration Superose 12 10/300 GL (GE Healthcare Bio-Sciences) column equilibrated with 20 mM Tris–HCl buffer, pH 7.6, and 0.2 M NaCl. The calibration curve was performed using the retention times vs. molecular mass of Protein Standards preparation (Bio-Rad Laboratories, CA, USA).

2.4. Stability assays

The thermodynamic stability was assessed by steady-state fluorescence measured with a Carry Eclipse spectrofluorometer (Agilent Technologies) at excitation wavelengths of 296 nm and emission wavelength of 350 nm [18,23]. For the equilibrium unfolding studies, guanidine hydrochloride (GdnHCl) concentrations in the range of 0–2.5 M in 20 mM Tris–HCl, 200 mM NaCl, pH 7.6, were used to induce protein unfolding, which was measured at room temperature. For thermal stability, the samples containing the enzymes (20 μM) in 20 mM Tris-HCl with 200 mM NaCl, pH 7.6, were placed onto a thermostatically controlled thermal block and then heated at a rate of 1 °C/min until 100 °C. Protein aggregation was monitored by static light scattering at 500 nm as excitation and emission wavelengths. The thermodynamic and thermal stability of enzymes was analyzed based on a two-state process using the equations previously described [23].

2.5. Crystallization and cryoprotection

Preliminary crystallization trials of *Pp*DyP wild-type, 6E10 and 29E4 variants (in 20 mM Tris-HCl pH 7.6 and 200 mM NaCl (10–15 mg/mL)) were performed using a mosquito crystallization robot (TTP Labtech) and Molecular Dimensions commercial screenings, namely JCSG+, BCS, and Morpheus. Vapor-diffusion crystallization trials were set using round-bottom 96-well CrystalQuickTM plates (Greiner Bio-One). The JCSG + screen produced *Pp*DyP wild-type crystals within two days at 20 °C in a crystallization solution com-

posed of 0.2 M magnesium formate dihydrate and 20 % (*w*/*v*) PEG 3350, using 0.1 μl of protein and reservoir solutions. The *Pp*DyP 6E10 crystals appeared at 20 °C after seven days in a condition part of the BCS screen: 0.08 M sodium bromide, 0.05 M sodium fluoride, 0.08 M sodium iodide, 0.1 M HEPES pH 7.8, and 22.5 % (*w*/*v*) PEG Smear Broad, using 0.1 μl of protein and reservoir solutions. *Pp*DyP 29E4 crystals were grown at 20 °C in a condition from the Morpheus screen: 0.02 M magnesium chloride hexahydrate, 0.02 M calcium chloride dehydrate, 0.1 M BICINE pH 8.5 and 37.5 % (*v*/*v*) of a precipitant mixture composed of 12.5 % (*v*/*v*) MPD, 12.5 % (*w*/*v*) PEG 1000, and 12.5% (*w*/*v*) PEG 3350, using 0.1 μl of protein and 0.2 μl of reservoir solutions. The crystallization hits were optimized at microliter-scale using hanging drop vapor diffusion trials in XRL 24-well crystallization plates (Molecular Dimensions). Best wild-type crystals were obtained in 0.3 M magnesium formate dihydrate, 17 % (*w*/*v*) PEG 3350 and 40 % (*v*/*v*) acetone. Wild-type and 6E10 crystals were transferred to their reservoir solutions supplemented with 25% (*v*/*v*) glycerol for flash-cooling in liquid nitrogen. Since the 29E4 crystallization condition is known to sustain "cryo-protective" properties [24], its crystals were directly flash-cooled in liquid nitrogen.

2.6. Data collection and processing

*Pp*DyP wild-type data sets were collected at 100 K in the European Synchrotron Radiation Facility (ESRF, Grenoble, France) at beamline ID30A-3. 6E10 and 29E4 data sets were collected in ALBA synchrotron (Barcelona, Spain) at beamline BL13-XALOC. Diffraction data were indexed, integrated, and scaled with the XDS program package [31]. Data collection and processing statistics are listed in Table 1.

2.7. Structure determination and refinement

Wild-type *Pp*DyP data were initially processed in space group $P3_221$, but data analyses with programs XPREP [61], XTRIAGE [83] and POINTLESS [22] indicated probable twinning defects in the crystal and suggested possible merohedral twin laws. As twinned crystals may show spurious symmetries, due to the imbedded twinned crystals, data were processed also in space groups C2 and P3, and tried to solve the phase problem with MORDA [77] testing all produced data sets, including possible alternative space groups. Two putative molecular replacement

Table 1
Data collection, processing, and refinement statistics.

	wild-type	6E10	29E4
Data collection			
Beamline	ID30A-3	BL13-XALOC	BL13-XALOC
Wavelength (Å)	0.9677	0.9792	0.9792
Space group	*$P3_221$*	*P23*	*P1*
Unit cell parameters (Å; °)	*a* = *b* = 141.9, *c* = 177.4;α = β = 90.0, γ = 120.0	*a* = *b* = *c* = 108.7; α = β = γ = 90.0	*a* = 72.9, *b* = 78.8, *c* = 98.9;α = 91.0, β = 92.9, γ = 94.0
Resolution (Å)	101.05–2.60 (2.76–2.60)	48.62–2.45 (2.54–2.45)	98.78–2.70 (2.80–2.70)
Number of observations	592,150 (68900)	73,015 (12699)	110,187 (14483)
Unique reflections	81,145 (12914)	14,987 (2546)	54,918 (7338)
Completeness (%)	99.8 (99.2)	93.5 (99.9)	90.9 (89.1)
Multiplicity	7.3 (5.3)	4.9 (5.0)	2.0 (2.0)
Mosaicity (°)	0.07	0.18	0.13
$CC_{1/2}$ (%)[a]	99.6 (31.6)	99.9 (46.1)	99.8 (49.7)
R_{sym} (%)[b]	13.7 (77.3)	5.5 (64.6)	4.4 (40.6)
R_{meas} (%)[c]	16.6 (196.5)	6.8 (138.8)	6.3 (70.1)
R_{pim} (%)[d]	5.4 (38.1)	2.8 (31.9)	4.0 (36.9)
<I/σ(I)>	9.49 (0.72)	17.03 (1.20)	11.55 (0.97)
Wilson B-factor ($Å^2$)	49	66	71
V_M ($Å^3$ Da^{-1})	4.1	3.4	2.3
Estimated solvent content (%)	70	64	46
Refinement statistics			
R_{factor} (%)[e]	22.0	22.5	18.9
R_{work} (%)[f]	21.9	22.2	18.1
R_{free} (%)[f]	23.5	25.5	23.6
RMSD for bond lengths (Å)	0.003	0.002	0.004
RMSD for bond angles (°)	0.615	0.553	0.752
Average chain B-factor ($Å^2$)	50, 54, 52, 55	80	71, 71, 76, 82, 76, 76, 76, 79
Number of residues	283	282	283
Ramachandran plot			
Residues in favored regions (%)	97.2	98.0	97.8
Residues in allowed regions (%)	2.8	2.0	2.2
Residues in disallowed regions (%)	0	0	0
PDB code	7QYQ	7QYZ	7QZA

[a] $CC_{1/2}$ = Percentage of correlation between intensities from random half-datasets [33].

[b] $R_{sym} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - <I(hkl)>| / \Sigma_{hkl} \Sigma_i I_i (hkl)$, where $I_i(hkl)$ is the observed intensity and < I(hkl) > is the average intensity of multiple observations from symmetry-related reflections [4].

[c] $R_{meas} = \Sigma_{hkl} [N/(N(hkl)-1)]^{1/2} \Sigma_i |I_i(hkl) - <I(hkl)>| / \Sigma_{hkl} \Sigma_i I_i (hkl)$, where N(hkl) is the data multiplicity, $I_i(hkl)$ is the observed intensity and < I(hkl) > is the average intensity of multiple observations from symmetry-related reflections. It is an indicator of the agreement between symmetry related observations [17].

[d] $R_{p.i.m.} = \Sigma_{hkl} [1/(N(hkl)-1)]^{1/2} \Sigma_i |I_i(hkl) - <I(hkl)>| / \Sigma_{hkl} \Sigma_i I_i (hkl)$, where N(hkl) is the data multiplicity, $I_i(hkl)$ is the observed intensity and < I(hkl) > is the average intensity of multiple observations from symmetry-related reflections. It is an indicator of the precision of the final merged and averaged data set [80].

[e] $R_{factor} = \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$, where F_{obs} and F_{calc} are the amplitudes of the observed and the model calculated structure factors, respectively. It measures the agreement between the experimental X-ray diffraction data and the crystallographic model.

[f] R_{work} refers to the actual working data set used in refinement, while R_{free} refers to a cross-validation set that is not directly used in refinement and is therefore free from refinement bias.

solutions were produced, showing similar probabilities and Rwork values, in space groups $P3_2$ or $P3_221$, with 99 % and 0.395, or 99 % and 0.388 values, respectively. In each model was inserted the corresponding heme B groups fitting them with COOT [19] in the electron density. Twinned refinement of each model using REFMAC5 [45] with intensities data and twin laws H, K, L and -H, -K, -L led to Rwork/Rfree = 0.394/0.395 for space group $P3_2$, while space group $P3_221$ showed Rwork/Rfree = 0.233/0.254, thus solving the previous space group ambiguity. Further three cycles of REFMAC5 twin refinement and manual improvement with COOT led to Rwork/Rfree = 0.227/0.247 and twin fractions 0.181/0.819.The two variants 6E10 and 29E4 structures were determined by molecular replacement using PHASER [41] within the PHENIX suite [3] and the previously solved *Pp*DyP wild-type structure as a search model. For cross-validation purposes, approximately 4% of reflections were randomly excluded from refinement. The TLSMD server (http://skuld.bmsc.washington.edu/~tlsmd) [47] was used to define polypeptide chain regions for translation, libration, and screw refinement of anisotropic atomic displacement parameters (*a.d.p.s*). PHENIX.REFINE [2] was used to proceed with the twinned structure refinement of *Pp*DyP wild-type, as well as with the structure refinements of the two variants. Cycles of iterative structure refinements were followed by inspection of σ_A-weighted 2|Fo|-|Fc| and |Fo|-|Fc| electron density Fourier maps for manual model improvement and completion using COOT [19]. Standard stereochemical dictionary [20] assisted the refinement program, except for inter-atomic distances involving iron sites that were refined without target restraints. The stereochemistry of the refined structures was analysed with MOLPROBITY [13]. All figures were prepared with PyMOL [16]. Analysis of molecular tunnels was performed with CAVER [14]. Three-dimensional superposition of polypeptide chains was performed with MODELLER [79]. Structure factors and atomic coordinates were deposited in the Protein Data Bank (PDB) [8] with accession codes 7QYQ, 7QYZ, and 7QZA for the *Pp*DyP wild-type, 6E10, and 29E4 structures, respectively. The refinement statistics are present in Table 1.

2.8. Computational methods for pH titration

The *Pp*DyP systems were set up using the X-ray structures of wild-type, and 6E10 (E188K/A142V/H125Y) proteins at resting state (RS) and Cpd I states from the PDB as a starting point. First, all titrating residues were renamed to match the CpHMD method nomenclature [38]). Each was assigned an initial protonation state at physiological pH, based on their pK_a. The MD simulations were performed using GROMACS 5.1.5 [1], the GROMOS 54A7 force field [56], and the SPC water model [26]. Each system was minimized (10^4 steps) using the steepest descent method. The initialization procedure consisted of a three-step protocol, starting with a restraint placed on all protein atoms: initially, the system is coupled to a v-rescale thermal bath at 310 K in a 1 ns step [10]; then, a 1 ns step on the NPT ensemble is performed using the Parrinello-Rahman barostat [46,48] until the average pressure is converged to 1 bar with a relaxation time of 5 ps and isothermal compressibility of 4.5×10^5 bar^{-1}; a final 1 ns MD step without atom restraints allowed the proteins to equilibrate to the system temperature and pressure. All non-bonded interactions within a 1.4 nm single cutoff scheme (list updated every five steps) were included. Beyond this cutoff, the Van der Waals interactions were truncated. The electrostatic interactions were treated with the Generalized Reaction-Field method (dielectric constant of 61 and ionic strength of 0.1 M) [72]. Protein bonds were constrained with the P-LINCS methods [27], while the water molecules were used and treated with the SETTLE algorithm [44]. The integrator time-step was 2 fs.

The Poisson-Boltzmann (PB) calculations were performed using Delphi V5.1 [53] using radii calculated using the Lennard-Jones parameters [71] obtained from the GROMOS 54A7 force field. The molecular surface of the protein was defined using a probe with a radius of 1.4 Å. At the same time, the ion exclusion layer was 2.0 Å, and the ionic strength was set to 0.1 M. The dielectric constants used were 2 and 80 for the solute and solvent, respectively. Each PB calculation requires a two-step focusing procedure. Initially, the titrating group was centered in a cubic grid with 81 points with a 1 Å spacing (coarse grid), followed by a focus grid with a smaller spacing between grid points (0.25 Å). The relaxation factors used in the linear and nonlinear iteration processes for the coarse grid were 0.75. A convergence threshold of 0.01 was used for the electrostatic potential. Monte Carlo (MC) calculations were performed with the PETIT program [5] to sample the protonation states of each residue using the free-energy terms obtained from the PB calculations. We performed five replicates for each system at both RS (Fe^{III}) and Cpd I (Fe^{IV} = O) states. Each replicate consisted of 50 ns CpHMD simulations ranging from pH 2.75 to 9.50 with 0.75 pH steps. Each MM/MD step was 20 ps ($\tau_{pr,t}$), and each solvent relaxation step was 0.2 ps (τ_{rlx}). All 61 titratable residues were allowed to titrate, including the termini and the heme propionate groups. To study the electrostatics and evaluate the local charge of the heme pocket, we also created a subset of charged/titrable sites restricted to residues with side chains located within ~15 Å of the heme Fe atom, namely: H/Y125, D126, E131, D132, E134, H162, D176, R181, E/K188, K199, R200, R214, R215, R245, D255, PA, PD, and the heme group. All error values shown were obtained using the standard error of the mean over the five replicates. For the pK_a error values, a modified resampling method based on the leave-one-out strategy (jackknife) was used [64].

2.9. Molecular dynamics and ensemble docking

MD simulations of the wild-type enzyme and the 6E10 variant were set up and ran with Yasara [34] using the AMBER force field [30] and TIP3P [39] water model. The protonation states of titratable residues correspond to those predicted at pH 8 according to the pK_a values observed in the CpHMD simulations. Periodic boundary conditions were used with a solvation cubic box with a 0.9 nm water buffer, and the system was neutralized with Na^+ and Cl^- ions. The system's energy was minimized, with the steepest descendent and simulated annealing, and equilibrated for 5 ns with a time step of 1 fs. In the first half of equilibration, the temperature was gradually increased to 300 K, then kept constant (together with the box volume). The temperature was controlled with the Berendsen thermostat variation [7] as implemented in Yasara. Production of MD runs was carried out at the same temperature and volume, updating the bonded- and non-bonded forces every 5 and 2 fs, respectively. We run 300 ns simulations for each enzyme, taking into account the tetrameric oligomerization state. Thus, conformations of the binding site and the surrounding loops for subsequent docking calculations are obtained for 4×300 ns. The van der Waals forces cutoff was set at 0.8 nm, while long-range electrostatics forces were treated with the Particle Mesh Ewald algorithm [21]. LINCS [28] and SETTLE [44] were adopted to restrain stretching and bending terms involving hydrogen atoms and water molecules in the system. All docking simulations were set up and run with Yasara on 166 equally spaced simulation frames from wild-type and 6E10 variant MD trajectories mimicking pH 8. Docking calculations were computed over each chain of every MD snapshot, giving 664 structures per system. DMP was prepared fully protonated (neutral), which is the most predominant form at pH 8. The prepared DMP was first docked on the complete protein with Autodock Vina [74], finding the heme and propionate sites (essential for catalysis) as the significant binding locations. Docking of ABTS was also carried out. Then, the second round of docking calculations was centered on each of these sites

to enhance the sampling of those regions of the protein. Molecular graphics showing the docking results were created with PyMOL [57].

3. Results and discussion

3.1. PpDyP displays a tetrameric structure

A structure-based sequence alignment revealed that *Pp*DyP is homologous to the available DyPs structures deposited in PDB, with root-mean-square-deviation (RMSD) values ranging from 1.90 to 2.30 Å (defined within a radius cut-off of 1.75 Å) for equivalent Cα atoms (Fig. 1). The monomeric unit shows the typical α + β ferredoxin-like fold, consisting of two similar domains containing four and three-strand antiparallel β-sheets (β1-β4 and β5-β7) flanked on either side by two groups of α-helices (α1-α4 and α5-α9) (Fig. 1 and Fig S1 a, b) [29,67,82]. The crystal structures of 6E10 (E188K, H125Y, and A142V) and 29E4 (E188K and H125Y) variants share the same three-dimensional architecture and are superimposable with RMSDs of 0.40–0.48 Å (Fig. 1).

In the wild-type structure, calculations using the Matthews coefficient estimated the presence of four monomers in the asymmetric unit (*a.u.*), indicating a crystallographic tetramer (Fig. 2a) with a solvent content of 70% [32,40] (Table 1). The *a.u.* in the 6E10 variant structure displays a single monomer, but the crystallographic symmetry also suggests a tetrameric assembly. In the 29E4 variant structure, the *a.u.* consists of eight monomers, four subunits forming one tetramer, and the remaining four by crystallographic symmetry led to tetramer units. Overall, the dimeric interfaces (structurally conserved in other DyPs) exhibit a slightly higher number of interactions (20) than the tetrameric interfaces (17); however, the salt bridges, hydrogen bonds, and hydrophobic interactions between *Pp*DyP interfaces are comparable among the three structures (Fig. 1 and Fig. 2b-e). The dimeric interfaces involve two groups of α-helices (α3 and α5) and one β-strand (β4), while the tetrameric ones include two groups of α-helices (α2 and α9) and two β-strands (β1 and β3) (Fig. 1 and Fig S1 a, b). The total buried surface area upon tetrameric or dimeric oligomerization is similar in wild-type and 29E4 (∼30 and ∼8%, respectively) but significantly lower in the 6E10 variant (∼10 and ∼4%, respectively) [35]; however, in solution the three proteins accumulate in the tetrameric and dimeric states as assessed by size-exclusion chromatography [55]) (Fig S2). Fungal DyPs are exclusively monomers as C-terminal insertion regions seem to prevent their dimerization [82], whereas bacterial DyPs are mainly described as assembled in dimers [11]; and high order oligomerization as a hexameric state was observed in *Bt*DyP from *Bacteroides thetaiotaomicron* (PDB 2GVK) [82] and DyPB from *Rhodococcus jostii* RHA1 (PDB 3QNS) [52].

3.2. Variant 6E10 shows an activity-stability trade-off

The heme access to the solvent, as defined using a 1.4 Å rolling probe, is made through two molecular tunnels (T1 and T2) and one cavity with similar dimensions in wild-type and variants' structures (Fig. 3a and Table S2). T1 is considered the main entrance to H_2O_2, and it is where the conserved distal catalytic residues D132 and R214 are located (Fig. 1 and Fig. 3b). T2 has a similar diameter (∼3.0 Å) to T1 and a slightly lower length (11.5 Å vs. 14.0 Å) (Fig. 3c,d,e). T2 is not present in other DyPs with known structures because the H125 residue of *Pp*DyP is replaced in others DyPs by a conserved arginine, which long side chain appears to occlude this tunnel (Fig. 1). The open cavity that gives access to the solvent-exposed heme propionate has an area of approximately 275 $Å^2$, with a depth of 7.5 Å (Fig. 3a,f), slightly smaller than cavities in other characterized DyPs that show areas of ∼400 $Å^2$ and depths of ∼12 Å [36,49,52,62,76]. The cavity represents an electron-transfer route from the porphyrin radical to the bound substrate in peroxidases [50]. In *Pp*DyP, it is delimited by an extensive network of hydrogen bonds, including four charged residues, E135, R181, E188, and E204 (Fig. 3f).

The heme pocket of *Pp*DyP is delimited by three conserved loop regions, L1 (120–139), L2 (179–205), and L3 (206–226) (Fig. 3a,b). Loops L1 and L3 are located on the distal side of the heme and include the conserved catalytic residues D132 and R214, and L2 is at the heme proximal side and contains the conserved H197 heme ligand (Fig. 1 and Fig. 3a). The mutated residues H125Y and E188K present in variants 6E10 and 29E4 variants, are in L1

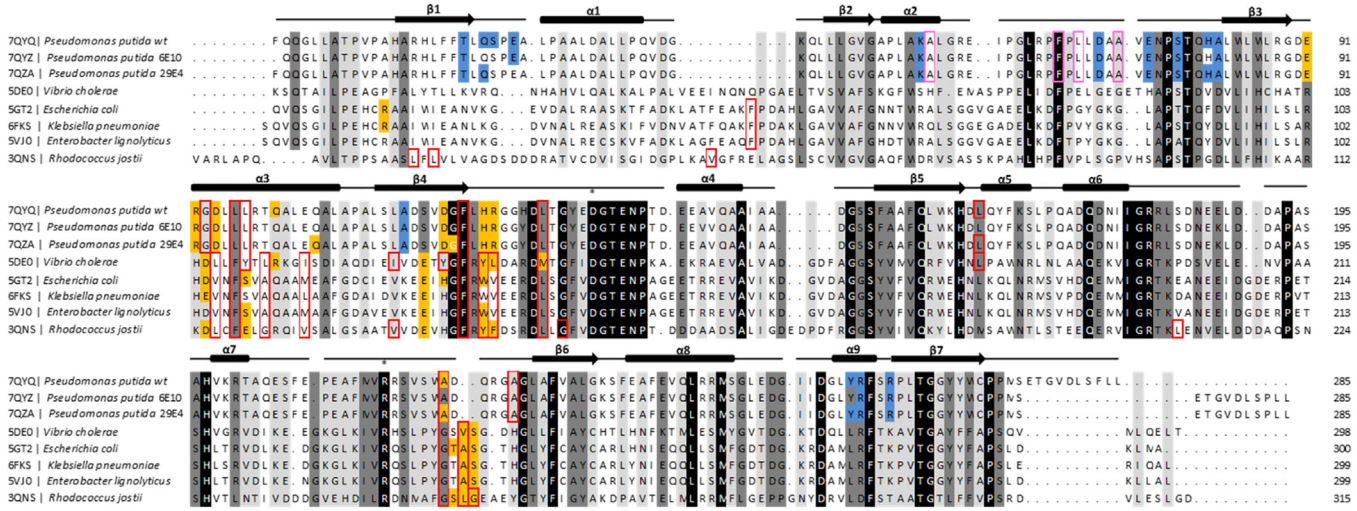

Fig. 1. Amino acid sequence alignment based on the known X-ray crystal structures of DyPs belonging to class P. Three-dimensional multi-body superposition of polypeptide chains was performed with MODELLER [79]. The *Pp*DyP secondary structure is shown above the alignment. The dimeric polar and apolar interactions are shown as yellow and red boxes, respectively. The tetrameric polar and apolar interactions are shown as blue and pink boxes, respectively. Distal catalytic residues are highlighted with *. Strictly conserved amino acids are represented as black boxes, whereas dark and light grey boxes represent the most and less conserved residues among the selected sequences. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

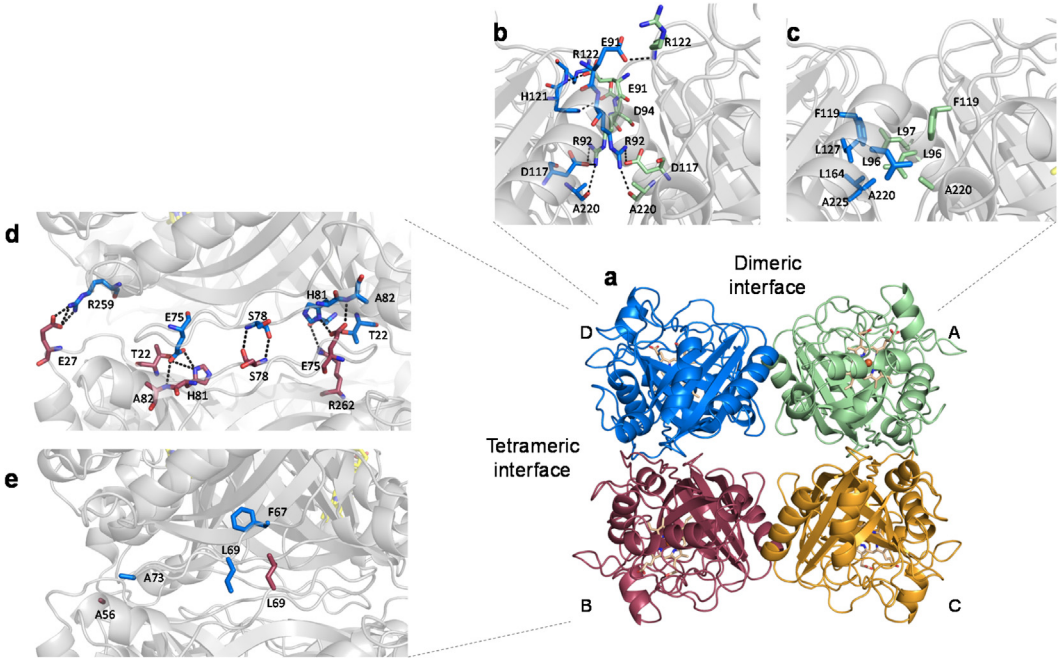

Fig. 2. Overall structure and heme access pathways in *Pp*DyP. **(a)** Dimeric and tetrameric interface interactions are involved in *Pp*DyP oligomerization. Analysis of the *Pp*DyP interfaces: the subunits AD/BC and BD/AC are the dimeric and tetrameric interfaces, respectively. Polar **(b)** and apolar **(c)** interactions in the canonical dimer, respectively. Polar **(d)** and apolar **(e)** interactions of the tetrameric interface, respectively. The amino acid residues are shown as sticks with carbon atoms in the same color as the corresponding chain. The oxygen and nitrogen atoms are shown in red and blue, respectively. The heme group is shown as sticks with carbon, nitrogen, and oxygen atoms in light pink, blue and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

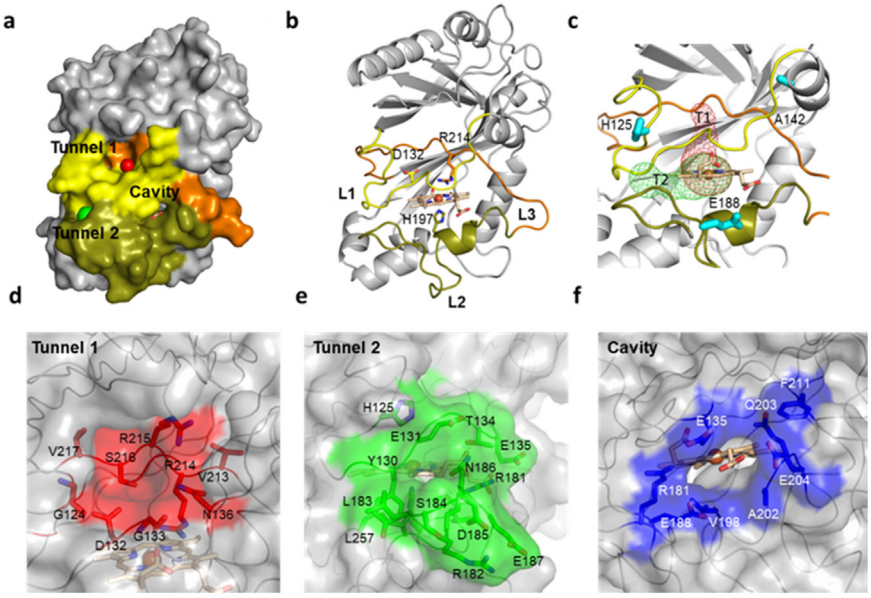

Fig. 3. Molecular pathways that give access to the heme cofactor in *Pp*DyP: tunnels 1 (T1) and 2 (T2), and cavity, surrounded by loops L1 (120–139, yellow), L2 (179–205, green), and L3 (206–226, orange) represented as solvent accessible surface area (**a**) and cartoon (**b**). The heme group is shown as sticks with carbon, nitrogen, and oxygen in light pink, blue and red, respectively. The iron atom is shown as an orange sphere. The distal catalytic residues, D132 and R214, are shown as sticks. The heme proximal ligand H197 is shown as sticks with carbon atoms colored in green. The oxygen and nitrogen atoms are shown in red and blue, respectively. In **(c),** T1 and T2 were calculated with CAVER program and are shown in red and green mesh, respectively. The mutated residues (H125Y, A142V, and E188K in 6E10 and H125Y and E188K in 29E4) are shown as sticks and colored in cyan. **(d-f)** Solvent accessible surface area representing residues limiting the T1 (**d**), T2 (**e**), and cavity (**f**) are colored in red, green, and blue, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and L2, respectively, whereas mutated residue A142V of 6E10 is located in a sort of hinge region in a small α-helix (α4) adjacent to the L1 area (Fig. 1 and Fig. 3c).

The comparison of normalized B-factors [70] in the three structures revealed a significantly different profile pattern (Fig. 4). Variant 6E10 shows the highest overall B-factor value (~80 Å^2),

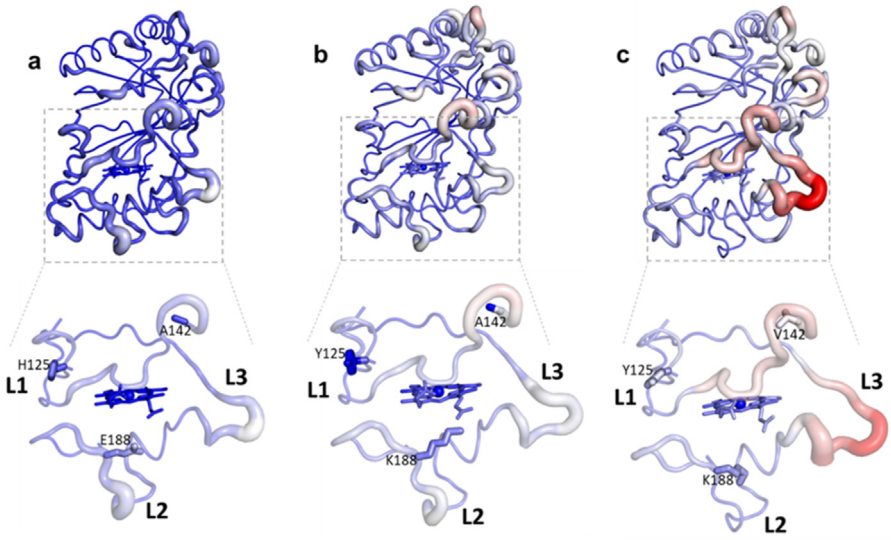

Fig. 4. Cartoon representations of the main chain of **(a)** wild-type, **(b)** 29E4, and **(c)** 6E10 variant structures with thickness proportional to normalized B-factor values color-coded as a blue-white-red color ramp with blue indicating the most negative value (less flexibility) and red the most positive value (more flexibility). The zoomed-view represents the loop regions L1 (120–139), L2 (179–205), and L3 (206–226), limiting the heme cofactor. The heme and mutated residues are shown as sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

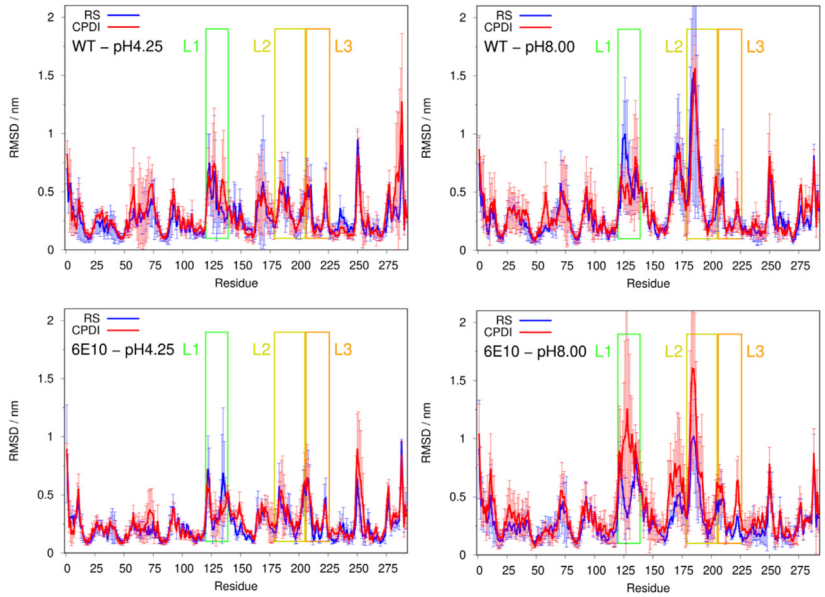

Fig. 5. Root mean square deviation (RMSD) per residue in wild-type *Pp*DyP and 6E10 variant at pH values 4.25 and 8.0. The RMSD values were calculated relative to the initial X-ray structure for resting-state (RS) and compound I (Cpd I). The loops are colored differently for visual clarity: L1 (green), L2 (yellow), and L3 (orange). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

followed by 29E4 (~71 Å^2) and wild-type (~50 Å^2) (Table 1). The insertion of mutations, in particular of A142V, triggered structural changes leading to fluctuations in the solvent exposure of residues in the loops regions (Fig. 4a-c and Table S3). This view is supported by molecular dynamics simulations, where a higher displacement is observed in the loop regions (Fig. 5). A tendency to higher RMSD values was observed along with pH increase and enzyme activation, i.e., after the H_2O_2-dependent reactions that change Fe(III) in RS to Fe(IV) in Cpd I during the catalytic mechanism. Notably, in the X-ray structures, the entrance to the T1 is more prominent in 6E10, and the conformations of its distal catalytic residues show slightly longer D132^{OD1}-R214^{NH1} interatomic distances (Fig. S3). The cavity entrance area in 6E10 is also enlarged, probably due to the higher flexibility of delimiting residues (Fig. S4). The extra plasticity of the molecular access pathways to the heme can positively impact enzyme catalysis facilitating substrate binding and product release and contributing to the highest measured catalytic efficiency (k_{cat}/K_m) of 6E10 for the substrates tested, particularly at the neutral to the alkaline range (Table 2) [6,9].

To explore this further, the binding of the lignin-related phenolic DMP substrate - for which a significantly higher activity was measured in the 6E10 variant at pH 8 (Table 2) - was investigated with ensemble docking calculations using snapshots from the Cpd I MD trajectories. The docking results confirm that DMP can bind

Table 2
Apparent steady-state kinetic parameters for the oxidation of ABTS and DMP of wild-type and variants.

		ABTS			DMP		
	pH	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild-type	4.3	28 ± 3	300 ± 20	(5.5 ± 0.1) × 10^4	0.10 ± 0.01	80 ± 20	(1.3 ± 0.2) × 10^3
	7-8[a]	6 ± 1	nd	nd	0.08 ± 0.01	60 ± 11	(1.2 ± 0.2) × 10^3
29E4	4.3	60 ± 5	810 ± 190	(7.4 ± 0.9) × 10^4	0.10 ± 0.01	40 ± 10	(2.5 ± 0.2) × 10^3
	7–8 [a]	4.1 ± 0.3	nd	nd	0.07 ± 0.02	43 ± 9	(1.6 ± 0.5) × 10^3
6E10	4.3	12 ± 3	44 ± 5	(2.7 ± 0.3) × 10^5	0.0035 ± 0.0001	nd	nd
	7-8[a]	20 ± 2	100 ± 30	(12.0 ± 0.1) × 10^4	3.5 ± 0.1	80 ± 10	(4.3 ± 0.4) × 10^4

[a] reactions using ABTS were performed at pH 4.3 and 7, and reactions using DMP at pH 4.3 and pH 8; nd-not determined

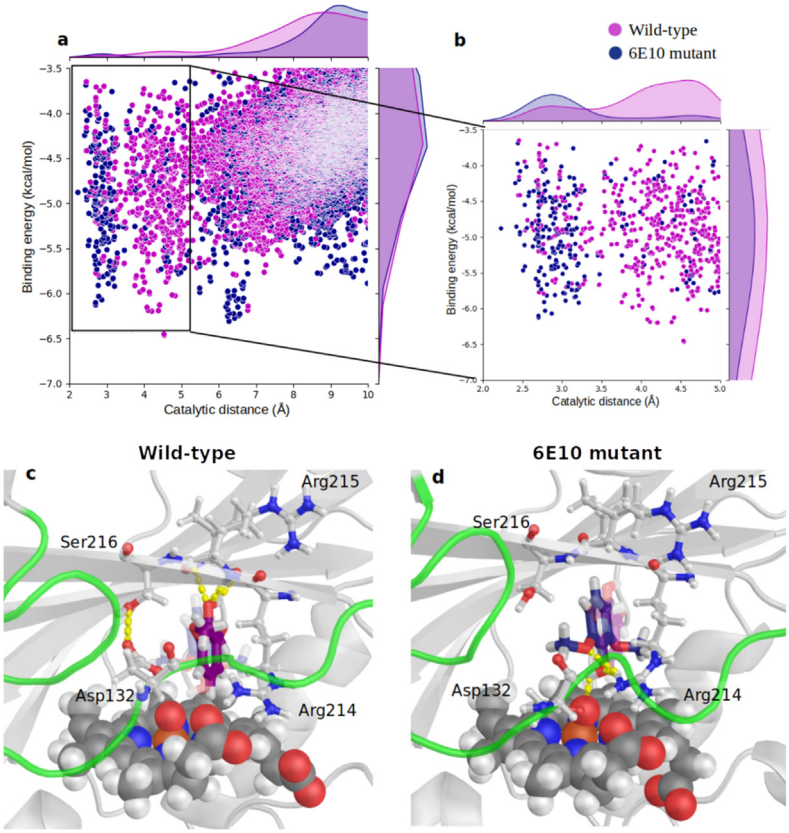

Fig. 6. Docking of DMP into wild-type (magenta) and 6E10 (dark blue) heme sites. (a) Binding energy vs. catalytic distances (distance between DMP hydroxyl oxygen atom and heme-bound oxygen). Each point represents a simulated conformation of the enzyme-substrate complex. **(b)** Binding energy vs. catalytic distances plot close-up showing docking poses with catalytic distances below 5 Å. **(c,d)** Molecular representation of the lowest energy enzyme-substrate complex for DMP in the heme-binding sites of the enzymes. The L1 loop is shown in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

into the heme pocket. Moreover, the average volume of this pocket along the MDs is significantly larger for 6E10 than for wild-type (408.2 $Å^3$ vs. 326.8 $Å^3$), in part due to the presence of a hydrogen bond between Asp132 and Ser216. These differences may explain the appearance of a well-defined energy minimum for DMP in a catalytic position (catalytic distance ~2.7 Å) at the heme site in the 6E10 variant (Fig. 6a,b). In the wild-type enzyme, DMP is stabilized at a farther position (catalytic distance ~4 Å) and tends to bind closer to the propionate (Fig. 6c,d). While catalysis can also occur at this longer distance, an optimal substrate positioning, like in the 6E10 variant, can have positively impacted the catalytic machinery and led to the higher activity observed experimentally. We hypothesize that entrance to the heme pocket is facilitated in the 6E10 variant thanks to the enlarged entrance cavity and increased loop flexibility (especially for L1 in Cpd I, Fig. 5) commented above. This idea is also supported by the global docking results, where DMP binding is observed all along T1, a tunnel that connects the protein surface with the heme pocket and shows an enlarged entrance in the case of 6E10 (Fig. S3c,f and Fig. S5).

In contrast, variant 6E10 displayed the lowest overall stability compared with 29E4 and wild-type enzymes as assessed by fluorescence spectroscopy that monitored its chemical and thermal unfolding (Table 3, Fig. S6 and S7). Thermal unfolding profiles were measured to calculate the mid-point thermal temperatures, i.e. the melting temperatures (T_m) (Fig S7a). However, protein aggregation assessed by static light scattering reveals the onset of protein aggregation (T_{agg}) at lower temperatures than T_m values invalidating the measured thermal unfolding mid-points (Table 3 and Fig S7b).

Nevertheless, variant 6E10 displayed the lowest T_{agg} compared with 29E4 and wild-type enzymes. These results indicate the

Table 3
Thermodynamic stability of the tertiary structure of *Pp*DyP wild-type and variants 29E4 and 6E10 as assessed by fluorescence spectroscopy.

	ΔG^{water} (kJ.mol^{-1})	*m* (kJ.mol^{-1}.M^{-1})	$[GdnHCl]_{50\%}$ (M)	T_{agg} (°C)
Wild-type	25.9 ± 2.1	−18.0 ± 0.8	1.4 ± 0.1	60.9 ± 0.2
29E4	10.5 ± 0.8	−8.4 ± 0.8	1.23 ± 0.02	58 ± 3
6E10	9.6 ± 0.8	−10.4 ± 0.4	0.9 ± 0.1	53 ± 3

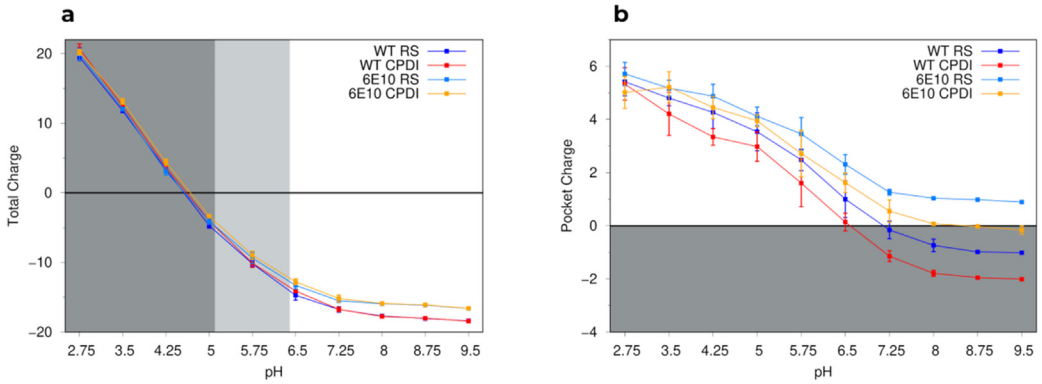

Fig. 7. **(a)** Total charge titration curve for Resting-State (RS) and Compound I (CPDI) systems for wild-type and 6E10 variants. Each colored region indicates a different average ΔCharge between the 6E10 variant and wild-type. These region limits are defined by the pK_a values of the mutated residues, E188 (∼5.1) and H125 (∼6.4). **(b)** Partial titration curve including only residues with side chains located within ∼15 Å of the heme Fe atom. The gray-shaded region corresponds to the negative charges that should generate an overall negative electrostatic potential.

occurrence of a stability-activity trade-off at the level of A142V residue, as variant 6E10 displays 2- and 50-fold increased catalytic efficiency for ABTS and the lignin-phenolic substrate DMP, respectively, as compared to variant 29E4 and the wild-type (Table 2). These observations corroborate previous data obtained with an engineered variant of *B. subtilis Bs*DyP that displays higher activity but reduced stability upon introducing mutations in the loops surrounding the heme pocket [54].

3.3. Variations in the electrostatic network in the heme vicinity affect the pH dependence of 6E10

The molecular details of the electrostatic network were investigated considering that two out of the three mutations in the 6E10 variant (H125Y and E188K) are pH-titratable residues located near the heme pocket (Fig. 3c). This is an electrostatically-rich environment that changes conformation and suffers charge fluctuations during the catalytic mechanism. The pH titration curves in the RS and Cpd I, in both wild-type and 6E10, show a total charge at pH 2.75 of ∼ +20 and an almost linear decrease in the acidic region between pH 3 and 5, where most Asp and Glu residues are titrating (Fig. 7a). The titration behavior in this pH range is comparable in all analyzed systems and is in agreement with the estimated isoelectric point values (4.6–4.7). At pH values higher than ∼5.1 and, in particular, higher than ∼6.4, the differences observed can be attributed mainly to H125Y and E188K mutations since when these residues were excluded, the titration curves almost overlap (Fig. S8). At neutral pH, the titration curves for RS and Cpd I show nearly double the negatively charged residues (Asp, Glu, and propionates) as compared with positively charged ones (Lys and Arg) in both wild-type and 6E10 (Fig. 7a). However, when only residues that have their side-chain within a ∼15 Å distance from the heme are included, a more clear impact of mutations H125Y and E188K in the electrostatic potential of the heme cavity is observed: the negative charge of the wild-type is mitigated or even inverted in 6E10 systems at pH 8.0 (Fig. 7b).

Notably, the pK_a values of pH-sensitive residues surrounding the heme group and located in the nearby loop regions are comparable in both enzymes (Table 4), except for the distal D132 catalytic residue and H162 near the T2 and buried under L1 (Fig. 8).

Table 4
pK_a values of selected key residues near the heme pocket and the surrounding loops of wild-type and 6E10 variant proteins. RS: resting state; Cpd I: compound I active state; PA/PD: propionate A/D.

wild-type			**6E10**		
Residue	**RS**	**Cpd I**	**Residue**	**RS**	**Cpd I**
H125	6.3 ± 0.2	6.5 ± 0.2	**Y125**	>9.5	>9.5
D126	<2.8	<2.8	**D126**	<2.8	<2.8
D132	<2.8	<2.8	**D132**	4.1 ± 0.8	2.8 ± 1.1
E135	4.2 ± 0.4	4.3 ± 0.2	**E135**	3.8 ± 0.1	3.8 ± 0.2
H162	<2.8	3.6 ± 0.5	**H162**	<2.8	<2.8
E188	4.9 ± 0.1	5.3 ± 0.1	**K188**	>9.5	>9.5
E204	4.0 ± 0.1	4.0 ± 0.1	**E204**	3.6 ± 0.3	3.7 ± 0.1
D255	<2.8	<2.8	**D255**	<2.8	< 2.8
Y258	>9.5	>9.5	**Y258**	>9.5	>9.5
PA	6.7 ± 0.2	6.6 ± 0.2	**PA**	6.2 ± 0.2	6.6 ± 0.3
PD	6.6 ± 0.1	6.9 ± 0.2	**PD**	5.8 ± 0.1	6.7 ± 0.1

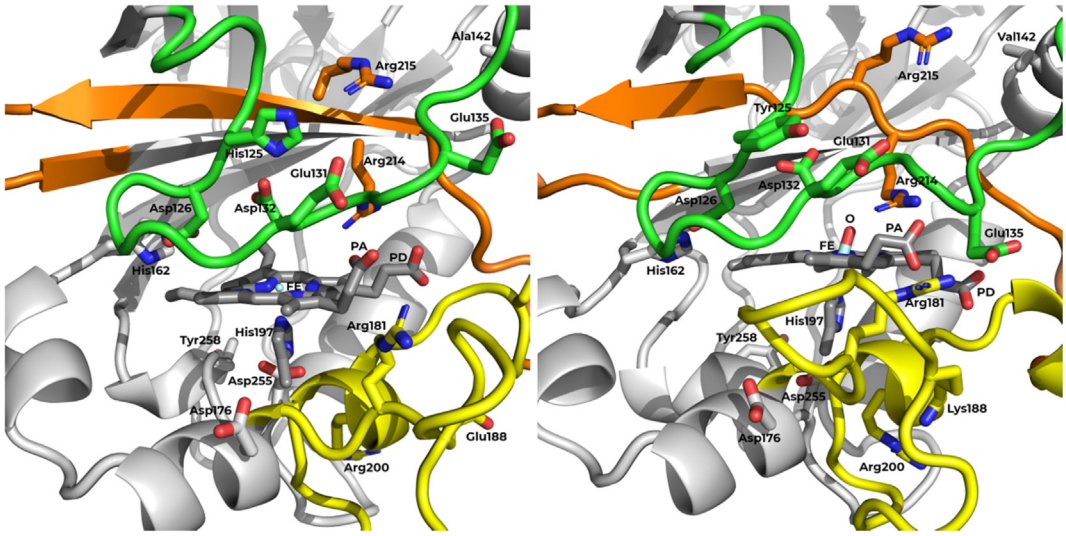

Fig. 8. Structural highlights of the heme group region in wild-type *Pp*DyP **(left)** and 6E10 variant **(right)**. The wild-type structure is shown in its resting state (RS), and the 6E10 variant is shown in its activated form (Cpd I). In labeled sticks are shown several critical residues, including residue A142, mutated from Ala to Val in the 6E10 variant. The green, yellow, and orange regions correspond to loops L1, L2, and L3, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

D132 often establishes interactions with nearby cationic residues expected to lower its pK_a value, including the distal R214 catalytic residue and H125. The replacement of H125 residue with the neutral tyrosine in the 6E10 variant, results in a less positive environment and increased pK_a values. Furthermore, the slightly longer interatomic distances (5.1 Å) between $D132^{OD1}$ and $R214^{NH1}$ in the 6E10 X-ray structure when compared to wild-type (4.1 Å) (Fig. S3) might also contribute to the protonation of D132 [25]. The limited access to the solvent of H162 stabilizes its neutral state, leading to very low pK_a values (Table 4). Still, during activation at low pH, there is an increase in the population where the H162 side chain is facing away from the heme (Fig. S9), resulting in increased solvation and a shift in the pK_a value towards their water reference (~6.5).

The negative electrostatic potential of the wild-type heme pocket at pH 8.0 can impair the access and oxidation of, *e.g.*, ABTS, a substrate with a formal charge of −2. In contrast, this barrier would be reverted by the positive electrostatic potential in the 6E10 variant. Furthermore, docking calculations of ABTS to the enzymes suggest that, contrary to what is observed for DMP, the larger substrate cannot fully access the heme cavity in wild-type and in the 6E10 variant (Fig. S10). Instead, ABTS reaches the heme by interacting from the propionates side, indicating an important role of the local electrostatics in this initial binding step. The more positive pocket of variant 6E10 at neutral to alkaline values, mentioned above is also in good agreement with its redox transition ($E^{0'}_{Fe2+/Fe3+}$ = −60 mV) that is >0.2 V higher than the respective value for wild-type *Pp*DyP ($E^{0'}_{Fe2+/Fe3+}$ = −260 mV) [9]. This is in excellent agreement with the deviation of the optimum pH from acidic in wild-type (pH_{op} ~4) to more neutral and alkaline values in the 6E10 variant for both ABTS (pH_{op} = 7) and DMP (pH_{op} = 8) oxidation (Table 2) [9]. The distal Asp residue is generally assumed to be the key in defining the optimum acidic pH observed in DyPs, even if recognizing a putative role of other acidic residues around the heme center [49,62,63]. In variant 6E10, D132 would be protonated at pH around 4, which may prevent activation by H_2O_2 at this pH in contrast to the wild-type. Furthermore, our results point to an essential role of other charged residues in the heme pocket, in line with the observation that the hydrogen bond network between carboxylate and His residues in the heme vicinity is pivotal in the enzyme pH dependence of *Vc*DyP [75]. Overall, the data obtained indicate that changes in loops' flexibility and conformation have a crucial impact on the electrostatics of the heme microenvironment, which in turn influences the enzyme's optimal pH for activity.

4. Concluding remarks

In this work, we solved the crystal structure of the *Pp*DyP wild-type and engineered variants 29E4 and 6E10. Both variants harbor pH-titrable mutations H125Y and E188K, whereas mutation A142V is only present in variant 6E10, which shifts the optimal pH from pH ~4 to pH ~8. Structural analysis, molecular dynamics simulations, and substrate docking show that insertion of A142V leads to changes in the conformational flexibility of loop regions of *Pp*DyP that, despite resulting in the lower overall stability of the enzyme, have facilitated the access of substrates to the heme cofactor promoting increased catalytic rates. Furthermore, analysis of 6E10 points to alterations in the electrostatic and hydrogen bond network in the heme pocket: a more positive electrostatic environment emerged, and a deviation of the optimal pH from acidic in wild-type enzyme to more neutral and alkaline in this variant. Understanding the importance of local flexibility around the DyPs heme pocket in modulating enzymes' activity, specificity, and stability is critical for advancing fundamental biochemical insights and allowing for the more efficient design of novel enzymes with tailored physicochemical properties. It is expected that DyPs will be an essential tool to achieve selective oxidation of lignin-derived phenolics and afford high-value products in the lignocellulose biorefinery realm.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank Bernardo Rodrigues and Joana Pedroso for their help in preliminary crystallographic experiments, and the staff of beamlines ID23-1 from ESRF, and XALOC from ALBA for the support during the synchrotron data collections.

Funding

This work was supported by the Fundação para a Ciência e Tecnologia, Portugal, grants CEECIND/02300/2017, PTDC/BBBEBB/0122/2014, PTDC/BII-BBF/29564/2017, MOSTMICRO-ITQB (UIDB/04612/2020 and UIDP/04612/2020), LS4FUTURE Associated Laboratory (LA/P/0087/2020), BioISI (UIDB/04046/2020 and UIDP/04046/2020), and projects UIDB/04326/2020, UIDP/04326/2020, LA/P/0101/2020, and the operational programs CRESC Algarve 2020 and COMPETE through project EMBRC.PT ALG-01-0145-FEDER-022121. LA thanks the Spanish Ministerio de Ciencia Innovación y Universidades (PGC2018-098592-B-100 and PID2021-126897NB-100) and Universitat Autònoma de Barcelona Talent Program. DS and TFDS acknowledge FCT Ph.D. Fellowships SFRH/BD/132702/2017 and SFRH/BD/140886/2018.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.07.032.

References

[1] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1–2:19–25.
[2] Adams P, Afonine P, Echols N, Headd J, Grosse-Kunstleve R, Moriarty N. New tools for structure refinement in Phenix. Acta Crystallogr. A 2010;66:S15–S.
[3] Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D 2010;66:213–21.
[4] Arndt UW, Crowther RA, Mallett JF. A computer-linked cathode-ray tube microdensitometer for x-ray crystallography. J Sci Instrum 1968;1:510–6.
[5] Baptista AM, Soares CM. Some theoretical and computational aspects of the inclusion of proton isomerism in the protonation equilibrium of proteins. J Phys Chem. B 2001;105:293–309.
[6] Barbosa C, Silveira CM, Silva D, Brissos V, Hildebrandt P, Martins LO, et al. Immobilized dye-decolorizing peroxidase (DyP) and directed evolution variants for hydrogen peroxide biosensing. Bios Bioelectronics 2020;153:112055.
[7] Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. Molecular-dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–90.
[8] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res 2000;28:235–42.
[9] Brissos V, Tavares D, Sousa AC, Robalo MP, Martins LO. Engineering a bacterial DyP-type peroxidase for enhanced oxidation of lignin-related phenolics at alkaline pH. *ACS Catal* 2017;7:3454–65.
[10] Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys. 2007;126(1).
[11] Chen C, Li T. Bacterial dye-decolorizing peroxidases: biochemical properties and biotechnological opportunities. Phys Sci Rev 2016;1(9).
[12] Chen C, Shrestha R, Jia K, Gao PF, Geisbrecht BV, Bossmann SH, et al. Characterization of Dye-decolorizing peroxidase (DyP) from *Thermomonospora curvata* reveals unique catalytic properties of A-type DyPs. J Biol Chem 2015;290:23447–63.
[13] Chen VB, Arendall 3rd WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D 2010;66:12–21.
[14] Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput Biol. 2012;8(10).
[15] Colpa DI, Fraaije MW, van Bloois E. DyP-type peroxidases: a promising and versatile class of enzymes. J Ind Microbiol Biotechnol 2014;41:1–7.
[16] DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. Curr. Opin. Struct. Biol. 2002;12:14–20.
[17] Diederichs K, Karplus PA. Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat Struct Biol 1997;4:269–75.
[18] Durao P, Bento I, Fernandes AT, Melo EP, Lindley PF, Martins LO. Perturbations of the T1 copper site in the CotA laccase from Bacillus subtilis: structural, biochemical, enzymatic and stability studies. J Biol Inorg Chem 2006;11:514–26.
[19] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr. D 2010;66:486–501.
[20] Engh RA, Huber R. Accurate bond and angle parameters for X-ray protein-structure refinement. Acta Crystallog A 1991;47:392–400.
[21] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J. Chem. Phys. 1995;103:8577–93.
[22] Evans PR. Scaling and assessment of data quality. Acta Crystallog D 2006;62:72–82.
[23] Fernandes AT, Martins LO, Melo EP. The hyperthermophilic nature of the metallo-oxidase from *Aquifex aeolicus*. Biochim Biophys Acta 2009;1794:75–83.
[24] Gorrec F. The MORPHEUS protein crystallization screen. J. Appl. Crystallogr. 2009;42:1035–42.
[25] Harris TK, Turner GJ. Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. IUBMB Life 2002;53:85–98.
[26] Hermans J, Berendsen HJC, Vangunsteren WF, Postma JPM. A consistent empirical potential for water-protein interactions. Biopolymers 1984;23:1513–8.
[27] Hess B. P-LINCS: A parallel linear constraint solver for molecular simulation. J. Chem. Theory Comput 2008;4:116–22.
[28] Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem 1998;18:1463–72.
[29] Hofbauer S, Pfanzagl V, Michlits H, Schmidt D, Obinger C, Furtmuller PG. Understanding molecular enzymology of porphyrin-binding α + β barrel proteins - One fold, multiple functions. BBA-Proteins Proteom 2020;1869:1400536.
[30] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–25.
[31] Kabsch W. Xds. *Acta Crystallogr D* 2010;66:125–32.
[32] Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Prot. Sci. 2003;12:1865–71.
[33] Karplus PA, Diederichs K. Linking crystallographic model and data quality. Science 2012;336:1030–3.
[34] Krieger E, Vriend G. New ways to boost molecular dynamics simulations. J Comput Chem 2015;36:996–1007.
[35] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 2007;372:774–97.
[36] Liu X, Yuan Z, Wang J, Cui Y, Liu S, Ma Y, et al. Crystal structure and biochemical features of dye-decolorizing peroxidase YfeX from *Escherichia coli* O157 Asp(143) and Arg(232) play divergent roles toward different substrates. Biochem Biophys Res Commun 2017;484:40–4.
[37] Lucic M, Wilson MT, Svistunenko DA, Owen RL, Hough MA, Worrall JAR. Aspartate or arginine? Validated redox state X-ray structures elucidate mechanistic subtleties of Fe(IV) = O formation in bacterial dye-decolorizing peroxidases. J Biol Inorg Chem 2021;26:743–61.
[38] Machuqueiro M, Baptista AM. Constant-pH molecular dynamics with ionic strength effects: Protonation-conformation coupling in decalysine. J. Phys. Chem. Chem. B 2006;110:2927–33.
[39] Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J. Phys. Chem. B 2001;105:9954–60.
[40] Matthews BW. Solvent content of protein crystals. J. Mol. Biol. 1968;33:491–7.
[41] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40(Pt 4):658–74.
[42] Mendes S, Brissos V, Gabriel A, Catarino T, Turner DL, Todorovic S, et al. An integrated view of redox and catalytic properties of B-type PpDyP from *Pseudomonas putida* MET94 and its distal variants. Arch Biochem Biophys 2015;574:99–107.
[43] Mendes S, Catarino T, Silveira C, Todorovic S, Martins LO. Catalytic mechanism of BsDyP an A-type dye-decolourising peroxidase: neither aspartate nor arginine is individually essential for peroxidase activity. Catal Sci Technol 2015;5:5196–207.
[44] Miyamoto S, Kollman PA. Settle – an analytical version of the shake and rattle algorithm for rigid water models. J. Comput. Chem. 1992;13:952–62.
[45] Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D 2011;67:355–67.
[46] Nosé S, Klein ML. Constant pressure molecular dynamics for molecular systems. Mol Phys 1983;50:1055–76.
[47] Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D 2006;62:439–50.
[48] Parrinello M, Rahman A. Polymorphic transitions in single-crystals – a new molecular-dynamics method. J Appl Phys 1981;52:7182–90.
[49] Pfanzagl V, Nys K, Bellei M, Michlits H, Mlynek G, Battistuzzi G, et al. Roles of distal aspartate and arginine of B-class dye-decolorizing peroxidase in heterolytic hydrogen peroxide cleavage. J. Biol. Chem. 2018;293:14823–38.
[50] Poulos TL. Heme enzyme structure and function. Chem. Rev. 2014;114:3919–62.
[51] Poulos TL. Thirty years of heme peroxidase structural biology. Arch Biochem Biophys 2010;500:3–12.
[52] Roberts JN, Singh R, Grigg JC, Murphy ME, Bugg TD, Eltis LD. Characterization of dye-decolorizing peroxidases from *Rhodococcus jostii* RHA1. Biochemistry 2011;50:5108–19.
[53] Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J Comput Chem 2002;23:128–37.
[54] Rodrigues CF, Borges PT, Scocozza MF, Silva D, Taborda A, Brissos V, et al. Loops around the heme pocket have a critical role in the function and stability of BsDyP from Bacillus subtilis. Int J Mol Sci 2021;22:10862.
[55] Santos A, Mendes S, Brissos V, Martins LO. New dye-decolorizing peroxidases from Bacillus subtilis and Pseudomonas putida MET94: towards biotechnological applications. Appl Microbiol Biotechnol 2014;98:2053–65.

[56] Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark, A.E., van Gunsteren WF. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J Biophys Lett*, **40**; 2011, 843-856.
[57] Schrodinger LLC. The PyMOL Molecular Graphics System; 2010.
[58] Scocozza MF, Martins LO, Murgida DH. Direct electrochemical generation of catalytically competent oxyferryl species of classes I and P dye decolorizing peroxidases. Int J Mol Sci 2021;22(22).
[59] Sezer M, Genebra T, Mendes S, Martins LO, Todorovic S. A DyP-type peroxidase at a bio-compatible interface: structural and mechanistic insights. Soft Matter 2012;8:10314–21.
[60] Sezer M, Santos A, Kielb P, Pinto T, Martins LO, Todorovic S. Distinct structural and redox properties of heme active in bacterial DyP-type peroxidases from two subfamilies: resonance Raman and electrochemistry study. Biochemistry 2013;52:3074–84.
[61] Sheldrick GM. XPREP Version 6.10. Bruker AXS Inc., Madison, Wisconsin, USA; 2000.
[62] Shrestha R, Huang GC, Meekins DA, Geisbrecht BV, Li P. Mechanistic insights into dye-decolorizing peroxidase revealed by solvent isotope and viscosity effects. ACS Catal 2017;7:6352–64.
[63] Shrestha R, Jia KM, Khadka S, Eltis LD, Li P. Mechanistic insights into DyPB from Rhodococcus jostii RHA1 via kinetic characterization. ACS Catal 2021;11:5486–95.
[64] Silva TFD, Vila-Vicosa D, Machuqueiro M. Improved protocol to tackle the pH effects on membrane-inserting peptides. J. Chem. Theory Comput. 2021;17:3830–40.
[65] Singh R, Eltis LD. The multihued palette of dye-decolorizing peroxidases. Arch Biochem Biophys 2015;574:56–65.
[66] Sugano Y. DyP-type peroxidases comprise a novel heme peroxidase family. Cell Mol Life Sci 2009;66:1387–403.
[67] Sugano Y, Muramatsu R, Ichiyanagi A, Sato T, Shoda M. DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family: ASP171 replaces the distal histidine of classical peroxidases. J Biol Chem 2007;282:36652–8.
[68] Sugano Y, Yoshida T. DyP-type peroxidases: recent advances and perspectives. Int J Mol Sci 2021;22(11).
[69] Sun Z, Fridrich B, de Santi A, Elangovan S, Barta K. Bright side of lignin depolymerization: toward new platform chemicals. Chem Rev 2018;118:614–78.
[70] Sun ZT, Liu Q, Qu G, Feng Y, Reetz MT. Utility of B-factors in protein science: interpreting rigidity, flexibility, and internal motion and engineering thermostability. Chem Rev 2019;119(3):1626–65.
[71] Teixeira VH, Cunha CA, Machuqueiro M, Oliveira ASF, Victor BL, Soares CM, et al. On the use of different dielectric constants for computing individual and pairwise terms in Poisson-Boltzmann studies of protein ionization equilibrium. J Phys Chem B 2005;109:14691–706.
[72] Tironi IG, Sperb R, Smith PE, Vangunsteren WF. A generalized reaction field method for molecular-dynamics simulations. J Chem Phys 1995;102:5451–9.
[73] Todorovic S, Hildebrandt P, Martins LO. Surface enhanced resonance Raman detection of a catalytic intermediate of DyP-type peroxidase. Phys Chem Chem Phys 2015;17(18):11954–7.
[74] Trott O, Olson AJ. Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 2010;31(2):455–61.
[75] Uchida T, Omura I, Umetsu S, Ishimori K. Radical transfer but not heme distal residues is essential for pH dependence of dye-decolorizing activity of peroxidase from Vibrio cholerae. J Inorg Biochem 2021;219.
[76] Uchida T, Sasaki M, Tanaka Y, Ishimorit K. A dye-decolorizing peroxidase from Vibrio cholerae. Biochemistry 2015;54(43):6610–21.
[77] Vagin A, Lebedev A. MoRDa, an automatic molecular replacement pipeline. Acta Crystallogr a-Foundation Adv 2015;71:S19–S.
[78] Van den Bosch S, Koelewijn SF, Renders T, Van den Bossche G, Vangeel T, Schutyser W, et al. Catalytic strategies towards lignin-derived chemicals. Top Curr Chem (Cham) 2018;376(5):36.
[79] Webb B, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol 2014;1137:1–15.
[80] Weiss M. Global indicators of X-ray data quality. J Appl Crystallogr 2001;34 (2):130–5.
[81] Yoshida T, Sugano Y. A structural and functional perspective of DyP-type peroxidase family. Arch Biochem Biophys 2015;574:49–55.
[82] Zubieta C, Krishna SS, Kapoor M, Kozbial P, McMullan D, Axelrod HL, et al. Crystal structures of two novel dye-decolorizing peroxidases reveal a beta-barrel fold with a conserved heme-binding motif. Proteins 2007;69:223–33.
[83] Zwart PH, Grosse-Kunstleve RW, Adams PD. XTRIAGE: Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. *CCP4 newsletter 43*; 2005.