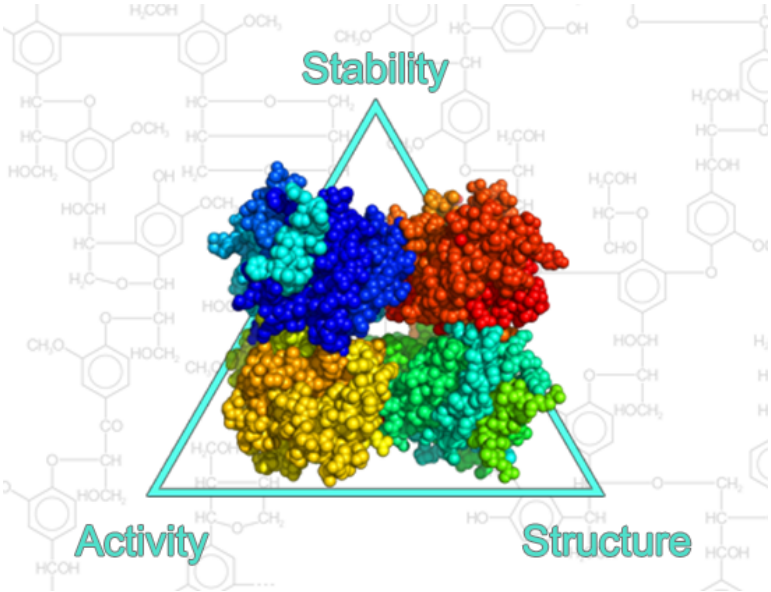
2022

A closer look at a potential biocatalyst:

Unravelling the catalytic, stability and structural features of *Pp*DyP, a DyP-type Peroxidase

Diogo André Ribeiro Silva



Dissertation presented to obtain the Ph.D degree in Molecular Biosciences

Oeiras, September, 2022

INSTITUTO DE TECNOLOGIA QUÍMICA E BIOLÓGICA ANTÓNIO XAVIER | UNIVERSIDADE NOVA DE LISBOA

A Closer Look at a Potential Biocatalyst: Unravelling the Catalytic, Stability and Structural Features of *Pp*DyP, a DyP-type Peroxidase

Diogo André Ribeiro Silva

Dissertation presented to obtain the Ph.D. degree in Molecular Biosciences

Instituto de Tecnologia Química e Biológica António Xavier

Universidade Nova de Lisboa

Supervisor: Prof. Lígia Oliveira Martins

Oeiras, September 2022

First Edition, September 2022
Microbial and Enzyme Technology Laboratory
Instituto de Tecnologia Quimica e Biologica Antonio Xavier
Universidade Nova de Lisboa
Av. da República, Oeiras

Acknowledgements

I would like to start by expressing my gratitude to my supervisor Lígia O. Martins, whose passion, hard work and dedication allowed for me to dedicate more than four years of my life to this work and to my growth as a scientist and a person. Through many lessons and hardships, Lígia has given me various opportunities to continue to develop and explore not only science itself, but myself as a scientist, providing all the conditions to test, explore and understand my own work, and even going on a limb to experiment with something new and unprecedented in the lab, something I really am thankful for. Lígia's guidance has been of invaluable, and without her help, I am certain I would not be in this position. Thank you Lígia*, e irei continuar a trabalhar na minha concentração!*

The members of my thesis committee, Eduardo Pinho e Melo and Maria Paula Robalo, were also very important for the completion of this thesis. Their feedback throughout this work was crucial and the collaborative effort enriched this work, and for that I am thankful.

I would also like to thank by lab partners: Vânia, André, Mario and, in particular, Carolina. Your companionship and friendship have helped me these last years, and allowed me to continue through any problems I found, scientific or not. I believe a good environment and good relationships between people is extremely important for whatever work one is doing, and you have certainly contributed to that.

I can't mention friendship without talking about those brave souls who have joined me as PhD students, in particular, Marta and Anaísa. Your friendship is truly invaluable, and all the hours we have spent together in the last four years is certainly a personal highlight. Through laughs or headbutts, we managed to get it done! A special thanks to Ricardo, Filipe, Rafael, Rita, Carolina, Inês and Patrícia as well!

Lastly, I would like to thank my parents and my brother for their continuous, relentless, unwavering support since I have moved to Oeiras. They have provided everything I have ever needed and have helped me way too many times to even count, but I will never forget any of those times. Even though you don't fully understand what I am doing and the path I am on, you have never left me behind, and for that, I am forever thankful.

I dedicate this thesis to Paula Leal (1991-2016). I know somehow, someway, you've helped me through this.

Abbreviations

***Aau*DyP** – Dye-decolorizing peroxidase from *Auricularia auricula-judae*

Abs – Absorbance

ABTS – 2,2'-azino bis (3-ethylbenzthiazoline-6-sulfonic acid)

Ala or **A** – Alanine

***Agro*DyP** – Dye-decolorizing peroxidase from *Agrobacterium* sp. B1

***Ana*PX** – Dye-decolorizing peroxidase from *Nostoc* sp.

AncDyPD-b – Ancestral D-type Dye-decolorizing peroxidase

Arg or **R** – Arginine

Asn or **N** – Asparagine

Asp or **D** – Aspartate

***Bad*DyP** – Dye-decolorizing peroxidase from *Bjerkandera adusta*

BR – Britton-Robinson

***Bs*DyP** – Dye-decolorizing peroxidase from *Bacillus subtilis*

***Bt*DyP** – Dye-decolorizing peroxidase from *Bacteroides thetaiotaomicron*

***Cbo*DyP** – Dye-decolorizing peroxidase from *Cellulomonas bogoriensis*

ChitO -- Chitooligosaccharide oxidase

CiP – Dye-decolorizing peroxidase from *Coprinus cinereus*

Clds – Chlorite dismutases

CotA-laccase – Laccase from *Bacillus subtilis*

Cpd I – Compound I

Cpd II – Compound II

CpHMD – Constant pH molecular dynamics

CslA – Cellulose synthase-like protein

***Ct*DyP** – Dye-decolorizing peroxidase from *Comamonas testosteroni*

Cys or **C** – Cysteine

Cyt c – Cytochrome C

DHLDH – Dihydrolipoamide dehydrogenase

DMP – 2,6-dimethoxyphenol

DtpA – Dye-decolorizing peroxidase from *Streptomyces lividans*

DtpAa – Dye-decolorizing peroxidase from *Streptomyces lividans*

DtpB – Dye-decolorizing peroxidase from *Streptomyces lividans*

DTT – Dithiothreitol

DyP2 – Dye-decolorizing peroxidase from *Amycolatopsis sp.* 75iv2

DypA – Dye-decolorizing peroxidase from *Rhodococcus jostii*

DypB – Dye-decolorizing peroxidase from *Rhodococcus jostii*

ε – Molar extinction coefficient

EDTA – Ethylenediaminetetraacetic acid

EfeB/YcdB – Dye-decolorizing peroxidase from *Escherichia coli*

***El*DyP** – Dye-decolorizing peroxidase from *Enterobacter lignolyticus*

EncMh – Encapsulin from *Mycolicibacterium hassiacum*

$\mathbf{E_{pa}}$ – Anodic peak potential

$\mathbf{E_{pc}}$ – Cathodic peak potential

epPCR – Error Prone PCR

ESI-MS – Electrospray ionization-mass spectrometry

EUGO – Eugenol oxidase

***Ftr*DyP** – Dye-decolorizing peroxidase from *Funalia trogii*

***Ftr*Lcc** – Laccase from *Funalia trogii*

GdnHCl – Guanidine hydrochloride

GGE – Guaiacylglycerol-β-guaiacyl ether

Gln or **Q** – Glutamine

Glu or **E**– Glutamic Acid

GlxA – Radical copper oxidase

Gly or **G** – Glycine

h – hour(s)

HCl – Hydrochloric acid

His or **H** – Histidine

HotAldO – Alditol oxidase

HPLC – High-performance liquid chromatography

HRP – Horseradish peroxidase

***Il*DyP** – Dye-decolorizing peroxidase from *Irpex lacteus*

Ile or **I** – Isoleucine

IPTG – Isopropyl-β-D-thiogalactopyranoside

k_{cat} – Turnover number

KCl – Potassium chloride

K_m – Michaelis constant

***Kp*DyP** – Dye-decolorizing peroxidase from *Klebsiella pneumoniae*

LB – Luria-Bertani

LC-MS – Liquid chromatography-mass spectrometry

Leu or **L** – Leucine

LiP – Lignin peroxidase

LRET – Long-Range Electron Transfer

Lys or **K** – Lysine

MD – Molecular dynamics

Met or **M** – Methionine

min – minute(s)

MnP – Mangenese peroxidase

***Msc*DyP** – Dye-decolorizing peroxidase from *Marasmius scorodonius*

***Mt*DyP** – Dye-decolorizing peroxidase from *Mycobacterium smegmatis*

NaCl – Sodium chloride

NHE – Standard hydrogen electrode

NMR – Nuclear magnetic resonance

PCR – Polymerase chain reaction

PDB – Protein data bank

PEG – Poly(ethylene glycol)

***Pf*DyP** – Dye-decolorizing peroxidase from *Pseudomonas fluorescens*

Phe or **F** – Phenylalanine

***Pos*DyP4** – Dye-decolorizing peroxidase from *Pleurotus ostreatus*

***Pp*DyP** – Dye-decolorizing peroxidase from *Pseudomonas putida* MET94

Pro or **P** – Proline

***Psa*DyP** – Dye-decolorizing peroxidase from *Pleurotus sapidus*

RB5 – Reactive Blue 5

RMSD – Root-mean-square-deviation

RS – Resting State

Rt – Retention times

s – Second(s)

SEM – Scanning electron microscopy

Ser or **S** – Serine

***St*DyP** – Dye-decolorizing peroxidase from *Streptomyces thermocarboxydus*

***Svi*DyP** – Dye-decolorizing peroxidase from *Saccharomonospora viridis* DSM43017

$\mathbf{T_{agg}}$ – Heat-induced aggregation temperature

TcDyP – Dye-decolorizing peroxidase from *Thanatephorus cucumeris*

***Tfu*DyP** – Dye-decolorizing peroxidase from *Thermobifida fusca*

Thr or **T** – Threonine

$\mathbf{T_m}$ – Melting temperature

Trp or **W** – Tryptophan

Tyr or **Y** – Tyrosine

TyrA – Dye-decolorizing peroxidase from *Shewanella oneidensis*

Val or **V** – Valine

***Vc*DyP** – Dye-decolorizing peroxidase from *Vibrio cholerae*

VGE – Veratryl glycerol-b-guaiacyl ether

VP – Versatile Peroxidase

***Xgr*DyP** – Dye-decolorizing peroxidase from *Xylaria grammica*

YfeX – Dye-decolorizing peroxidase from *Escherichia coli* O157

Abstract

Lignin is the largest reserve of aromatics on Earth and is a key renewable source of chemicals and materials; however, it remains mostly unexplored due to its recalcitrance to degradation. Recently, new state-of-the-art methods have been developed to depolymerize lignin and generate a portfolio of well-defined, industrial-relevant compounds, however, a lot of work remains to be done on the transformation of these compounds.

Laccases and peroxidases are part of the ligninolytic enzymatic system commonly found in wood-degrading fungi and have the potential to operate as biocatalysts in the transformation of lignin-derived compounds, providing a green and sustainable alternative to conventional chemical reactions, due to the selectivity, controllability, and economy of their reactions.

Dye-decolorizing Peroxidases (DyP) are the most recently discovered family of heme peroxidases, exhibiting differentiating features from classical peroxidases, both in their sequence and structure. These enzymes display an α+β ferredoxin-like fold that is analogous to chlorite dismutases and a carboxylate residue as an acid-base catalyst that differs from the histidine found in classical peroxidases. DyPs have a broad range of substrates, including synthetic azo and anthraquinone dyes, phenolic and nonphenolic lignin units, wheat straw lignocellulose, kraft lignin, iron and manganese ions. DyPs were firstly classified into four classes, based on their primary sequence, with subclasses A-C containing bacterial DyPs and subclass D containing fungal enzymes. More recently, a tertiary structure-based classification was proposed: class I, intermediate, corresponds to former class A; class P, primitive, corresponds to former class B; and class V, advanced, contains members from former classes C and D.

Bacterial DyPs are easier to engineer than fungal enzymes since bacterial systems benefit from a wide range of commercial molecular biology tools that facilitate their engineering towards improved activity, stability, resistance to solvents and inhibitors. A DyP from *Pseudomonas putida* was engineered, by Directed Evolution towards improved activity for the lignin-related phenolic compound 2,6-dimethoxyphenol. The hit variant 6E10 was fully characterized and exhibits 100-fold higher activity than wild type and an upshift of optimum pH towards pH 8.

In the first part of my thesis, we investigated the biocatalyst potential of variant 6E10 by exploring its substrate scope of lignin-related phenolic compounds. We performed the catalytic characterization of 6E10 using twenty-four phenol representatives of syringyl, guaiacyl, and hydroxybenzene derivatives. By combining spectrophotometric enzymatic assays and HPLC substrate conversion analysis we observed up to 100-fold higher specific activity of 6E10 and higher substrate conversion yields at pH 8 for all tested substrates after 24 h of reaction, when compared to the wild type reactions at pH 4. Additionally, by combining mass spectrometry and nuclear magnetic resonance, we identified the products obtained in the reactions of 6E10 with the best substrates. We showed that the lignin-related monomers were oxidized to their dimeric form, with industrially relevant compounds such as syringaresinol and di-vanillin being identified.

The activity improvements observed for 6E10 came at the cost of overall stability, as analysed in the second part of this thesis. The X-ray structural characterization of *Pp*DyP and and variants 29E4 (E188K, H125Y) and 6E10 (E188K, A142V, H125Y), in combination with molecular dynamics and docking experiments, allowed the unveiling of the dynamics behind the activity-stability trade-off of 6E10, and the understanding of the relation between catalysis and improved

optimal pH. The analysis of X-ray crystal structures revealed the typical ferredoxin-like fold, with three heme access pathways, two tunnels, and one cavity, limited by three long loops (L1, L2 and L3), which include the catalytic residues. The trade-off between activity and stability was pinpointed to the flexibility of these loops: variant 6E10 displayed significantly increased flexibility in the three loops, contrasting with the rigidity of the wild type enzyme and 29E4, and this favors function over stability. Using constant-pH MD simulations, in the neutral to the alkaline pH range, a more positively charged microenvironment near the heme pocket of variant 6E10 was identified, and this modulates the pK_a of essential residues in the heme vicinity. These features should account for the improved activity of 6E10 at pH 7-8, compared to the wild type and 29E4 that show optimal enzymatic activity close to pH 4.

The combination of structural and *in* silico analysis forwarded our understanding of the structure-function relationships of DyPs, implicating a role of conserved loops in the activity and stability, which can be modulated for the engineering of the catalytic properties of DyPs for future biotechnological applications. Additionally, we have elucidated the molecular determinants of pH-dependence of *Pp*DyP, which contributed to the understanding of the acidic preference of DyPs, a hallmark feature of this family of peroxidases that, in the past, has been attributed to the catalytic aspartate.

In the third and last part of my thesis, we aimed to stabilize *Pp*DyP using computational tools. While *Pp*DyP and its DE variants have great catalytic potential, these enzymes lack reliable and strong stability features. The stabilization of potential biocatalysts is paramount for their utilization in an industrial setup, where conditions tend to be harsh. The *in silico* design of proteins is a growing field that relies on computer-aided engineering to design new proteins or improve

features in an existing protein. These methods provide a fast, easy, and cheap alternative to the laborious laboratory evolution, using the protein structure and ever-improving algorithms to target the best position to mutate. We have used the webservers PROSS and FireProt to design thermostable variants of *Pp*DyP. These web servers suggested variants with 29 (PROSS) and 21 (FireProt) mutations with each variant gene being synthesized, expressed in *Escherichia coli*, and the proteins produced and purified. Both variants were produced in higher yields, up to 4-times higher than wild type, and exhibited higher activity for ABTS/H_2O_2 on a temperature optimum of 65-70 °C, a 45-50 °C increase from *Pp*DyP. Notably, the stability increases for both variants: both exhibited a 400-fold higher half-life time (20 h) than wild type (0.5 h) after incubation at 60°C and are more resistant to chemical denaturants. Both wild type and FireProt variant have an aggregation temperature (T_{agg}) around 60-65 °C which prevents an accurate measurement of their melting temperatures. However, the PROSS variant avoided the temperature-induced aggregation of its tertiary structure in a 20 to 100 °C temperature range and has a melting temperature of 88 °C, a putative 25 °C increase from the wild type.

DyP-type peroxidases provide an interesting alternative to classical peroxidases, from a fundamental point of view but also from an application point of view. The broadness of their substrate scope, the catalytic potential and ease to engineer is well documented through this thesis: different techniques were used, random mutagenesis and computational engineering, to obtain *Pp*DyP variants with substantially better activity and stability. These features strengthen DyPs as targets for industrial exploration.

Resumo

A lenhina é a maior reserva de aromáticos na Terra e é uma fonte renovável de químicos e materiais; no entanto, continua a ser pouco explorada devido à sua resistência à degradação. Recentemente, novos métodos têm sido desenvolvidos para degradar lenhina e gerar um portfolio de compostos bem definidos e relevantes para várias indústrias, no entanto, ainda existe muito trabalho por fazer na transformação destes compostos.

As lacases e peroxidases fazem parte do sistema enzimático ligninolítico que é frequentemente encontrado em fungos que degradam matéria vegetal e madeira. Estas enzimas têm o potencial de funcionar como biocatalisador na transformação de compostos derivados de lenhina, providenciando uma alternativa verde e sustentável para as reações, devido à seletividade, controlo e economia destas reações.

Dye-decolorizing peroxidases (DyP) são a mais recente família de peroxidases hémicas, e têm características que as diferenciam das peroxidases clássicas, tanto em termos de sequência como de estrutura. Estas enzimas possuem um *α+β ferredoxin-like fold* análogo a dismutases de clorito, e um resíduo carboxilato como catalista acido-base que difere da histidina encontrada em peroxidases clássicas. As DyPs têm uma ampla gama de substratos, incluindo corantes sintéticos azo e de antraquinona, unidades de lenhina fenólicas e não fenólicas, lignocelulose de palha de trigo e *kraft lignin*, iões de ferro e de manganês. As DyPs foram inicialmente classificadas em quatro classes, com base na sua sequência primária, com as subclasses A-C contendo DyPs bacterianas e a subclasse D contendo enzimas fúngicas. Recentemente, foi proposta uma nova classificação baseada na estrutura terciária destas enzimas: a classe I, intermediária,

corresponde à antiga classe A; a classe P, primitiva, corresponde à antiga classe B; e a classe V, avançada, engloba as antigas classes C e D.

As DyPs bacterianas são mais facilmente alteradas do que as enzimas fúngicas, uma vez que os sistemas bacterianos beneficiam de uma gama maior de ferramentas comerciais de biologia molecular que facilitam a sua engenharia para melhorar a atividade, estabilidade, resistência a solventes, entre outras propriedades. Uma DyP de *Pseudomonas putida* foi mutada, por *Directed Evolution*, para melhorar a atividade para o 2,6-dimetoxifenol, um composto fenólico encontrado na lenhina. O melhor variante, 6E10, foi completamente caracterizado e exibe atividade 100 vezes maior do que o *wild type* e uma mudança de pH ótimo para pH 8.

Na primeira parte da minha tese, investigamos o potencial do variante 6E10 como biocatalisador, explorando a sua capacidade de oxidar uma vaga gama de compostos fenólicos derivados de lenhina. Realizámos a caracterização catalítica do 6E10 usando vinte e quatro fenóis representativos de derivados de lenhina dos grupos siringil, guaiacil e hidroxibenzeno. Ao combinar ensaios enzimáticos espectrofotométricos e a analise de níveis de conversão de substrato por HPLC, observamos no 6E10 atividades específicas até 100 vezes maiores do que na enzima *wild type* e rendimentos de conversão mais altos, a pH 8, para todos os substratos fenólicos testados, após 24 h de reação, comparativamente às reações do *wild type* a pH 4. Além disso, através da identificação dos produtos das reações com os melhores substratos, por espectrometria de massa e RMN, mostramos que os monómeros derivados da lenhina foram oxidados para a sua forma dimérica, tendo identificado compostos industrialmente relevantes, como siringaresinol e di-vanilina.

As melhorias de atividade observadas para o 6E10 vieram à custa da estabilidade geral desta enzima, algo que foi analisado na segunda parte desta tese. A caracterização estrutural por raios-X da enzima *Pp*DyP e variantes 29E4 (E188K, H125Y) e 6E10 (E188K, A142V, H125Y), em combinação com dinâmicas moleculares e experiências de *docking*, permitiu desvendar a dinâmica por detrás do *trade-off* atividade-estabilidade observado no 6E10, e entender melhor a relação entre catálise e pH ótimo. A análise das estruturas cristalinas de raios X revelou um *ferredoxin-like fold*, com três vias de acesso ao hemo, dois túneis e uma cavidade, limitada por três *loops* longos (L1, L2 e L3), que incluem os resíduos catalíticos. O *trade-off* entre atividade e estabilidade está relacionado com a flexibilidade destes *loops*: no variante 6E10 estes três *loops* são significativamente mais flexíveis, contrastando com a rigidez da enzima *wild type* e do variante 29E4, o que favorece a atividade à custa da estabilidade. Usando simulações de *molecular docking* a pH constante, na faixa de pH neutro a alcalino, foi identificado um microambiente mais positivo próximo ao *pocket* do hemo no variante 6E10, e isso modula o pK_a dos resíduos essenciais presentes na vizinhança do hemo. Isto explica a melhor atividade do variante 6E10 a pH 7-8 em comparação com o *wild type* e 29E4 que têm uma atividade enzimática ideal próxima a pH 4.

A combinação da análise estrutural e *in silico* permitiu entender as relações estrutura-função nas DyPs, evidenciando o papel de *loops* conservados na atividade e estabilidade, abrindo a porta à sua modelação para a engenharia das propriedades catalíticas das DyPs para futuras aplicações biotecnológicas. Para além disso, elucidamos os determinantes moleculares da dependência do pH da *Pp*DyP, contribuindo para a compreensão da preferência acídica das DyPs,

uma característica chave desta família de peroxidases que inicialmente foi atribuída ao aspartato catalítico.

Na terceira e última parte da minha tese, tínhamos como objetivo estabilizar a *Pp*DyP usando ferramentas computacionais. Embora a *Pp*DyP e os seus variantes tenham um grande potencial catalítico, estas enzimas carecem de uma estabilidade forte. A estabilização de potenciais biocatalisadores é fundamental para viabilizar a sua utilização em ambiente industrial, onde as condições tendem a ser adversas. O *design in silico* de proteínas é um tópico cujo interesse tem crescido consideravelmente e que depende da engenharia de enzimas, auxiliada por computadores, para desenhar novas proteínas ou melhorar as características de uma proteína existente. Estes métodos fornecem uma alternativa rápida, fácil e barata para a laboriosa evolução laboratorial, usando a estrutura da proteína e algoritmos cada vez melhores para identificar a melhor posição para mutar. Usamos os servidores PROSS e FireProt para desenhar variantes termoestáveis da *Pp*DyP. Estes *webs servers* sugeriram variantes com 29 (PROSS) e 21 (FireProt) mutações. Cada variante foi sintetizado, expresso em *Escherichia coli* e as proteínas produzidas e purificadas. Ambos os variantes foram produzidas com rendimentos mais altos, até 4 vezes mais do que o *wild type*, e exibiram maior atividade para ABTS/H_2O_2 com uma temperatura ótima de 65-70 °C, um aumento de 45-50 °C relativamente à *Pp*DyP. A estabilidade foi aumentada para ambos os variantes: os dois exibiram um tempo de meia-vida 400 vezes maior (20 h) do que o *wild type* (0.5 h) após incubação a 60 °C e são mais resistentes a desnaturantes químicos. Tanto o *wild type* como o variante FireProt têm uma temperatura de agregação (T_{agg}) por volta dos 60-65 ºC, o que impede uma medição precisa das temperaturas de *melting*. No entanto, o variante PROSS evitou a agregação induzida pela temperatura numa rampa de

temperaturas entre 20 a 100 °C e tem uma temperatura de *melting* de 88 °C, um aumento putativo de 25 °C em relação ao *wild type*.

As DyP peroxidases fornecem uma alternativa interessante às peroxidases clássicas, do ponto de vista fundamental, mas também do ponto de vista da aplicação. A variedade de substratos oxidados, o potencial catalítico e a facilidade de engenharia, estão documentadas nesta tese: diferentes técnicas foram usadas, mutagénese aleatória e engenharia computacional, para obter variantes da *Pp*DyP com atividade e estabilidade significativamente melhores. Estas características fortalecem o papel das DyPs como alvos para exploração industrial.

List of Publications

Catarina Barbosa, Célia M. Silveira, **Diogo Silva**, Vânia Brissos, Peter Hildebrandt, Lígia O. Martins and Smilja Todorovic (2020). Immobilized dye-decolorizing peroxidase (DyP) and directed evolution variants for hydrogen peroxide biosensing. *Biosens. Bioelectron*. **153**, 112055.
https://doi.org/10.1016/j.bios.2020.112055

Carolina F. Rodrigues, Patrícia T. Borges, Magali F. Scocozza, **Diogo Silva**, André Taborda, Vânia Brissos, Carlos Frazão and Lígia O. Martins (2021). Loops around the Heme Pocket Have a Critical Role in the Function and Stability of *Bs*DyP from *Bacillus subtilis. Int J Mol Sci*. **22**, 10862
https://doi.org/10.3390/ijms221910862

Diogo Silva, Patrícia Borges, Tomás F. D. Silva, Vânia Brissos, Marina Cañellas, Maria Fátima Lucas, Laura Masgrau, Eduardo P. Melo, Miguel Machuqueiro, Carlos Frazão and Lígia O. Martins (2022). Unveiling Molecular Details behind Improved Activity at Neutral to Alkaline pH of an Engineered DyP-type Peroxidase. *Comput. Struct. Biotechnol. J.* **20:** 3899-3910.
https://doi.org/10.1016/j.csbj.2022.07.032

Diogo Silva, Ana Catarina Sousa, Vânia Brissos, Maria Paula Robalo and Lígia O. Martins (2022). A Wide Array of Lignin-Related Phenolics are Oxidized by an Evolved DyP an at Alkaline pH. *N Biotechnol*.
https://doi.org/10.1016/j.nbt.2022.12.003

Diogo Silva, Carolina F. Rodrigues, Patrícia T. Borges and Lígia O. Martins (2023). “Biocatalysis for Lignocellulose Biorefineries: The Case of Dye-Decolorizing Peroxidases (DyPs)”. (In preparation)

Diogo Silva, Patrícia Borges, Eduardo P. Melo and Lígia O. Martins (2023). Overcoming low stability of a DyP from *Pseudomonas putida* using a one-step computational approach. (In preparation)

Table of Contents

Chapter 1

General Introduction

This chapter contains data published in:

Diogo Silva, Carolina F. Rodrigues, Patrícia T. Borges and Lígia O. Martins (2023). “Biocatalysis for Lignocellulose Biorefineries: The Case of Dye-Decolorizing Peroxidases (DyPs)”. (In preparation)

1.1. Biocatalysis for Lignocellulose Biorefineries: Overcoming the Lignin Barrier

Green and white biotechnology can contribute to the successful implementation of biorefineries by providing bio-based feedstocks and more environmentally friendly biocatalysts tools. Lignocellulosic biomass is an essential source of renewable bulk and materials, fine chemicals, energy, and fuels. Lignin is the second most abundant carbon source and the largest source of natural phenolic compounds on Earth. The pulp and paper industry produces about 50 million tons of lignin annually, but most are burned for power; only 1 million tons reach the chemicals market (Sun et al., 2018). Lignin is currently used for low- and medium-value applications (e.g., binding and dispersing agents), with energy capturing about 89% of the market. Lignin is expected to become an alternative feedstock for aromatic chemicals, most currently derived from the BTX process in the petroleum industry (Chen and Wan, 2017; Hamalainen et al., 2018). Lignin depolymerization was problematic in the past, hindering its valorization. However, the recent implementation of novel strategies for lignin depolymerization, involving electrochemistry, photocatalysis, heterogeneous catalysis, and ionic liquids, allowed to derive well-defined compounds from lignin in acceptable quantities and implement a lignin-derived platform of chemicals **(Figure 1.1)** (Renders et al., 2017; Sun et al., 2018; Van den Bosch et al., 2018).

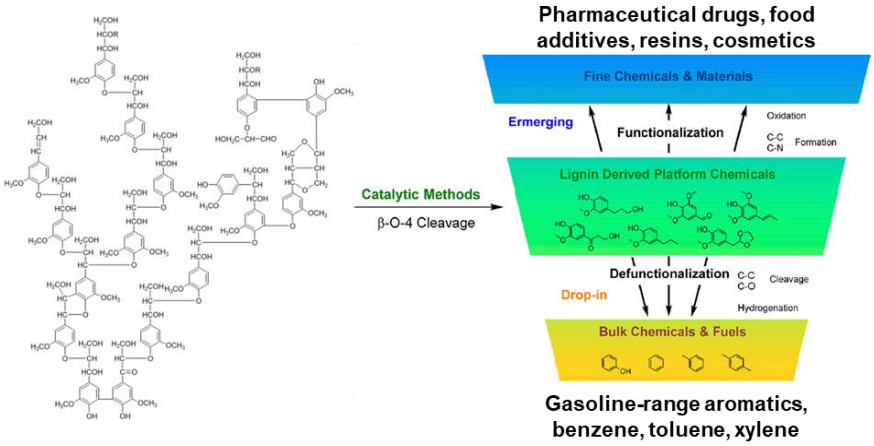


Figure 1.1. **Lignin-derived chemical platform**. The depolymerization of lignin leads to a set of aromatic chemicals that can be functionalized, resulting in added-value compounds such as drugs, food additives, and polymers, and defunctionalized, resulting in bulk chemicals such as benzene, toluene, and xylene. Adapted from (Sun et al., 2018).

The lignin platform of chemicals includes aromatics such as β-aryl ethers, biphenyls, diaryl propane, pinoresinol, phenylcoumaranes, and ferulic acid, among many others. Therefore, at present, the challenge is to set-up atom-economical and waste-free processes that allow the full implementation of lignin as a sustainable starting material for the production of drop-in chemicals, polymers, or emerging functional materials (Llevot et al., 2016; Natte et al., 2020; Runeberg et al., 2019; von Vacano et al., 2022; Zalesak et al., 2019). Biocatalysis is an eco-friendly and sustainable biotechnology. In nature, the reactions involved in depolymerization, and conversion of lignin are performed by white-rot fungi and bacteria. These microorganisms offer a toolbox of auxiliary enzymes, ligninolytic laccases, heme-peroxidases, copper-radical oxidases, glutathione-S-transferases, aldo-keto reductases, catalase-peroxidase, cytochrome P-450, esterases, capable of performing a set of reactions that can be replicated in biorefineries (Bissaro et al., 2018; Liu et al., 2022; Wong, 2009).

1.2. Classical Ligninolytic Peroxidases

Peroxidases are ubiquitous in living organisms. These enzymes play crucial roles in cells, e.g., are part of mammalians' immune response or remove H_2O_2 from plant chloroplasts. The report of the first peroxidase reaction dates as early as two centuries back when Planche observed the development of blue color in the tincture of *Guaiacum* plant resin upon soaking with fresh horseradish roots in 1810 (Planche, 1810). Peroxidases use one-equivalent H_2O_2 to oxidize two equivalents of substrate molecules, generating two equivalents of water molecules and reduced products. Animal and plant peroxidases differ in primary and tertiary structures, the nature of heme cofactor, and active site architecture. Fungal peroxidases are part of the plant peroxidase superfamily, classified as Class I and Class II plant peroxidases. Class II contains the ligninolytic peroxidases secreted by fungi, lignin (LiP), manganese (MnP), and versatile (VP) peroxidases that play a role in lignin decomposition **(Figure 1.2)** (Hammel and Cullen, 2008; Lundell et al., 2010; Martinez et al., 2009). LiP and VP are believed to directly oxidize benzenic rings of lignin irrespective of their methoxylation degree and interunit linkages (Martinez et al., 2009). The unstable aromatic cation radicals formed chemically evolve, leading to lignin depolymerization (Kirk, 1987). For its part, MnP (and VP) oxidize Mn(II) to Mn(III), which can diffuse and penetrate wood when the size of the cell wall pores are too small for the penetration of enzymes. Mn(III) is chelated by organic acids and oxidizes phenolic structures in the lignin (Hammel and Cullen, 2008). Mn(III) also generates peroxy radicals via lipid peroxidation, which oxidize the non-phenolic structures in lignin.

1.3 The New DyP-type Peroxidases Family of Enzymes

In 1999, a new heme-containing enzyme was identified in the fungus *Geotrichum candidum* Dec 1, later renamed *Thanatephorus cucumeris* Dec 1, capable of efficiently degrading synthetic anthraquinone industrial dyes (Kim, 1999), hence the name attributed to dye-decolorizing peroxidase (DyP) (Sugano, 2009). Interestingly, no homology in the sequence was found between this peroxidase and classical peroxidases. Since then, DyP-like sequences have been found in bacteria, fungi, and archaeal genomes, resulting in a new family of microbial peroxidases (Linde et al., 2015b; Singh and Eltis, 2015; Sugano and Yoshida, 2021; Yoshida and Sugano, 2015).

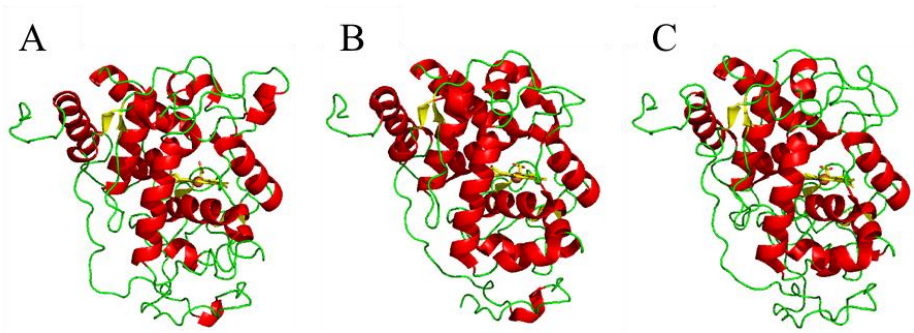


Figure 1.2. Structures of representative Lignin Peroxidase LiP (A), Versatile Peroxidase VP (B), and Manganese Peroxidase MnP (C). LiP from *Phanerodontia chrysosporium* (PDB 1LLP), VP from *Pleurotus eryngii* (PDB 3FMU), and MnP from *Phanerodontia chrysosporium* (PDB 1MNP) colored by their secondary structure: red for α-helices, yellow for β-sheets, and green for loops. The heme is represented as sticks with carbon, oxygen, and nitrogen atoms colored yellow, red, and blue, respectively. The iron atom is shown as an orange sphere.

Every characterization of these enzymes points to their peroxidase function. However, other functions were reported. In the case of *T. cucumeris* Dec 1 *Tc*DyP, the degradation of reactive blue 5, an anthraquinone dye, resulted in the production of phthalic acid. This implicates the degradation of the anthraquinone core by hydrolase- or

oxidase-catalyzed reactions (Sugano, 2009). In the case of hydrolase function, the water molecule may be formed from hydrogen peroxide. The role of a water molecule in the heme pocket has been previously investigated in classical peroxidases, with the hypothesis that a dry (non-H_2O-containing) and wet (H_2O-containing) heme pocket may promote different reactions (Jones, 2001). An oxidase mechanism is possibly present in *Marasmius scorodonius Msc*DyP, which oxidizes β-carotene without H_2O_2 (Scheibner et al., 2008) and, likewise, *Psa*DyP from *Pleurotus sapidus* was found to transform β-carotene in the presence and absence of H_2O_2 (Krahe et al., 2020; Lauber et al., 2017).

Anthraquinone dyes used in industrial setups are mostly synthetic compounds and are not produced naturally, which suggests that natural anthraquinone compounds might be actual physiological DyP substrates. For example, alizarin, an anti-fungal compound produced by the plant *Rubia tinctorum*, is efficiently decomposed by the DyP from *Bjerkandera adusta* Dec 1, a parasitic fungus, indicating a possible physiological role of DyPs in degrading antifungal anthraquinones and symbiotically accelerate plant parasitism (Sugawara et al., 2019).

DyPs in the genome of lignin-degrading basidiomycetes (white-rot fungi) are good evidence of their role as ligninolytic enzymes (Fernandez-Fueyo et al., 2012; Ruiz-Duenas et al., 2013). Moreover, DyPs are widely expressed amongst degrading lignin fungi from forest habitats (Kellner et al., 2014) and can oxidize several lignin-related compounds, including syringyl and guaiacyl-type phenolics, kraft lignin and lignin models (Linde et al., 2015b). DyPs are versatile enzymes that also oxidize non-phenolic methoxylated aromatics (e.g., veratryl alcohol), Fe(II), and Mn(II) (Santos et al., 2014). In 2011, the transcriptomic and proteomic analysis showed that the expression of a

DyP from *Enterobacter lignolyticus* was up-regulated when this organism was grown with lignin as the sole carbon source, suggesting a link between DyP and lignin depolymerization (DeAngelis et al., 2013).

In *Streptomyces lividans*, the developmental switch between vegetative mycelium and aerial hyphae is dependent on a morphogenetic pathway that involves a DyP-type peroxidase (DtpA) and a radical copper oxidase (GlxA) (Petrus et al., 2016). DtpA allows the maturation of GlxA, i.e., the formation of a cross-linked tyrosyl-cysteine cofactor essential for enzymatic activity (Chaplin et al., 2015). DtpA activates the enzyme, which, in turn, generates H_2O_2 from O_2, providing the H_2O_2 substrate for DtPA-mediated oxidation. GlxA will then act with a cellulose synthase-like protein, CslA, forming a protective extracellular glycan specific for aerial hypha (Petrus et al., 2016).

DyPs have been identified as the cargo of encapsulins (Putri et al., 2017; Sutter et al., 2008). These are a recently discovered family of conserved prokaryotic proteinaceous nanocompartments (Sutter et al., 2008) which are important for increasing the local concentration of functionally related enzymes and confining unstable reaction intermediates within the cell. The DypB from *Rhodococcus jostii* was shown to assemble with encapsulin *in vitro* (Rahmanpour and Bugg, 2013). In this case, one possibility is that the encapsulin nanocompartment is a dynamic structure that can disassemble and localize DypB on the surface of lignin or lignocellulose that will generate short-lived oxidants (Mn(II) or phenolic radicals), which can diffuse in bulky lignin substrates to decompose the recalcitrant compounds. *Mycobacterium tuberculosis* encapsulin and DyP were found in a two-gene operon, suggesting tight translational coupling, and are putatively involved in the oxidative stress response to evade

the host immune system (Contreras et al., 2014). Recently, the structure of the *Mycobacterium smegmatis* dodecameric DyP (two hexamers with a twofold axis of symmetry), the cargo of encapsulin, was solved by cryo-EM, at 3.7 Å resolution (PDB 7BOK) (Tang et al., 2021). The central channels of the DyP hexamers are highly positively charged, which may facilitate the entrance of negatively charged substrates (Tang et al., 2021), and transfer electrons through, e.g., long-range electron transfer (LRET) pathways to the heme (see below). The encapsulin shell plays a role in stabilizing the dodecameric DyP, and the DyP can protect *M. smegmatis* from oxidative stress.

1.4. Structural Insights

The first solved DyP-type peroxidase X-ray crystal structure was that of *T. cucumeris* DyP, renamed *B. adusta* (PDB 2D3Q) (Sugano et al., 2007). Since then, many more structures have been solved **(Table 1.1)**, showing that DyP-type peroxidases have significantly different overall structures compared to classical peroxidases, which are α-helix rich structures (Veitch, 2004; Zamocky et al., 2008). DyPs have two domains that adopt an α + β ferredoxin-like fold similar to the tertiary structure of chlorite dismutases (Clds) **(Figure 3 a)** (Hofbauer et al., 2021). Both DyPs and Clds belong to the dimeric α + β barrel superfamily. This superfamily encompasses 23 enzymes families; however, DyPs and Clds share striking structural and mechanistic similarities: both can bind heme within their ferredoxin-like fold, have a fully conserved arginine in the distal site, and have a histidine as the proximal ligand. Mechanistically, both families require initial oxidation by either hydrogen peroxide or chlorite. In contrast, other heme-binding members of the dimeric α + β barrel

superfamily do not require the activation of the heme cofactor. The protomer of DyPs is formed by two ferredoxin-like domains, one in each terminal, composed of four-stranded antiparallel β-sheets and two peripheral α-helices **(Figure 3 b,c).**

Table 1.1. Available DyP structures deposited at Protein Data Bank (PDB).

	Organism	PDB	Reference
Class P	*C. bogoriensis Cbo*DyP	6QZO	(Habib et al., 2019)
	*K. pneumoniae Kp*DyP	6FKS, 6FIY, 6FL2, 6FKT, 6RR6, 6RR8, 6RR4, 6RR1, 6RPD, 6RPE, 6RQY, 6RR5	(Pfanzagl et al., 2018)
	*E. lignolyticus El*DyP	5VJ0	(Shrestha et al., 2017)
	*V. cholerae Vc*DyP	5DE0	(Uchida et al., 2015)
	R. jostii RHA1 DyPB	4HOV, 3VEC, 3VED, 3VEE, 3VEF, 3VEG, 3QNR, 3QNS	(Ahmad et al., 2011; Singh et al., 2012; Singh et al., 2013)
	S. oneidensis TyrA	2HAG, 2IIZ	(Zubieta et al., 2007a; Zubieta et al., 2007b)
	S. coelicolor DyP	4GU7	-
	S. lividans DtpB	6YRJ, 6YRC, 6YR4, 6YRD	(Lučić et al., 2020b)
	E. coli O157 YfeX	5GT2	(Liu et al., 2017)
Class I	*T. curvata Tc*DyP	5JXU	(Shrestha et al., 2016)
	*T. fusca Tfu*DyP	5FW4	(Rahmanpour et al., 2016)
	T. cellulosilytica DyP	4GS1	-
	E. coli O157 EfeB/YcdB	2Y4D, 2Y4E, 2Y4F, 3O72	(Liu et al., 2011)
	*B.subtilis Bs*DyP	7PKX, 7PL0, 6KMM, 6KMN	(Dhankhar et al., 2020; Rodrigues et al., 2021)
	S. lividans DtpAa	6TB8	(Lučić et al., 2020b)
	S. lividans DtpA	6GZW, 5MJH, 5MAP	(Chaplin et al., 2019; Kekilli et al., 2017)
	S. coelicolor DyP	4GT2, 4GRC	-
Class V	*D. discoideum* DyPA	7O9J, 7ODZ, 7O9L	(Rai et al., 2021)
	*B. thetaiotaomicron Bt*DyP	2GVK	(Zubieta et al., 2007b)
	*M. smegmatis Mt*DyP	7BOK	(Tang et al., 2021)
	*Anabaena sp.Ana*PX	5C2I	(Yoshida et al., 2016)
	*P. ostreatus Pos*DyP4	6FSK, 6FSL	(Fernández-Fueyo et al., 2018)
	*A. auricula-judae Aau*DyP	5IKD, 5IKG, 4W7J, 4W7K, 4W7L, 4W7M, 4W7N, 4W7O, 4UZI, 4AU9, 5AG0, 5AG1	(Linde et al., 2016; Linde et al., 2015a; Strittmatter et al., 2013a)
	*B. adusta Bad*DyP	3VXI, 3VXJ, 2D3Q, 3MM1, 3MM2, 3MM3, 3AFV	(Yoshida et al., 2011)
	I. lacteus F17 *Il*DyP	7D8M	(Li et al., 2021)
	Amycolatopsis sp. 75iv2 DyP2	4G2C	(Brown et al., 2012)
	G. candidum AncDyPD-b	7ANV	(Zitare et al., 2021)

The molecular mass of DyPs monomers ranges between 30 to 60 kDa, with Class P DyPs ranging between 30-40 kDa, Class I between 40-50 kDa, and Class V between 45-60 kDa, explained by larger primary sequences that add on the tertiary structure. Bacterial DyPs are mainly assembled in dimers (Li et al., 2012; Liu et al., 2011; Rodrigues et al., 2021), while fungal DyPs are exclusively monomers (Chen and Li, 2016; Johjima et al., 2003; Sugano et al., 2007; Zitare et al., 2021). Sequence analysis indicates that fungal DyPs have numerous C-terminal insertion sequences, which are postulated to prevent the enzymes from dimerization. In bacterial enzymes, dimerization is usually caused by hydrophobic interaction between monomers. Moreover, it has been proposed that radical formation at the protein surface can lead to oligomerization; intermolecular interactions were observed in the reaction of *Vc*DyP with H_2O_2, and the mutation of two tyrosines positioned at the dimer interface disrupted oligomerization (Uchida et al., 2015). High-order oligomerization as a hexameric state was observed in *Bt*DyP from *Bacteroides thetaiotaomicron* (Zubieta et al., 2007b) and DyPB from *R. jostii* RHA1 (Ahmad et al., 2011).

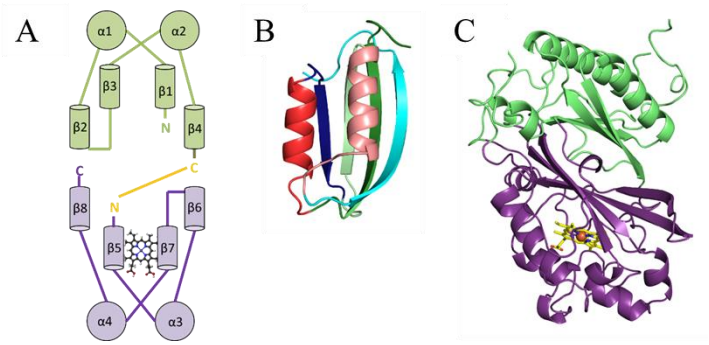


Figure 1.3. Structural domains of DyP-type peroxidases. Schematic representation of a subunit from the dimeric α + β barrel superfamily, consisting of two ferredoxin folds where the N- and C-terminal and linker are colored in green, purple, and yellow (a). Cartoon representation of a ferredoxin-like fold, with α1 and α2 colored in red and pink, respectively, and β1, β2, β3, and β4 colored in dark green, dark blue, light green, and cyan (b). Cartoon representation of *B. subtilis Bs*DyP monomer, with ferredoxin-like folds colored accordingly panel a). Adapted from (Hofbauer et al., 2020).

DyPs were initially grouped based on its primary structure from Class A to D (Sugano, 2009). Phylogenetic analysis indicated that the ancestor DyP was among Class A, particularly from the thermophilic facultatively anaerobic bacteria of the division *Firmicutes* (Zamocky et al., 2015). Evolution resulted in the shortened and divergent DyPs in Class B and elongated and convergent DyPs in Classes C and D clustered together **(Figure 1.4)**.

Interestingly, Class D is composed exclusively of fungal DyPs. At the same time, Class C harbors DyPs from proteobacteria, actinobacteria, and cyanobacteria, which indicates that throughout evolution, fungal peroxidases diverged from bacterial genes, most probably by horizontal gene transference. The DyPs from basidiomycetes are secreted to the extracellular space, where they can act on phenolic and non-phenolic lignin model compounds. Likewise, DyP-type B genes were found in lower eukaryotes and even in the human parasite *Schistosoma mansoni*, indicating that horizontal gene transfer events also occurred in Class B (Zamocky et al., 2015).

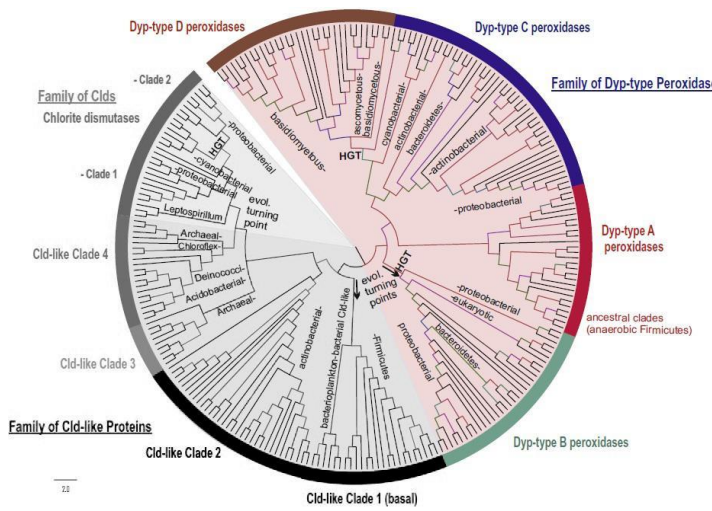


Figure 1.4. Phylogenetic tree of 250 members of the peroxidase-chlorite dismutate superfamily. The dye-decolorizing peroxidase main family is divided into the four-class system based on its primary sequence. (Adapted from Zamocky et al., 2015).

In 2015, a new classification scheme was proposed based on structure-based sequence alignments (Yoshida and Sugano, 2015) **(Figure 1.5)**: Classes P (primitive) and I (intermediate) replace Classes B and A, respectively, and Classes C and D are fused in Class V (advanced). Class P members have the smallest molecular size among the three classes, with members of class V having the largest size and also higher catalytic efficiencies (Catucci et al., 2020).

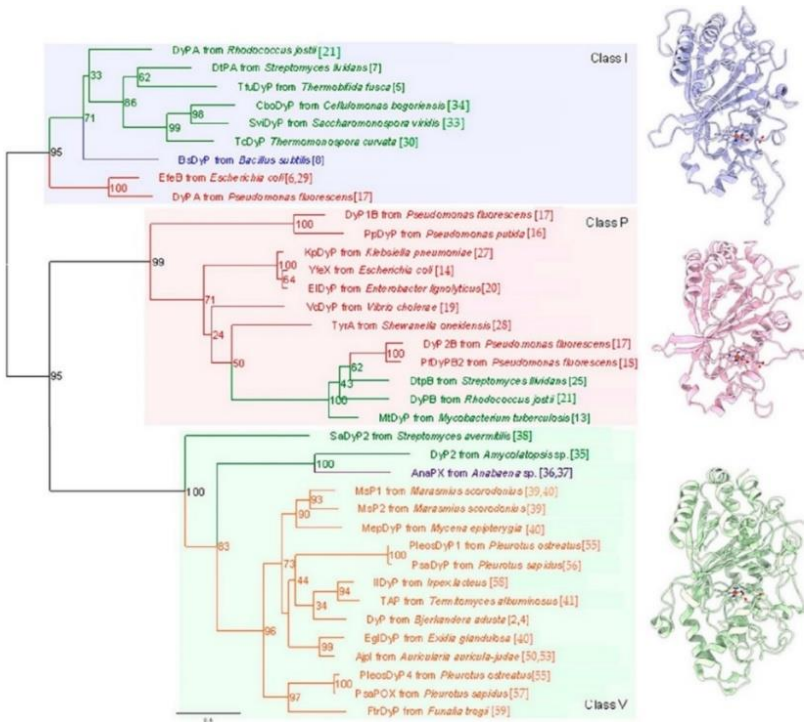


Figure 1.5. Structures of representative DyPs of the three classes are shown on the right: Class I, *Bs*DyP (PDB 7PKX); Class P, *Pp*DyP (PDB 7QYQ); Class V: AncDyPD-b (PDB 7ANV). From Sugano et al., 2021.

1.4.1. Active Site Structure and Catalytic Mechanism

The heme pocket of DyP-type peroxidases is located in the C-terminal domain of the protomer, buried within the folded α-helix and β-sheets **(Figure 1.6 a,c,e).** The loops and α-helix regions delimiting the heme pocket are structurally conserved in all DyPs, forming a molecular tunnel giving direct access of H_2O_2 to the distal side of the heme **(Figure 1.6 b,d,f)** (Habib et al., 2019; Rodrigues et al., 2021; Shrestha et al., 2017; Strittmatter et al., 2013a). Class I and P DyPs also have an open wide cavity that gives access to the solvent-exposed heme propionate p6 group allowing electron transfer from the secondary substrate to the heme **(Figure 1.6 a-d)** (Catucci et al., 2020; Perez-Boada et al., 2005; Pfanzagl et al., 2019). In class V DyPs this cavity seems to be occluded by longer (10 to 30 amino acids) loops **(Figure 1.6 e,f)**; however, the expected high flexibility of loops in this region may promote solvent access and oxidation of small substrates at the surface close to the heme.

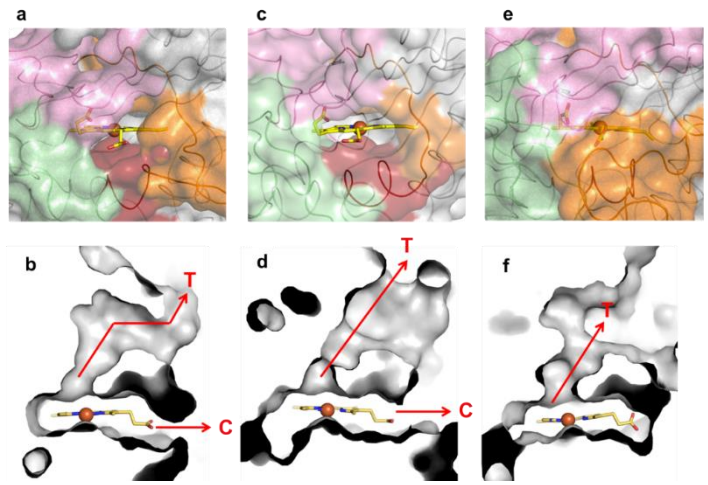


Figure 1.6. Loops, molecular tunnels, and cavities in Class P *K. pneumoniae* DyP (PDB 6FKS) (a, b), class I *B. subtilis Bs*DyP (PDB 7PKX) (c,d), and class V *A. auricula-judae* DyP (PDB 4AU9) (e,f). The regions lining the cavities are coloured light pink (loop 1), green (loop 2), dark red (small α-helix), and orange (loop 3). The tunnels (T) and cavities (C) corresponding to panels (a,c,e) are represented as grey accessible surface area in panels (b,d,f), respectively. From (Rodrigues et al., 2021).

The heme cofactor is non-covalently bound and is coordinated by a proximal axial histidine ligand **(Figure 1.7)**. At the distal site, DyP-type peroxidases have two conserved residues: a negatively charged residue, usually an aspartate, and a positively charged arginine that are part of the tunnel that allows the access of the H_2O_2 to the heme **(Figures 1.6 and 1.7)**. A polar residue, usually a distal asparagine, interacting with the iron-coordinated solvent species (Singh et al., 2012) is located nearby the two distal conserved residues. While it is not conserved in all DyPs classes, this position seems relevant for the geometry of the active site distal side (Silveira et al., 2020), and heme stabilization since similar distances in different classes are observed for the heme iron and the coordinated solvent species.

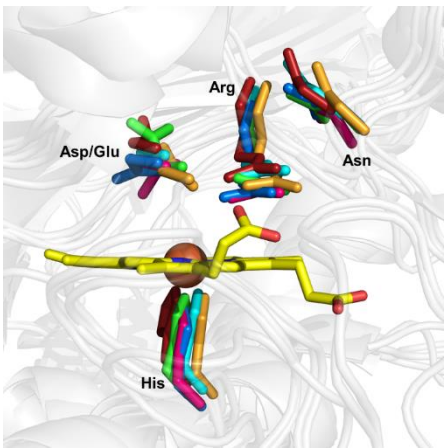


Figure 1.7. Representation of the heme binding pocket of DyPs indicating the proximal His and the conserved distal residues Asp/Glu, Asn, and Arg. The residues of DyPs from *Thermomonospora curvata* (PDB 5JXU), *C. bogoriensis (*PDB 6QZO), *E. coli* (PDB 5GT2), *V. cholerae* (PDB 5DE0), *E. lignolyticus* (PDB 5VJ0) and *Pseudomonas putida* (PDB 7QYQ) are shown as sticks with atoms colored in blue, pink, cyan, dark red, orange and green, respectively. The heme cofactor is shown as sticks with carbon, oxygen, and nitrogen atoms colored yellow, red, and blue, respectively.

The comparative resonance raman study of DyPs from 3 subfamilies reveals that (i) the heme active site architecture of DyPs in solution does not appear to be subfamily dependent; (ii) the nature of amino

acid residues that line up the distal side also cannot be directly correlated with spin configurations, since the DyPs that, e.g., house Asp, Arg, Asn and Phe in the distal site populate distinct spin states that follow very different pH dependences; (iii) there is a surprisingly high abundance of catalytic incompetent 6cLS populations in several DyPs, which was attributed to the specific H-bonding network in the vicinity of the heme active site (Silveira et al., 2020). The pronounced pH dependence of the spin state population distribution emphasizes the importance of the second sphere arrangements in tuning the pK_a of the distal residues and fine-tuning hydrogen peroxide reactivity.

It has become clear that there is mechanistic variation in Compound I (Cpd I) formation amongst DyPs, and the question of why Asp or Arg is selected to facilitate proton transfer and rate enhancement of O–O fission remains unanswered. The aspartate, which replaces the conserved distal histidine in classical peroxidases, is part of the characteristic GXXDG motif and functions as an acid-base catalyst in DyPs (Chen et al., 2015; Lučić et al., 2020a; Pfanzagl et al., 2019; Shrestha et al., 2017; Sugano et al., 2007). However, mutagenesis studies have shown that this residue is not crucial for activity in, e.g., *Bacillus subtilis Bs*DyP, *Pseudomonas putida Pp*DyP, *R. jostii* DypB, *S. lividans* DtpA and DtpAa (Jones, 2001; Lučić et al., 2020b; Lučić et al., 2022; Mendes et al., 2015a; Mendes et al., 2015b; Shrestha et al., 2021; Singh et al., 2012). On the other hand, removing the arginine in some DyPs leads to a significant decrease in activity, as shown in *R. jostii* RHA1 DypB, *P. putida Pp*DyP and *S. lividans* DtpB (Lučić et al., 2020b; Mendes et al., 2015a; Singh et al., 2012). For *Vibrio cholerae Vc*DyP, both aspartate and arginine are reported to be catalytically relevant (Uchida et al., 2015).

Recently, *S. lividans* DyPs structures of the Fe(III) and Fe(IV)=O redox states were determined, showing that in DtpA and DtpAa, a water

molecule is hydrogen-bonded to the catalytic aspartate facilitating the rapid formation of Cpd I (Lučić et al., 2020a; Lučić et al., 2020b). In contrast, in *S. lividans* DtpB, the presence of a distal asparagine leads to a water-free distal heme site where arginine functions as a catalytic base (Lučić et al., 2020b). Therefore, it was proposed that DyPs containing a 'wet' distal site choose aspartate, and a 'dry' distal site choose the arginine over the aspartate to facilitate Cpd I production. However, more recently, it was shown that the replacement of the distal asparagine by alanine in DtpB results in a 'wet' distal heme site that nonetheless does not facilitate proton transfer, indicating that the catalytic arginine is involved in Cpd I formation, whether the distal heme site is 'wet' or 'dry' (Lučić et al., 2022).

The first spectroscopic identification of the intermediates Cpd I and Compound II (Cpd II) was in *T. cucumeris* Dec 1 *Tc*DyP (Sugano et al., 2007), and the first estimation of rate constants assuming the two-step (classical) catalytic mechanism **(Figure 1.8)** was performed on DypB from *R. jostii* RHA1 (Ahmad et al., 2011). Since these early studies, the catalytic mechanism has been investigated in other DyPs such as *B. subtilis Bs*DyP (Mendes et al., 2015a), *Thermomonospora curvata Tc*DyP (Chen et al., 2015), *E. lignolyticus El*DyP (Shrestha et al., 2017), *P. putida Pp*DyP (Brissos et al., 2017; Mendes et al., 2015a), *R. jostii* RHA1 DypB (Shrestha et al., 2021), *S. lividans* DtpA, DtpAa, DtpB (Lučić et al., 2020a; Lučić et al., 2020b; Lučić et al., 2021; Lučić et al., 2022), and *Klebsiella pneumoniae Kp*DyP (Pfanzagl et al., 2018).

The first step of the classical peroxidase catalytic mechanism is the deprotonation of hydrogen peroxide by a distal acid-base catalyst, aspartate, or arginine in DyPs **(Figure 1.8)**.

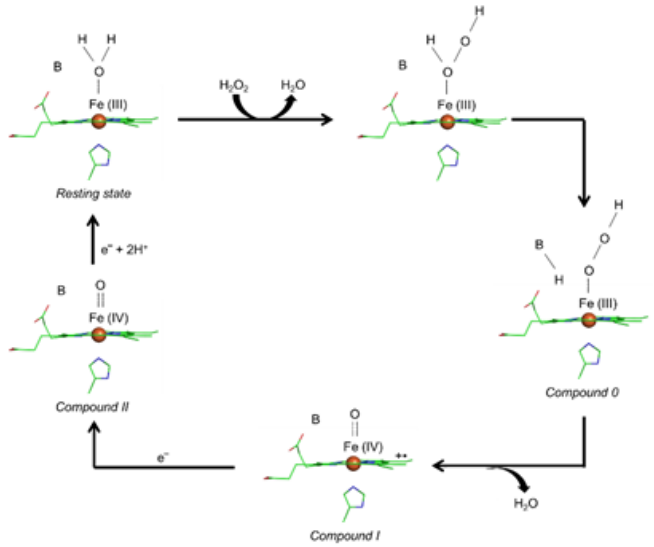


Figure 1.8. Proposed catalytic mechanism of DyPs based on the classical mechanism. The resting-state Fe(III) binds H_2O_2 and its deprotonation by an acid-base catalyst (B), forming Cpd 0. After O-O scission, the Cpd I is formed and will decay into Cpd II after one electron reduction. The Cpd II, after receiving a second electron from a substrate, will conduct to Fe(III) resting state.

The catalytic residue helps remove the proton from the proximal oxygen atom of the Fe(III)–O_2H_2 complex to generate an anionic precursor to Cpd I, Fe(III)-O-O-H, known as Compound 0 (Cpd 0). The proton removed is then transferred to the distal oxygen of Cpd 0, forming an oxy-water complex (Fe(III)–O–OH_2), promoting the heterolytic cleavage of the O-O a bond that results in the formation of $(Fe(IV)=O)^{+\bullet}$, designated as Cpd I, and the release of a water molecule. Cpd II, (Fe(IV)=O), is formed via the one-electron reduction of the protein by a reducing substrate, that by a second one-electron reduction regenerates the ferric iron to its resting state Fe(III). With the reaction cycle, two water molecules and two oxidized substrates are produced. Interestingly, the intermediate precursor of Cpd I, Cpd 0, was easily observed in the reaction of *B. subtilis Bs*DyP with H_2O_2 (Mendes et al., 2015b) and later identified in other DyP enzymes (Brissos et al., 2017; Pfanzagl et al., 2018; Shrestha et al., 2017; Shrestha et al., 2021).

In *E. lignolyticus El*DyP, a 'dry' Cpd I is more prone to direct two-electron reduction to the resting state (Shrestha et al 2017), similar to *Dictyostelium discoideum Dy*PA, where Cpd I is stable and does not decay into Cpd II, returning to the resting state as assessed by stopped-flow measurements (Rai et al., 2021). To test if the presence of water molecules in the distal heme pocket of DyPs influence the redox pathway of Cpd I, the *S. lividans* DtpB was investigated utilizing a combination of serial femtosecond crystallography and rapid kinetic studies: it was shown that a 'wet' site, characterized by distal bound water, is more prone to a one-electron reduction pathway, leading to a faster Cpd I one-electron reduction resulting in a high concentration of Cpd II that is then reduced to the ferric state **(Figure 1.9, left)** (Lučić et al., 2022), as it was observed in *R. jostii* DypB (Shrestha et al., 2021), *T. curvata Tc*DyP (Chen et al., 2015), *B. subtilis Bs*DyP (Mendes et al., 2015b) and *P. putida Pp*DyP (Brissos et al., 2017; Mendes et al., 2015a).

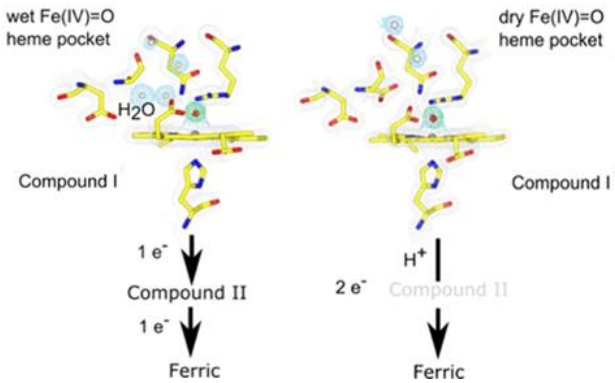


Figure 1.9. The proposed catalytic mechanisms for DyPs. Left: In a wet site, the reduction of Compound I is faster, has no deuterium effect, and yields highly populated Compound II, which is subsequently reduced to the ferric form. Right: Compound I reduction proceeds through a low Compound II concentration in a dry distal heme site. Reduction is coupled with proton uptake, and the resulting protonated Compound II (Fe(IV)–OH) rapidly reduces to the ferric state. Adapted from Lucic et al 2022.

Alternatively, a ‘dry’ distal pocket favors the Cpd I two-electron reduction to a protonated Cpd II (Fe(IV)−OH) that is promptly reduced to the ferric state **(Figure 1.9 right)**, similar to the mechanistic pathway observed in the P-class *El*DyP from *E. lignolyticus* and *D. discoideum Dy*PA (Rai et al., 2021; Shrestha et al., 2017).

1.4.2. Long-Range Electron Transfer

DyP-type peroxidases can oxidize bulky substrates as anthraquinone dyes and lignin models at distal sites. Electrons are transferred from the substrate via surface-exposed oxidation sites to the heme through LRET. These solvent-exposed regions usually enclose redox-active tryptophan and tyrosine residues that are known to form radicals and thus been responsible for the oxidation of the substrate (Gray and Winkler, 2021; Linde et al., 2015a; Linde et al., 2015b; Shrestha et al., 2016; Sugano and Yoshida, 2021; Uchida et al., 2015). Generally, the surface substrate binding sites in DyPs seem to fall under two categories: (i) radical-forming aromatic residues that allow for the oxidation of veratryl alcohol, bulky substrates, and dyes, and (ii) negatively charged patches that allow for the oxidation of Mn(II).

These surface oxidation sites have been previously identified in fungal LiP, VP, and MnP (Gelpke et al., 2002; Ruiz-Duenas et al., 2009; Ruiz-Duenas et al., 2007; Saez-Jimenez et al., 2016). A unique Cβ-hydroxylated tryptophan residue on the surface of a LiP from *Phanerochaete chrysosporium* was implicated in electron transfer from natural substrates to the heme cofactor (Doyle et al., 1998; Gelpke et al., 2002; Ruiz-Duenas et al., 2007). Similarly, the molecular architecture of *Pleurotus eryngii* VP exhibits an exposed tryptophan that is the putative responsible for aromatic substrate oxidation (Ruiz-

Duenas et al., 2007), connecting to the heme group via LRET (Fernandez-Fueyo et al., 2014).

Interestingly, while VP, LiP, and MnP enzymes include 1 to 5 aromatic residues in their sequence, DyPs are richer in containing between 7 and 20 tyrosine and tryptophan residues (Linde et al., 2015a; Linde et al., 2015b; Rodrigues et al., 2021; Shrestha et al., 2016). DyPs from class P show higher aromatics residues, followed by class V and class I **(Figure 1.10)**.

Classes P and V display more tyrosines, whereas class I have more tryptophans **(Figure 1.10)**. Class V DyPs are the ones that present the higher number of aromatics residues that are solvent-exposed (Rodrigues et al., 2021). A combination of multifrequency electron paramagnetic resonance (EPR) spectroscopy, site-directed mutagenesis, QM/MM (quantum mechanics/molecular mechanics), and molecular dynamics (MD) simulations have been applied to identify radical sites in surface-exposed aromatic residues and predict LRET pathways in, e.g., *Auricularia auricula-judae* DyP (*Aau*DyP) (Baratto et al., 2015; Linde et al., 2015a; Strittmatter et al., 2013a; Strittmatter et al., 2015; Strittmatter et al., 2013b), *S. lividans* DtpA (Chaplin et al., 2019), *K. pneumoniae* DyP (*Kp*DyP) (Nys et al., 2021), and *D. discoideum Dy*PA (Rai et al., 2021).

In agreement with the negative patches involved in Mn(II) oxidation in classical MnPs, in *R. jostii* RHA1 DypB, a constellation of surface negatively charged residues putatively constitutes the binding site for Mn(II) (Ahmad et al., 2011). *Amycolatopsis sp*. 75iv2 DyP2 oxidizes Mn(II) at the same efficiency as fungal MnP and VP, and an Mn(II) binding site was identified in the structure of this enzyme, with a potential redox active Tyr serving in the communication between the Mn binding site and the heme pocket (Brown et al., 2011).

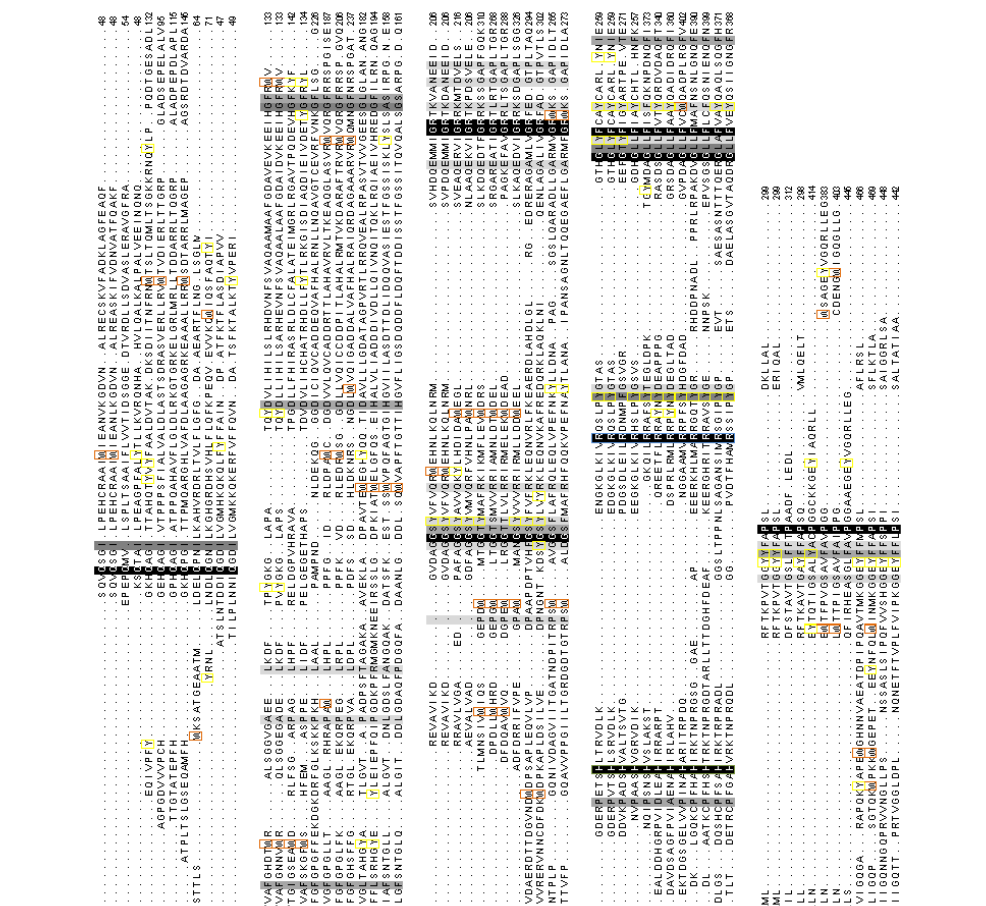


Figure 1.10. Alignment based on 3D structural superposition of DyP-type peroxidases (identified by classes and PDB codes). The tyrosine and tryptophan residues are underlined as yellow and orange boxes, respectively. Catalytic aspartate and arginine residues are highlighted with red and blue boxes, respectively. The proximal histidine ligand is marked with a green box. Strictly conserved amino acids are represented in black boxes, whereas dark and light grey boxes represent the most and less conserved residues among the selected sequences.

Substrate diffusion simulations, EPR, and site-directed mutagenesis performed at *Pleurotus ostreatus Pos*DyP4 allowed the identification of four surface negatively charged residues (three Asp and one Glu) involved in Mn(II) binding and oxidation in the presence of H_2O_2 **(Figure 1.11)** (Fernández-Fueyo et al., 2018). Since the Mn(II) ion is far from the heme cofactor, a tyrosine residue (Tyr339) was identified as a seemingly crucial residue for the electron transfer mechanism since its removal resulted in total loss of Mn(II) oxidation. A second pathway for electron transfer was identified by the formation of a tryptophanyl radical (Trp405) at the enzyme's surface. It was associated with the oxidation of aromatic substrates and dyes (Fernández-Fueyo et al., 2018).

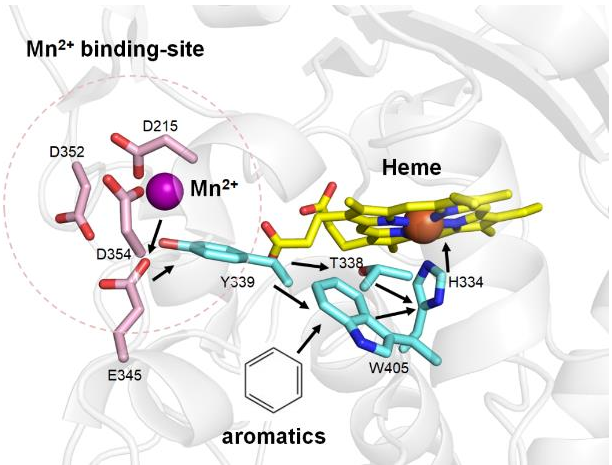


Figure 1.11. Proposed long-range electron transfer pathways in *Pos*DyP4.The acidic residues that are part of the Mn^{2+} site are represented as sticks with carbon atoms colored in light pink. The aromatic residues Y339 and W405, T338, and H334, are represented as sticks with carbon atoms colored in cyan. The LRET paths from the Mn^{2+} and aromatics oxidation are shown as arrows. The heme is shown as sticks with carbon, oxygen, and nitrogen atoms colored yellow, red, and blue. The manganese and iron ions are shown as purple and orange spheres, respectively. Adapted from (Fernandez-Fueyo et al., 2018).

1.5. General Biochemical Properties

The efficiency of DyPs to oxidize anthraquinone dyes varies depending on the classes, with class V exhibiting the highest catalytic efficiencies, ranging from 10^5 to 10^7 $M^{-1}s^{-1,}$ followed by Class I, ranging from 10^4 to 10^6 $M^{-1}s^{-1}$ and Class P DyPs ranging from 10^2 to 10^5 $M^{-1}s^{-1}$ (Shrestha, 2017).

DyPs have low activity towards azo dyes, a feature observed with *Aau*DyPI, *Bs*DyP and *Pp*DyP that oxidized azo dyes such Reactive Black 5 and Mordant Black 9 with catalytic efficiencies of 3 to 4 orders of magnitude lower (Linde et al., 2015a; Santos et al., 2014). One exception seems to be *Ana*PX, that oxidized Reactive Black 5 with a remarkable efficiency of 1.2 x 10^7 (Ogola et al., 2009). Curiously, the addition of syringaldehyde as a mediator to reaction with azo dyes further enhanced decolorization.

This wide spread of catalytic efficiencies is also observed for the oxidation of the standard peroxidase substrates, such as 2,2-diaza-bis (3-ethyl-benzothiazole-6-sulfonic acid) diammonium salt (ABTS). The efficiencies reported for ABTS range from 10^3 to 10^7 $M^{-1}s^{-1}$ with fungal DyPs exhibiting in general the high activity values.

As previously mentioned, DyPs also oxidize phenolic substrates, as well non-phenolic methoxylated aromatics (e.g., veratryl alcohol). Numerous studies have shown that DyPs will maintain a high activity under acidic conditions, between 4.0 and 5.0. The optimal temperature of DyPs is around 40–60 °C. *Tfu*DyP maintained maximum reaction rates between 45–65 °C (Rahmanpour et al., 2016) which goes in line with *Thermobifida fusca*'s growth temperature of 55 ºC.

The enhancing effect of Mn^{2+} was observed for DyPB, that exhibited increased activity towards various substrates such as Kraft lignin and ABTS upon addition of $MnCl_2$ (Ahmad et al., 2011) showing that this

compound can act as a redox mediator. On the other hand, the effect of several inhibitors such as of aminotriazole, EDTA, imidazole, DTT, Cys, and sodium azide on enzyme activity was showed (Lončar et al., 2013; Lončar et al., 2016). Aminotriazole is less inhibitory since DyP can still maintain around 60% of the activity. Still, the final complete inhibitory concentration of aminotriazole, EDTA, and imidazole has not been tested, and the tolerance of DyPs from different sources to inhibitors varies greatly. The stability of DyP is still poorly characterized, with different, non-comparable assays done on very few enzymes. At this moment, the only melting temperatures reported for DyPs are 44 ºC for *Cellulomonas bogoriensis Cbo*DyP (Habib et al., 2019) and 62 ºC for *Bs*DyP (Rodrigues et al., 2021).

1.6. Engineering for Improved Properties

The large substrate scope of DyP-type peroxidases makes these enzymes an excellent starting point for generating interesting biocatalysts for the biorefinery and bioremediation fields. Furthermore, bacterial systems are largely unexplored but may provide a rich source of new ligninolytic and auxiliary enzymes. They hold a high potential considering the ease of gene cloning, protein production, and the many molecular tools available for enzyme engineering. The generation of tailor-made enzymes represents a proficient way to design highly efficient and stable biocatalysts to support the major limiting factors for their industrial application; thermostability, resistance to organic solvents, extremes of pH (acid or alkaline), and inhibitors. A summary of engineering studies of DyPs can be found in **Table 1.2**.

In 1999, Cherry et al. applied a combination of directed evolution, with site-directed mutagenesis and gene recombination, to engineer the

fungal DyP CiP from *Coprinus cinereus* to laundry detergent conditions. The last hit variant exhibited 174-fold higher thermal stability and 100-fold higher oxidative stability at pH 10.5 and 50ºC as compared with the wild type enzyme that was rapidly inactivated at high pH (10.5), high temperatures (50ºC) and increased hydrogen peroxide concentrations (5-10 mM) (Cherry et al., 1999). In a DyP from *Anabaena sp.,* the replacement of methionine residues in the heme cavity resulted in mutants with up to 8-fold higher H_2O_2 resistance (Ogola et al., 2010). DyP4, a peroxidase from *P. ostreatus,* was fused with the *Escherichia coli* osmotically-inducible protein Y for extracellular secretion and engineered by directed evolution and saturation mutagenesis, resulting in a variant with higher protein yields, higher catalytic efficiency and also a higher hydrogen peroxide resistance (Alessa et al., 2019).

The biosynthesis of molecules containing stereogenic centers is of high bio-industrial interest; for example, chiral sulfoxides, containing a sulfur atom as a chiral center, have a wide range of applications, from chiral auxiliaries to pharmaceuticals, e.g., s-omeprazole is a multi-billion-dollar drug produced by a modified cyclohexanone monooxygenase. Therefore, using a structural-based design, the access to the heme pocket of *Aau*DyiP from *A. auricula-judae* was engineered to perform sulfoxidation better (Linde et al., 2016); the access to the heme was enlarged for better accommodation of sulfide substrates, which become positioned closer to the reactive cofactor with more favorable interaction energies. Site-saturation mutagenesis was also used to change the active site of *Pseudomonas fluorescens* Dyp1B, resulting in variants capable of (i) oxidizing 2-chlorophenol with up to 8-fold higher efficiency, (ii) oxidizing Mn(II) with up to 7-fold higher efficiency and (iii) exhibiting higher thermostability (Rahman Pour et al., 2019). Variants of *P. sapidus Psa*POX with increased

activity were also identified in a library of 101 basidiospore-derived monokaryons, which was traced to a single mutation resulting in higher activity and stability (Krahe et al., 2021). DyPB from *R. jostii* RHA1 activity for Mn(II) was improved 80-fold after the replacement of an asparagine, found within a pocket of negatively charged residues at the heme edge, to an alanine (Singh et al., 2013).

Taking myoglobin as a starting point, Zhang and colleagues engineered this enzyme to become an artificial DyP by inserting a Tyr/Trp chain on the surface near the heme pocket (Zhang et al., 2019). After three rational designs, the authors obtained an artificial DyP with catalytic efficiency compared with other native DyPs for Reactive Blue 19 (*circa* $1.2x10^6$ M^{-1} s^{-1}). Molecular docking studies suggested that the replacement of a Pro by a Trp at the protein surface was crucial for substrate binding and oxidation, acting as a catalytic site. Likewise, a cytochrome *c* was converted into a DyP-type peroxidase with the goal of optimizing the residual dye-decolorizing features of cyt c towards the levels of a DyP. By focusing surface residues, usually connected with the binding of large molecules such as dyes, and the axial histidine, the authors generated artificial enzymes with 3 and 80-fold improved activity towards Reactive Blue 19 (Omura et al., 2022).

A main feature of DyPs is the acidic optimal pH (3-4) for enzymatic activity, which represents one of the most relevant drawbacks of applying these enzymes in industrial setups. Uchida *et al*. designed an artificial DyP with an optimal pH shifted from pH 4.5 to 7 for Reactive Blue 19. Site-directed mutagenesis studies of DyP-type from *V. cholerae* (*Vc*DyP) showed that distal amino acid residues do not play a key role in the pH dependence of activity, but rather the hydrogen bond between His and Asp residues placed at around 3 Å from the heme (Uchida et al., 2021). Similarly, the *Pp*DyP from *P. putida* was

engineered using directed evolution for improved activity towards 2,6-dimethoxyphenol (DMP), a lignin-related phenolic compound, using random mutagenesis by epPCR followed by high-throughput screening. A hit variant (6E10) showed 100-fold higher activity towards DMP, higher hydrogen peroxide resistance, and an up-shift of pH optimum towards pH 8 (Brissos et al., 2017). This is so far the only known peroxidase that shows an optimal pH at the neutral to alkaline pH range.

Similarly, an engineered *Bs*DyP variant revealing 7-fold higher activity for DMP showed an increased loops flexibility and a widening of the heme cavity entrance that resulted in a concomitantly decreased overall stability, indicating a trade-off between functionality and stability (Rodrigues et al., 2021). As previously mentioned, DyPs can function as physiologic cargo proteins for encapsulin, a bacterial nanocompartment protein. Taking advantage of this, Lončar and colleagues packaged *Svi*DyP from *Saccharomonospora viridis* DSM43017 with the encapsulin from mesothermophile bacterium *Mycolicibacterium hassiacum* (EncMh) and showed that the packaged DyP showed comparable activity to the non-packaged version, but a significantly higher operational stability at 40 ºC: the unpackaged DyP lost its activity within 30 min, while the packaged showed a decrease only after 25 h (Lončar et al., 2020). his study proves that DyPs can be packaged in a functional form in encapsulin and be used as highly stable encapsulated biocatalysts, an appealing vehicle to be exploited, for example, as a drug delivery system.

Table 1.2. Enzyme engineering tools applied to DyPs for improved properties.

	Enzyme	Enzyme Engineering	Improved Features	Reference
Class P	*R. jostii* RHA1 DyPB	Rational Design	80-fold higher activity for Mn^{2+}	(Singh et al., 2013)
	P. putida MET94 *Pp*DyP	Directed Evolution	100-fold higher activity towards DMP, higher hydrogen peroxide resistance, up-shift of pH optimum towards pH 8	(Brissos et al., 2017)
	P. fluorescens DyP1B	Site-Saturation Mutagenesis	8-fold higher catalytic efficiency for 2-chlorophenol; 7-fold higher catalytic efficiency for Mn (II); higher thermostability	(Rahman Pour et al., 2019)
	P. putida MET94 *Pp*DyP	Directed Evolution	Higher sensitivity and shorter response times compared to other H_2O_2 biosensors reported	(Barbosa et al., 2020)
	*V. cholerae Vc*DyP	Rational Design	Optimum pH for degradation of RB19 shifted to 7	(Uchida et al., 2021)
Class I	*B. subtilis Bs*DyP	Directed Evolution	7-fold higher activity towards DMP and higher protein yield	(Rodrigues et al., 2021)
	*S. viridis Svi*DyP	Molecular packaging in encapsulin EncMh	Higher activity for ABTS and higher stability at 40 ºC	(Lončar et al., 2020)
Class V	*C. cinereus* CIP	Directed Evolution, DNA Shuffling, and Rational Design	174-fold higher thermal stability and 100-fold higher stability for H_2O_2 at pH 10.5 and 50 ºC	(Cherry et al., 1999)
	*Anabaena sp Ana*PX	Rational Design	8-fold higher H_2O_2 resistance	(Ogola et al., 2010)
	*A. auricula-judae Aau*DyP	Rational Design and Computational Modelling	Sulfoxidation of substrates thioanisole (methyl-phenyl sulfide, MPS) and methyl-p-tolyl sulfide (MTS)	(Linde et al., 2016)
	*P. ostreatus Pos*DyP4	Directed Evolution and Site-Saturation Mutagenesis	Higher protein yield and secretion, higher catalytic efficiency for ABTS and higher H_2O_2 tolerance	(Alessa et al., 2019)
	*P. sapidus Psa*POX	Basidiospore-derived monokaryons with intraspecific variability	2-fold higher peroxidase and alkene cleavage activity (using *trans*-anethol), and higher stability in sub-cultivation	(Krahe et al., 2021)

1.7. Biotechnological Applications

Due to their versatility, multi-functionality, and ability to react with many substrates DyPs have been associated with multiple biotechnological applications. This has been explored and the results mentioned in this chapter are summarized in **Table 1.3.**

Synthetic dyes are chemically stable xenobiotics used in the manufacture of textiles, cosmetics, foods, and pharmaceuticals. Azo and anthraquinone dyes are the most frequent chemicals in dye-contaminated wastewater that must be treated before being discharged into the environment. In addition to aesthetic problems, this represents a liability since dyes and their breakdown products are toxic and potentially mutagenic to living organisms (Rawat et al., 2016). Many enzymes from different plants and microorganisms have been reported to play an essential role in various waste treatment applications (Mendes et al., 2015c). This is mainly because, unlike chemical catalysts, enzymes successfully convert complex chemical structures under mild environmental conditions with high efficiency. For instance, *Ana*PX from *Anabaena* sp. decolorized anthraquinone Reactive Blue 5 (RB5), Reactive Blue 4, Reactive Blue 114, Reactive Blue 119, and Acid Blue 45 (Ogola et al., 2009); DyP from *T. cucumeris* Dec 1 was able to decolorize RB5 to light red-brown compounds (Sugano et al., 2009); *Xgr*DyP *Xylaria grammica* oxidizes the synthetic dye RB5 (Kimani et al., 2021); *Pp*DyP decolorized nearly 80% of the azo dye Acid Red 299, while *Bs*DyP exhibited a decolorization efficiency of 76% for Direct Red R (Santos et al., 2014); *Vc*DyP from *V. cholerae* degraded Reactive Blue 19 (Uchida et al., 2015); *T. fusca Tfu*DyP reacted with Eosin Y, Acid Blue 129, and Crocetin (Lončar et al., 2016); *Sa*DyP1 and *Sa*DyP2 from *Streptomyces avermitilis* degraded Acid Blue 234 (Sugawara et al., 2017; Sugawara et al., 2021); *Ftr*DyP from *Funalia trogii* was reactive

towards the anthraquinone dye Reactive Blue 19 and the azo dye Reactive Black 5 (Kolwek et al., 2018); the SUMO-fused *Cbo*DyP wild type and E201D variant from *C. bogoriensis* showed the ability to decolorize Reactive Blue 19, Acid Red 14 and Disperse Blue 1 (Habib et al., 2019); *P. fluorescens* Pf0-1 *Pf*DyP B2 oxidized Reactive Blue 4 (Lončar et al., 2019). All these results point to the potential application of these enzymes in the bioremediation of dye-contaminated wastewater.

P. putida MET94 *Pp*DyP and the variants generated by directed evolution were studied in solution and immobilized states to design biosensors for H_2O_2 detection (Barbosa et al., 2020; Sezer et al., 2012). Sezer and colleagues showed that *Pp*DyP was electronically coupled to the Ag electrode, with preserved structural integrity and efficient catalytic activity while in 2020, Barbosa et al revealed that both *Pp*DyP wild type and variant 29E4 showed superior stability, sensitivity and shorter response times compared to other H_2O_2 biosensors reported in the literature, making them excellent candidates for developing biosensors in disposable and single-use configurations for various biological applications. Additionally, it was shown that *B. subtilis Bs*DyP and *P. putida Pp*DyP form long-lived oxyferryl species by the simple application of a reductive electrode potential under aerobic conditions that can oxidize ABTS over a broad pH range with similar efficiencies (Scocozza et al., 2021). These findings pave the way for the design of DyP-based electrocatalytic reactors operable in an extended pH range without the need for harmful reagents such as H_2O_2. The potential use of *Cbo*DyP, *Sco*DyP*,* and *Tfu*DyP as H_2O_2 biosensors was tested after immobilization on biocompatible silver electrodes and functionalized with alkanethiols (Zuccarello et al., 2021).

Table 1.3. Biotechnological applications of characterized DyPs.

	Enzyme	Enzyme	Biotechnological Application	Reference
Class P	*R. jostii* RHA1 DyPB	Wild type	Biorefineries; lignin depolymerization: oxidation of β-aryl ether lignin model and Mn(II)	(Ahmad et al., 2011)
	*V. cholerae Vc*DyP	Wild type	Degradation of Reactive Blue 19; bioremediation of contaminated wastewater	(Uchida et al., 2015)
	P. fluorescens DyP1B	Wild type	Oxidation of Mn (II) and Kraft lignin; biorefineries; lignin depolymerization to aromatic products	(Rahmanpour and Bugg, 2015)
	P. putida MET94 *Pp*DyP	Wild type	Decolorization of Mordant Black 9, Reactive Blue 62 and Acid Blue 62 dyes; bioremediation of contaminated wastewater	(Santos et al., 2014)
		Variant 6E10	Lignin depolymerization: oxidation Kraft lignin, phenolics and aromatic amines	(Brissos et al., 2017)
		Wild type and variant 29E4	Design of biosensors with higher sensitivity and shorter response times compared to other H_2O_2 biosensors reported	(Barbosa et al., 2020)
	*C. bogoriensis Cbo*DyP	Wild type and E201D variant	Decolorization of Reactive Blue 19, Acid Red 14 and Disperse Blue 1 dyes	(Habib et al., 2019)
	*P. fluorescens Pf0-1 Pf*DyP B2	Wild type	Decolorization of Reactive Blue 19 dye	(Lončar et al., 2019)
Class I	*B. subtilis Bs*DyP	Wild type	Decolorization of Mordant Black 9, Reactive Blue 62 and Acid Blue 62 dyes; bioremediation of contaminated wastewater	(Santos et al., 2014)
	*T. fusca Tfu*DyP	Wild type	Decolorization of Eosin Y, Acid Blue 129 and Cocetin dyes; bioremediation of contaminated wastewater; divanillin production for food industry	(Lončar et al., 2016)
	S. viridis DSM43017 *Svi*DyP	Wild type	Biobleaching for paper and pulp industry	(Yu et al., 2014)
		Fusion proteins P-HotAldO and P-ChitO	Sugar detection	(Colpa et al., 2017)
	*T. curvata Tc*DyP	Wild type	Lignin valorization: enzymatic activity towards four stereospecific β-O-4 dilignolçs	(Huang et al., 2017)
	S. thermocarboxydus 4129 *St*DyP	Wild type	Degradation of mycotoxins B_1(AFB_1), ZEN and DON in the presence of Mn^{2+}; decontamination of agricultural products	(Qin et al., 2021)
Class V	*C. cinereus* CIP	Variants 972 and 974	Laundry detergent environment	(Cherry et al., 1999)
	*Anabaena sp Ana*PX	Wild type	Decolorization of Reactive Blue 5, Reactive Blue 4, Reactive Blue 114, Reactive Blue 119, and Acid Blue 45; bioremediation of contaminated wastewater; lignin degradation: active towards several aromatic substrates such as guaiacol, 4-aminoantipyrine and pyrogallol	(Ogola et al., 2009)
	*F. trogii Ftr*DyP	Wild type	Decolorization of anthraquinone dye Reactive Blue 19 and the azo dye Reactive black 5.	(Kolwek et al., 2018)
		Wild type combined with FtrLcc	Food industry: grapefruit flavor constituent (+)-nootkatone	
	Amycolatopsis sp. 75iv2 DyP 2	Wild type	Modification of native and organosolv hardwood lignin to different extents;Lignin valorization	(Brown et al., 2012; Vuong et al., 2021)
	T. cucumeris Dec 1 DyP	Wild type	Decolorization of anthraquinone dyes namely Reactive Blue 5	(Sugano et al., 2009)
	*S. avermitilis Sa*DyP2	Wild type	Decolorization of anthraquinone dyes namely Acid Blue 234	(Sugawara et al., 2017)
	*S. avermitilis Sa*DyP1	Wild type	Decolorization of anthraquinone dyes namely Acid Blue 234	(Sugawara et al., 2021)
	*P. sapidus Psa*POX	Wild type	Fragrance and flavor industry: generation of aromatic aldehydes; food industry: β-carotene bleaching	(Krahe et al., 2020)
	*Agrobacterium sp. B1 Agro*DyP	Wild type combined with LigE	Lignin valorization: 2-methoxyhydroquinone, hydroxyquinol, vanillin and vanillic acid	(Rashid and Bugg, 2021)
	*X. grammica Xgr*DyP	Wild type	Lignin depolymerization: Oxidation of DMP, veratryl alcohol and Mn (II); Decolorization of Reactive Blue 5; Bioremediation	(Kimani et al., 2021)

The authors reported that only *Cbo*DyP preserved its native structure upon immobilizing the electrode. However, a great discrepancy was observed when the redox potentials of the enzyme in solution and immobilized states were compared, and *Cbo*DyP lost its catalytic and redox properties in the presence of H_2O_2. The bacterial peroxidase *Svi*DyP was fused with alditol oxidase (HotAldO) and chitooligosaccharide oxidase (ChitO). The bifunctional fused catalysts P-HotAldO and P-ChitO couple the oxidase-peroxidase activity, where the oxidase originates hydrogen peroxide that fuels the peroxidase, with the oxidase's specificity of substrates. Since HotAldO is active on alditols, for instance, xylitol and sorbitol, whereas P-ChitO is active towards mono-, di-, and oligosaccharides, such as glucose, lactose, cellobiose, and maltose, the fusion enzymes have potential as sugar detectors/biosensors (Colpa et al., 2017).

The role of DyP in the biodegradation and/or valorization of lignocellulosic material is grounded on the activity features observed for several enzymes. DyPs exhibited activity toward complex forms of lignin, such as β-aryl ether lignin (Rai et al., 2021). *Ana*PX was also active towards several aromatic substrates such as guaiacol, 4-aminoantipyrine and pyrogallol (Ogola et al., 2009). DypB from *R. jostii* RHA 1 exhibited oxidative activity towards β-aryl ether lignin model and Mn(II) (Ahmad et al., 2011). Similarly, the fungal DyP from *X. grammica Xgr*DyP, was shown to oxidize several substrates, including DMP, veratryl alcohol and Mn(II), revealing its possible involvement in lignin depolymerization and conversion, particularly in fungi that lack the typical ligninolytic peroxidases. (Kimani et al., 2021) The DyP from *B. subtilis* KCTC2023 produced veratryl aldehyde and guaiacol from the oxidation of veratryl glycerol-b-guaiacyl ether (VGE) (Min et al., 2015), while DyP2 and 6E10, a *Pp*DyP variant, oxidized guaiacylglycerol-β-guaiacyl ether (GGE) (Brissos et al., 2017; Brown et

al., 2012). *Tc*DyP from *T. curvata* displayed significant activity for different β-O-4 dilignols used as lignin models (Huang et al., 2017), showing that *Tc*DyP will not discriminate against the stereochemistry of the source biomass. Syringaresinol, a lignan formed from two sinapyl alcohol units, was produced in a one-pot conversion with 81% yield by coupling a eugenol oxidase (EUGO) with horseradish peroxidase (HRP). Starting from 2,6-dimetoxy-4-allyphenol, EUGO converts it into sinapyl alcohol, generating H_2O_2, which is afterward used by HRP to produce syringaresinol. The last step in this one-pot conversion reaction can also be performed by a DyP, since they show activity for phenolic compounds, such as sinapyl alcohol (Habib et al., 2018). DyPs were tested for their potential in converting lignin into low molecular weight aromatic products in the presence of the accessory enzymes *Sphingobacterium* sp. T2 dihydrolipoamide dehydrogenase 1 and 2 (DHLDH1 and DHLDH2), *Burkholderia cenocepacia* peroxiredoxin, *Agrobacterium* sp. LigE, and *Desulfitobacterium hafniense* arylsulfotransferase (Rashid and Bugg, 2021). This was expected to prevent repolymerization or recondensation of the phenoxy radicals that originated from lignin oxidation. DyP1B from *P. fluorescens*, *Ct*DyP from *Comamonas testosteroni* TK102, and *Agro*DyP from *Agrobacterium* sp. B1 were incubated with a lignin model, and it was clear that the presence of accessory enzymes promoted the formation of monomeric products, including the monocyclic aromatics 2-methoxyhydroquinone, hydroxyquinol, vanillin, and vanillic acid. The DyP2 from *Amycolatopsis sp.* 75iv2 alone or in combination with small laccase from the same microorganism transforms a variety of lignins, including native and modified organosolv hardwood lignin, generating product profiles that depend on the enzyme type and presence of mediator (Vuong et al., 2021). Biobleaching of kraft pulp is a process of high biotechnological interest for the pulp and paper industry. *Svi*DyP showed increased pulp

brightness when eucalyptus kraft pulp was prebleached at 65ºC and pH 7 (Yu et al., 2014). Additionally, scanning electron microscopy (SEM) studies highlighted the existence of local fractures and large areas containing holes that generated a looser fiber structure with the lignin internal structure exposed. This structure breakdown helps the permeation of bleaching agents and facilitates delignification to obtain molecules of interest. Combined, these results show that DyPs have the potential to be included in processes of the functionalization and valorization of lignin monomers in, for example, a biorefinery setup.

DyPs were also reported to have promising applications in the food industry, namely in producing vanilla, a flavor whose production is challenging and in high demand. Divanillin, a molecule valued as a taste enhancer, was identified with LC-MS as the main product resulting from the oxidation of vanillin mediated by *Tfu*DyP (Lončar et al., 2016). The fusion of the bacterial peroxidase *Svi*DyP with eugenol oxidase (EUGO), and 5-hydroxymethylfurfural oxidase (HMFO), two oxidases that act on monophenolic substrates, resulted in P-EUGO and P-HMFO (Colpa et al., 2017). These fusion enzymes were able to dimerize vanillyl alcohol into divanillin in a cascade reaction that includes the oxidation of vanillyl alcohol to vanillin by the oxidases followed by dimerization catalyzed by *Svi*DyP. The high-value grapefruit flavor constituent (+)-nootkatone was reported to be originated from (+)-valencene through a one-pot reaction catalyzed by a system involving a dye-decolorizing peroxidase (*Ftr*DyP) and a laccase (*Ftr*Lcc), obtained from the basidiomycete *F. trogii*, in the presence of Mn^{2+} and p-coumaric acid (Kolwek et al., 2018). In fact, (+)-nootkatone was the predominant volatile product. The basidiomycete *P. sapidus* PsaPOX was reported to cleave trans-anethole, an aryl alkene model (Krahe et al., 2020), an activity useful for the generation of aromatic aldehydes, such as *p*-anisaldehyde,

veratraldehyde, and acetophenone, that are commonly used as fragrance and flavor in the food industry. Additionally, *Psa*POX displayed the ability to bleach β-carotene, which can be applied in the bleaching of whey and wheat dough. Recently, DyPs from *B. subtilis* SCK6 and *Streptomyces thermocarboxydus* 4129 were shown to degrade different types of mycotoxins in the presence of Mn^{2+}, using the laccase/mediator system, such as aflatoxin B_1(AFB_1), zearalenone and deoxynivalenol in animal feed and food (Qin et al., 2021). These mycotoxins are widely present in contaminated agricultural products and foods, resulting in significant economic loss and a health threat to humans and animals.

1.8. Concluding Remarks and Thesis Outline

Since the discovery of the first DyP in *T. cucumeris* Dec 1, the curtain behind DyP biological functions and catalytic mechanism has been opening, revealing interesting properties such as its role in the life cycle of bacteria, an encapsulation mechanism, and a wet-dry dichotomy associated with their mechanism that needs further studies to paint a clear picture of how these enzymes operate.

While several studies have highlighted uniting features among the different classes of DyPs, such as the crucial role of the loops encircling the active site in the activity and stability of DyPs, others have shown the differences between these classes, e.g., the differences in catalytic efficiency. Regardless, these enzymes have been identified as potential catalysts in the emerging biorefinery industry due to their ability to oxidize lignin-related compounds and models. Recent works have tackled DyPs main limitations regarding its potential use: low activity to lignin-related compounds compared to other peroxidases and its preference towards acidic pH. By

engineering these enzymes, e.g., using directed evolution, DyP variants with improved activity towards phenolic compounds such as syringol and a neutral pH optimum have been generated, increasing the compatibility with other enzymatic systems, for the creation of multi-enzyme systems and/or fusion proteins to be used in lignin degradation and valorization.

This doctoral thesis is composed by six chapters, with the state of the art on DyPs being presented in this first one.

In **Chapter 2** we have studied the substrate scope of *Pp*DyP and its variant 6E10, with the aim of understanding how capable are these enzymes to oxidize lignin related phenolic compounds and what industrially-relevant products can we obtain from these reactions. We observed that the variant 6E10 oxidized 24 phenolic compounds with higher rates and identified interesting reaction products such as syringaresinol and dyapocynin.

In **Chapter 3** we perform the kinetic, biochemical and structural characterization of wild type *Pp*DyP and of two engineered variants, 6E10 and 29E4. This characterization is complemented with conformational and protonation analysis from constant-pH molecular dynamics simulations (CpHMD), that contribute to understanding the features that underlie the activity and stability of these enzymes. And throught the use of molecular docking, we obtain further details on enzyme interaction with substrates. We showed that that insertion of the mutations A142V leads to changes in the conformational flexibility of loop regions of PpDyP 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.

In **Chapter 4** we use *in silico* tools to design thermostable mutants of *Pp*DyP. We have performed the cloning, expression, production, purification and characterization of two new variants obtained through the web servers PROSS and FireProt and showed that the introduction of 29 and 21 mutations, respectively, resulted in significant improvements, such as higher production yields and activity and considerable improvements in thermodynamic and functional stability.

Chapter 2

A Wide Array of Lignin-Related Phenolics are Oxidized by an Evolved DyP at Alkaline pH

This chapter contains data published in:

Diogo Silva, Ana Catarina Sousa, Vânia Brissos, Maria Paula Robalo and Lígia O. Martins (2022). A Wide Array of Lignin-Related Phenolics are Oxidized by an Evolved DyP an at Alkaline pH. *N Biotechnol*.
https://doi.org/10.1016/j.nbt.2022.12.003

2.1. Abstract

In this chapter, we have investigated the catalytic potential of variant 6E10 using twenty-four phenol representative derivatives of lignin syringyl, guaiacyl, and hydroxybenzene derivatives. Electrochemical data using cyclic voltammetry was first assessed for substrates of interest. Then, a comparative analysis was performed combining spectrophotometric and HPLC enzymatic assays. 6E10 variant showed up to 100-fold higher specific activity and overall higher substrates conversion yields at pH 8 all tested phenolic substrates after 24 h of reaction compared with the wild type enzyme (reactions at pH 4). The identification of products of reactions was attempted by mass spectrometry and by NMR. Taken together, we demonstrate the potential of the evolved variant 6E10 for oxidation and conversion of lignin-related substrates for biotechnological purposes, showing that these enzymes can tackle the heterogeneity of lignin monomers while simultaneously yielding valuable products.

Keywords: Dye-decolorizing peroxidases, phenolic compounds, lignin valorization, biorefineries, added-value compounds

2.2. Introduction

The pulp and paper industry produces about 50 million tons of lignin annually, but most of this is burned for energy; only 1 million tons of lignin reach the chemicals market (Sun et al., 2018). Lignin is currently used for low- and medium-value applications (e.g., binding and dispersing agents), with energy capturing about 89% of the market. Recently, well-defined fractions in acceptable quantities were obtained by the implementation of novel strategies for lignin depolymerization, involving electrochemistry, photocatalysis, heterogeneous catalysis, and ionic liquids (Abu-Omar et al., 2021; Stone et al., 2018; Sun et al., 2018; Van den Bosch et al., 2018), and started new value-chains from the lignin-derived platform of chemicals **(Figure 2.1)**.

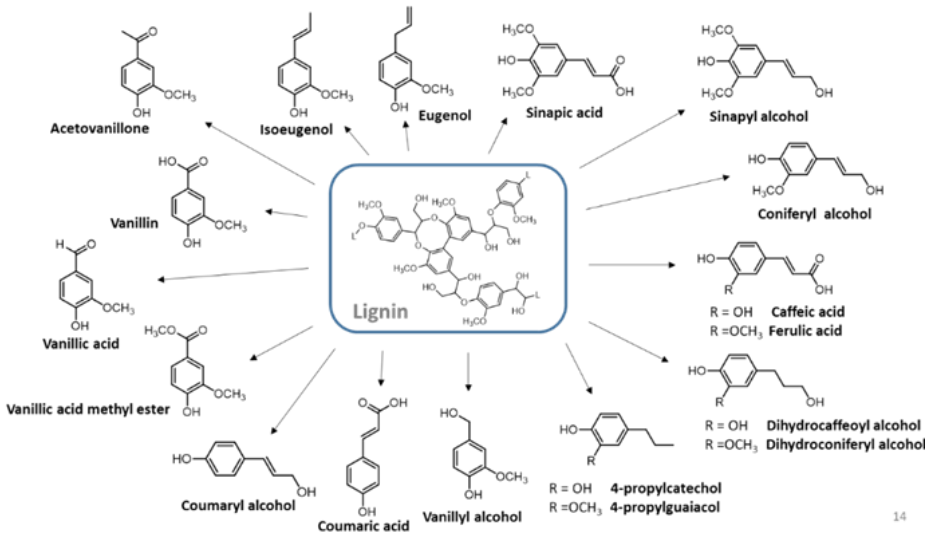


Figure 2.1. Phenolic lignin units available from biorefineries. Adapted from (Sun, Fridrich et al. 2018)

Therefore, at present, the challenge is the setup of economic and waste-free processes that allow the full implementation of lignin as sustainable starting material for the production of drop-in chemicals, polymers, or emerging functional materials (Borchert et al., 2022; Gall et al., 2017; Linger et al., 2014; Ragauskas et al., 2014; Sun et al., 2018).

Lignin-derived chemicals can significantly impact every aspect of our life (Borchert et al., 2022; Gall et al., 2017; Schutyser et al., 2018). For instance, lignans and neolignans are secondary plant metabolites extracted at very low yields that have been used as antioxidants, antitumors, anti-inflammatory, anti-neurodegenerative, antiviral, and antimicrobial agents (Zalesak et al., 2019). The central motif of these multifunctional compounds can be obtained by dimerization of phenylpropanoid units via oxidative coupling, a process mediated by laccases and peroxidases (Adelakun et al., 2012a; Hamalainen et al., 2018; Hammel and Cullen, 2008; Kudanga et al., 2017; Natte et al., 2020; Runeberg et al., 2019). For example, dimers formed upon oxidation of ferulic acid show antioxidant and anticancer agents (Adelakun et al., 2012b; Aljawish et al., 2014); diapocynin, a dimer of acetovanillone, has important anti-inflammatory and neuroprotective effects (Ismail et al., 2014); phellinsin A, a dimeric product of caffeic acid, is an antioxidant agent (Cardullo et al., 2022; Nemadziva et al., 2018); divanillin, a dimer of vanillin, is a taste enhancer (Krings et al., 2015); and pinoresinol, a dimer of coniferyl alcohol, is a hypoglycemic and chemopreventive agent (Wan et al., 2007). Furthermore, the functionalization of the hydroxyl groups of monolignols has been used to produce polymers, with vanillin, ferulic acid, guaiacol, syringaldehyde, or 4-hydroxybenzoic acid topping the list of most used lignin derivatives for these purposes (Sun et al., 2018). Indeed, the aromatic units offer rigidity, hydrophobicity, and fire resistance in a

polymeric backbone. These include thermosets, such as vinyl ester resins, cyanate ester, epoxy, and benzoxazine resins, as well as thermoplastic polymers, including polyesters, polyanhydrides, Schiff base polymers, polyacetals, polyoxalates, polycarbonates, and acrylate polymers (Llevot et al., 2016). Syringaresinol, formed upon oxidation of synapil alcohol, displays interesting bioactivity properties (Habib et al., 2018; Jaufurally et al., 2016b) and is also of great interest to polymer research (Chio et al., 2019; Janvier et al., 2016; Janvier et al., 2017; Llevot et al., 2016).

*P. putida Pp*DyP is an enzyme that oxidizes anthraquinone and azo dyes with high efficiency and, to a lower extent, metal ions and phenolic and non-phenolic lignin-related compounds (Santos et al., 2014). A directed evolution approach resulted in an engineered variant, 6E10, displaying improved catalytic efficiency towards the lignin-related phenol DMP (2,6-dimethoxyphenol, also known as syringol), a shift in pH optima from 4.3 to 8 and an increased resistant to catalytic inhibition by hydrogen peroxide (Brissos et al., 2017). In this work, we elucidate the substrate scope of 6E10 for lignin-related phenolic substrates by screening twenty-four phenolic compounds from the p-hydroxyphenyl, syringl, and guaicyl types and evaluate its potential for biotechnological applications.

2.3 Material and Methods

2.3.1. Enzymes and Chemicals

Recombinant wild type *Pp*DyP and evolved variant 6E10 were produced and purified as previously described (Brissos et al., 2017; Santos et al., 2014). Acetosyringone, syringaldehyde, acetovanillone, vanillyl alcohol, guaiacol, vanillin, ferulic acid, coniferyl aldehyde, 4-hydroxyacetophenone, caffeic acid and p/m/o- coumaric acids were purchased from Sigma-Aldrich. Syringol, methylsyringate, synapic acid, and syringic acid were purchased from Alfa Aesar. Gallic acid was purchased from Merck, and Vanillic acid was purchased from Fluka analytics.

2.3.2. Cyclic Voltammetry measurements

The redox potentials of the phenolic were measured by cyclic voltammetry using an EG&G Princeton Applied Research (PAR) Model 273A potentiostat/galvanostat monitored with Electrochemistry PowerSuite v2.51 software from PAR. Cyclic voltammograms of 1 mM of phenolics preparations in 0.1 M acetate, pH 4 and phosphate buffer, pH 8 were obtained using a three-electrode configuration cell with a glassy carbon (GC) disk working electrode (3.0 mm diameter, Radiometer analytical, SAS, France), a platinum wire counter electrode and an aqueous Ag/AgCl/saturated KCl reference electrode (Radiometer analytical, SAS, France). For aqueous insoluble compounds, 10% of ethanol (v/v) was added. The working electrode was carefully polished with alumina powder before each measurement. The potential was scanned from -0.3 to 1.4 V at a scan rate of 100 mV/s. All measurements were performed at room temperature and the solutions were flushed with dinitrogen before use. The measured

potentials were corrected by +0.197 V to the standard hydrogen electrode (NHE).

2.3.3. UV-visible spectra and determination of molar extinction coefficients of phenolic compounds

The UV-visible absorption spectra of phenolic compounds were recorded in 0.1 M citrate buffer, pH 4, and 0.1 M phosphate buffer, pH 8, in the wavelength range 200-900 nm, using a Synergy 2 (Bioteck) microtiter plate reader. The molar extinction coefficients (in the concentration range of 10-200 µM) were determined at their maximal wavelengths using the Lambert-Beer law ($A = \varepsilon \times l \times c$; where A is the absorbance, l the optical length (0.65 cm) and c the compound concentration) at pH 4 and pH 8 using a Synergy 2, Biotek, 96-well microplate reader.

2.3.4. Assays of enzymatic activity of PpDyP and 6E10

Enzymatic reactions were performed at 30 °C in 96 well plates with 0.5 mM of each substrate at 0.1 M citrate buffer, pH 4 for reactions with wild type *Pp*DyP (0.2 mg/mL in the reaction) and 0.1 M phosphate buffer, pH 8 for reactions with variant 6E10 (0.1 mg/mL in the reaction). Reactions were started by adding 0.5 mM H_2O_2 and were monitored at each substrate's maximal wavelength except for syringol, where product formation was monitored. Apparent steady-state kinetic parameters were measured at substrate concentrations from 0.01 to 0.5 mM in the presence of 0.5 mM H_2O_2.

2.3.5. High-performance liquid chromatography (HPLC)

Reactions were set up using 0.1 M citrate buffer, pH 4 for wild type and 0.1 M phosphate buffer pH 8.0 for 6E10, with 0.5 mM of each substrate. The addition of 0.5 mM H2O2 started reactions, and after 24 h, these were stopped with 50% methanol before injection. Enzymatic oxidation was monitored by high-performance liquid chromatography (HPLC) performed using a Waters Alliance 2695 HPLC System with a Purospher STAR RP-18e column (125x4 mm), 5 µm particle size (Merck KGaA, Germany), at 40 ºC and a flow rate of 1.0 $mL \cdot min^{-1}$. Reaction products were eluted with a linear gradient of H2O/methanol plus 0.5% acetic acid (solvent A) from 30% to 80% of methanol over 25 min and maintained isocratic for 10 min, then returned to initial conditions for 2 min and maintained isocratic for 8 min. The absorption was monitored between 200 and 500 nm by a Waters Photodiode Array Detector 2996 operated by Empower Pro, version 5, 2002 (Waters Chromatography).

2.3.6. Set-up for enzymatic reactions using 6E10 for identification of products by Mass Spectrometry

The enzymatic reactions (0.5 mM substrate) were performed at room temperature and in a total volume of 1 mL, using 100 mM Ammonium bicarbonate buffer at pH 8, 6E10 (1 U/mL) and 0.5 mM H_2O_2, for 1 hour. The reactions were stopped by adding methanol in a similar volume to the volume of reaction, centrifugated, and the soluble fraction was stored at 4 °C until MS analysis. A 20 mg/ml sample solution was prepared in water (Optima™ LC/MS Grade, Fisher Scientific) and further diluted to 1:100 in 50% methanol (Optima™ LC/MS Grade, Fisher Scientific), 1 µL were injected on column. Chromatographic analysis was performed on an UltiMate 3000 UHPLC

(Thermo Scientific). The separation was performed using a Waters XBridge column C18 (2.1x150 mm, 3.5 µm particle size, P/N 186003023). The mobile phase A was water with 0.1 % formic acid (v/v), and mobile phase B was acetonitrile with 0.1 % formic acid (v/v) (Optima™ LC/MS Grade, Fisher Scientific). The gradient program is listed in **Table 2.1**. The column temperature was maintained at 30 ºC, and a flow rate of 400 µl/mim was used. The data was acquired on a Q Exactive Focus (Thermo Scientific) coupled to UHPLC, using Xcalibur software v.4.0.27.19 (Thermo Scientific). The method consisted of several cycles of Full MS scans (R=70000) in positive mode. External calibration was performed using LTQ ESI Positive Ion Calibration Solution (Thermo Scientific). The raw MS data was analysed using Qual Browser Xcalibur software v4.0 (Thermo Scientific). Data provided by the Mass Spectrometry Unit (UniMS), ITQB/iBET, Oeiras, Portugal.

Table 2.1. Gradient program of the method.

Time (mim)	Mobile phase A (%)	Mobile phase B (%)
1	99	1
13	1	99
15	1	99
16	99	1
20	99	1

2.3.7. Enzymatic reactions for identification of products by 1H NMR

The enzymatic reactions (5 mM substrate) were performed at room temperature in a total volume of 10 mL, using 100 mM potassium phosphate buffer at pH 8, 6E10 (5 U/mL), and 5 mM H_2O_2, for 1 h. The reactions were stopped by adding methanol and the solvent was evaporated using a Büchi Rotavapor R-205. The residues were resuspended in methanol, filtered and the solvent evaporated for 1^H NMR analysis. 1H NMR spectra were recorded on a Bruker Avance (400 MHz) spectrometer in CD_3OD-d_4 as solvent with a 5 mm probe. Signals were referenced to the residual signal of the deuterated solvent.

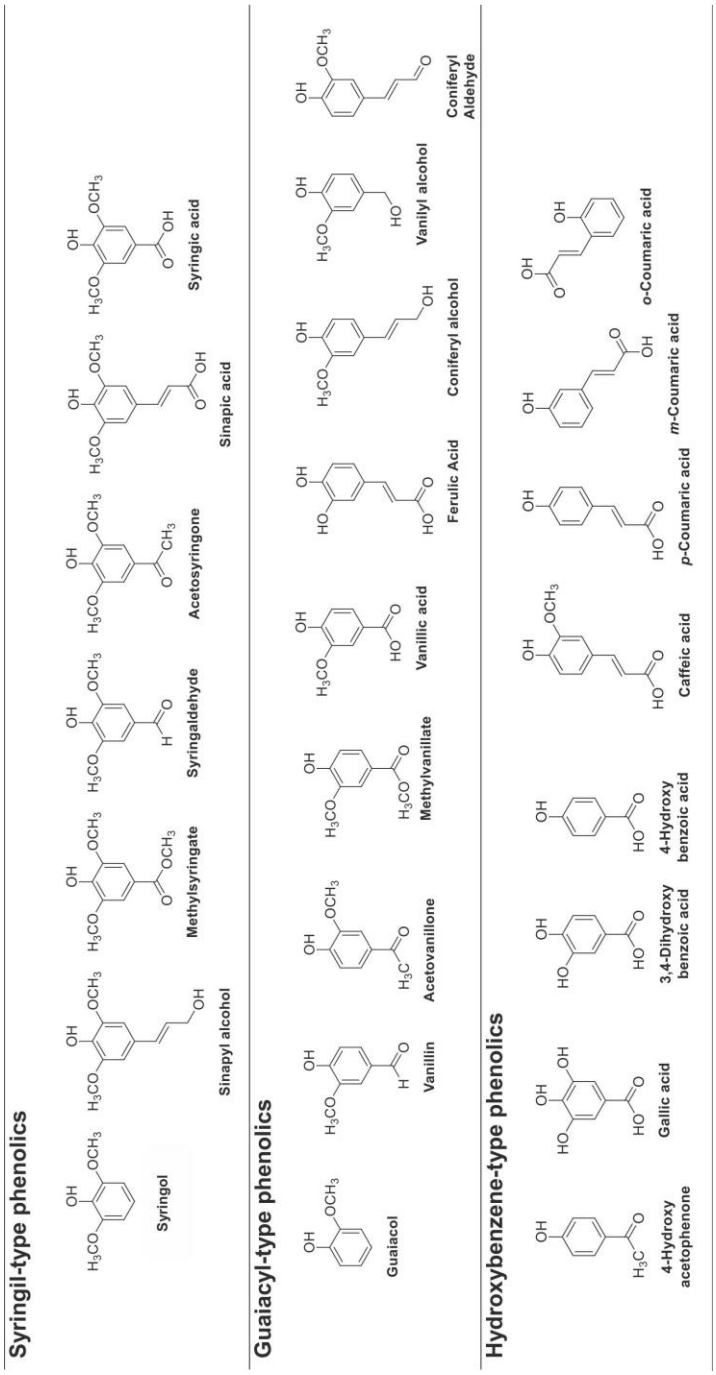
2.4. Results and Discussion

2.4.1. Electrochemical characterization of the lignin-related phenolic substrates

The twenty-four phenolic compounds selected are good representatives of the three classes of lignin-related phenolics: syringyl-type phenolics, with methoxy substituents in both *ortho* positions, guaiacyl-type phenolics, with a single methoxy group, and hydroxyl benzenes **(Table 2.2)**. The electrochemical data found in the literature for these compounds are hardly comparable insofar as they were measured at variable scan rates, pH values, and in the presence of different organic solvents and electrolyte concentrations. Therefore, we have determined the redox potential for all compounds tested, considering their structural diversity and the redox potentials' dependence on phenolics' structural characteristics.

The one-electron oxidation of phenols is a dissociative electron transfer and involves the loss of a proton: $PhOH + H_2O \rightarrow PhO^{\cdot} + H_3O^+ + e^-$, and therefore, their redox potentials are pH-dependent. The oxidation of phenolate ions results in the formation of phenoxyl radicals: $PhO^- \rightarrow PhO^{\cdot} + e^-$. The electrochemical studies were performed for all phenolic compounds at pH 4 and 8, the optimal values for wild type *Pp*DyP and variant 6E10, respectively **(Table 2.3)**. The cyclic voltammograms of the phenolic compounds displayed one well-defined anodic peak, and after the oxidation, no reverse peak was observed **(Figure SI 2.1)**. The absence of a cathodic peak in the reverse scan indicates that the oxidized generated species are involved in chemical reactions and rapidly removed from the medium. The exceptions are with syringaldehyde, acetosyringone, and methyl syringate, where a cathodic wave is observed, although in an irreversible process **(Table 2.2** and **Figure SI 2.2)**.

Table 2.2. Structure of the lignin-phenolic compounds investigated in this work.



The oxidation potentials of phenolics at pH 8 are lower than at pH 4, which relates to the proximity to the pK_a values of most phenolics **(Table 2.3)**; due to the higher concentrations of the phenolate ions at pH 8, they are more prone to be oxidized than the corresponding phenolic group (Rosado et al., 2012). The oxidation potentials follow the trend: of syringyl < guaiacyl < hydroxybenzene **(Table 2.3)**, and the presence and positions of substituents on the aromatic ring directly affect redox potentials. The oxidation is easier when electron-donating groups (e.g., Me, OMe, OH) are present; these groups might induce radical stabilization resulting in a decrease in the redox potential, while the presence of electron-withdrawing functionalities (e.g., COR, R=H, CH_3, OH, OCH_3) has the opposite effect. The electron-withdrawing effect of COR groups at the *para* position is evident in both syringyl and guaiacyl-type groups **(Tables 2.2 and 2.3)**; they release electron density from the aromatic ring, deactivating the phenol group that becomes less available for electron release and more difficult to oxidize. Different substituents show different acceptor capabilities depending on the phenolic family. However, no significant differences were found; the syringyl type show anodic peak potential (E_{pa}) values between 0.87-1.06 V, and the guaiacyl type from 1.09-1.17 V **(Table 2.3)**. The relative position of the other substituents to the phenol group can also be decisive, as illustrated by the coumaric acid series: *o*- and *m*-coumaric acids showed higher anodic potentials than the corresponding *para* isomer, which is the easiest to oxidize, probably due to hyperconjugation between the aromatic ring and the CH=CHCOOH group. The introduction of methoxy groups in the *ortho* positions lowers the E_{pa} values **(Table 2.3)**, as is evident by the comparison of values for acetosyringone (1.04 V), acetovanillone (1.12 V), and 4-hydroxyacetophenone (1.3 V). This is probably due to the conjugative electron donation of the OCH_3 group and the stabilization of the intermediate. The effect of a conjugated double

bond is evident by the lowest oxidation potentials measured for coniferyl alcohol, coniferyl aldehyde, and sinapic acid when compared with vanillyl alcohol, vanillin, and syringic acid, respectively **(Table 2.3)**. The syringyl-type counterparts, sinapyl alcohol and sinapic acid, are the most easily oxidized compounds.

Table 2.3 Electrochemical data of lignin-related phenolics vs. NHE in buffered solutions (100 mM acetate, pH 4 and 100 mM phosphate, pH 8) at a scan rate of 100 mV s^{-1}.

Lignin-related phenolics	**pK_a**	**pH 4**		**pH 8**	
		E_{pa} (V)	E_{pc} (V)	E_{pa} (V)	E_{pc} (V)
Syringil type phenolics					
Syringol	9.9[a]	0.76	---	0.69	---
Sinapyl alcohol	9.4[b]	0.77	---	---	---
Acetosyringone	7.9[a]	1.04 ---	--- ---	0.71[3)] 1.31	0.56[3)] ---
Methyl syringate	8.7[a]	0.99	---	0.68[1)]	0.60[1)]
Syringaldehyde	7.3[a]	1.06	0.78	0.76[2)]	0.60[2)]
Sinapic acid	9.2[a]	0.82	---	0.63	---
Syringic acid	7.8[c]	0.87	---	---	---
Guaiacyl type phenolics					
Coniferyl alcohol	9.5[a]	0.82	---	0.69	---
Vanillin	7.4[a]	1.14	---	0.86	---
Coniferyl aldehyde	7.9[a]	0.96	---	0.68	---
Ferulic acid	9.4[a]	0.94	---	0.73	---
Methyl vanillate	8.4[d]	1.17	---	0.86	---
Vanillyl alcohol	9.8[a]	1.01	---	0.9	---
Acetovanillone	7.8[a]	1.12	---	0.89	---
Guaiacol	9.9[a]	1.06	---	0.83	---
Vanillic acid	8.5[c]	1.09	---	---	---
Hydroxybenzene type phenolics					
Caffeic acid	8.6[e]	---	---	0.72	---
m-Coumaric acid	10.4[f]	1.22	---	1.05	---
p-Coumaric acid	9.9[f]	1.06	---	0.89	---
o-Coumaric acid	9.6[f]	1.12	---	0.94	---
3,4-hydroxybenzoic acid	-	1.00	---	0.94	---
4-hydroxybenzoic acid	8.7[c]	---	---	---	---
Gallic acid	12.2[c]	---	---	---	---
4-hydroxyacetophenone	8.1[d]	1.3	---	1.17	---

[1)] $E_{1/2}$ = 0.63; ΔE =180 mV; I_c/I_a =0.9 [2)] $E_{1/2}$ = 0.68 V; ΔE =170 mV; I_c/I_a =0.5 [3)] $E_{1/2}$ = 0.64 V; ΔE =150 mV; I_c/I_a =1.0 *several weak oxidation waves; ** insufficient sample; [a] (Ragnar et al., 2000); [b] https://foodb.ca/compounds/; [c] (Erdemgil et al., 2007); [d] https://www.chemicalbook.com/ ; [e] (Genaro-Mattos et al., 2015); [f] (Benvidi et al., 2019);

2.4.2. Conversion yields of lignin-related phenolics

HPLC estimated the substrate conversions after 24 h of reaction at room temperature. Both enzymes converted all substrates, but 6E10 showed improved yields (from 10 to 75%) across the syringyl, guaiacyl, and hydroxybenzene-type families **(Figure 2.1** and **Table SI 2.1)**. Wild type converted ≥ 98% of sinapic acid and coniferyl alcohol, for example. In contrast, the 6E10 variant converted ≥ 98% of those in addition to syringol, sinapyl alcohol, coniferyl alcohol, coniferyl aldehyde, ferulic acid, vanillin, methyl vanillate, vanillyl alcohol, and acetovanillone. Both enzymes show high conversion yields towards guaiacyl-type phenols. Wild type and 6E10 averaged 47 and 77% conversion yields for syringyl-type phenolics, 69 and 91% for guaiacyl-type phenolics, and 33 and 58% for hydroxybenzene-type phenolic compounds, respectively.

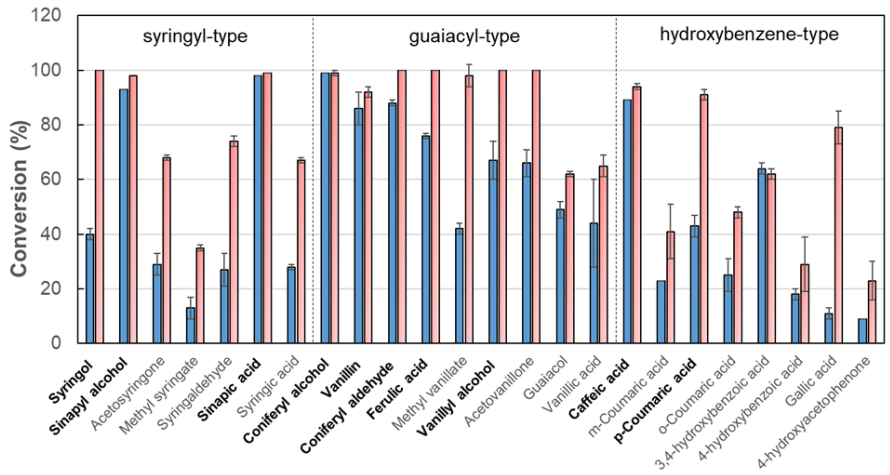


Figure 2.2. Substrate conversion yields measured by HPLC after 24 h of reaction with wild type (blue) and 6E10 variant (red), using 0.5 mM of phenolics at pH 4 and pH 8, respectively. Reactions started by addition of 0.5 mM H_2O_2. The substrates selected for product identification are in bold.

2.4.3. Catalytic parameters for the lignin-related phenolic substrates

Each substrate's maximal wavelength and molar extinction coefficient were determined **(Table SI 2.2)** and used afterward to measure enzymatic activities and estimate catalytic parameters **(Table 2.4)**. We could not set up colorimetric assays to measure the activity for seven of the twenty-four phenolic substrates (vanillyl alcohol, acetovanillone, *m*-coumaric acid, 3,4-hydroxybenzoic, 4-hydroxybenzoic acid, gallic acid, and 4-hydroxy acetophenone). The catalytic parameters were estimated for seventeen substrates using wild type and 6E10 at the optimal pH of the enzymes. Wild type oxidizes all phenolics at the same order of magnitude; the average k_{cat} values for syringyl, guaiacyl, and hydroxybenzene phenolics are comparable: 0.056 s^{-1}, 0.063 s^{-1}, and 0.08 s^{-1}, whereas the K_m values range from 0.01 to 1 mM **(Table 2.4)**. In contrast, 6E10 shows a clear preference for syringyl-type compounds, which are oxidized at considerably higher rates showing an average k_{cat} of 1.44 s^{-1}, as compared to 0.4 s^{-1} for guaiacyl and 0.12 s^{-1} for the hydroxybenzene phenolics. Therefore, the activity improvements in the evolved variant are more prominent for the syringyl-type compounds **(Table 2.4)**, which might be related to the fact that syringol was the substrate used during the activity screenings in the directed evolution of this enzyme, confirming the rule "*you get what you screen for*." In general, the increased activity of 6E10, as compared to the wild type, was accompanied by a decrease in substrate affinity as assessed by the K_m values **(Table 2.4)**.

Table 2.4. Apparent steady-state kinetic parameters for the oxidation of phenolic substrates by *Pp*DyP WT and 6E10, in the presence of 0.5 mM H_2O_2, at the optimal pH of each enzyme.

	Wild type (pH 4)			**6E10 (pH 8)**		
	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
Sinapyl-type phenolics						
Syringol	0.05 ± 0.01	0.70 ± 0.20	7.1×10^2	3.5 ± 0.1	0.08 ± 0.06	4.3×10^4
Sinapyl Alcohol	0.14 ± 0.01	0.04 ± 0.01	3.5×10^3	2.2 ± 0.3	0.32 ± 0.09	6.8×10^3
Methylsyringate	0.05 ± 0.01	0.08 ± 0.02	2.5×10^2	1.7 ± 0.7	1.20 ± 0.70	1.4×10^3
Acetosyringone	0.03 ± 0.01	0.05 ± 0.02	6.0×10^2	0.9 ± 0.1	0.22 ± 0.05	4.1×10^3
Syringaldehyde	0.01 ± 0.01	0.04 ± 0.01	2.5×10^2	0.7 ± 0.03	0.010 ± 0.001	7.0×10^4
Sinapic Acid	0.05 ± 0.01	0.03 ± 0.05	1.6×10^3	0.4 ± 0.02	0.08 ± 0.01	5.0×10^3
Syringic Acid	0.060± 0.01	0.15 ± 0.05	4.0×10^2	0.7 ± 0.03	0.08 ± 0.01	8.8×10^3
Guaiacyl-type phenolics						
Coniferyl alcohol	0.12 ± 0.03	0.3 ± 0.1	4.0×10^2	0.3 ± 0.01	0.019 ± 0.004	1.6×10^4
Vanillin	-	-	-	0.7 ± 0.5	0.8 ± 0.5	8.8×10^2
Coniferyl aldehyde	-	-	-	0.10 ± 0.01	0.050 ± 0.009	2.0×10^3
Ferulic acid	0.020 ± 0.002	0.015 ± 0.01	1.3×10^4	0.10 ± 0.01	0.11 ± 0.03	0.9×10^4
Methyl vanillate	-	-	-	0.20 ± 0.04	0.33 ± 0.13	6.1×10^2
Guaiacol	0.05 ± 0.01	0.01 ± 0.001	0.9×10^3	0.50 ± 0.02	0.030 ± 0.003	1.9×10^4
Vanillic acid	-	-	-	1.0 ± 0.2	0.24 ± 0.1	4.7×10^3
Hydroxybenzene-type phenolics						
Caffeic acid	0.080± 0.008	0.09 ± 0.03	8.9×10^2	0.27 ± 0.05	0.3 ± 0.1	9.0×10^2
p-coumaric acid	nd	-	-	0.06 ± 0.003	0.09 ± 0.02	6.6×10^2
o-coumaric acid	nd	-	-	0.02 ± 0.001	0.15 ± 0.02	1.3×10^2

Nevertheless, 2 to 100-fold improved catalytic efficiencies (k_{cat}/K_m) of 6E10 were estimated for all substrates except for ferulic acid and caffeic, which show comparable efficiencies to the wild type. In that context, 6E10 is the only a unique DyP that offer maximal rates at pH 8-8.5. It is also a better biocatalyst for lignin-related phenolics (3.5 s^{-1}

for syringol *versus* wild type's 0.05 s^{-1}) than other bacterial DyPs reported in the literature. For example, activities of around 0.03 to 0.5 U/mg were measured for syringol in DyPs from *S. viridis (Yu et al., 2014), S. avermitilis* (Sugawara et al., 2017), *E. lignolyticus* (for guaiacol an activity of 0.34 U/mg (Shrestha et al., 2017), *T. fusca* (Rahmanpour et al., 2016), and *B. subtilis (Santos et al., 2014)*. Other fungal DyPs, such as the ones found in *Irpex lacteus, A. auricula-judae* or *P. ostreatus,* exhibit k_{cat} values 10 to 100 orders of magnitude higher than 6E10, however, these are highly acidic enzymes, operating at an optimum pH between 2 and 3.5, which is a setback for their industrial application (de Eugenio et al., 2021).

There is no clear trend between chemical and enzymatic oxidation of the lignin-related phenolics tested besides the preference for pH 8.0 and syringyl-type substrates **(Tables 2.3** and **2.4)**. Sinapic acid with the lowest value of E_{pa} at pH 8.0 (0.63 V) displays the lowest turnover rate of the syringyl type family (0.4 s^{-1}), and substrates such as syringol, coniferyl alcohol, and coniferyl aldehyde although sharing similar redox potentials (0.68-0.69 V) are oxidized at different rates: 6E10 oxidizes syringol 20 and 60-fold faster than coniferyl alcohol and coniferyl aldehyde. The coumaric acid series (*p*-, *o*- and *m*-coumaric acid) indicates well the effect of the position of the substituents of the phenolic ring on the enzymatic oxidation rates following the order *p*-coumaric acid > *o*-coumaric acid > *m*-coumaric acid, with 6E10 oxidizing the *p*-substituent at 3-fold higher rates than the *o*-substituent and showing no measurable activity for the *m*-substituent. The observed differences are therefore expected to result from specific substrate-enzyme interactions, e.g., accommodation of substrates to the binding site(s), distance to the heme, electron transfer rates, products dissociation.

2.4.5. Identification of products of reaction

A set of eleven substrates were selected among those that were oxidized at higher rates and showed > 90% conversion yields (syringol, sinapyl alcohol, sinapic acid, coniferyl alcohol, vanillin, vanillyl alcohol, acetovanillone, ferulic acid, caffeic acid, and *p*-coumaric acid) for reactions to identify and characterize the products of reactions. Reactions monitored after 1h by HPLC revealed no differences when compared to the 24 h reactions in substrate conversion yields **(Figure SI 2.3)**. The ESI-MS results confirmed that the transformation was incomplete for some substrates (vanillin, vanillyl alcohol, acetovanillone, ferulic acid, and coumaric acid). In general, the products' signals obtained are consistent with the formation of dimeric structures with molecular weights of (2 × $MW_{substrate}$ - 2H), except for coniferyl aldehyde, for which no relevant signals were found **(Table 2.4)**. DyPs mediate the one-electron oxidation of phenolics resulting in the formation of radical intermediates or quinones that are subsequently coupled with other substrate molecules. The direction taken by the radical coupling reactions depends on the substitution pattern of the aromatic ring, reflecting the influence of electronic features on the stabilization and chemical reactivity of the first formed radical. Various stereo and regioisomers can be considered, as the same *m/z* values were found at different retention times, indicating the presence of different dimeric isomers. We have also observed the oxidation of the non-phenolic hydroxy group to the aldehyde in the reaction of 6E10 with sinapyl alcohol, coniferyl alcohol, and vanillyl alcohol.

The oxidation of syringol showed a peak at *m/z* 307, which is consistent with the molecular formula $C_{16}H_{18}O_6$ and can be attributed to C-C, C-O, or O-O dimers **(Figure SI 2.4)**, although the fragmentation pattern is inconsistent with C-C and C-O dimers

(Adelakun et al., 2012a; Schirmann et al., 2018). The MS data for reaction products with guaiacyl-type derivatives, vanillyl alcohol, vanillin, and acetovanillone showed the formation of at least one dimeric structure. Furthermore, the oxidation of the non-phenolic hydroxyl group of vanillyl alcohol was observed with the formation of vanillin. In the reaction with vanillin, the peak at m/z 303 is consistent with the molecular formula $C_{16}H_{14}O_6$ and was attributed to divanillin (5,5' dimer) or the 4-O-5 isomer. The ^{1}H NMR analysis of the scale-up reaction with vanillin confirmed the incomplete oxidation of the substrate and the identification of both 5,5' and 4-O-5 dimers at low yields (~10%) (see Table S3). Considering these results, for acetovanillone, the peak at m/z 331 is consistent with the molecular formula $C_{18}H_{18}O_6$ and can be assigned to diapocynin or its C-O isomer. Divanillin is a taste enhancer, and flavoring, and diapocynin, the dimer of acetovanillone, has important anti-inflammatory and neuroprotective effects (Ismail et al., 2014; Krings et al., 2015). Furthermore, both compounds are valuable intermediates for the synthesis of polymers, such as polyvanillin, Schiff bases, and conjugated polymers (Fang et al., 2020; Llevot et al., 2016; Sun et al., 2018). The production of both divanillin and diapocynin using laccases and horseradish peroxidase was previously reported (Krings et al., 2015; Nishimura et al., 2010). The production of divanillin was also observed with the fusion of an oxidase with the DyP from *S. viridis* DSM 43017. The fusion enzyme was able to oxidize vanillyl alcohol to vanillin, due to the oxidase, while the DyP dimerized vanillin to divanillin (Colpa et al., 2017).

In the case of phenylpropanoid type monomers (for example, coniferyl alcohol, sinapic acid, caffeic acid, and *p*-coumaric acid), the *m/z* values found are compatible with a higher number of dimeric structures **(Table 2.5** and **Figures SI 2.5-2.9**). The oxidation of these lignan and neolignan precursors to the corresponding phenoxy radicals can undergo delocalization along the unsaturated backbone leading to different resonance forms involving carbon atoms **(Figure 2.3 A)**. Different pathways involving free-radical coupling at the O-4, 5- and 8-positions after that lead to the formation of lignan and neolignan dimers and oligomers, where linkages between subunits involve either C-C (5,5'-, 8,5'- and 8,8'-dimers) or C-O (8-O-4 and 4-O-5-dimers) bonds **(Figure 2.3 B)**. Moreover, the ratio of the different coupled dimers is strongly dependent on the pH and other experimental conditions, such as the rate of substrate supply, the pH, the presence of co-solvents, and reaction time (Gruz et al., 2015; Jaufurally et al., 2016a; Jaufurally et al., 2016b).

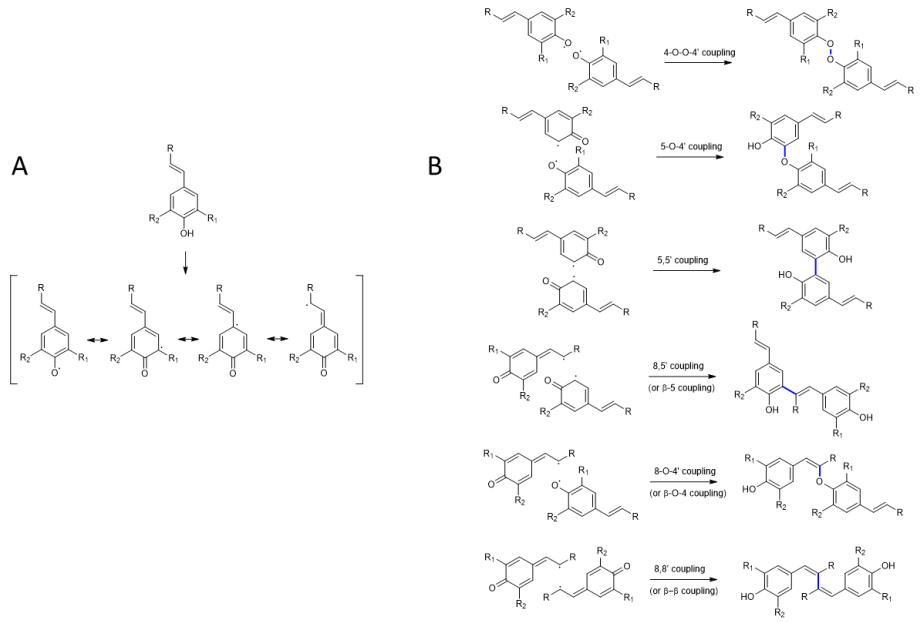
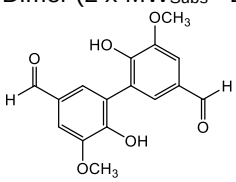
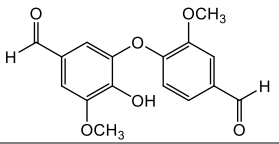
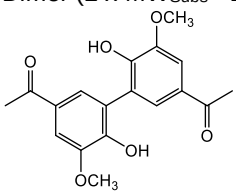
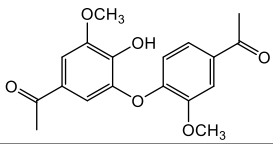
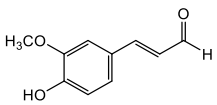
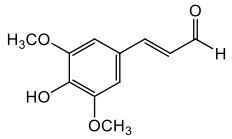
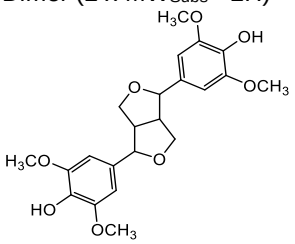
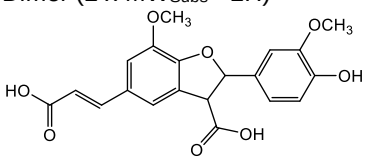


Figure 2.3. (A) Proposed radical species originated from substituted phenylpropanoid derivatives oxidation (R1, R2 = H, OCH3). (B) Overview of possible lignan and neolignan type dimers formed after DyP-type peroxidase oxidation of phenylpropanoid derivatives.

After enzymatic oxidation of coniferyl alcohol, four peaks can be distinguished after one-hour of reaction and attributed to vanillin (rt = 7.07 min, m/z 153), coniferyl aldehyde (rt = 7.83 min, m/z 179) resulting from the oxidation of the non-phenolic group, and a dimeric product eluting at 8.09 min for which different radical coupling isomers have been considered **(**see **Figure SI 2.5)**. A similar behavior was found for sinapyl alcohol oxidation where several peaks with varying times of retention were detected at 6.65 min (m/z 209; $C_{11}H_{12}O_4$) attributed to sinapyl aldehyde, 7.32 min (m/z 405; $C_{21}H_{24}O_8$); 8.36 min (m/z 419; $C_{22}H_{26}O_8$) and 9.61 min (m/z 417; $C_{22}H_{24}O_8$). Considering the principal radical coupling reactions within the syringyl type monomers, the 8,8' and 8-O-4 couplings, the peaks at m/z 419 and m/z 417 can be assigned to lignan dimeric structures (2×MW-2H) and (2×MW-4H) respectively. The ^{1}H NMR analysis of a reaction with sinapyl alcohol using 6E10 indicates the complete oxidation of the substrate and the formation of the lignan syringaresinol as the main product, confirming the structure of the 8,8' (or β-β dimer (2 × MW - 2H) proposed from the mass spectrometry analysis **(Table SI 2.3)**. The neolignan 8-O-4 dimer was not detected by neither MS or NMR analysis. The enzymatic synthesis of syringaresinol, a furofuran lignan dimer formed upon oxidation of sinapyl alcohol, was previously described using the *Trametes versicolor* laccase (Jaufurally et al., 2016b) and one-pot conversion using eugenol oxidase and horseradish peroxidase (Habib et al., 2018). Syringaresinol displays interesting bioactivity properties (Bajpai et al., 2018; Habib et al., 2018; Jaufurally et al., 2016b) and is of interest to replace bisphenol A for polymer synthesis (Chio et al., 2019; Janvier et al., 2016; Janvier et al., 2017; Llevot et al., 2016). Despite being available naturally in plants, syringaresinol is isolated from natural sources in meager yields (Monthong et al., 2011).

Table 2.5. ESI-MS (positive mode) data and retention times (rt) for 6E10 reactions with substrates at pH 8 and putative products

Substrate	rt (min)	m/z (+)	Attribution
Guaiacyl and Syringyl type derivatives			
Syringol	8.32	307	Dimer (2 x MW_{Subs} - 2H) [(1)]
Vanillyl alcohol	5.83	137	Vanillyl alcohol
	6.64	289	Dimer (2 x MW_{Subs} - 2H)-18
	7.08	153	Vanillin
	9.42	549	---
Vanillin	7.05	153	Vanillin
	7.99	468	---
	8.52	303	Dimer (2 x MW_{Subs} - 2H)  and 
	8.65	469	---
Acetovanillone	7.31	167	Acetovanillone
	8.65	331	Dimer (2 x MW_{Subs} - 2H)  or 
	8.86	511	---
Lignan and neolignan type derivatives			
Coniferyl alcohol	7.07	153	Vanillin
	7.83	179	 Coniferyl aldehyde
	8.09	373	Dimer (2 x MW_{Subs} – 4H + O)

	8.11	359	Dimer (2 x MW_{Subs} - 2H)[2]
Sinapyl alcohol	6.65	209	 Sinapyl aldehyde
	7.32	405	---
	8.36	419	Dimer (2 x MW_{Subs} - 2H) 
	9.61	417	Dehydrodimer (2 x MW_{Subs} - 4H)
p-Coumaric acid	7.03	165	p-coumaric acid
	7.61	327	Dimer (2 x MW_{Subs} - 2H) [3]
	8.37	327	Dimer (2 x MW_{Subs} - 2H) [3]
	8.53	327	Dimer (2 x MW_{Subs} - 2H) [3]
Caffeic acid	7.40	359	Dimer (2 x MW_{Subs} - 2H) [4]
	7.60	359	Dimer (2 x MW_{Subs} - 2H) [4]
	7.71	359	Dimer (2 x MW_{Subs} - 2H) [4]
	7.93	359	Dimer (2 x MW_{Subs} - 2H) [4]
Ferulic acid	7.25	179	Ferulic acid
	7.39	369	---
	8.37	387	Dimer (2 x MW_{Subs} - 2H) [5] 
	9.37	373	---
Sinapic acid	7.20	447	Dimer (2 x MW_{Subs} - 2H) [6]
	8.07	447	Dimer (2 x MW_{Subs} - 2H) [6]

(1) see **Figure SI. 2.4**; [2] see **Figure SI. 2.5**; [3] see **Figure SI. 2.6**; [4] see **Figure SI. 2.7**; [5] see **Figure SI. 2.8**; [6] see **Figure SI. 2.9**

The enzymatic reactions with hydroxyphenyl and guaiacyl type derivatives *p*-coumaric, caffeic and ferulic acids are characterized by the production of multiple dimeric isomers, eluted at different retention times, and showing different mass fragmentation patterns, and reflect the coupling possibilities of the primary formed radicals **(Figure 2.3)**, their relative stability, and their dependence on critical parameters and reaction conditions (Flourat et al., 2022; Gruz et al., 2015). The product profile of the *p*-coumaric acid reaction is complex. The presence of three different dimer isomers m/z 327 (consistent with the molecular formula $C_{18}H_{14}O_6$) at different retention times indicates the formation of different isomers **(**see **Figure SI 2.6)**. Mass fragmentation data showed, in all the cases, losses of one or two neutral water, CO, and CO_2 molecules. Still, the differences in the MS fragmentation patterns did not allow for distinguishing a particular isomer. The same result was obtained for caffeic acid, where four products can be determined, all isomers of m/z 359 **(**see **Figure SI 2.7** for the dimeric structures considered**)**. For sinapic acid, two isomers of the dimer m/z 447 are eluting at RT 7.20 and 8.07 min. Although several structural isomers can be regarded as possible **(**see **Figure SI 2.9)**, the MS fragmentation patterns in both cases did not allow the attribution of a particular isomer. A different result was obtained for ferulic acid since the reaction was incomplete after one- hour, and only one dimer at rt 8.37 min (m/z 387 consistent with the molecular formula $C_{20}H_{18}O_8$) was present. Although a high number of dimeric structures can be considered **(**see **Figure SI 2.8)**, the comparison between the mass fragmentation pattern of this dimer and the literature data for diferulates strongly points out to the neolignan 8,5' benzofurandehydroferulic acid dimer as the oxidation product (Vismeh et al., 2013). This dimer is a potential biomarker of several food products, such as wheat, corn, and rye, and shows antioxidant properties (Garcia-Conesa et al., 1997).

2.5. Concluding remarks

All lignin-derived phenolics tested were oxidized by *Pp*DyP wild type and 6E10 variant at pH 4 and 8, respectively. Kinetic measurements show that the turnover rate of 6E10 was higher across the board, resulting in higher catalytic efficiency for all substrates tested. When substrate conversion after 24 h was analysed, differences were not so pronounced, and both enzymes could oxidize the tested substrates to comparable extents. Reactions of 6E10 with phenolic substrates tend to form dimers by radical coupling, as assessed after identifying reaction products by MS; dimers such as divanillin, diapocynin, and the lignan syringaresinol are interesting biological active molecules with potential applications such as a building block for the resins industry or food additive (divanillin), as pharmaceuticals (diapocynin) or as a starting material in the polymer industry (syringaresinol). These are natural products in plant extracts that show interesting biological activity (e.g, anti-inflammatory and anti-cancer properties) but whose extraction suffers from low yields. Structurally diverse products can be obtained through electron delocalizing phenolic radicals before coupling. Future studies need to focus on testing reaction conditions that lead to the control of the reaction into the desired path, i.e., that allow the control of the distribution of the products and the concomitant improvement of the reaction yields. Unveiling the molecular basis of chemical stereo- and region-specificity of monolignol-derived radical-radical coupling leading to lignans, neolignans, and other dimeric structures will represent a robust knowledge and significant advances behind the state-of-the-art in combinatory lignin chemistry with critical reflection in eco-friendly lignin valorization that supports the economic sustainability of lignocellulose biorefineries.

Chapter 3

Unveiling Molecular Details behind Improved Activity at Neutral to Alkaline pH of an Engineered DyP-type Peroxidase

This chapter contains data published in:

Diogo Silva, Patrícia Borges, Tomás F. D. Silva, Vânia Brissos, Marina Cañellas, Maria Fátima Lucas, Laura Masgrau, Eduardo P. Melo, Miguel Machuqueiro, Carlos Frazão and Lígia O. Martins (2022). Unveiling Molecular Details behind Improved Activity at Neutral to Alkaline pH of an Engineered DyP-type Peroxidase. *Comput. Struct. Biotechnol. J.* **20:** 3899-3910.
https://doi.org/10.1016/j.csbj.2022.07.032

Carolina F. Rodrigues, Patrícia T. Borges, Magali F. Scocozza, **Diogo Silva**, André Taborda, Vânia Brissos, Carlos Frazão and Lígia O. Martins (2021). Loops around the Heme Pocket Have a Critical Role in the Function and Stability of *Bs*DyP from *Bacillus subtilis. Int J Mol Sci*. **22**, 10862
https://doi.org/10.3390/ijms221910862

Catarina Barbosa, Célia M. Silveira, **Diogo Silva**, Vânia Brissos, Peter Hildebrandt, Lígia O. Martins and Smilja Todorovic (2020). Immobilized dye-decolorizing peroxidase (DyP) and directed evolution variants for hydrogen peroxide biosensing. *Biosens. Bioelectron*. **153**, 112055.
https://doi.org/10.1016/j.bios.2020.112055

3.1. Abstract

DyP-type peroxidases (DyPs) are microbial enzymes with enormous biotechnological potential in biorefineries due to their enlarged substrate scope, but 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 when compared to wild type and 29E4 variants. A more positively charged microenvironment, near the heme pocket of variant 6E10, was unraveled by constant-pH MD simulations, particularly in the neutral to the 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, including their pH-dependent conformational plasticity and function at the molecular level. They are essential for understanding and engineering the catalytic properties of DyPs for future biotechnological applications.

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

3.2. Introduction

DyPs (EC 1.11.1.19) contain a α + β 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 (Hofbauer et al., 2021; Singh and Eltis, 2015; Sugano and Yoshida, 2021). Furthermore, the heme pocket of DyP-type peroxidases is buried within loops and helix regions that are structurally conserved in all DyPs (Rodrigues et al., 2021). Loops and channels are among the most flexible structural elements of an enzyme and critically affect the enzymes' motions and dynamic properties, critical for biological functions.

Loops were first identified as a secondary structure category in 1986 (Leszczynski and Rose, 1986), and are a diverse class of secondary structures comprising turns, random coils, and strands that connect the main secondary structures (α-helices and β-strands). They are often located at the protein surface and their flexibility affects enzyme activity and helps intervene in the binding of substrates and release of products. Their role in enzyme catalysis remains unclear but they can, for example, protect the reaction space from the aqueous environment, and stabilize reaction intermediates, and the interactions between catalytic residues and binding subsites (Kress et al., 2018). Likewise, channels are highly dynamic structural elements and their aminoacidic composition can influence the behaviour of substrates on their way to a buried active site (Malabanan et al., 2010; Petrovic et al., 2018).

A link between conformational diversity and the emergence of new enzyme functionalities has been recognized for many years (Petrovic et al., 2018; Tokuriki and Tawfik, 2009). However, it was only recently that state-of-the-art computational and experimental approaches are revealing the crucial molecular details of this link. With the improvement of structure prediction tools and MD simulations, a better

understanding of how altered loop flexibilities and compositions influence different aspects of biocatalysts has been brought to light (Bhabha et al., 2011; Heinemann et al., 2021; Hoque et al., 2017; Kress et al., 2018). Specifically, evolutionary trajectories leading to functional optimization for a given host environment or to the emergence of a new function typically involve enriching catalytically competent conformations and/or the freezing out of non-competent conformations of an enzyme (Heinemann et al., 2021; Pabis et al., 2018). In some cases, these evolutionary changes are achieved through distant mutations that shift the protein ensemble towards productive conformations (Campbell et al., 2018; Yang et al., 2020). Conformational diversity can assist the emergence of a completely new active site through a single mutation by facilitating transition-state binding (Heinemann et al., 2021; Kress et al., 2018).

In this work, X-ray structural information of wild type *P. putida Pp*DyP and the two engineered 6E10 and 29E4 variants, obtained by directed evolution after insertion of mutations in loop segments, close to the heme pocket, was complemented with conformational and protonation analysis from constant-pH molecular dynamics simulations (CpHMD). This analysis provided details into the features that underlie the activity and stability of these enzymes. Furthermore, details on enzyme interaction with the substrate 2,6-dimethoxyphenol were also studied using an ensemble docking protocol. Our findings have fundamental importance in understanding the molecular dynamics and catalytic features of DyPs and evaluating their biotechnological potential.

3.3 Materials and Methods

3.3.1 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 (Santos et al., 2014) or its evolved variants (p6E10 (Brissos et al., 2017), p29E4 (Barbosa et al., 2020). 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 (Barbosa et al., 2020). 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 ~ 1-2 **(Table 3.1)**.

Table 3.1. Production levels and properties of proteins. The concentration of purified proteins was estimated using the molar absorption coefficient of *Pp*DyP (ε_{280} = 34,850 $M^{-1} \cdot cm^{-1}$). The heme content was determined by the pyridine ferrohemochrome method using an extinction coefficient of $\varepsilon_{R\text{-}O\ 556nm}$ (28.32 $mM^{-1} \cdot cm^{-1}$).

	Enzyme production ($mg \cdot L^{-1}$)	**Rz**	**λ_{max}** (nm)	**ε** ($mM^{-1}\ cm^{-1}$)	**Heme content** [heme](mM) / [protein](mM)
***Pp*DyP**	14	1.3	405	46	1.0 ± 0.1
29E4	13	1.3	405	44	1.1 ± 0.3
6E10	11	1.2	405	43	1.2 ± 0.1

3.3.2 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).

3.3.3 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).

3.3.4 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 (Durao et al., 2006; Fernandes et al., 2009). 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 (Fernandes et al., 2009).

3.3.5 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 ºC in a crystallization solution composed 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 ºC 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 ºC 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 (Gorrec, 2009), its crystals were directly flash-cooled in liquid nitrogen.

3.3.6 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 (Kabsch, 2010). Data collection and processing statistics are given in **Table SI. 3.1**.

3.3.7 Structure determination and refinement

Wild type *Pp*DyP data were initially processed in space group $P3_221$, but data analyses with programs XPREP (Sheldrick, 2000), XTRIAGE (Zwart, 2005) and POINTLESS (Evans, 2006) 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 (Vagin and Lebedev, 2015) testing all produced data sets, including possible alternative space groups. Two putative molecular replacement solutions were produced, showing similar probabilities and R_{work} 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 (Emsley et al., 2010) in the electron density. Twinned refinement of each model using REFMAC5 (Murshudov et al., 2011) with intensities data and twin laws H, K, L and -H, -K, -L led to R_{work}/R_{free} = 0.394/0.395 for space group $P3_2$, while space group $P3_221$ showed R_{work}/R_{free} = 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 R_{work}/R_{free} = 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 (McCoy et al., 2007) within the PHENIX suite (Adams, P.D. et al., 2010), 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) (Painter and Merritt, 2006) was used to define polypeptide chain regions for translation, libration, and screw refinement of anisotropic atomic displacement parameters

(*a.d.p.s*). PHENIX.REFINE (Adams, P. et al., 2010) 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 (Emsley et al., 2010). Standard stereo-chemical dictionary (Engh and Huber, 1991) 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 (Chen et al., 2010). All figures were prepared with PyMOL (DeLano, 2002). Analysis of molecular tunnels was performed with CAVER (Chovancova et al., 2012). Three-dimensional superposition of polypeptide chains was performed with MODELLER (Webb and Sali, 2014). Structure factors and atomic coordinates were deposited in the Protein Data Bank (PDB) (Berman et al., 2000) 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 SI. 3.1**.

3.3.8 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 (Machuqueiro and Baptista, 2006). 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 (Abraham et al., 2015), the GROMOS 54A7 force field (Schmid et al., 2011), and the SPC water model (Hermans et al., 1984). 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 (Bussi et al., 2007); then, a 1 ns step on the NPT ensemble is performed using the Parrinello-Rahman barostat (Nosé and Klein, 1983; Parrinello and Rahman, 1981) 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) (Tironi et al., 1995). Protein bands were constrained with the P-LINCS methods (Hess, 2008), while the water molecules were used and treated with the SETTLE algorithm (Miyamoto and Kollman, 1992). The integrator time-step was 2 fs.

The Poisson-Boltzmann (PB) calculations were performed using Delphi V5.1 (Rocchia et al., 2002) using radii calculated using the Lennard-Jones parameters (Teixeira et al., 2005) 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 centred 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 (Baptista and Soares, 2001) 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 2 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/titratable 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 (Silva et al., 2021).

3.3.9 Molecular Dynamics and Ensemble Docking

MD simulations of the wild type enzyme and the 6E10 variant were set up and ran with Yasara (Krieger and Vriend, 2015) using the AMBER force field (Hornak et al., 2006) and TIP3P (Mark and Nilsson, 2001) 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 (Berendsen et al., 1984) 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 (Essmann et al., 1995). LINCS (Hess et al., 1998) and SETTLE (Miyamoto and Kollman, 1992) 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 (Trott and Olson, 2010), 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 (Schrodinger, 2010).

3.4 Results and Discussion

3.4.1 *Pp*DyP has 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 **(Figure SI. 3.1)**. 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) **(Figure SI. 3.1** and **Figure 3.1 a, b)** (Hofbauer et al., 2021; Sugano et al., 2007; Zubieta et al., 2007b). 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 **(Figure SI 3.1)**.

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 **(Figure SI. 3.2a)** with a solvent content of 70% (Kantardjieff and Rupp, 2003; Matthews, 1968) **(Table SI. 3.1)**. 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 **(Figure SI. 3.1** and **Figure SI. 3.2 b-e)**.

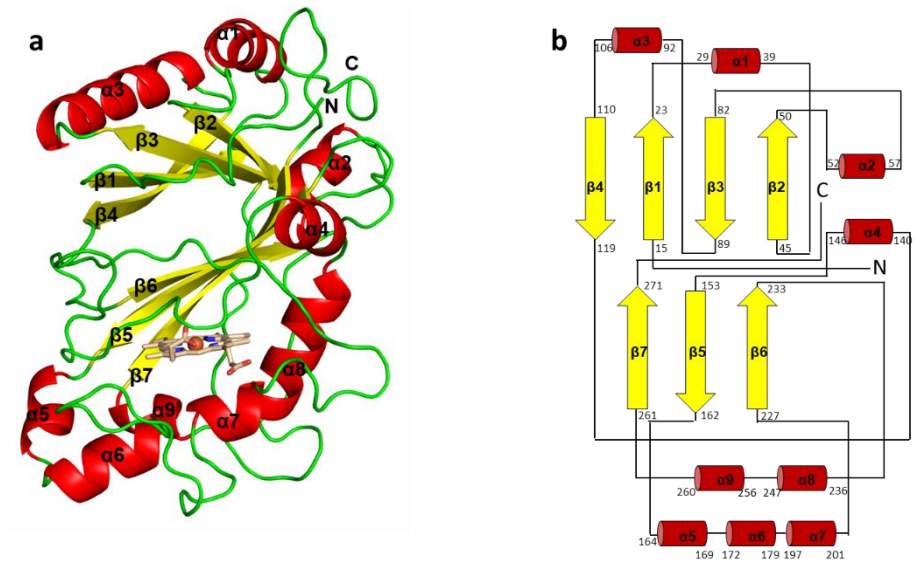


Figure 3.1. PpDyP monomer. (a) Cartoon representation of PpDyP secondary structure elements. The α-helices and β-chains are colored in red and yellow, respectively. Loops are colored in green, and heme cofactor is shown as sticks with carbon, oxygen, and nitrogen colored in light pink, blue and red, respectively. Iron is represented as an orange sphere. (b) Topology diagram of the monomer, with cylinders and arrows representing α-helices and β -chains, respectively, and color-coded as in (a). The residues part of these secondary structure elements is also depicted.

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) **(Figure SI. 3.1** and **Figure 3.1 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) (Krissinel and Henrick, 2007); however, in solution, the three proteins accumulate in the tetrameric and dimeric states as assessed by size-exclusion chromatography (Santos et al., 2014) **(Figure SI. 3.3)**. Fungal DyPs are exclusively monomers as C-terminal insertion regions seem to prevent their dimerization (Zubieta et al., 2007b), whereas bacterial DyPs are mainly described as assembled in dimers (Chen and Li, 2016); and high order oligomerization as a

hexameric state was observed in *Bt*DyP from *B. thetaiotaomicron* (PDB 2GVK) (Zubieta et al., 2007b) and DyPB from *R. jostii* RHA1 (PDB 3QNS) (Roberts et al., 2011).

3.4.2 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 **(Figure 3.2 a** and **Table SI. 3.2)**. T1 is considered the main entrance to H_2O_2, and it is where the conserved distal catalytic residues D132 and R214 are located **(Figure SI. 3.1** and **Figure 3.2 b).** T2 has a similar diameter (~ 3.0 Å) to T1 and a slightly lower length (11.5 Å vs. 14.0 Å) **(Figure 3.2 c,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 **(Figure SI. 3.1)**. 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 **(Figure 3.2 a,f)**, slightly smaller than cavities in other characterized DyPs that show areas of ~400 $Å^2$ and depths of ~12 Å (Liu et al., 2017; Pfanzagl et al., 2018; Roberts et al., 2011; Shrestha et al., 2017; Uchida et al., 2015). The cavity represents an electron-transfer route from the porphyrin radical to the bound substrate in peroxidases (Poulos, 2014). In *Pp*DyP, it is delimited by an extensive network of hydrogen bonds, including four charged residues, E135, R181, E188, and E204 **(Figure 3.2 f)**.

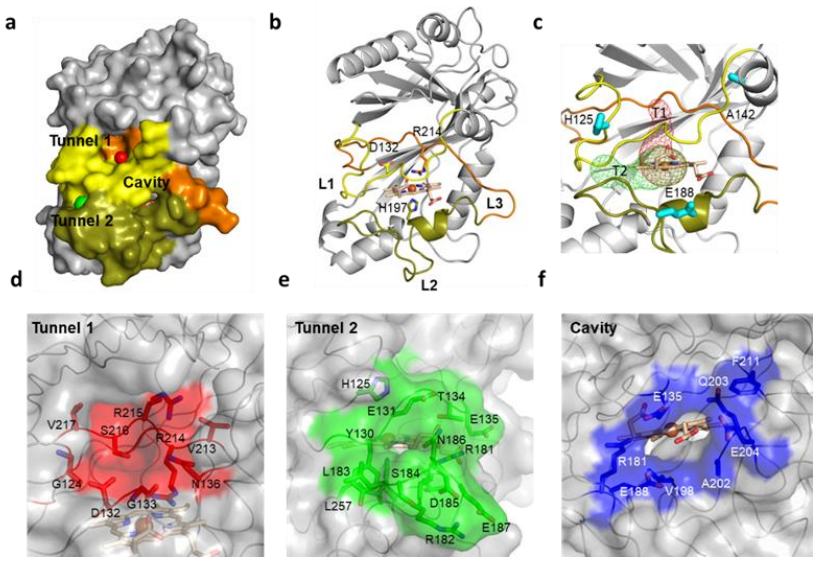


Figure 3.2. Molecular pathways that give access to the heme cofactor in *Pp*DyP: tunnels 1 (T1) and 2 (T2), and cavity, surrounded by loops L1 (120-139, yellow), L2 (179-205, green), and L3 (206-226, orange) represented as solvent accessible surface area (**a**) and cartoon (**b**). The heme group is 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.

The heme pocket of *Pp*DyP is delimited by three conserved loop regions, L1 (120-139), L2 (179-205), and L3 (206-226) **(Figure 3.2 a,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 **(Figure SI. 3.1** and **Figure 3.2 a)**. The mutated residues H125Y and E188K present in variants 6E10 and 29E4 variants, are in L1 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 **(Figure SI. 3.1** and **Figure 3.2 c)**.

The comparison of normalized B-factors (Sun et al., 2019) in the three structures revealed a significantly different profile pattern **(Figure 3.3)**. Variant 6E10 shows the highest overall B-factor value (~80 $Å^2$), followed by 29E4 (~71 $Å^2$) and wild type (~50 $Å^2$) **(Table SI. 3.1)**. The insertion of mutations, in particular of A142V, triggered structural changes leading to fluctuations in the solvent exposure of residues in the loops' regions **(Figure 3.3** and **Table SI. 3.3)**.

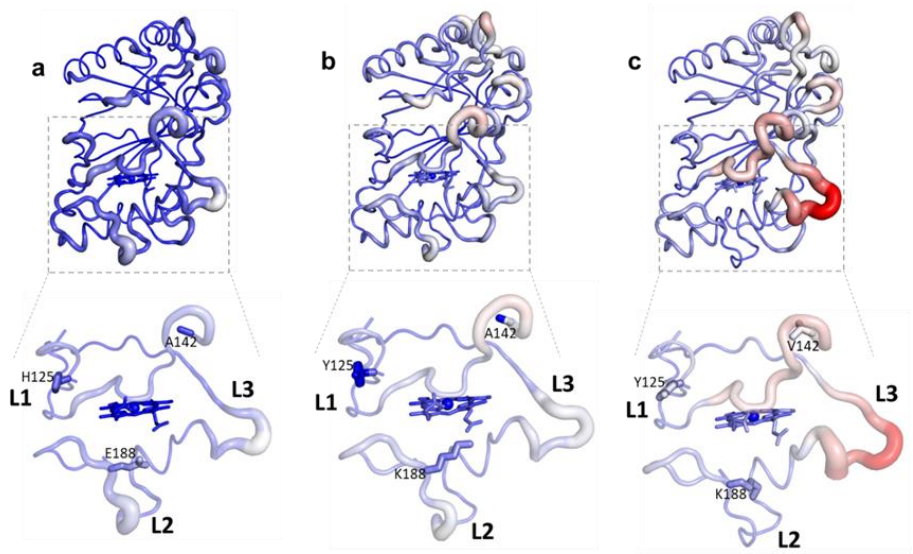


Figure 3.3. Cartoon representations of the main chain of **(a)** wild type, **(b)** 29E4, and **(c)** 6E10 variant structures with thickness proportional to normalized B-factor values color-coded as a blue-white-red color ramp with blue indicating the most negative value (less flexibility) and red the 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

This view is supported by molecular dynamics simulations, where a higher displacement is observed in the loop regions **(Figure 3.4)**. 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 **(Figure SI. 3.4)**. The cavity entrance area in 6E10 is also enlarged, probably due to the higher flexibility of delimiting residues **(Figure 3.5)**.

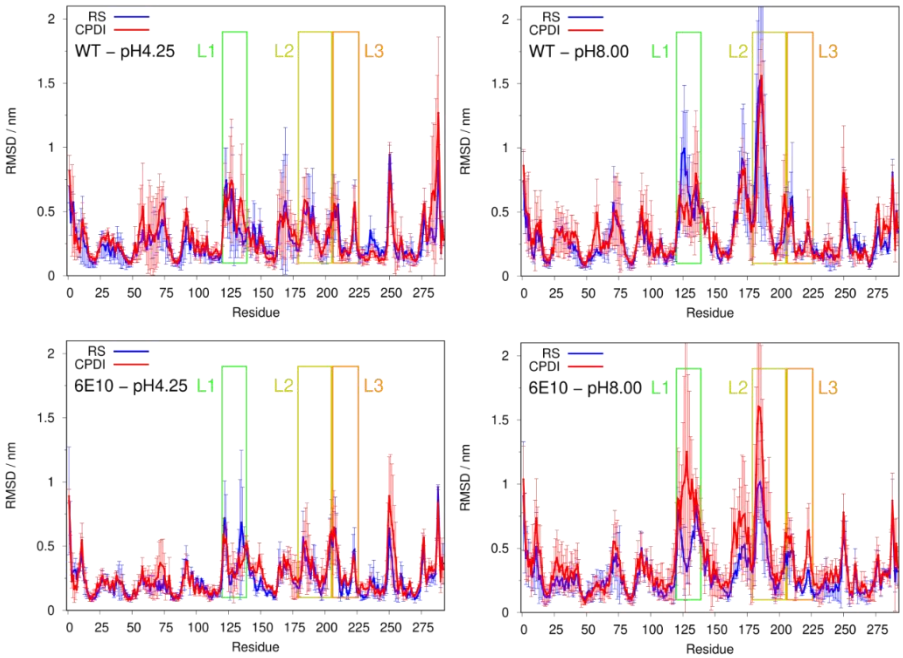


Figure 3.4. Root mean square deviation (RMSD) per residue in wild type PpDyP and 6E10 variant at pH values 4.25 and 8.0. The RMSD values were calculated relative to the initial X-ray structure for resting-state (RS) and compound I (Cpd I). The loops are colored differently for visual clarity: L1 (green), L2 (yellow), and L3 (orange).

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 3.2)** (Barbosa et al., 2020; Brissos et al., 2017).

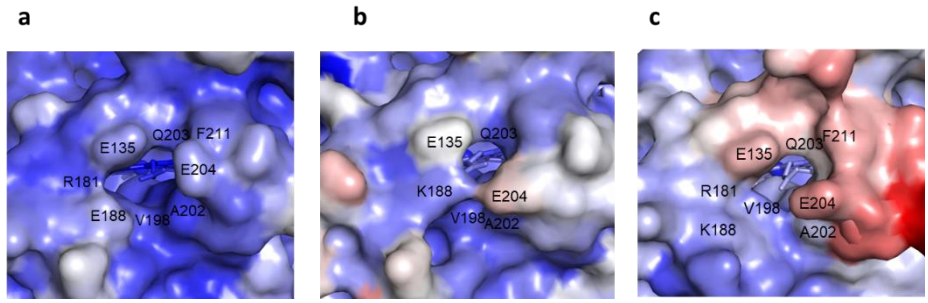


Figure 3.5. Representation of the solvent-accessible surface area of residues that define the cavity in *Pp*DyP wild type**(a)**, 29E4 **(b)**, and 6E10 **(c)** variants. Representation of *a.d.p* values color-coded ramping from blue (28 Å^2) to red (203 Å^2).

Table 3.2. 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

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 3.2)** - was investigated with ensemble docking calculations using snapshots from the Cpd I MD trajectories. The docking results confirm that DMP can bind 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 **(Figure 3.6 a,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 **(Figure 3.6 c,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.

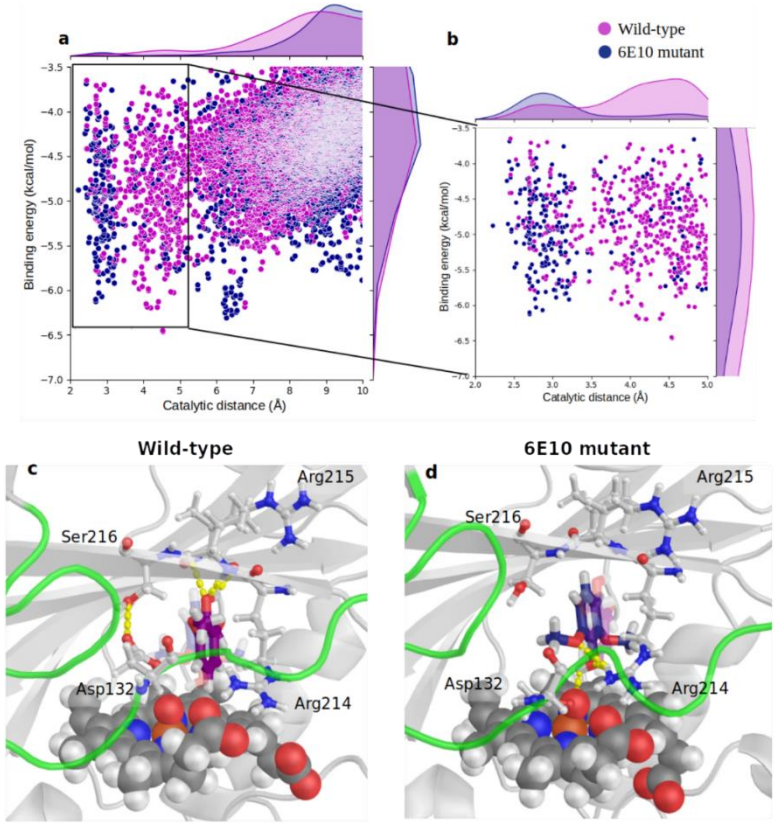


Figure 3.6. Docking of DMP into wild type (magenta) and 6E10 (dark blue) heme sites. (a) Binding energy vs. catalytic distances (distance between DMP hydroxyl oxygen atom and heme-bound oxygen). Each point represents a simulated conformation of the enzyme-substrate complex. (b) Binding energy vs. catalytic 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.

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, **Figure 3.4**) 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 **(Figure 3.6** and **Figure SI. 3.5)**.

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 3.3, Figures 3.7** and **3.8)**. Thermal unfolding profiles were measured to calculate the mid-point thermal temperatures, *i.e.* the melting temperatures (T_m) **(Figure 3.8 b)**. 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 3.3** and **Figure 3.8 a).**

Table 3.3. 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} (ºC)
Wild-type	25.9 ± 2.1	-18. ± 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

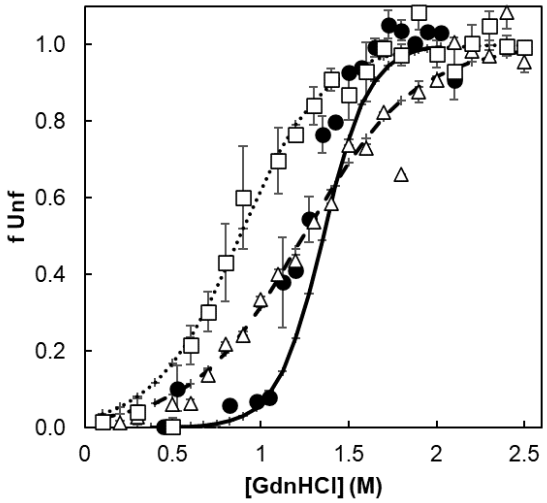


Figure 3.7. Thermodynamic stability of the tertiary structure of *Pp*DyP wild type (filled circles, thick line), 29E4 (open triangles, dashed line), and 6E10 (open squares, dotted line) variants as assessed by fluorescence spectroscopy.

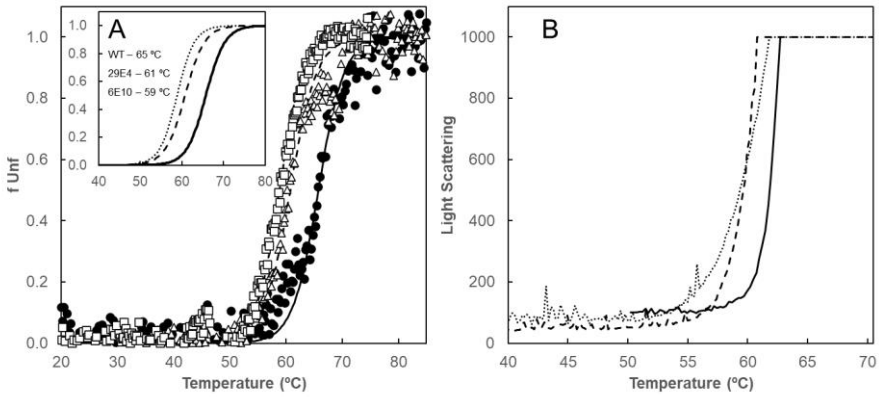


Figure 3.8. Aggregation precedes unfolding. Thermal unfolding profiles (A) of *Pp*DyP wild type (filled circles, thick line), 29E4 (open triangles, dashed line), and 6E10 (open squares, dotted line). Inset in (A) – Fit of the fraction of *Pp*DyP wild type (thick line), 29E4 (dotted line), and 6E10 (dashed line) variants unfolded (f Unf) by the temperature at pH 7.6 as measured by tryptophan emission. Temperature-dependent aggregation profiles (B) of wild type (solid line), 29E4 (dashed line) and 6E10 (dotted line) variants.

Nevertheless, variant 6E10 displayed the lowest T_{agg} compared with 29E4 and wild type enzymes. These results indicate the 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 3.2)**. 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 (Rodrigues et al., 2021) **(Figure 3.9)**.

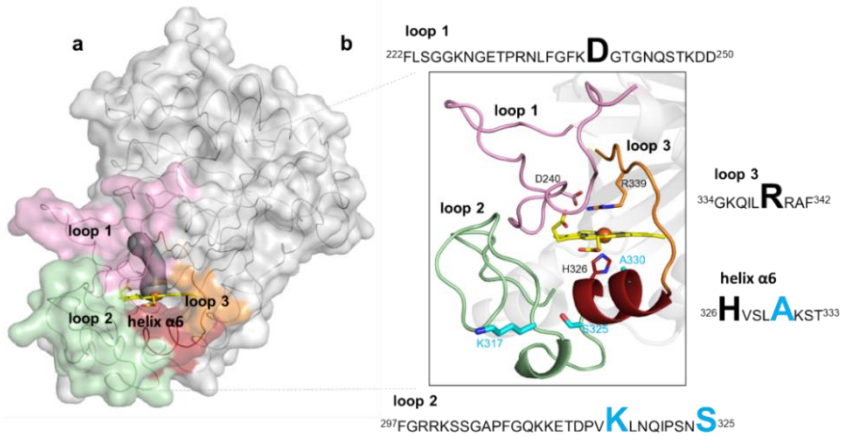


Figure 3.9. Representation of loops and the small helix delimiting the heme pocket in *Bs*DyP. (a) The *Bs*DyP monomeric form is shown as solvent accessible surface colored in gray. The loops that surround the heme pocket are shown in pink (loop 1), green (loop 2), and orange (loop 3) and a small helix α6 is represented in dark red. The access to the heme, as defined using a 1.4-Å rolling probe, is made through a distal tunnel (in dark gray) and one cavity. (b) Zoomed view of the cartoon representation as showed in (a). The catalytic residues, D240 and R339, are shown as sticks with carbon atoms colored in pink and orange, respectively. The heme proximal ligand H326 is shown as sticks with carbon atoms colored in red. The residues K317, S325 and A330 that were replaced in the evolved variant 5G5 are shown as sticks with carbon atoms colored in cyan. The nitrogen and oxygen atoms are shown as sticks colored in blue and red, respectively. Highlighted in the text boxes are the catalytic residues (D240 and R339) and the heme proximal ligand (H326) in black, and the residues (K317, S325 and A330) that were replaced during evolution in light blue.

3.4.3 *Pp*DyP WT and 29E4 higher stability leads to outperform when immobilized in Ag electrodes

The potential of *Pp*DyP WT, 29E4 and 6E10 as 3rd generation electrochemical biosensor for H_2O_2 detection was assessed by the immobilization of these enzymes to Ag electrodes and their characterization. Both WT and 29E4 performed exceptionally, exhibiting a good dynamic response range (1–200 μM H_2O_2), short response times (2 s) and a superior sensitivity (1.3–1.4 $A \cdot M^{-1} \cdot cm^{-2}$) for H_2O_2, outperforming 6E10 as evidenced in the kinetic characterization of the protein/ag electrodes. **(Figure 3.10)** as well as selectivity and long-term stability. *Pp*DyP WT and 29E4 represent excellent candidates for construction of H_2O_2 biosensors, as they show superior sensitivity and shorter response times compared to HRP based H_2O_2 biosensors reported so far in the literature.

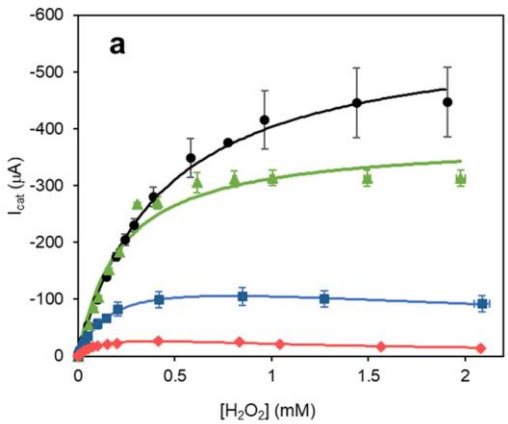


Figure 3.10. Michaelis–Menten plots of catalytic currents vs. H2O2 concentration for wild type (black circles), 6E10 (blue squares) and 29E4 (green triangles) *Pp*DyP/SAM/Ag electrodes.

3.4.4 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 **(Figure 3.2 c)**. 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 **(Figure 3.11 a)**.

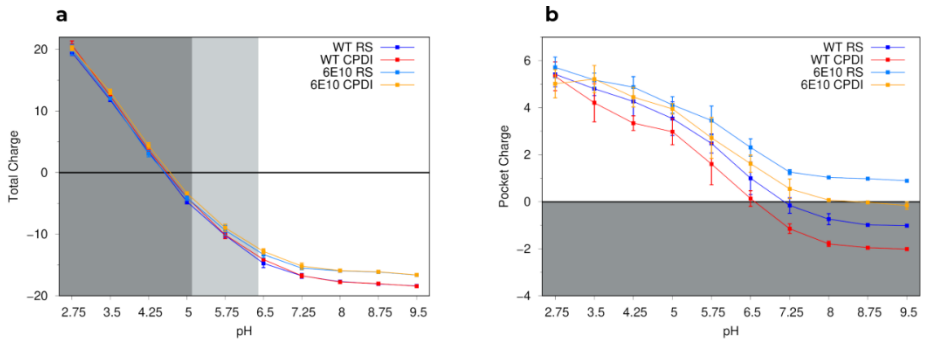


Figure 3.11. Total charge titration curve for Resting-State (RS) and Compound I (CPDI) systems for wild type and 6E10 variants. Each colored region indicates a different average ΔCharge between the 6E10 variant and wild type. These region limits are defined by the pK_a values of the mutated residues, 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.

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 **(Figure 3.12)**. 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 **(Figure 3.11 a)**. 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 **(Figure 3.11 b)**.

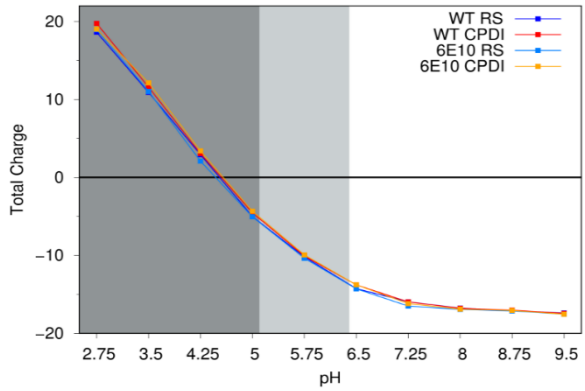


Figure 3.12. Total charge titration curve for all systems excluding residues 125 and 188 from the calculation. The colored regions indicating the average pK_a values of E188 and H125 are still shown to help the comparison with Figure 3.11a. The error values are similar to those shown in Figure 3.11b and were omitted for clarity.

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 3.4)**, except for the distal D132 catalytic residue and H162 near the T2 and buried under L1 **(Figure 3.13)**. 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) **(Figure 3.13)** might also contribute to the protonation of D132 (Harris and Turner, 2002).

The limited access to the solvent of H162 stabilizes its neutral state, leading to very low pK_a values **(Table 3.4)**. Still, during activation at low pH, there is an increase in the population where the H162 side chain is facing away from the heme **(Figure SI. 3.6)**, 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.

Table 3.4 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

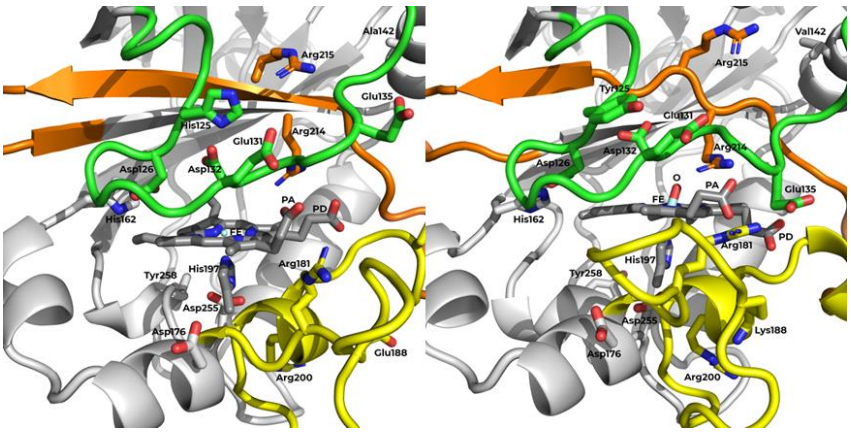


Figure 3.13. Structural highlights of the heme group region in wild type *Pp*DyP **(left)** and 6E10 variant **(right)**. The wild type structure is shown in its resting state (RS), and the 6E10 variant is shown in its activated form (Cpd I). In labeled sticks are shown several critical residues, including 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.

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 **(Figure 3.14)**. 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 more than 0.2 V higher than the respective value for wild type *Pp*DyP ($E^{0'}_{Fe2+/Fe3+}$ = -260 mV) (Brissos et al., 2017). This is in excellent agreement with the deviation of the optimum pH from acidic in wild type ($pH_{op} \sim 4$) to more neutral and alkaline values in the 6E10 variant for both ABTS (pH_{op} = 7) and DMP (pH_{op} = 8) oxidation **(Table 3.2)** (Brissos et al., 2017). 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 (Pfanzagl et al., 2018; Shrestha et al., 2017; Shrestha et al., 2021). 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 (Uchida et al., 2021). 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.

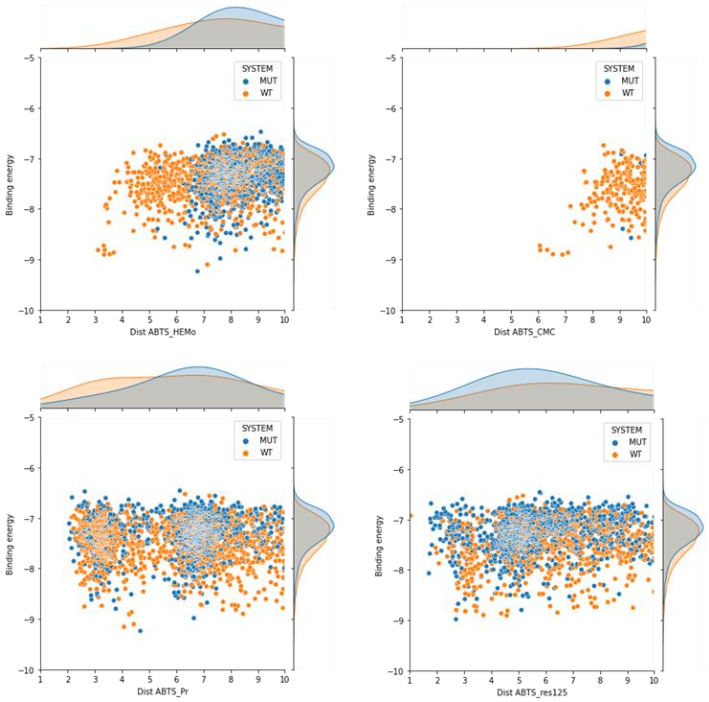


Figure 3.14. Docking of ABTS into wild type (orange) and 6E10 (blue). Binding energy vs. different distances are depicted to show where ABTS is binding: ABTS_HEMo (distance between any ABTS atom and the heme-bound oxygen); ABTS_CMC (minimum distance between ABTS and the porphyrin ring; distance values <3 Å would be indicative of ABTS fully accessing the heme cavity); ABTS_Pr (minimum distance between ABTS and both propionate groups); ABTS_res125 (minimum distance between ABTS and the H125Y residue). Each point represents a simulated conformation of the enzyme-substrate complex. The data shows that ABTS does not reach completely the heme cavity

3.5 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-titratable 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.

Chapter 4

Overcoming low stability of a DyP from *Pseudomonas putida* using a one-step computational approach

This chapter contains unpublished data:

Diogo Silva, Patrícia Borges, Eduardo P. Melo and Lígia O. Martins (2023).” Overcoming low stability of a DyP from *Pseudomonas putida* using a one-step computational approach.” (In preparation)

4.1. Abstract

Stabilizing industrially relevant enzymes is crucial to unlock their full potential for biotechnological applications. In this work, we have used the PROSS and FireProt algorithms to design more stable variants of *P. putida Pp*DyP. The genes coding for the variants suggested by the PROSS web server (with 29 mutations) and FireProt (with 21 mutations) were synthesized, expressed in *E. coli*, the proteins produced, and purified. The kinetic and biochemical characterization show that both enzymes exhibit remarkable improvements being produced at 3-4-fold higher yields than wild type and displaying a shift in their optimal temperature from 15-30 °C (in the wild type) to 60-70 °C. Fluorescent spectrometry shows that the FireProt variant has a melting temperature of 73 °C, while PROSS has a melting temperature of 88 °C, a 10 and 25 °C increase from the wild type. Notably the PROSS variant avoided temperature-induced aggregation of its tertiary structure in a 20 to 100 °C temperature range present in the wild type.

Keywords: Protein Stability, Protein Engineering, Computational Engineering, *in silico* design, DyP-type Peroxidases

4.2 Introduction

Enzyme stability is of paramount importance for biocatalysts in particular in industrial set-ups, where proteins must endure extreme temperatures, pH values, and denaturating solvents (Sanchez-Ruiz, 2010). Protein stability is a complex, multi-variable equation, but the relationship between protein structure and thermostability is evidenced by common features present in thermophilic enzymes and among those features are the significant presence of charged residues, hydrophobic residues, and prolines (Xu, Z. et al., 2020). These differences at the primary sequence level seem to translate into stronger structures, with a better-packed hydrophobic core where enthalpically favourable van der Waals interactions are formed, and higher overall rigidity probably due to the additional interactions: hydrogen bonds, salt bridges, hydrophobic interaction, aromatic interactions, and disulfide bonds (Makhatadze and Privalov, 1995). These features seem to protect proteins from irreversible degradation reactions at high temperatures, such as the deamidation of the amide side chains of Asn and Gln, succinimide formation at Glu and Asp, and oxidation of His, Met, Cys, Trp and Tyr residues (Kumar et al., 2000).

Tackling protein stability with single-point mutations and iterative rounds of random mutagenesis can be time and energy-consuming, but with the predictive power of computational tools improving, computational methods have become very attractive tools for engineering thermostability. These tools provide predictions for fields such as enzyme discovery, protein solubility, enzyme activity and specificity, protein stability or protein dynamics (Marques et al., 2021; Musil et al., 2019), but our focus is on protein stability.

In silico tools for protein stabilization are usually categorized by their engineering approach: Sequence-based engineering, structure-based

engineering, *de novo* design and a comprehensive, combined approach (Marques et al., 2021; Planas-Iglesias et al., 2021).

Sequence-based engineering relies on multiple sequence alignments and consensus analysis: the comparison of a target protein with sequences of thermostable proteins should indicate possible mutation position based on conserved residues among stable enzymes. While stability is complex, stabilization occurred through evolution and *''nothing in biology can be understood except in the light of evolution''* (Dobzhansky, 1973), and several enzymes have been stabilized by the mutation of a hotspot position to conserved residues in their thermophilic counterparts (Zhang et al. 2016, Denisenko et al. 2017, Yang et al. 2017).

Structure-based engineering is a more complex strategy that relies on protein structures and modelling to predict stabilizing predictions based on the classical stabilization mechanisms and energy calculations. For instance, flexible regions in a protein weaken the overall structure and are targets for engineering. Combining the identification of these regions by molecular dynamic simulations, followed by prediction of mutations that lead to higher rigidity (Bommarius and Paye, 2013; Xu, Zhe et al., 2020). Additionally, the calculation of the Gibbs free energy ($\Delta\Delta G_{Fold}$) via tools such as FoldX and Rosetta has allowed for very complete analysis of possibly stabilization mutations (Marques et al., 2021; Planas-Iglesias et al., 2021). By mimicking site-saturation or site-directed mutagenesis in the algorithm, this software can predict the effect of any mutation, at any position, in a time-scale significantly lower than following such protocols in the bench, and by calculation the free energy chance resultant of a mutation, can predict single or multiple-point mutations for thermostabilization.

De novo design has gained traction lately. This approach relies on the knowledge of the structure-function relationship, and bluntly adds a

feature to a protein, based on theoretical information. For instance, it is well established that a well packed hydrophobic core is paramount for protein stability, and by analysis of a target protein with tools such as $Rosetta_{VIP}$ mutations that would improve core packing can be predicted. Also, oxidation of free cysteines is one of irreversible degradation reactions occurring at high temperatures, and can be tackled by *de novo* engineering of dissulfide bonds that, not only protect these cysteines, but also increases the number of molecular interaction in a protein, strengthening its structure (Bender et al., 2016; Borgo and Havranek, 2012).

Lastly, a comprehensive computational analysis tackles the low efficiency of single-strategy approach by combining the aforementioned strategies to provide a complete, multi-approach prediction of stabilizing mutations. Tools such as FireProt (Bednar et al., 2015; Musil et al., 2017), PROSS (Goldenzweig et al., 2016), and FRESCO (Wijma et al., 2018; Wijma and Janssen, 2013) have a multilevel framework combining consensus analysis, energy calculations and in FRESCO, *de novo design*, to analyse possibly stabilizing mutations from sequence-based and energy-based analysis, calculate possible antagonistic effects among predicted mutations and provide a list of target positions for experimental verification.

In this work, we used FireProt and PROSS to stabilize the dye-decolorizing peroxidase from *P. putida* and elucidate the mechanisms for the stabilization of DyPs while providing thermostable *Pp*DyP variants, more suitable for industrial use.

4.3 Materials and Methods

4.3.1 *In silico* stability design

The monomeric structure (PDB Code; Chain A) and dimeric structure (PDB Code; Chain B and D) of PpDyP were submitted to PROSS (https://pross.weizmann.ac.il/) and FireProt (https://loschmidt.chemi.muni.cz/fireprotweb/) algorithms. Both submissions were performed in standards settings, except that the conserved residues 132, 197 and 214 were fixed in PROSS. The original mutation sets provided were manually reviewed **(Table SI. 4.1** and **4.2)**.

4.3.2 Bacterial strains, plasmids and media

E. coli Tuner (DE3, Novagen) was used to express the *ppDyP* gene previously cloned in pET21-a (+) plasmid (Novagen) (pRC-1 plasmid (Santos et al., 2014)) and the *in silico* variants (*Pp*DyP_PR and *Pp*DyP_FP). In the Tuner strain, genes are under the control of T7 promoter and its expression is induced by IPTG. *E. coli* strain DH5α (Novagen) was used for amplification of plasmid constructs. LB medium was used for the growth of *E. coli* strains, supplemented with 100 $\mu g \cdot mL^{-1}$ ampicillin.

4.3.3 Enzyme production and purification

Recombinant enzymes were produced in 1L of LB medium in 5L-Erlenmeyers and purified as previously described (Barbosa et al., 2020). Cell sediments were suspended in 20 mM Tris–HCl buffer, pH 7.6, containing DNase I (10 $\mu g mL^{-1}$ extract), $MgCl_2$ (5 mM), and a mixture of protease inhibitors, antipain, and leupeptin (2 $\mu g mL^{-1}$ extract). Cells were disrupted in a French press. Cell debris was

removed by centrifugation (18,000×g, 2 h, 4 °C) and crude extracts were used for protein purification in an AKTA purifier (GE Healthcare, BioSciences, Uppsala, Sweden) at room temperature. For purification of PpDyP, the crude extract was loaded onto a Q-Sepharose column equilibrated with 20 mM Tris–HCl buffer, pH 7.6. Elution was carried out with 1 M NaCl in buffer. The active fractions were pooled and concentrated before applying on a Superdex 200 HR 16/60 column (GE Healthcare, BioSciences) equilibrated with 20 mM Tris-HCl buffer, pH 7.6 with 0.2 M NaCl. The purified protein was stored at −20 °C until it was used.

4.3.4 Spectroscopic analysis

The UV–visible absorption spectra of purified enzymes were recorded on a Nicolet Evolution 300 spectrophotometer from Thermo Industries (Madison, USA) at room temperature. The heme content was determined by the pyridine ferrohemochrome method using an extinction coefficient of $\varepsilon_{R\text{-}O\ 556\ nm}$ (28.32 mM^{-1} cm^{-1}) (Berry and Trumpower 1987)

4.3.5 Apparent steady-state kinetic analysis

The apparent kinetic parameters (k_{cat} and K_m) were estimated for hydrogen peroxide, and for the reduced substrates 2,2'-azino bis (3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) and 2,6-dimethoxy phenol (DMP). Reactions were monitored at 420 nm (ABTS, ε_{420nm} = 36,000 M^{-1} cm^{-1}) and 468 nm (DMP, ε_{468nm} = 49,600 M^{-1} cm^{-1}) using the Synergy2 microplate reader (BioTek, Vermont, USA) at 25 °C. Reactions to estimate the kinetic parameters for ABTS (0.1 – 3 mM) were performed in the presence of 1.5 mM H_2O_2 (wild type) and 0.5

mM H_2O_2 (PROSS and FireProt), in 100 mM sodium acetate buffer at pH 4.3. Reactions to estimate kinetic parameters of H_2O_2 (0.01 - 6 mM) were performed in the presence of 3 mM ABTS in 100 mM sodium acetate buffer at pH 4.3. The reactions for DMP (0.01 - 2 mM) were performed in 100 mM sodium acetate or phosphate buffer, in the presence of 1.5 mM H_2O_2 (wild type) and 0.5 mM H_2O_2 (PROSS and FireProt), at the optimal pH of the enzymes. The kinetic data was fitted directly using the Michaelis-Menten equation or the equation for non-linear curve that fits enzyme kinetics affected by substrate inhibition (v = Vmax[S]/(Km+[S](1+[S]/Ki))) (Origin software).

4.3.6 Kinetic stability

Thermal inactivation assays were performed as previously described. (Santos et al., 2014) In brief, enzyme preparations in 20 mM Tris-HCl buffer, 200 mM NaCl, pH 7.6 were incubated at 60 °C, and at fixed time intervals, sample aliquots were withdrawn and tested for activity following ABTS oxidation at 25 °C. Thermal inactivation appears to obey first-order kinetics and the half-life values ($t_{1/2}$) were calculated using $t_{1/2}= \ln 2/k_{in}$.

4.3.7 Thermodynamic stability assays

The thermodynamic stability was assessed by steady-state fluorescence measured with a Carry Eclipse spectrofluorimeter (Agilent Technologies) at excitation wavelengths of 296 nm and emission wavelength of 350 nm (Durao, Bento et al., 2006; Fernandes, Martins et al., 2009a). For the equilibrium unfolding studies, guanidine hydrochloride (GdnHCl) concentrations in the range of 0 - 3.5 M in 20 mM Tris–HCl buffer, 200 mM NaCl, pH 7.6, were used to induce protein unfolding, that was measured at room temperature. For the

thermal unfolding assays, the samples containing the enzymes (20 µM in 20 mM Tris-HCl buffer, 200 mM NaCl, pH 7.6) were placed onto a controlled thermal block, and then heated at a rate of 1 °C/min up to 100 °C. Protein aggregation was monitored by measuring static light scattering at 500 nm as excitation and emission wavelengths. The thermodynamic and thermal stability of enzymes was analysed based on a two-state process using the equations previously described. (Fernandes, Martins et al., 2009b).

4.3.8 Other methods

The concentration of purified proteins was estimated using the molar absorption coefficient of *Pp*DyP ($\varepsilon_{280\ nm}$ = 34,850 M-1·cm-1), calculated from the protein sequence using the ExPASy Bioinformatics Resource Portal (http://web.expasy.org). 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).

4.4 Results and Discussion

4.4.1 Design of thermostable PpDyP variants

Although the recently solved structure of *Pp*DyP is homotetramer, in solution, a tetramer, dimer and monomer forms can be found as assessed by SEC **(Figure 3.4)**. Considering this information, monomeric and dimeric structures of *Pp*DyP were submitted in PROSS and FireProt webservers. The four sets of output variants were reviewed and mutations filtered based on rules such as (i) mutations found exclusively in the monomeric set, and (ii) close to the heme factor were excluded. *Pp*DyP encoding genes with 29 mutations for variant based in PROSS (included all mutations from the dimeric structure was accepted) and 21 mutations for a variant based in FireProt (five suggested mutations were discarded) were synthetized **(Table 4.1** and **4.2)**. The mutations suggested by PROSS and FireProt are similarly distributed in the structure **(Figure 4.1 A)**, without overloading a particular secondary structure. The distance to the heme is also very similar between the algorithms, with all mutations distal to the haem (> 10 Å), half of mutations located at 11 to 20 Å, and the other half distanced over 20 Å, at the surface of *Pp*DyP. **(Figure 4.1 B)**.

Table 4.1. Final list of mutations for PROSS algorithm. The mutations suggested by both algorithms are in bold.

Mutation	ASA (%)	Secondary structure position	Distance to heme (Cα-Fe) (Å)
L7I	3	Loop 1	18
R17V	29	Sheet 1	20
A28D	27	Helix 1	39
A35R	37	Helix 1	36
V49I	1	Sheet 2	24
E91D	51	Loop 6	26
L96F	17	Helix 3	28
L97H	57	Helix 3	31
T99A	7	Helix 3	28
E103I	79	Helix 3	32
L110F	7	Sheet 4	32
S111V	45	Sheet 4	31
S115E	40	Sheet 5	22
L120R	31	Loop 8	15
H121Y	5	Loop 8	15
G123D	68	Loop 8	16
T138Q	90	Loop 8	16
D139G	116	Loop 8	20
E140D	141	Helix 4	22
Q144E	109	Helix 4	20
A149D	117	Loop 9	26
Q165D	120	Helix 5	20
D174E	70	Helix 6	17
A194E	51	Loop 12	15
V217M	9	Sheet 5	12
Q222G	118	Loop 15	24
K234H	107	Loop 16	18
E237D	107	Helix 8	16
F283A	97	Loop 20	32

Table 4.2. Final list of mutations for FireProt algorithm. The mutations suggested by both algorithms are in bold.

Mutation	ASA (%)	Secondary structure position	Distance to heme (Cα-Fe) (Å)
A9L	37	Loop 1	20
S25F	34	Loop 2	37
A54V	6	Helix 2	29
E91D	51	Loop 6	26
D94E	47	Helix 3	29
Q100R	66	Helix 3	31
L110F	7	Sheet 4	32
G118A	0	Loop 8	17
L120R	31	Loop 8	15
H125R	28	Loop 8	11
D139G	116	Loop 8	20
A149L	117	Loop 9	26
A155V	1	Sheet 5	15
S169A	86	Loop 11	21
D174E	70	Helix 6	17
V217M	9	Sheet 5	12
S218P	2	Sheet 5	13
Q222G	118	Loop 15	24
L232F	1	Sheet 6	18
E237D	107	Helix 8	16
V279L	24	Loop 20	30

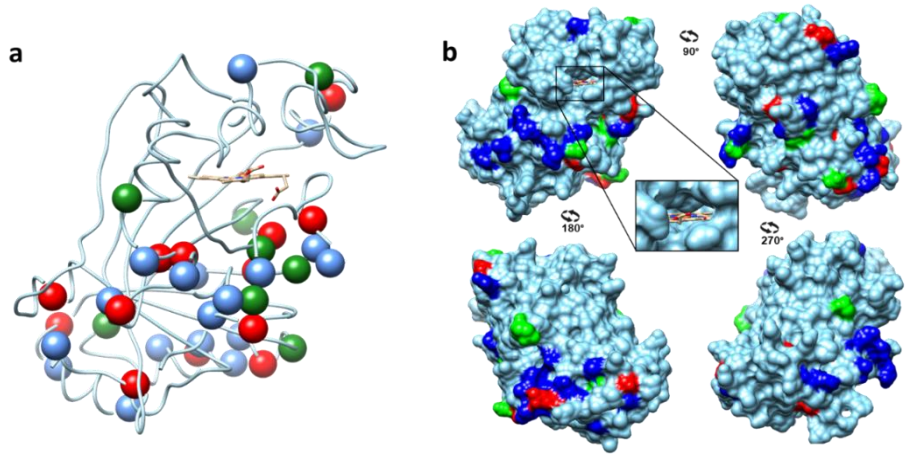


Figure 4.1. (a) Structure of *Pp*DyP with mutated positions highlighted; (b) Surface representation of *Pp*DyP highlighting mutation positions. Blue colored spheres/surfaces correspond to positions mutated by PROSS, red by FireProt, and green represents the aminoacids that both algorithms mutated.

4.4.2 Biochemical and kinetic characterization.

The genes for variants PROSS and FireProt were synthesized by GenScript (NJ, USA) and cloned in pET21b. The variants were overproduced, purified and characterized to evaluate the effect of mutations on proteins' function and stability. The protein yields doubled (27 and 35 mg L^{-1}) in relation to the wild type (14 mg L^{-1}) **(Table 4.3)**, however, the incorporation of the heme cofactor was affected in both variants, which incorporated half mole of heme per mole of protein, contrasting with the full-heme incorporation of wild type. All proteins exhibited the wild type spectroscopic fingerprints, with the characteristic Soret band at 404 nm, Q bands at 520 nm, and a charge transfer band at ~ 640 nm **(Figure. 4.2)**.

Structure stabilisation is known to impact heterologous expression, for instance, PROSS was used to generate mutants of different fungal versatile peroxidase which were successfully expressed in *S. cerevisiae* whereas their wild type counterparts were not (Barber-Zucker et al., 2022). The same was observed for the *Plasmodium*

falciparum reticulocyte-binding protein homolog 5 (PfRH5), a promising malaria vaccine candidate, that was firstly produced in bacteria upon introduction of 18 mutations suggested by PROSS, while the wild type could only be produced in a eukaryotic system (Campeotto et al., 2017).

Protein oligomerization of *Pp*DyP variants seem to be unaffected, since both variants show a combination of monomer, dimer and tetramer, when in solution, as observed for wild type **(Figure 4.3)**.

Table 4.3 Spectroscopic and redox properties of purified *Pp*DyP wild type, PROSS and FireProt variants.

	Enzyme Production (mg L^{-1})	Rz	Λmax (nm)	ε (mM^{-1} cm^{-1})	Heme Content ([heme] (mM)/[protein] (mM))
Wild-type	14 ± 3	1.7	404	32	1.0 ± 0.1
PROSS	27 ± 5	1.2	405	35	0.4 ± 0.02
FireProt	34 ± 4	1.3	404	32	0.5 ± 0.03

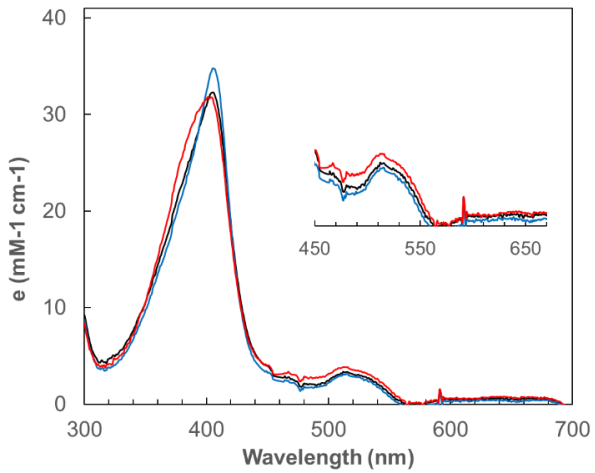


Figure 4.2. UV-vis spectra of purified *Pp*DyP wild type (black), PROSS (blue) and FireProt (red) variants with the Soret band at 404 nm and the Q bands at 520 nm.

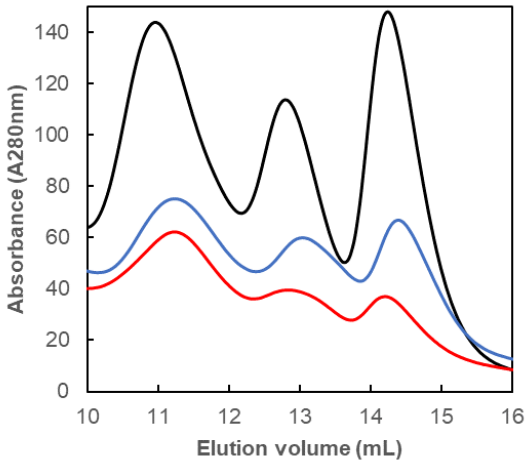


Figure 4.3. Size-exclusion chromatography of *Pp*DyP (black), PROSS (blue) and FireProt (red) variants exhibiting random ratios of the tetrameric (128 kDa), dimeric (64 kDa) and monomeric (32 kDa) forms.

The variants maintained maximum activity for ABTS at pH 4.3 however, in the case of DMP oxidation, the PROSS variant exhibited a 4-unit shift in the optimum pH to pH 8.0, a feature previously observed in a DE variant 6E10 (Brissos et al., 2017) (see chapter 3) **(Figure 4.4 b)**. Apparent steady-state kinetics at room temperature showed that PROSS and FireProt variants displayed ~2 and ~8-fold higher k_{cat} for ABTS, respectively, as compared to wild type **(Table 4.4)**. Importantly, the activity of PROSS variant for DMP was 7-fold improved, which in conjunction with a small increase in the affinity, resulted in a catalytic efficiency (k_{cat}/K_m) 9-fold higher than wild type, and at pH 8, which is more suitable for industrial processes involving lignin, which is more soluble at alkaline pH values (Hamalainen et al., 2018) **(Table 4.4** and **Figure 4.4b)**.

Notably, both PROSS and FireProt variants showed maximal activity around 70 °C, an increase of almost 40 °C in optimal temperature as compared to wild type **(Figure 4.5)**. Their optimal temperature is comparable to the thermophilic *Tfu*DyP (40-70 °C), *Svi*DyP (60-70 °C),

*Dy*PPa (20-70 °C) and *Ftr*DyP (60 °C), however, the combination of high Topt, alkaline pH preference and higher activity for phenolic compounds, is unprecedented amongst DyPs.

The thermodynamic stability of wild type and in *silico* variants was determined by probing the unfolding of the tertiary structure by fluorescence, after chemically inducing denaturation with guanidine hydrochloride (GdnHCl), and interestingly, both proteins require higher concentrations of GdnHCl to denaturate **(Table 4.6** and **Figure 4.6)**.

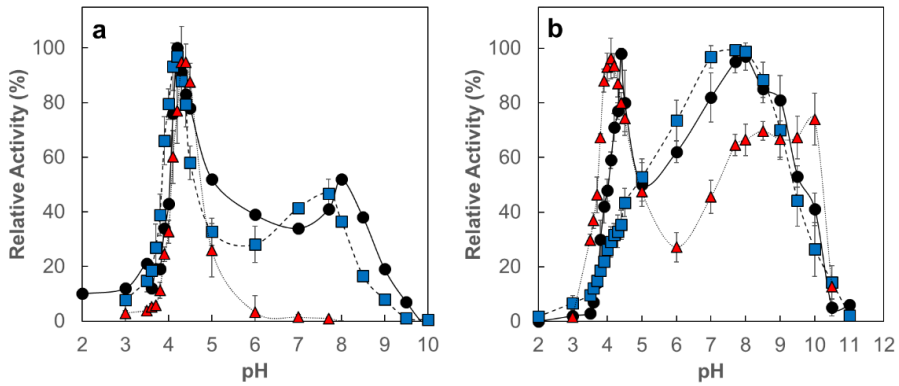


Figure 4.4. pH profiles of *Pp*DyP wild type (black, thick line), PROSS (blue, dashed line) and FireProt (red, dotted line) variants for ABTS (a) and DMP (b). The assays were performed in the presence of 1.5 mM H_2O_2 for wild type, and 0.5 mM H_2O_2 for the variants, using Briton and Robinson buffer (pH range 2 to 12).

Table 4.4. Apparent steady-state kinetic parameters for reduction of H_2O_2, using 3 mM ABTS as substrate, and oxidation of ABTS of *Pp*DyP wild type, in the presence of 1.5 mM H_2O_2, and of PROSS and FireProt variants, in the presence of 0.5 mM H_2O_2, at 25 °C and at the optimal pH of the enzymes

	H_2O_2					ABTS			
	pH	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	K_i (mM)	pH	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
Wild type	4.3	28 ± 3	0.76 ± 0.16	(3.7 ± 0.5) ×10^4	5.8 ± 1.5	4.3	28 ± 3	0.30 ± 0.02	(9.3 ± 0.1) × 10^4
PROSS	4.3	72 ± 3	0.14 ± 0.006	(5.1 ± 0.5) × 10^5	1.2 ± 0.5	4.3	72 ± 3	0.153 ± 0.03	(4.7 ± 0.1) × 10^5
FireProt	4.3	202 ± 16	0.09 ± 0.004	(2.2 ± 0.5) × 10^6	4.8 ± 0.5	4.3	240 ± 6	0.650 ± 0.04	(3.7 ± 0.3) × 10^5

Table 4.5. Apparent steady-state kinetic parameters for oxidation of DMP of *Pp*DyP wild type, in the presence of 1.5 mM H_2O_2, and of PROSS and FireProt variants, in the presence of 0.5 mM H_2O_2, at 25 °C and at the optimal pH of the enzymes.

	DMP			
	pH	**k_{cat} (s^{-1})**	**K_m (μM)**	**k_{cat}/K_m ($M^{-1}s^{-1}$)**
Wild type	4	0.09 ± 0.01	80 ± 20	(1.1 ± 0.2) × 10^3
PROSS	8	0.65 ± 0.01	65 ± 5	(1.0 ± 0.1) × 10^4
FireProt	4	0.06 ± 0.01	100 ± 9	(0.6 ± 0.2) × 10^3

Table 4.6. Thermodynamic stability of the tertiary structure of *Pp*DyP wild type, PROSS and FireProt variants as assessed by fluorescence spectroscopy.

	ΔG^{water} (kJ/mol)	**m ($kJ.mol^{-1}.M^{-1}$)**	**$[GdnHCl]_{50\%}$ (M)**
Wild-type	25.7 ± 4.1	-18.5 ± 0.18	1.40 ± 0.05
PROSS	28.7 ± 4.4	-12.9 ± 0.47	2.24 ± 0.47
FireProt	16.0 ± 4.1	-8.7 ± 0.38	1.85 ± 0.11

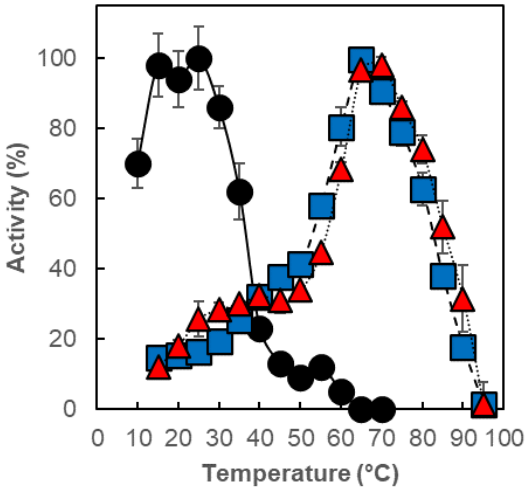


Figure 4.5. Temperature profile of *Pp*DyP wild type (black, thick line), PROSS (blue, dashed line) and FireProt (red, dotted line) variants. The assay was performed with ABTS, in the presence of 1.5 mM H_2O_2 for wild type, and 0.5 mM H_2O_2 for the variants, at pH 4.3.

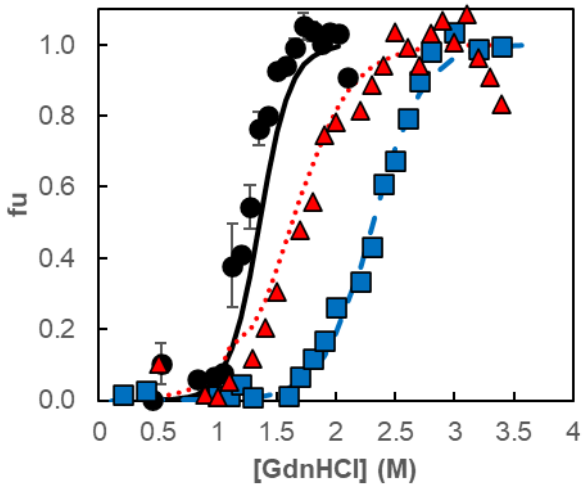


Figure 4.6. Unfolded fraction (fU) of *Pp*DyP wild type (black, thick line) and the PROSS (blue, dashed line) and FireProt variants (red, dotted line) after inducing denaturation with guanidine chloride at pH 7.6. The solid lines are the fit according to the equation fU=exp(-ΔG°/RT)/1+exp(-ΔG°/RT), which assumes the equilibrium N ⇄ U.

4.4.3 Thermal stability features

Enzyme thermostability encompasses thermodynamic and kinetic stabilities (Vieille and Zeikus, 2001). Thermodynamic stability is defined by the enzyme's free energy of stabilization (ΔG_{stab}) and by its melting temperature (T_m, the temperature at which 50% of the protein is unfolded). For the enzymes that unfold irreversibly, only *Tm* can be determined. Kinetic stability depends on the energy barrier to unfolding (i.e., the activation energy [*Ea*] of unfolding). An enzyme's kinetic stability is often expressed as its half-life ($t_{1/2}$) at defined temperatures. The thermal stability features of wild type and the *in silico* variants were evaluated by combining the calculation of melting temperature, temperature-induced aggregation and half-life time at 60 ºC. *Pp*DyP starts aggregating at 61 °C, whereas the FireProt variant aggregates beyond 64 °C, a 3 °C increase that is eclipsed by the lack of aggregation observed in PROSS, up to 100 °C **(Table 4.7** and **Figure 4.7 B,C)**, that could be the result of the new charged residues introduced around the protein surface.

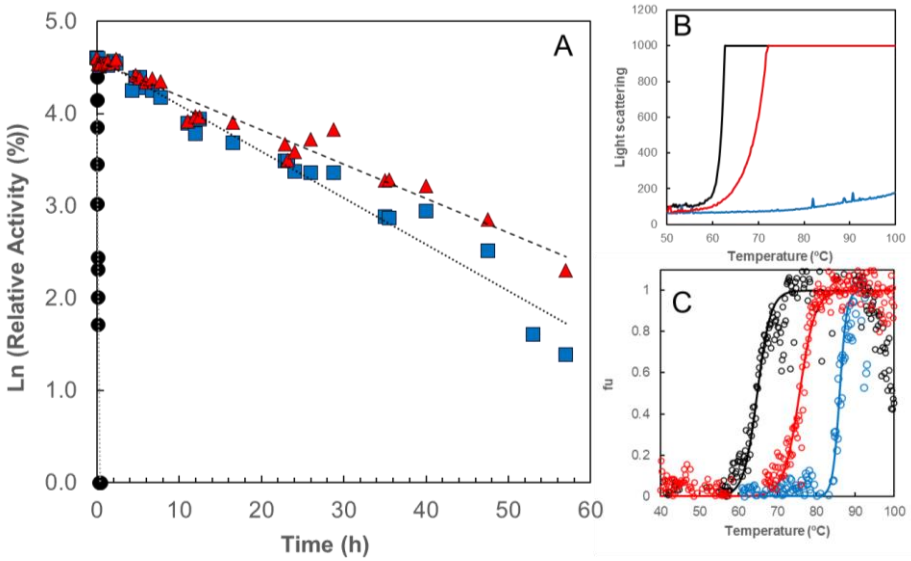


Figure 4.7. Thermal stability of *Pp*DyP wild type (black) and the PROSS (blue) and FireProt (red) variants. The thermal inactivation or kinetic stability was measured at 60 ºC and was fitted to an exponential decay (A). The temperature-induced aggregation (B) was estimated by following particles in solution in the fluorimeter (emission and excitation at 500 nm) upon a temperature ramp from 20 to 100 ºC. A similar temperature ramp was used to evaluate the melting temperature (C), where 50% of the protein in solution is denaturated, was estimated by following the unfolding of tertiary structure in a fluorimeter (emission at 296 nm and excitation at 350 nm)

From the 29 mutations introduced in the PROSS variant, three removed wild type charged residues (R17V, E103I and D139G), but ten mutations led to the introduction of new charged residues (A28D, A35 R, L97H, S115E, L120R, G132D, Q144E, A149D, Q165D and A194E), while FireProt removed one native charged residue (D139G) and introduced two new surface arginines (Q100R and L120R) **(Figure 4.8)**. The introduction of these charges likely disrupts hydrophobic-rich areas, which, upon denaturation, are the core of aggregates, but also might form a repulsive network of electrostatic interactions between molecules (Wayne and Bolon, 2010). Since the temperature-dependent aggregation of wild type and FireProt precede their

unfolding, it is not possible to accurately measure thermal unfolding mid-points. However, in the case of PROSS, that does not apply, and this variant exhibited a melting temperature of 88 °C, which could be a 20 °C from the apparent unfolding of *Pp*DyP. Similar improvements have been previously reported, in particular for PROSS variants. For instance, PROSS generated mutants of two CplS proteins from *Agrobacterium tumefaciens* exhibited 25 to 29 °C increased melting temperature (Tullman et al., 2020) and an acetylcholinesterase (AChE) variant bearing 51 mutations suggested by PROSS had a 16.8 °C improvements in its melting temperature (Goldenzweig et al., 2016).

Table 4.7. Thermal stability of the tertiary structure of *Pp*DyP wild type, PROSS and FireProt variants as assessed by fluorescence spectroscopy and half-life measurements at 60 ºC.

	T_{agg} (ºC)	T_m (ºC)	$t_{1/2}$ (h) 60 ºC
Wild-type	61 ± 0.2	65 ± 1	0.05
PROSS	-	88 ± 2	16 ± 3 h
FireProt	64 ± 2	73 ± 2	23 ± 2 h

These results reveal the improved thermal resistance of the protein structures, but the benefits are also observed in the protein function. FireProt and PROSS have a massive increase (360- and 450-fold, respectively) in the half-life time after incubation at 60 °C **(Table 4.7** and **Figure 4.7 A)**, maintaining high levels of activity in the first hours, while wild type completely loses activity after twenty minutes. Improved half-life times were observed in FireProt generated mutants of hyperstable fibroblast growth factor 2, with one of the three variants having a half-life time of over 24 h when incubated at 65 °C (Dvorak et al., 2018), and PROSS generated mutants of versatile peroxidases (Barber-Zucker et al., 2022). PROSS exhibit Tm values, 30 and 40 °C

higher than the only melting temperature (44 °C) reported for e.g the DyP, *Cbo*DyP.

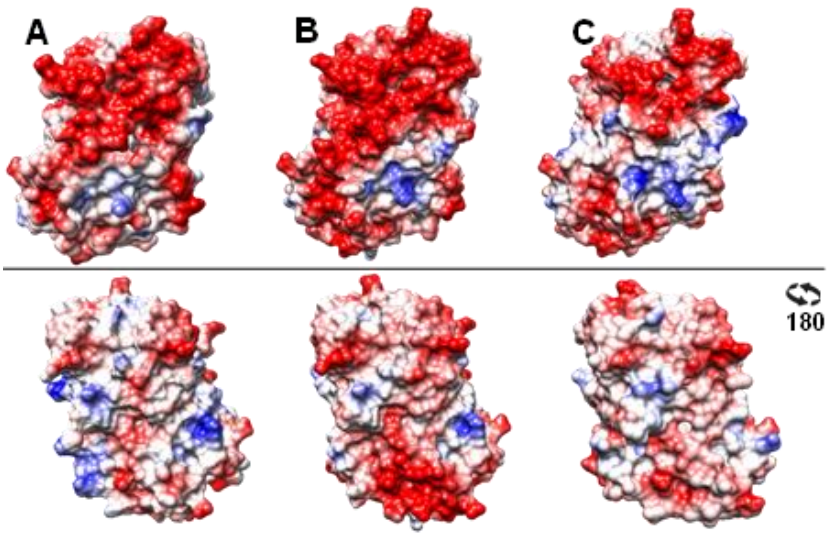


Figure 4.8. Comparison of the surface electrostatic potential distribution of *Pp*DyP (A), PROSS (B) and FireProt (C). Surface electrostatic potential was calculated with coulombic surface colouring in Chimera UCSF. Potential is colored on a scale from -10 kcal/mol (red) to +10.0 kcal/mol (blue).

4.5 Concluding Remarks

With the ever-growing use of enzymes in industrial setups, the modulation and understanding of enzyme activity, specificity, and stability has become more critical. Directed evolution and rational approaches to enzyme engineering have been successful to improve enzyme activity and specificity but underperform when a larger number of mutations is required to target broad dynamic features that need to be improved for stability. Webservers such as PROSS and FireProt provide a one-step, simple solution for improved stability, using energy and sequence computational analysis of protein structure. In this work

we submitted the *Pp*DyP structure to both servers to generate two variants with 29 and 21 mutations, from PROSS and FireProt respectively. Both variants exhibited outstanding stability features, with increases in protein expression (2-fold), higher optimum temperature of activity (+ 40 ºC), higher resistance to denaturants, higher melting temperatures (ΔT_m > 10-20 °C) and delayed aggregation. The improved stability came with no cost to activity, with both variants oxidizing ABTS and H_2O_2 2 to 10 times better than wild type. Interestingly, PROSS exhibited ~7-fold higher activity towards the lignin-related phenolic compound DMP ~7-fold, with a pH optimum of pH 8.0, a 4-unit shift from *Pp*DyP. This property, coupled with the remarkable stability, makes this variant a potential catalyst for the valorisation of lignin and lignin monomer in the biorefinery realm.

Chapter 5

Conclusions and Future Perspectives

Enzymes provide a sustainable and economical solution for several problems in the world, such as cleaner energy and materials, food properties enhancement, bioremediation, etc.

Dye-decolorizing peroxidases are a group of bacterial heme peroxidases with a wealth of potential as biocatalysts due to their capability to oxidize a broad range of substrates, including phenolic and non-phenolic lignin-related compounds, and its ease to be engineered by techniques such as the Nobel winning prize Directed Evolution.

In this thesis, we aimed to further our understanding of the catalytic, stability and structural dynamics of a dye-decolorizing peroxidase from *P. putida* and its directed evolution engineered variants, with a focus on the improvement of the stability of these proteins, a field still understudied.

We started by investigating the catalytic ability of wild type DyP from *P. putida* and its evolved variant 6E10 over a range of lignin-related phenolic compounds. The improved activity for the lignin-related DMP by 6E10 was translated to 24 other phenolic compounds, showing the power of directed evolution and our screening method, but also the potential of these enzymes to tackle lignin heterogeneity and valorize lignin monomers. The *Pp*DyP-mediated oxidation of monomers such as syringol, vanillin, or acetovanillone resulted in the production of syringaresinol, divanillin, and diapoxynin dimeric phenolic compounds with industrial relevance that, at this moment, are obtained chemically. These results can serve as proof of concept for using 6E10 to functionalize lignin-related monomers derived from lignin depolymerization and the consequent production of added-value compounds. However, further optimizing these reactions seems like a good next step to improve specificity and product yields, since the formation of radicals due to DyP activity, combined with the high

turnover rate, can lead to secondary products and product polymerization that negatively impact the process.

While the catalytic power of 6E10 is a good outlook on the system, this variant's lackluster stability features may hinder its use as an industrial catalyst. To better understand the underlying determinants of DyP stability, the structures of *Pp*DyP, 6E10, and 29E4, a DE variant engineered from 6E10, for improved stability, were determined. The characterization of the three proteins highlighted a trade-off between activity and stability: the improved activity of 6E10 also resulted in a decrease in stability, while regaining stability in 29E4 led to the loss of the activity improvements from 6E10. The analysis of the structures indicates that the structures' flexibility played a role in the balance between activity and stability, particularly the flexibility of the three loops surrounding the heme pocket, which contain the catalytic residues. The high flexibility of these loops in 6E10 favored catalysis, facilitating the reactions' turnover but weakening its overall stability. The same was observed for another DyP's variant from *B. subtilis*; a DE variant with 7-fold higher activity for DMP exhibited an increase in the flexibility of these conserved loops, with adverse effects on protein stability. The work we have done in *Pp*DyP and *Bs*DyP highlights, for the first time, the relevance of these conserved loops for the function of DyP, showing that these secondary structures are essential targets for DyP engineering and modulating of its activity and stability features.

Additionally, computational analysis of the pK_a of titratable residues on the heme pocket helped answer a long-lasting question regarding DyPs. Contrary to classical peroxidases, this family has an optimum activity pH in the acidic range, with most enzymes showing maximum activity between pH 3 and 5. However, this property was believed to be related to the catalytic aspartate, which replaces a histidine in classical peroxidases. Our results indicate that the determining cause of this

property seems to be the overall electrostatic environment of the heme; the more positive environment in 6E10 accounts for its alkaline catalytic preference. At the same time, 29E4 and wild type have a more acidic heme pocket, hence the pH optimum at pH 4.3. While these results are crucial for our fundamental understanding of DyPs, they also provide another target for enzyme engineering, as one could modulate the overall pocket charge to tune the pH optimum of the enzyme.

Enzyme stability is a crucial factor in the application of proteins in biotechnological industries. Regardless of its activity, an unstable enzyme cannot survive the harsh conditions usually found in reactors, from extreme pHs to high temperatures and high concentrations of chemicals. While the activity of an enzyme can be finely tuned by one or two mutations, tackling protein stability requires a broader approach. The directed evolution approach we previously used was not appropriate for stabilizing DyP since only one mutation was introduced per generation, and we have observed a conflict between activity and stability. Therefore, we attempted to engineer the thermostability feature in *Pp*DyP using a novel, fast and straightforward computational approach. We uploaded the crystal structure of *Pp*DyP to the webservers PROSS and FireProt, which suggested variants with 21 and 29 mutations, respectively. Both variants exhibited remarkable stability, with no negative impact observed in activity: a 40 ºC increase in Top of activity, 10 to 20 ºC increase in T_m, higher resistance to denaturants, and 300 to 400 times higher half-life time when incubated at 60 ºC.

Interestingly, the PROSS variant also oxidized DMP nearly ten times better than *wild type* and at a pH optimum of 8.0, resembling what we observed in 6E10. The stability properties of PROSS are unprecedented in the DyP family and, combined with the higher activity

for DMP, make this variant a perfect catalyst for industrial applications. This work provided an excellent enzyme for possible industrial exploration and an interesting enzyme to further study and understand the stabilization mechanisms that underlie such a stability improvement. We expect to use modelling and other computation approaches to understand the changes better. While we aim to solve the structure of both variants, we also plan to engineer PROSS further to remove potentially damaging mutations, improve stability, and strengthen its catalytic features.

During this thesis, we advanced our knowledge of DyPs while providing different improved enzyme variants for industrial exploration.

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1-2, 19-25.

Abu-Omar, M.M., Barta, K., Beckham, G.T., Luterbacher, J.S., Ralph, J., Rinaldi, R., Roman-Leshkov, Y., Samec, J.S.M., Sels, B.F., Wang, F., 2021. Guidelines for performing lignin-first biorefining. Energy Environ Sci 14, 262-292.

Adams, P., Afonine, P., Echols, N., Headd, J., Grosse-Kunstleve, R., Moriarty, N., 2010. New tools for structure refinement in Phenix. Acta Crystallogr A 66, S15.

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., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213-221.

Adelakun, O.E., Kudanga, T., Green, I.R., le Roes-Hill, M., Burton, S.G., 2012a. Enzymatic modification of 2,6-dimethoxyphenol for the synthesis of dimers with high antioxidant capacity. Process Biochem 47, 1926-1932.

Adelakun, O.E., Kudanga, T., Parker, A., Green, I.R., le Roes-Hill, M., Burton, S.G., 2012b. Laccase-catalyzed dimerization of ferulic acid amplifies antioxidant activity. J Mol Catal 74, 29-35.

Ahmad, M., Roberts, J.N., Hardiman, E.M., Singh, R., Eltis, L.D., Bugg, T.D., 2011. Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase. Biochemistry 50, 5096-5107.

Alessa, A.H.A., Tee, K.L., Gonzalez-Perez, D., Omar Ali, H.E.M., Evans, C.A., Trevaskis, A., Xu, J.H., Wong, T.S., 2019. Accelerated directed evolution of dye-decolorizing peroxidase using a bacterial extracellular protein secretion system (BENNY). Bioresour Bioprocess 6, 20.

Aljawish, A., Chevalot, I., Jasniewski, J., Paris, C., Scher, J., Muniglia, L., 2014. Laccase-catalysed oxidation of ferulic acid and ethyl ferulate in aqueous medium: A green procedure for the synthesis of new compounds. Food Chem 145, 1046-1054.

Arndt, U.W., Crowther, R.A., Mallett, J.F., 1968. A computer-linked cathode-ray tube microdensitometer for x-ray crystallography. Journal of scientific instruments 1, 510-516.

Bajpai, V.K., Alam, M.B., Quan, K.T., Ju, M.K., Majumder, R., Shukla, S., Huh, Y.S., Na, M., Lee, S.H., Han, Y.K., 2018. Attenuation of inflammatory responses by (+)-syringaresinol via MAP-Kinase-mediated suppression of NF-kappaB signaling *in vitro* and *in vivo*. Sci Rep 8, 9216.

Baptista, A.M., Soares, C.M., 2001. Some theoretical and computational aspects of the inclusion of proton isomerism in the protonation equilibrium of proteins. J Phys Chem B 105, 293-309.

Baratto, M.C., Sinicropi, A., Linde, D., Saez-Jimenez, V., Sorace, L., Ruiz-Duenas, F.J., Martinez, A.T., Basosi, R., Pogni, R., 2015. Redox-Active Sites in *Auricularia auricula-judae* Dye-Decolorizing Peroxidase and Several Directed Variants: A Multifrequency EPR Study. J Phys Chem B 119, 13583-13592.

Barber-Zucker, S., Mindel, V., Garcia-Ruiz, E., Weinstein, J.J., Alcalde, M., Fleishman, S.J., 2022. Stable and Functionally Diverse Versatile

Peroxidases Designed Directly from Sequences. J Am Chem Soc 144, 3564-3571.

Barbosa, C., Silveira, C.M., Silva, D., Brissos, V., Hildebrandt, P., Martins, L.O., Todorovic, S., 2020. Immobilized dye-decolorizing peroxidase (DyP) and directed evolution variants for hydrogen peroxide biosensing. Biosens Bioelectron 153, 112055.

Bednar, D., Beerens, K., Sebestova, E., Bendl, J., Khare, S., Chaloupkova, R., Prokop, Z., Brezovsky, J., Baker, D., Damborsky, J., 2015. FireProt: Energy- and Evolution-Based Computational Design of Thermostable Multiple-Point Mutants. Plos Comput Biol 11.

Bender, B.J., Cisneros, A., 3rd, Duran, A.M., Finn, J.A., Fu, D., Lokits, A.D., Mueller, B.K., Sangha, A.K., Sauer, M.F., Sevy, A.M., Sliwoski, G., Sheehan, J.H., DiMaio, F., Meiler, J., Moretti, R., 2016. Protocols for Molecular Modeling with Rosetta3 and RosettaScripts. Biochemistry 55, 4748-4763.

Benvidi, A., Dadras, A., Abbasi, S., Tezerjani, M., Rezaeinasab, M., Tabaraki, R., Namazian, M., 2019. Experimental and computational study of the pKa of coumaric acid derivatives. J Chin Chem Soc 66, 589-593.

Berendsen, H.J.C., Postma, J.P.M., Vangunsteren, W.F., Dinola, A., Haak, J.R., 1984. Molecular-Dynamics with Coupling to an External Bath. J Chem Phys 81, 3684-3690.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Res 28, 235-242.

Bhabha, G., Lee, J., Ekiert, D.C., Gam, J., Wilson, I.A., Dyson, H.J., Benkovic, S.J., Wright, P.E., 2011. A Dynamic Knockout Reveals That Conformational Fluctuations Influence the Chemical Step of Enzyme Catalysis. Science 332, 234-238.

Bissaro, B., Varnai, A., Rohr, A.K., Eijsink, V.G.H., 2018. Oxidoreductases and Reactive Oxygen Species in Conversion of Lignocellulosic Biomass. Microbiol Mol Biol Rev 82.

Bommarius, A.S., Paye, M.F., 2013. Stabilizing biocatalysts. Chem Soc Rev 42, 6534-6565.

Borchert, A.J., Henson, W.R., Beckham, G.T., 2022. Challenges and opportunities in biological funneling of heterogeneous and toxic substrates beyond lignin. Curr Opin Biotechnol 73, 1-13.

Borgo, B., Havranek, J.J., 2012. Automated selection of stabilizing mutations in designed and natural proteins. PNAS 109, 1494-1499.

Brissos, V., Tavares, D., Sousa, A.C., Robalo, M.P., Martins, L.O., 2017. Engineering a Bacterial DyP-Type Peroxidase for Enhanced Oxidation of Lignin-Related Phenolics at Alkaline pH. ACS Catal 7, 3454-3465.

Brown, M.E., Barros, T., Chang, M.C., 2012. Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium. ACS Chem Biol 7, 2074-2081.

Brown, M.E., Walker, M.C., Nakashige, T.G., Iavarone, A.T., Chang, M.C., 2011. Discovery and characterization of heme enzymes from unsequenced bacteria: application to microbial lignin degradation. J Am Chem Soc 133, 18006-18009.

Bussi, G., Donadio, D., Parrinello, M., 2007. Canonical sampling through velocity rescaling. J Chem Phys 126.

Campbell, E.C., Correy, G.J., Mabbitt, P.D., Buckle, A.M., Tokuriki, N., Jackson, C.J., 2018. Laboratory evolution of protein conformational dynamics. Curr Opin Struct Biol 50, 49-57.

Campeotto, I., Goldenzweig, A., Davey, J., Barfod, L., Marshall, J.M., Silk, S.E., Wright, K.E., Draper, S.J., Higgins, M.K., Fleishman, S.J., 2017. One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen. PNAS 114, 998-1002.

Cardullo, N., Muccilli, V., Tringali, C., 2022. Laccase-mediated synthesis of bioactive natural products and their analogues. Rsc Chem Biol 3, 614-647.

Catucci, G., Valetti, F., Sadeghi, S.J., Gilardi, G., 2020. Biochemical features of dye-decolorizing peroxidases: Current impact on lignin degradation. Biotechnol Appl Biochem 67, 751-759.

Chaplin, A.K., Chicano, T.M., Hampshire, B.V., Wilson, M.T., Hough, M.A., Svistunenko, D.A., Worrall, J.A.R., 2019. An Aromatic Dyad Motif in Dye Decolourising Peroxidases Has Implications for Free Radical Formation and Catalysis. Chemistry 25, 6141-6153.

Chaplin, A.K., Petrus, M.L., Mangiameli, G., Hough, M.A., Svistunenko, D.A., Nicholls, P., Claessen, D., Vijgenboom, E., Worrall, J.A., 2015. GlxA is a new structural member of the radical copper oxidase family and is required for glycan deposition at hyphal tips and morphogenesis of *Streptomyces lividans*. Biochem J 469, 433-444.

Chen, C., Li, T., 2016. Bacterial dye-decolorizing peroxidases: Biochemical properties and biotechnological opportunities. Phys Sci Rev 1.

Chen, C., Shrestha, R., Jia, K., Gao, P.F., Geisbrecht, B.V., Bossmann, S.H., Shi, J., Li, P., 2015. Characterization of Dye-decolorizing Peroxidase (DyP) from *Thermomonospora curvata* Reveals Unique Catalytic Properties of A-type DyPs. J Biol Chem 290, 23447-23463.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66, 12-21.

Chen, Z., Wan, C., 2017. Biological valorization strategies for converting lignin into fuels and chemicals. Renew Sust Energ Rev 73, 610-621.

Cherry, J.R., Lamsa, M.H., Scheider, P., Vind, J., Svendsen, A., Jones, A., Pedersen, A.H., 1999. Directed evolution of a fungal peroxidase. Nat Biotechnol 17.

Chio, C.L., Sain, M., Qin, W.S., 2019. Lignin utilization: A review of lignin depolymerization from various aspects. Renew Sust Energ Rev 107, 232-249.

Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., Damborsky, J., 2012. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. Plos Comput Biol 8.

Colpa, D.I., Lončar, N., Schmidt, M., Fraaije, M.W., 2017. Creating Oxidase-Peroxidase Fusion Enzymes as a Toolbox for Cascade Reactions. ChemBioChem 18, 2226-2230.

Contreras, H., Joens, M.S., McMath, L.M., Le, V.P., Tullius, M.V., Kimmey, J.M., Bionghi, N., Horwitz, M.A., Fitzpatrick, J.A., Goulding, C.W., 2014. Characterization of a *Mycobacterium tuberculosis* nanocompartment and its potential cargo proteins. J Biol Chem 289, 18279-18289.

de Eugenio, L.I., Peces-Perez, R., Linde, D., Prieto, A., Barriuso, J., Ruiz-Duenas, F.J., Martinez, M.J., 2021. Characterization of a Dye-

Decolorizing Peroxidase from *Irpex lacteus* Expressed in *Escherichia coli*: An Enzyme with Wide Substrate Specificity Able to Transform Lignosulfonates. J Fungi 7.

DeAngelis, K.M., Sharma, D., Varney, R., Simmons, B., Isern, N.G., Markilllie, L.M., Nicora, C., Norbeck, A.D., Taylor, R.C., Aldrich, J.T., Robinson, E.W., 2013. Evidence supporting dissimilatory and assimilatory lignin degradation in *Enterobacter lignolyticus* SCF1. Front Microbiol 4.

DeLano, W.L., 2002. Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 12, 14-20.

Dhankhar, P., Dalal, V., Mahto, J.K., Gurjar, B.R., Tomar, S., Sharma, A.K., Kumar, P., 2020. Characterization of dye-decolorizing peroxidase from *Bacillus subtilis*. Arch Biochem Biophys 693, 108590-108605.

Diederichs, K., Karplus, P.A., 1997. Improved R-factors for diffraction data analysis in macromolecular crystallography. Nature structural biology 4, 269-275.

Dobzhansky, T., 1973. Nothing in Biology Makes Sense except in Light of Evolution. Am Biol Teach 35, 125-129.

Doyle, W.A., Blodig, W., Veitch, N.C., Piontek, K., Smith, A.T., 1998. Two Substrate Interaction Sites in Lignin Peroxidase Revealed by Site-Directed Mutagenesis. Biochemistry 37.

Durao, P., Bento, I., Fernandes, A.T., Melo, E.P., Lindley, P.F., Martins, L.O., 2006. Perturbations of the T1 copper site in the CotA laccase from *Bacillus subtilis*: structural, biochemical, enzymatic and stability studies. J Biol Inorg Chem 11, 514-526.

Dvorak, P., Bednar, D., Vanacek, P., Balek, L., Eiselleova, L., Stepankova, V., Sebestova, E., Kunova Bosakova, M., Konecna, Z., Mazurenko, S., Kunka, A., Vanova, T., Zoufalova, K., Chaloupkova, R., Brezovsky, J., Krejci, P., Prokop, Z., Dvorak, P., Damborsky, J., 2018. Computer-assisted engineering of hyperstable fibroblast growth factor 2. Biotechnol Bioeng 115, 850-862.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66, 486-501.

Engh, R.A., Huber, R., 1991. Accurate Bond and Angle Parameters for X-Ray Protein-Structure Refinement. Acta Crystallogr A 47, 392-400.

Erdemgil, F.Z., Sanli, S., Sanli, N., Ozkan, G., Barbosa, J., Guiteras, J., Beltran, J.L., 2007. Determination of pK_a values of some hydroxylated benzoic acids in methanol-water binary mixtures by LC methodology and potentiometry. Talanta 72, 489-496.

Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., 1995. A Smooth Particle Mesh Ewald Method. J Chem Phys 103, 8577-8593.

Evans, P.R., 2006. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr, 72-82.

Fang, Z., Nikafshar, S., Hegg, E.L., Nejad, M., 2020. Biobased Divanillin As a Precursor for Formulating Biobased Epoxy Resin. Acs Sustain Chem Eng 8, 9095-9103.

Fernandes, A.T., Martins, L.O., Melo, E.P., 2009. The hyperthermophilic nature of the metallo-oxidase from *Aquifex aeolicus*. Biochim Biophys Acta Proteins Proteom 1794, 75-83.

Fernández-Fueyo, E., Davó-Siguero, I., Almendral, D., Linde, D., Baratto, M.C., Pogni, R., Romero, A., Guallar, V., Martínez, A.T., 2018. Description of

a Non-Canonical Mn(II)-Oxidation Site in Peroxidases. ACS Catal 8, 8386-8395.

Fernandez-Fueyo, E., Ruiz-Duenas, F.J., Ferreira, P., Floudas, D., Hibbett, D.S., Canessa, P., Larrondo, L.F., James, T.Y., Seelenfreund, D., Lobos, S., Polanco, R., Tello, M., Honda, Y., Watanabe, T., Watanabe, T., Ryu, J.S., Kubicek, C.P., Schmoll, M., Gaskell, J., Hammel, K.E., St John, F.J., Vanden Wymelenberg, A., Sabat, G., Splinter BonDurant, S., Syed, K., Yadav, J.S., Doddapaneni, H., Subramanian, V., Lavin, J.L., Oguiza, J.A., Perez, G., Pisabarro, A.G., Ramirez, L., Santoyo, F., Master, E., Coutinho, P.M., Henrissat, B., Lombard, V., Magnuson, J.K., Kues, U., Hori, C., Igarashi, K., Samejima, M., Held, B.W., Barry, K.W., LaButti, K.M., Lapidus, A., Lindquist, E.A., Lucas, S.M., Riley, R., Salamov, A.A., Hoffmeister, D., Schwenk, D., Hadar, Y., Yarden, O., de Vries, R.P., Wiebenga, A., Stenlid, J., Eastwood, D., Grigoriev, I.V., Berka, R.M., Blanchette, R.A., Kersten, P., Martinez, A.T., Vicuna, R., Cullen, D., 2012. Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. PNAS 109, 5458-5463.

Fernandez-Fueyo, E., Ruiz-Duenas, F.J., Martinez, M.J., Romero, A., Hammel, K.E., Medrano, F.J., Martinez, A.T., 2014. Ligninolytic peroxidase genes in the oyster mushroom genome: heterologous expression, molecular structure, catalytic and stability properties, and lignin-degrading ability. Biotechnol biofuels 7.

Flourat, A.L., Zeaiter, N., Vallée, E., Nguyen, V.P.T., Fadlallah, S., Allais, F., 2022. A sustainable preparative-scale chemo-enzymatic synthesis of 6-hydroxy-5,7-dimethoxy-2-naphthoic acid (DMNA) from sinapic acid. Green Chem Lett Rev 15, 557-571.

Gall, D.L., Ralph, J., Donohue, T.J., Noguera, D.R., 2017. Biochemical transformation of lignin for deriving valued commodities from lignocellulose. Curr Opin Biotechnol 45, 120-126.

Garcia-Conesa, M.T., Plumb, G.W., Kroon, P.A., Wallace, G., Williamson, G., 1997. Antioxidant properties of ferulic acid dimers. Redox Rep 3, 239-244.

Gelpke, M.D.S., Lee, J., Gold, M.H., 2002. Lignin Peroxidase Oxidation of Veratryl Alcohol: Effects of the Mutants H82A, Q222A, W171A, and F267L. Biochemistry 41, 8.

Genaro-Mattos, T.C., Maurício, Â.Q., Rettori, D., Alonso, A., Hermes-Lima, M., 2015. Antioxidant Activity of Caffeic Acid against Iron-Induced Free Radical Generation—A Chemical Approach. PloS One 10, 12.

Goldenzweig, A., Goldsmith, M., Hill, S.E., Gertman, O., Laurino, P., Ashani, Y., Dym, O., Unger, T., Albeck, S., Prilusky, J., Lieberman, R.L., Aharoni, A., Silman, I., Sussman, J.L., Tawfik, D.S., Fleishman, S.J., 2016. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. Mol Cell 63, 337-346.

Gorrec, F., 2009. The MORPHEUS protein crystallization screen. J Appl Crystallogr 42, 1035-1042.

Gray, H.B., Winkler, J.R., 2021. Functional and protective hole hopping in metalloenzymes. Chem Sci 12, 13988-14003.

Gruz, J., Pospisil, J., Kozubikova, H., Pospisil, T., Dolezal, K., Bunzel, M., Strnad, M., 2015. Determination of free diferulic, disinapic and dicoumaric acids in plants and foods. Food Chem 171, 280-286.

Habib, M., Trajkovic, M., Fraaije, M.W., 2018. The Biocatalytic Synthesis of Syringaresinol from 2,6-Dimethoxy-4-allylphenol in One-Pot Using a Tailored Oxidase/Peroxidase System. ACS Catal 8, 5549-5552.

Habib, M.H., Rozeboom, H.J., Fraaije, M.W., 2019. Characterization of a New DyP-Peroxidase from the Alkaliphilic Cellulomonad, *Cellulomonas bogoriensis*. Molecules 24.

Hamalainen, V., Gronroos, T., Suonpaa, A., Heikkila, M.W., Romein, B., Ihalainen, P., Malandra, S., Birikh, K.R., 2018. Enzymatic Processes to Unlock the Lignin Value. Front Bioeng Biotechnol 6.

Hammel, K.E., Cullen, D., 2008. Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11, 349-355.

Harris, T.K., Turner, G.J., 2002. Structural basis of perturbed pK_a values of catalytic groups in enzyme active sites. IUBMB Life 53, 85-98.

Heinemann, P.M., Armbruster, D., Hauer, B., 2021. Active-site loop variations adjust activity and selectivity of the cumene dioxygenase. Nat Commun 12.

Hermans, J., Berendsen, H.J.C., Vangunsteren, W.F., Postma, J.P.M., 1984. A Consistent Empirical Potential for Water-Protein Interactions. Biopolymers 23, 1513-1518.

Hess, B., 2008. P-LINCS: A parallel linear constraint solver for molecular simulation. J Chem Theory Comput 4, 116-122.

Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., 1998. LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18, 1463-1472.

Hofbauer, S., Pfanzagl, V., Michlits, H., Schmidt, D., Obinger, C., Furtmuller, P.G., 2021. Understanding molecular enzymology of porphyrin-binding α + β barrel proteins - One fold, multiple functions. Biochim Biophys Acta Proteins Proteom 1869, 17.

Hoque, M.A., Zhang, Y., Chen, L.Q., Yang, G.Y., Khatun, M.A., Chen, H.F., Hao, L., Feng, Y., 2017. Stepwise Loop Insertion Strategy for Active Site Remodeling to Generate Novel Enzyme Functions. ACS Chem Biol 12, 1188-1193.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C., 2006. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 65, 712-725.

Huang, G., Shrestha, R., Jia, K., Geisbrecht, B.V., Li, P., 2017. Enantioselective Synthesis of Dilignol Model Compounds and Their Stereodiscrimination Study with A Dye-Decolorizing Peroxidase. Org Lett 19.

Ismail, H.M., Scapozza, L., Ruegg, U.T., Dorchies, O.M., 2014. Diapocynin, a Dimer of the NADPH Oxidase Inhibitor Apocynin, Reduces ROS Production and Prevents Force Loss in Eccentrically Contracting Dystrophic Muscle. PLoS One 9.

Janvier, M., Hollande, L., Jaufurally, A.S., Ducrot, P.H., Allais, F., 2016. Syringaresinol: a new bio-based bisphenolic building-block for polymers synthesis. Am Chem Soc 252.

Janvier, M., Hollande, L., Jaufurally, A.S., Pernes, M., Menard, R., Grimaldi, M., Beaugrand, J., Balaguer, P., Ducrot, P.H., Allais, F., 2017. Syringaresinol: A Renewable and Safer Alternative to Bisphenol A for Epoxy-Amine Resins. Chemsuschem 10, 738-746.

Jaufurally, A.S., Hollande, L., Teixeira, A.R.S., Ducrot, P.H., Allais, F., 2016a. Chemo-enzymatic synthesis and functionalization of syringaresinol: a

promising biobased antiradical additive and platform for bisphenolic monomers. Abstr Pap Am Chem Sci, 252.

Jaufurally, A.S., Teixeira, A.R.S., Hollande, L., Allais, F., Ducrot, P.H., 2016b. Optimization of the Laccase-Catalyzed Synthesis of (+/-)-Syringaresinol and Study of its Thermal and Antiradical Activities. Chemistryselect 1, 5165-5171.

Johjima, T., Ohkuma, M., Kudo, T., 2003. Isolation and cDNA cloning of novel hydrogen peroxide-dependent phenol oxidase from the basidiomycete *Termitomyces albuminosus*. Appl Microbiol Biotechnol 61, 220-225.

Jones, P., 2001. Roles of Water in Heme Peroxidase and Catalase Mechanisms. J Biol Chem 276, 13791-13796.

Kabsch, W., 2010. Xds. Acta Crystallogr D Biol Crystallogr 66.

Kantardjieff, K.A., Rupp, B., 2003. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Prot Sci 12, 1865-1871.

Karplus, P.A., Diederichs, K., 2012. Linking Crystallographic Model and Data Quality. Science 336, 1030-1033.

Kekilli, D., Moreno-Chicano, T., Chaplin, A.K., Horrell, S., Dworkowski, F.S.N., Worrall, J.A.R., Strange, R.W., Hough, M.A., 2017. Photoreduction and validation of haem-ligand intermediate states in protein crystals by in situ single-crystal spectroscopy and diffraction. IUCrJ 4, 263-270.

Kellner, H., Luis, P., Pecyna, M.J., Barbi, F., Kapturska, D., Kruger, D., Zak, D.R., Marmeisse, R., Vandenbol, M., Hofrichter, M., 2014. Widespread occurrence of expressed fungal secretory peroxidases in forest soils. PLoS One 9.

Kim, S.J.S., M., 1999. Purification and Characterization of a Novel Peroxidase from *Geotrichum candidum* Dec 1 Involved in Decolorization of Dyes. Appl Environ Microbiol 65.

Kimani, V., Ullrich, R., Buttner, E., Herzog, R., Kellner, H., Jehmlich, N., Hofrichter, M., Liers, C., 2021. First Dye-Decolorizing Peroxidase from an Ascomycetous Fungus Secreted by *Xylaria grammica*. Biomolecules 11.

Kirk, T.K., 1987. Enzymatic "Combustion": The Microbial Degradation of Lignin. Annu Rev Microbiol 41, 40.

Kolwek, J., Behrens, C., Linke, D., Krings, U., Berger, R.G., 2018. Cell-free one-pot conversion of (+)-valencene to (+)-nootkatone by a unique dye-decolorizing peroxidase combined with a laccase from *Funalia trogii*. J Ind Microbiol Biotechnol 45, 89-101.

Krahe, N.K., Berger, R.G., Ersoy, F., 2020. A DyP-Type Peroxidase of *Pleurotus sapidus* with Alkene Cleaving Activity. Molecules 25.

Krahe, N.K., Berger, R.G., Witt, M., Zorn, H., Omarini, A.B., Ersoy, F., 2021. Monokaryotic *Pleurotus sapidus* Strains with Intraspecific Variability of an Alkene Cleaving DyP-Type Peroxidase Activity as a Result of Gene Mutation and Differential Gene Expression. Int J Mol Sci 22.

Kress, N., Halder, J.M., Rapp, L.R., Hauer, B., 2018. Unlocked potential of dynamic elements in protein structures: channels and loops. Curr Opin Chem Biol 47, 109-116.

Krieger, E., Vriend, G., 2015. New ways to boost molecular dynamics simulations. J Comput Chem 36, 996-1007.

Krings, U., Esparan, V., Berger, R.G., 2015. The taste enhancer divanillin: a review on sources and enzymatic generation. Flavour Fragr J 30, 362-365.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J Mol Biol 372, 774-797.

Kudanga, T., Nemadziva, B., Le Roes-Hill, M., 2017. Laccase catalysis for the synthesis of bioactive compounds. Appl Microbiol Biotechnol 101, 13-33.

Kumar, S., Tsai, C.J., Nussinov, R., 2000. Factors enhancing protein thermostability. Protein engineering 13, 179-191.

Lauber, C., Schwarz, T., Nguyen, Q.K., Lorenz, P., Lochnit, G., Zorn, H., 2017. Identification, heterologous expression and characterization of a dye-decolorizing peroxidase of *Pleurotus sapidus*. AMB Express 7, 164.

Leszczynski, J.F., Rose, G.D., 1986. Loops in Globular-Proteins - a Novel Category of Secondary Structure. Science 234, 849-855.

Li, J., Liu, C., Li, B., Yuan, H., Yang, J., Zheng, B., 2012. Identification and molecular characterization of a novel DyP-type peroxidase from *Pseudomonas aeruginosa* PKE117. Appl Biochem Biotechnol 166, 774-785.

Li, L., Wang, T., Chen, T., Huang, W., Zhang, Y., Jia, R., He, C., 2021. Revealing two important tryptophan residues with completely different roles in a dye-decolorizing peroxidase from *Irpex lacteus* F17. Biotechnol Biofuels 14.

Linde, D., Cañellas, M., Coscolín, C., Davó-Siguero, I., Romero, A., Lucas, F., Ruiz-Dueñas, F.J., Guallar, V., Martínez, A.T., 2016. Asymmetric sulfoxidation by engineering the heme pocket of a dye-decolorizing peroxidase. Catal Sci Technol 6, 6277-6285.

Linde, D., Pogni, R., Canellas, M., Lucas, F., Guallar, V., Baratto, M.C., Sinicropi, A., Saez-Jimenez, V., Coscolin, C., Romero, A., Medrano, F.J., Ruiz-Duenas, F.J., Martinez, A.T., 2015a. Catalytic surface radical in dye-decolorizing peroxidase: a computational, spectroscopic and site-directed mutagenesis study. Biochem J 466, 253-262.

Linde, D., Ruiz-Duenas, F.J., Fernandez-Fueyo, E., Guallar, V., Hammel, K.E., Pogni, R., Martinez, A.T., 2015b. Basidiomycete DyPs: Genomic diversity, structural-functional aspects, reaction mechanism and environmental significance. Arch Biochem Biophys 574, 66-74.

Linger, J.G., Vardon, D.R., Guarnieri, M.T., Karp, E.M., Hunsinger, G.B., Franden, M.A., Johnson, C.W., Chupka, G., Strathmann, T.J., Pienkos, P.T., Beckham, G.T., 2014. Lignin valorization through integrated biological funneling and chemical catalysis. PNAS 111, 12013-12018.

Liu, H., Liu, Z.H., Zhang, R.K., Yuan, J.S., Li, B.Z., Yuan, Y.J., 2022. Bacterial conversion routes for lignin valorization. Biotechnol Adv 60.

Liu, X., Du, Q., Wang, Z., Zhu, D., Huang, Y., Li, N., Wei, T., Xu, S., Gu, L., 2011. Crystal structure and biochemical features of EfeB/YcdB from *Escherichia coli* O157: ASP235 plays divergent roles in different enzyme-catalyzed processes. J Biol Chem 286, 14922-14931.

Liu, X., Yuan, Z., Wang, J., Cui, Y., Liu, S., Ma, Y., Gu, L., Xu, S., 2017. 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 484, 40-44.

Llevot, A., Grau, E., Carlotti, S., Grelier, S., Cramail, H., 2016. From Lignin-derived Aromatic Compounds to Novel Biobased Polymers. Macromol Rapid Commun 37, 9-28.

Lončar, N., Božić, N., Vujčić, Z., 2013. Cloning and characterization of a new dye degrading laccase from *Bacillus amyloliquefaciens* 12B1. FEBS J 280, 599-600.

Lončar, N., Colpa, D.I., Fraaije, M.W., 2016. Exploring the biocatalytic potential of a DyP-type peroxidase by profiling the substrate acceptance of *Thermobifida fusca* DyP peroxidase. Tetrahedron 72, 7276-7281.

Lončar, N., Drašković, N., Božić, N., Romero, E., Simić, S., Opsenica, I., Vujčić, Z., Fraaije, M.W., 2019. Expression and Characterization of a Dye-Decolorizing Peroxidase from *Pseudomonas fluorescens* Pf0-1. Catalysts 9.

Lončar, N., Rozeboom, H.J., Franken, L.E., Stuart, M.C.A., Fraaije, M.W., 2020. Structure of a robust bacterial protein cage and its application as a versatile biocatalytic platform through enzyme encapsulation. Biochem Biophys Res Commun 529, 548-553.

Lučić, M., Chaplin, A.K., Moreno-Chicano, T., Dworkowski, F.S.N., Wilson, M.T., Svistunenko, D.A., Hough, M.A., Worrall, J.A.R., 2020a. A subtle structural change in the distal haem pocket has a remarkable effect on tuning hydrogen peroxide reactivity in dye decolourising peroxidases from *Streptomyces lividans*. Dalton Trans 49, 1620-1636.

Lučić, M., Svistunenko, D.A., Wilson, M.T., Chaplin, A.K., Davy, B., Ebrahim, A., Axford, D., Tosha, T., Sugimoto, H., Owada, S., Dworkowski, F.S.N., Tews, I., Owen, R.L., Hough, M.A., Worrall, J.A.R., 2020b. Serial Femtosecond Zero Dose Crystallography Captures a Water-Free Distal Heme Site in a Dye-Decolorising Peroxidase to Reveal a Catalytic Role for an Arginine in Fe(IV)=O Formation. Angew Chem Int Ed Engl 59, 21656-21662.

Lučić, M., Wilson, M.T., Svistunenko, D.A., Owen, R.L., Hough, M.A., Worrall, J.A.R., 2021. 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 26, 743-761.

Lučić, M., Wilson, M.T., Tosha, T., Sugimoto, H., Shilova, A., Axford, D., Owen, R.L., Hough, M.A., Worrall, J.A.R., 2022. Serial Femtosecond Crystallography Reveals the Role of Water in the One- or Two-Electron Redox Chemistry of Compound I in the Catalytic Cycle of the B-Type Dye-Decolorizing Peroxidase DtpB. ACS Catal 12, 13349-13359.

Lundell, T.K., Makela, M.R., Hilden, K., 2010. Lignin-modifying enzymes in filamentous basidiomycetes-ecological, functional and phylogenetic review. J Basic Microbiol 50, 5-20.

Machuqueiro, M., Baptista, A.M., 2006. Constant-pH molecular dynamics with ionic strength effects: Protonation-conformation coupling in decalysine. J Phys Chem B 110, 2927-2933.

Makhatadze, G.I., Privalov, P.L., 1995. Energetics of protein structure. Adv Protein Chem 47, 307-425.

Malabanan, M.M., Amyes, T.L., Richard, J.P., 2010. A role for flexible loops in enzyme catalysis. Curr Opin Struct Biol 20, 702-710.

Mark, P., Nilsson, L., 2001. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. J Phys Chem B 105, 9954-9960.

Marques, S.M., Planas-Iglesias, J., Damborsky, J., 2021. Web-based tools for computational enzyme design. Curr Opin Struct Biol 69, 19-34.

Martinez, A.T., Ruiz-Duenas, F.J., Martinez, M.J., Del Rio, J.C., Gutierrez, A., 2009. Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20, 348-357.

Matthews, B.W., 1968. Solvent Content of Protein Crystals. J Mol Biol 33, 491-497.

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

Mendes, S., Brissos, V., Gabriel, A., Catarino, T., Turner, D.L., Todorovic, S., Martins, L.O., 2015a. An integrated view of redox and catalytic properties of B-type PpDyP from *Pseudomonas putida* MET94 and its distal variants. Arch Biochem Biophys 574, 99-107.

Mendes, S., Catarino, T., Silveira, C., Todorovic, S., Martins, L.O., 2015b. The catalytic mechanism of A-type dye-decolourising peroxidase BsDyP: neither aspartate nor arginine is individually essential for peroxidase activity. Catal Sci Technol 5, 5196-5207.

Mendes, S., Robalo, M.P., Martins, L.O., 2015c. Bacterial Enzymes and Multi-enzymatic Systems for Cleaning-up Dyes from the Environment, Microbial Degradation of Synthetic Dyes in Wastewaters. pp. 27-55.

Min, K., Gong, G., Woo, H.M., Kim, Y., Um, Y., 2015. A dye-decolorizing peroxidase from *Bacillus subtilis* exhibiting substrate-dependent optimum temperature for dyes and beta-ether lignin dimer. Sci Rep 5.

Miyamoto, S., Kollman, P.A., 1992. Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. J Comput Chem 13, 952-962.

Monthong, W., Pitchuanchom, S., Nuntasaen, N., Pompinom, W., 2011. (+)-Syringaresinol Lignan from New Species *Magnolia Thailandica*. Am J Appl Sci 8, 1268-1271.

Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67, 355-367.

Musil, M., Konegger, H., Hong, J., Bednar, D., Damborsky, J., 2019. Computational Design of Stable and Soluble Biocatalysts. ACS Catal 9, 1033-1054.

Musil, M., Stourac, J., Bendl, J., Brezovsky, J., Prokop, Z., Zendulka, J., Martinek, T., Bednar, D., Damborsky, J., 2017. FireProt: web server for automated design of thermostable proteins. Nucleic Acids Res 45, W393-W399.

Natte, K., Narani, A., Goyal, V., Sarki, N., Jagadeesh, R.V., 2020. Synthesis of Functional Chemicals from Lignin-derived Monomers by Selective Organic Transformations. Adv Synth Catal 362, 5143-5169.

Nemadziva, B., Le Roes-Hill, M., Koorbanally, N., Kudanga, T., 2018. Small laccase-catalyzed synthesis of a caffeic acid dimer with high antioxidant capacity. Process Biochem 69, 99-105.

Nishimura, R.T., Giammanco, C.H., Vosburg, D.A., 2010. Green, Enzymatic Syntheses of Divanillin and Diapocynin for the Organic, Biochemistry, or Advanced General Chemistry Laboratory. J Chem Educ 87, 526–527.

Nosé, S., Klein, M.L., 1983. Constant pressure molecular dynamics for molecular systems. Mol Phys 50, 1055-1076.

Nys, K., Furtmuller, P.G., Obinger, C., Van Doorslaer, S., Pfanzagl, V., 2021. On the Track of Long-Range Electron Transfer in B-Type Dye-Decolorizing Peroxidases: Identification of a Tyrosyl Radical by Computational Prediction and Electron Paramagnetic Resonance Spectroscopy. Biochemistry 60, 1226-1241.

Ogola, H.J., Hashimoto, N., Miyabe, S., Ashida, H., Ishikawa, T., Shibata, H., Sawa, Y., 2010. Enhancement of hydrogen peroxide stability of a novel *Anabaena sp.* DyP-type peroxidase by site-directed mutagenesis of methionine residues. Appl Microbiol Biotechnol 87, 1727-1736.

Ogola, H.J., Kamiike, T., Hashimoto, N., Ashida, H., Ishikawa, T., Shibata, H., Sawa, Y., 2009. Molecular characterization of a novel peroxidase from the cyanobacterium *Anabaena* sp. strain PCC 7120. Appl Environ Microbiol 75, 7509-7518.

Omura, I., Ishimori, K., Uchida, T., 2022. Converting cytochrome c into a DyP-like metalloenzyme. Dalton Trans 51, 12641-12649.

Pabis, A., Risso, V.A., Sanchez-Ruiz, J.M., Kamerlin, S.C.L., 2018. Cooperativity and flexibility in enzyme evolution. Curr Opin Struct Biol 48, 83-92.

Painter, J., Merritt, E.A., 2006. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D Biol Crystallogr 62, 439-450.

Parrinello, M., Rahman, A., 1981. Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. J Appl Phys 52, 7182-7190.

Perez-Boada, M., Ruiz-Duenas, F.J., Pogni, R., Basosi, R., Choinowski, T., Martinez, M.J., Piontek, K., Martinez, A.T., 2005. Versatile peroxidase oxidation of high redox potential aromatic compounds: site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways. J Mol Biol 354, 385-402.

Petrovic, D., Risso, V.A., Kamerlin, S.C.L., Sanchez-Ruiz, J.M., 2018. Conformational dynamics and enzyme evolution. J R Soc Interface 15.

Petrus, M.L., Vijgenboom, E., Chaplin, A.K., Worrall, J.A., van Wezel, G.P., Claessen, D., 2016. The DyP-type peroxidase DtpA is a Tat-substrate required for GlxA maturation and morphogenesis in *Streptomyces*. Open Biol 6.

Pfanzagl, V., Bellei, M., Hofbauer, S., Laurent, C., Furtmuller, P.G., Oostenbrink, C., Battistuzzi, G., Obinger, C., 2019. Redox thermodynamics of B-class dye-decolorizing peroxidases. J Inorg Biochem 199.

Pfanzagl, V., Nys, K., Bellei, M., Michlits, H., Mlynek, G., Battistuzzi, G., Djinovic-Carugo, K., Van Doorslaer, S., Furtmuller, P.G., Hofbauer, S., Obinger, C., 2018. Roles of distal aspartate and arginine of B-class dye-decolorizing peroxidase in heterolytic hydrogen peroxide cleavage. J Biol Chem 293, 14823-14838.

Planas-Iglesias, J., Marques, S.M., Pinto, G.P., Musil, M., Stourac, J., Damborsky, J., Bednar, D., 2021. Computational design of enzymes for biotechnological applications. Biotechnol Adv 47.

Planche, L.A., 1810. Note sur la sophistication de la résine de jalap et sur les moyens de la reconnaître. Bulletin de Pharmacie 2, 2.

Poulos, T.L., 2014. Heme Enzyme Structure and Function. Chem Rev 114, 3919-3962.

Putri, R.M., Allende-Ballestero, C., Luque, D., Klem, R., Rousou, K.A., Liu, A., Traulsen, C.H., Rurup, W.F., Koay, M.S.T., Caston, J.R., Cornelissen, J., 2017. Structural Characterization of Native and Modified Encapsulins as Nanoplatforms for *in vitro* Catalysis and Cellular Uptake. ACS Nano 11, 12796-12804.

Qin, X., Xin, Y., Su, X., Wang, X., Wang, Y., Zhang, J., Tu, T., Yao, B., Luo, H., Huang, H., 2021. Efficient Degradation of Zearalenone by Dye-

Decolorizing Peroxidase from *Streptomyces thermocarboxydus* Combining Catalytic Properties of Manganese Peroxidase and Laccase. Toxins 13.

Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., Davison, B.H., Dixon, R.A., Gilna, P., Keller, M., Langan, P., Naskar, A.K., Saddler, J.N., Tschaplinski, T.J., Tuskan, G.A., Wyman, C.E., 2014. Lignin valorization: improving lignin processing in the biorefinery. Science 344.

Ragnar, M., Lindgren, C.T., Nilvebrant, N., 2000. pK_a-Values of Guaiacyl and Syringyl Phenols Related to Lignin. J Wood Chem Technol 20, 277-305.

Rahman Pour, R., Ehibhatiomhan, A., Huang, Y., Ashley, B., Rashid, G.M., Mendel-Williams, S., Bugg, T.D.H., 2019. Protein engineering of *Pseudomonas fluorescens* peroxidase Dyp1B for oxidation of phenolic and polymeric lignin substrates. Enzyme Microb Technol 123, 21-29.

Rahmanpour, R., Bugg, T.D., 2013. Assembly in vitro of *Rhodococcus jostii* RHA1 encapsulin and peroxidase DypB to form a nanocompartment. FEBS J 280, 2097-2104.

Rahmanpour, R., Bugg, T.D., 2015. Characterisation of Dyp-type peroxidases from *Pseudomonas fluorescens* Pf-5: Oxidation of Mn(II) and polymeric lignin by Dyp1B. Arch Biochem Biophys 574, 93-98.

Rahmanpour, R., Rea, D., Jamshidi, S., Fulop, V., Bugg, T.D., 2016. Structure of *Thermobifida fusca* DyP-type peroxidase and activity towards Kraft lignin and lignin model compounds. Arch Biochem Biophys 594, 54-60.

Rai, A., Klare, J.P., Reinke, P.Y.A., Englmaier, F., Fohrer, J., Fedorov, R., Taft, M.H., Chizhov, I., Curth, U., Plettenburg, O., Manstein, D.J., 2021. Structural and Biochemical Characterization of a Dye-Decolorizing Peroxidase from *Dictyostelium discoideum*. Int J Mol Sci 22.

Rashid, G.M.M., Bugg, T.D.H., 2021. Enhanced biocatalytic degradation of lignin using combinations of lignin-degrading enzymes and accessory enzymes. Catal Sci Technol 11, 3568-3577.

Rawat, D., Mishra, V., Sharma, R.S., 2016. Detoxification of azo dyes in the context of environmental processes. Chemosphere 155, 591-605.

Renders, T., Van den Bosch, S., Koelewijn, S.F., Schutyser, W., Sels, B.F., 2017. Lignin-first biomass fractionation: the advent of active stabilisation strategies. Energy Environ Sci 10, 1551-1557.

Roberts, J.N., Singh, R., Grigg, J.C., Murphy, M.E., Bugg, T.D., Eltis, L.D., 2011. Characterization of dye-decolorizing peroxidases from *Rhodococcus jostii* RHA1. Biochemistry 50, 5108-5119.

Rocchia, W., Sridharan, S., Nicholls, A., Alexov, E., Chiabrera, A., Honig, B., 2002. 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 23, 128-137.

Rodrigues, C.F., Borges, P.T., Scocozza, M.F., Silva, D., Taborda, A., Brissos, V., Frazao, C., Martins, L.O., 2021. Loops around the Heme Pocket Have a Critical Role in the Function and Stability of BsDyP from *Bacillus subtilis*. Int J Mol Sci 22.

Rosado, T., Bernardo, P., Koci, K., Coelho, A.V., Robalo, M.P., Martins, L.O., 2012. Methyl syringate: an efficient phenolic mediator for bacterial and fungal laccases. Bioresour Technol 124, 371-378.

Ruiz-Duenas, F.J., Lundell, T., Floudas, D., Nagy, L.G., Barrasa, J.M., Hibbett, D.S., Martinez, A.T., 2013. Lignin-degrading peroxidases in

Polyporales: an evolutionary survey based on 10 sequenced genomes. Mycologia 105, 1428-1444.

Ruiz-Duenas, F.J., Morales, M., Garcia, E., Miki, Y., Martinez, M.J., Martinez, A.T., 2009. Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60, 441-452.

Ruiz-Duenas, F.J., Morales, M., Perez-Boada, M., Choinowski, T., Martinez, M.J., Piontek, K., Martinez, A.T., 2007. Manganese Oxidation Site in *Pleurotus eryngii* Versatile Peroxidase: A Site-Directed Mutagenesis, Kinetic, and Crystallographic Study. Biochemistry 46.

Runeberg, P.A., Brusentsev, Y., Rendon, S.M.K., Eklund, P.C., 2019. Oxidative Transformations of Lignans. Molecules 24.

Saez-Jimenez, V., Acebes, S., Garcia-Ruiz, E., Romero, A., Guallar, V., Alcalde, M., Medrano, F.J., Martinez, A.T., Ruiz-Duenas, F.J., 2016. Unveiling the basis of alkaline stability of an evolved versatile peroxidase. Biochem J 473, 1917-1928.

Sanchez-Ruiz, J.M., 2010. Protein kinetic stability. Biophys Chem 148, 1-15.

Santos, A., Mendes, S., Brissos, V., Martins, L.O., 2014. New dye-decolorizing peroxidases from *Bacillus subtilis* and *Pseudomonas putida* MET94: towards biotechnological applications. Appl Microbiol Biotechnol 98, 2053-2065.

Scheibner, M., Hulsdau, B., Zelena, K., Nimtz, M., de Boer, L., Berger, R.G., Zorn, H., 2008. Novel peroxidases of *Marasmius scorodonius* degrade beta-carotene. Appl Microbiol Biotechnol 77, 1241-1250.

Schirmann, J.G., Dekker, R.F.H., Borsato, D., Barbosa-Dekker, A.M., 2018. Selective control for the laccase-catalyzed synthesis of dimers from 2,6-dimethoxyphenol: Optimization of 3,3',5,5'-tetramethoxy-biphenyl-4,4'-diol synthesis using factorial design, and evaluation of its antioxidant action in biodiesel. Appl Catl A-Gen 555, 88-97.

Schmid, N., Eichenberger, A.P., Choutko, A., Riniker, S., Winger, M., Mark, A.E., van Gunsteren, W.F., 2011. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur Biophys J 40, 843-856.

Schrodinger, L.L.C., 2010. The PyMOL Molecular Graphics System.

Schutyser, W., Renders, T., Van den Bosch, S., Koelewijn, S.F., Beckham, G.T., Sels, B.F., 2018. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. Chem Soc Rev 47, 852-908.

Scocozza, M.F., Martins, L.O., Murgida, D.H., 2021. Direct Electrochemical Generation of Catalytically Competent Oxyferryl Species of Classes I and P Dye Decolorizing Peroxidases. Int J Mol Sci 22.

Sezer, M., Genebra, T., Mendes, S., Martins, L.O., Todorovic, S., 2012. A DyP-type peroxidase at a bio-compatible interface: structural and mechanistic insights. Soft Matter 8.

Sheldrick, G.M., 2000. XPREP Version 6.10, Bruker AXS Inc., Madison, Wisconsin, USA.

Shrestha, R., 2017. Molecular mechanism and enzymological studies of dye-decolorizing peroxidases (DyPs) from *Thermomonospora curvata* and *Enterobacter lignolyticus*, Department of Chemistry. Kansas State University.

Shrestha, R., Chen, X., Ramyar, K.X., Hayati, Z., Carlson, E.A., Bossmann, S.H., Song, L., Geisbrecht, B.V., Li, P., 2016. Identification of Surface-Exposed Protein Radicals and A Substrate Oxidation Site in A-Class

Dye-Decolorizing Peroxidase from *Thermomonospora curvata*. ACS Catal 6, 8036-8047.

Shrestha, R., Huang, G., Meekins, D.A., Geisbrecht, B.V., Li, P., 2017. Mechanistic Insights into Dye-Decolorizing Peroxidase Revealed by Solvent Isotope and Viscosity Effects. ACS Catal 7, 6352-6364.

Shrestha, R., Jia, K., Khadka, S., Eltis, L.D., Li, P., 2021. Mechanistic Insights into DyPB from *Rhodococcus jostii* RHA1 Via Kinetic Characterization. ACS Catal 11, 5486-5495.

Silva, T.F.D., Vila-Vicosa, D., Machuqueiro, M., 2021. Improved Protocol to Tackle the pH Effects on Membrane-Inserting Peptides. J Chem Theory Comput 17, 3830-3840.

Silveira, C.M., Moe, E., Fraaije, M., Martins, L.O., Todorovic, S., 2020. Resonance Raman view of the active site architecture in bacterial DyP-type peroxidases. RSC Adv 10, 11095-11104.

Singh, R., Eltis, L.D., 2015. The multihued palette of dye-decolorizing peroxidases. Arch Biochem Biophys 574, 56-65.

Singh, R., Grigg, J.C., Armstrong, Z., Murphy, M.E.P., Eltis, L.D., 2012. Distal heme pocket residues of B-type dye-decolorizing peroxidase: arginine but not aspartate is essential for peroxidase activity. J Biol Chem 287, 10623-10630.

Singh, R., Grigg, J.C., Qin, W., Kadla, J.F., Murphy, M.E., Eltis, L.D., 2013. Improved manganese-oxidizing activity of DypB, a peroxidase from a lignolytic bacterium. ACS Chem Biol 8, 700-706.

Stone, M.L., Anderson, E.M., Meek, K.M., Reed, M., Katahira, R., Chen, F., Dixon, R.A., Beckham, G.T., Roman-Leshkoy, Y., 2018. Reductive Catalytic Fractionation of C-Lignin. Acs Sustain Chem Eng 6, 11211-11218.

Strittmatter, E., Liers, C., Ullrich, R., Wachter, S., Hofrichter, M., Plattner, D.A., Piontek, K., 2013a. First crystal structure of a fungal high-redox potential dye-decolorizing peroxidase: substrate interaction sites and long-range electron transfer. J Biol Chem 288, 4095-4102.

Strittmatter, E., Serrer, K., Liers, C., Ullrich, R., Hofrichter, M., Piontek, K., Schleicher, E., Plattner, D.A., 2015. The toolbox of *Auricularia auricula-judae* dye-decolorizing peroxidase - Identification of three new potential substrate-interaction sites. Arch Biochem Biophys 574, 75-85.

Strittmatter, E., Wachter, S., Liers, C., Ullrich, R., Hofrichter, M., Plattner, D.A., Piontek, K., 2013b. Radical formation on a conserved tyrosine residue is crucial for DyP activity. Arch Biochem Biophys 537, 161-167.

Sugano, Y., 2009. DyP-type peroxidases comprise a novel heme peroxidase family. Cell Mol Life Sci 66, 1387-1403.

Sugano, Y., Matsushima, Y., Tsuchiya, K., Aoki, H., Hirai, M., Shoda, M., 2009. Degradation pathway of an anthraquinone dye catalyzed by a unique peroxidase DyP from *Thanatephorus cucumeris* Dec 1. Biodegradation 20, 433-440.

Sugano, Y., Muramatsu, R., Ichiyanagi, A., Sato, T., Shoda, M., 2007. DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family: ASP171 replaces the distal histidine of classical peroxidases. J Biol Chem 282, 36652-36658.

Sugano, Y., Yoshida, T., 2021. DyP-Type Peroxidases: Recent Advances and Perspectives. Int J Mol Sci 22.

Sugawara, K., Igeta, E., Amano, Y., Hyuga, M., Sugano, Y., 2019. Degradation of antifungal anthraquinone compounds is a probable physiological role of DyP secreted by *Bjerkandera adusta*. AMB Express 9.

Sugawara, K., Nishihashi, Y., Narioka, T., Yoshida, T., Morita, M., Sugano, Y., 2017. Characterization of a novel DyP-type peroxidase from *Streptomyces avermitilis*. J Biosci Bioeng 123, 425-430.

Sugawara, K., Yoshida, T., Hirashima, R., Toriumi, R., Akiyama, H., Kakuta, Y., Ishige, Y., Sugano, Y., 2021. Characterization of Class V DyP-Type Peroxidase SaDyP1 from *Streptomyces avermitilis* and Evaluation of SaDyPs Expression in Mycelium. Int J Mol Sci 22.

Sun, Z., Fridrich, B., de Santi, A., Elangovan, S., Barta, K., 2018. Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. Chem Rev 118, 614-678.

Sun, Z.T., Liu, Q., Qu, G., Feng, Y., Reetz, M.T., 2019. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. Chem Rev 119, 1626-1665.

Sutter, M., Boehringer, D., Gutmann, S., Gunther, S., Prangishvili, D., Loessner, M.J., Stetter, K.O., Weber-Ban, E., Ban, N., 2008. Structural basis of enzyme encapsulation into a bacterial nanocompartment. Nat Struct Mol Biol 15, 939-947.

Tang, Y., Mu, A., Zhang, Y., Zhou, S., Wang, W., Lai, Y., Zhou, X., Liu, F., Yang, X., Gong, H., Wang, Q., Rao, Z., 2021. Cryo-EM structure of *Mycobacterium smegmatis* DyP-loaded encapsulin. PNAS 118.

Teixeira, V.H., Cunha, C.A., Machuqueiro, M., Oliveira, A.S.F., Victor, B.L., Soares, C.M., Baptista, A.M., 2005. 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 109, 14691-14706.

Tironi, I.G., Sperb, R., Smith, P.E., Vangunsteren, W.F., 1995. A Generalized Reaction Field Method for Molecular-Dynamics Simulations. J Chem Phys 102, 5451-5459.

Tokuriki, N., Tawfik, D.S., 2009. Protein Dynamism and Evolvability. Science 324, 203-207.

Trott, O., Olson, A.J., 2010. 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 31, 455-461.

Tullman, J., Christensen, M., Kelman, Z., Marino, J.P., 2020. A ClpS-based N-terminal amino acid binding reagent with improved thermostability and selectivity. Biochem Eng J 154.

Uchida, T., Omura, I., Umetsu, S., Ishimori, K., 2021. Radical transfer but not heme distal residues is essential for pH dependence of dye-decolorizing activity of peroxidase from *Vibrio cholerae*. J Inorg Biochem 219.

Uchida, T., Sasaki, M., Tanaka, Y., Ishimori, K., 2015. A Dye-Decolorizing Peroxidase from *Vibrio cholerae*. Biochemistry 54, 6610-6621.

Vagin, A., Lebedev, A., 2015. MoRDa, an automatic molecular replacement pipeline. Acta Crystallogr A 71.

Van den Bosch, S., Koelewijn, S.F., Renders, T., Van den Bossche, G., Vangeel, T., Schutyser, W., Sels, B.F., 2018. Catalytic Strategies Towards Lignin-Derived Chemicals. Top Curr Chem 376, 36.

Veitch, N.C., 2004. Horseradish peroxidase: a modern view of a classic enzyme. Phytochemistry 65, 249-259.

Vieille, C., Zeikus, G.J., 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65, 1-43.

Vismeh, R., Lu, F., Chundawat, S.P., Humpula, J.F., Azarpira, A., Balan, V., Dale, B.E., Ralph, J., Jones, A.D., 2013. Profiling of diferulates (plant cell wall cross-linkers) using ultrahigh-performance liquid chromatography-tandem mass spectrometry. Analyst 138, 6683-6692.

von Vacano, B., Mangold, H., Vandermeulen, G.W.M., Battagliarin, G., Hofmann, M., Bean, J., Kunkel, A., 2022. Sustainable Design of Structural and Functional Polymers for a Circular Economy. Angew Chem Int Ed Engl.

Vuong, T.V., Singh, R., Eltis, L.D., Master, E.R., 2021. The Comparative Abilities of a Small Laccase and a Dye-Decoloring Peroxidase From the Same Bacterium to Transform Natural and Technical Lignins. Front Microbiol 12.

Wan, Y.Y., Lu, R., Kazuhiro, A., Tetsuo, M., Du, Y.M., 2007. Enzymatic synthesis of bioactive compounds by *Rhus* laccase from Chinese *Rhus vemicifera*. Sci China Chem 50, 179-182.

Wayne, N., Bolon, D.N., 2010. Charge-rich regions modulate the anti-aggregation activity of Hsp90. J Mol Biol 401, 931-939.

Webb, B., Sali, A., 2014. Protein structure modeling with MODELLER. Methods Mol Biol 1137, 1-15.

Weiss, M., 2001. Global indicators of X-ray data quality. J Appl Crystallogr 34, 130-135.

Wijma, H.J., Furst, M., Janssen, D.B., 2018. A Computational Library Design Protocol for Rapid Improvement of Protein Stability: FRESCO. Methods Mol Biol 1685, 69-85.

Wijma, H.J., Janssen, D.B., 2013. Computational design gains momentum in enzyme catalysis engineering. FEBS J 280, 2948-2960.

Wong, D.W., 2009. Structure and action mechanism of ligninolytic enzymes. Appl Biochem Biotechnol 157, 174-209.

Xu, Z., Cen, Y.K., Zou, S.P., Xue, Y.P., Zheng, Y.G., 2020. Recent advances in the improvement of enzyme thermostability by structure modification. Crit Rev Biotechnol 40, 83-98.

Xu, Z., Xue, Y.-P., Zou, S.-P., Zheng, Y.-G., 2020. Enzyme engineering strategies to confer thermostability, Biofuel Bioprod Biorefin pp. 67-89.

Yang, G.R., Miton, C.M., Tokuriki, N., 2020. A mechanistic view of enzyme evolution. Protein Sci 29, 1724-1747.

Yoshida, T., Ogola, H.J., Amano, Y., Hisabori, T., Ashida, H., Sawa, Y., Tsuge, H., Sugano, Y., 2016. *Anabaena sp.* DyP-type peroxidase is a tetramer consisting of two asymmetric dimers. Proteins 84, 31-42.

Yoshida, T., Sugano, Y., 2015. A structural and functional perspective of DyP-type peroxidase family. Arch Biochem Biophys 574, 49-55.

Yoshida, T., Tsuge, H., Konno, H., Hisabori, T., Sugano, Y., 2011. The catalytic mechanism of dye-decolorizing peroxidase DyP may require the swinging movement of an aspartic acid residue. FEBS J 278, 2387-2394.

Yu, W., Liu, W., Huang, H., Zheng, F., Wang, X., Wu, Y., Li, K., Xie, X., Jin, Y., 2014. Application of a novel alkali-tolerant thermostable DyP-type peroxidase from *Saccharomonospora viridis* DSM 43017 in biobleaching of eucalyptus kraft pulp. PLoS One 9.

Zalesak, F., Bon, D.J.D., Pospisil, J., 2019. Lignans and Neolignans: Plant secondary metabolites as a reservoir of biologically active substances. Pharmacol Res 146.

Zamocky, M., Hofbauer, S., Schaffner, I., Gasselhuber, B., Nicolussi, A., Soudi, M., Pirker, K.F., Furtmuller, P.G., Obinger, C., 2015. Independent evolution of four heme peroxidase superfamilies. Arch Biochem Biophys 574, 108-119.

Zamocky, M., Jakopitsch, C., Furtmuller, P.G., Dunand, C., Obinger, C., 2008. The peroxidase-cyclooxygenase superfamily: Reconstructed evolution of critical enzymes of the innate immune system. Proteins 72, 589-605.

Zhang, P., Xu, J., Wang, X.-J., He, B., Gao, S.-Q., Lin, Y.-W., 2019. The Third Generation of Artificial Dye-Decolorizing Peroxidase Rationally Designed in Myoglobin. ACS Catal 9, 7888-7893.

Zitare, U.A., Habib, M.H., Rozeboom, H., Mascotti, M.L., Todorovic, S., Fraaije, M.W., 2021. Mutational and structural analysis of an ancestral fungal dye-decolorizing peroxidase. FEBS J 288, 3602-3618.

Zubieta, C., Joseph, R., Krishna, S.S., McMullan, D., Kapoor, M., Axelrod, H.L., Miller, M.D., Abdubek, P., Acosta, C., Astakhova, T., Carlton, D., Chiu, H.J., Clayton, T., Deller, M.C., Duan, L., Elias, Y., Elsliger, M.A., Feuerhelm, J., Grzechnik, S.K., Hale, J., Han, G.W., Jaroszewski, L., Jin, K.K., Klock, H.E., Knuth, M.W., Kozbial, P., Kumar, A., Marciano, D., Morse, A.T., Murphy, K.D., Nigoghossian, E., Okach, L., Oommachen, S., Reyes, R., Rife, C.L., Schimmel, P., Trout, C.V., van den Bedem, H., Weekes, D., White, A., Xu, Q., Hodgson, K.O., Wooley, J., Deacon, A.M., Godzik, A., Lesley, S.A., Wilson, I.A., 2007a. Identification and structural characterization of heme binding in a novel dye-decolorizing peroxidase, TyrA. Proteins 69, 234-243.

Zubieta, C., Krishna, S.S., Kapoor, M., Kozbial, P., McMullan, D., Axelrod, H.L., Miller, M.D., Abdubek, P., Ambing, E., Astakhova, T., Carlton, D., Chiu, H.J., Clayton, T., Deller, M.C., Duan, L., Elsliger, M.A., Feuerhelm, J., Grzechnik, S.K., Hale, J., Hampton, E., Han, G.W., Jaroszewski, L., Jin, K.K., Klock, H.E., Knuth, M.W., Kumar, A., Marciano, D., Morse, A.T., Nigoghossian, E., Okach, L., Oommachen, S., Reyes, R., Rife, C.L., Schimmel, P., van den Bedem, H., Weekes, D., White, A., Xu, Q., Hodgson, K.O., Wooley, J., Deacon, A.M., Godzik, A., Lesley, S.A., Wilson, I.A., 2007b. Crystal structures of two novel dye-decolorizing peroxidases reveal a beta-barrel fold with a conserved heme-binding motif. Proteins 69, 223-233.

Zuccarello, L., Barbosa, C., Galdino, E., Lončar, N., Silveira, C.M., Fraaije, M.W., Todorovic, S., 2021. SERR Spectroelectrochemistry as a Guide for Rational Design of DyP-Based Bioelectronics Devices. Int J Mol Sci 22.

Zwart, P.H., Grosse-Kunstleve, R.W., Adams, P.D. , 2005. XTRIAGE: Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. . CCP4 newsletter 43.

Chapter 6

Annexes

Annex I

Table SI. 2.1. Molar extinction coefficients of the tested lignin-related phenolic compounds at their maximal wavelength at pH 4 and pH 8.

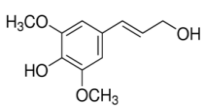
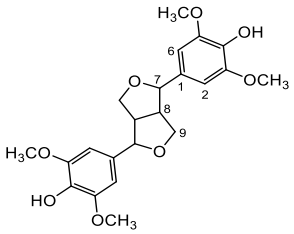
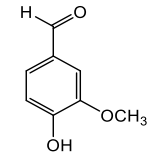
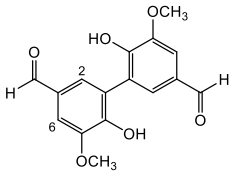
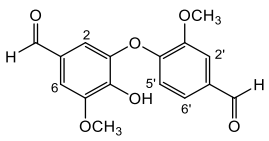
Lignin-related phenolics	λ (nm)	ε (M^{-1} cm^{-1}) pH 4	λ (nm)	ε (M^{-1} cm^{-1}) pH 8
Syringyl type phenolics				
Syringol	468*	49600	468*	49600
Sinapyl alcohol	275	4692	275	2503
Acetosyringone	300	9861	360	11162
Methyl syringate	280	10509	280	8056
Syringaldehyde	300	13020	360	21812
Sinapic acid	310	16097	310	14820
Syringic acid	270	8510	260	7473
Guaiacyl type phenolics				
Coniferyl alcohol	260	11022	260	9976
Vanillin	310	8833	340	16881
Coniferyl aldehyde	340	22936	340	15773
Ferulic acid	300	14163	300	13495
Methyl vanillate	300	4593	300	8898
Vanillyl alcohol	280	2526	280	2570
Acetovanillone	275	10136	340	12022
Guaiacol	470*	26600	470*	26600
Vanillic acid	260	9570	250	8021
Hydroxybenzene type phenolics				
Caffeic acid	320	5954	320	3646
p-Coumaric acid	300	16888	300	13642
o-Coumaric acid	275	14542	260	11566
m-Coumaric acid	275	16083	275	13766
3,4-hydroxybenzoic acid	260	7740	250	7080
4-hydroxybenzoic acid	250	13511	250	11569
Gallic acid	275	6986	260	7040
4-hydroxyacetophenone	270	20769	320	9406

* The wavelengths used in reactions with syringol and guaiacol allow to follow the products of the reaction.

Table SI. 2.2. Conversion yields of lignin-related phenolics after 24 h reaction as assessed by HPLC.

Substrates	Conversion (%)	
	WT (pH 4)	**6E10 (pH 8)**
Syringol	40 ± 2	100 ± 0
Sinapyl alcohol	93 ± 0	98 ± 0
Acetosyringone	29 ± 4	68 ± 1
Methyl syringate	13 ± 4	35 ± 1
Syringaldehyde	27 ± 6	74 ± 2
Sinapic acid	98 ± 0	99 ± 0
Syringic acid	28± 1	67 ± 1
Coniferyl alcohol	99 ± 0	99 ± 1
Vanillin	87 ± 6	98 ± 2
Coniferyl aldehyde	88 ± 1	100 ± 0
Ferulic acid	76 ± 1	99 ± 0
Methyl vanillate	42 ± 2	98 ± 4
Vanillyl alcohol	67 ± 7	98 ± 0
Acetovanillone	66 ± 5	100 ± 0
Guaiacol	49 ± 3	62 ± 1
Vanillic acid	44 ± 16	65 ± 4
Caffeic acid	89 ± 0	94 ± 1
p-Coumaric acid	43 ± 4	91 ± 2
o-Coumaric acid	25 ± 6	48 ± 2
m-Coumaric acid	23 ± 0	41 ± 10
3,4-hydroxybenzoic aci	64 ± 2	62 ± 2
4-hydroxybenzoic ac	18 ± 2	29 ± 10
Gallic acid	11 ± 2	79 ± 6
4hydroxyacetophenone	9 ± 0	23 ± 7

Table SI. 2.3. ^{1}H NMR data for substrates sinapyl alcohol and vanillin and reaction products syringaresinol, divanillin and 4-O-5 divanillin.

Structure	**^{1}H NMR data (MeOD)[a)]**
 **Sinapyl alcohol**	3.84 (s, 6H, OCH_3); 4.20 (d, 2H, J=6.0 Hz, H_9, **CH_2**OH); 6.21 (dt, 1H, J=16.0; 5.6 Hz, H_8, =**CH**CH_2OH); 6.50 (d, 1H, J=16.0 Hz, H_7, Ph**CH**=); 6.69 (s, 2H, H_2,H_6)
 **Syringaresinol**	3.14 (m, 2H, H_8); 3.85 (s, 12H, OCH_3); 3.87 (m, 2H, H_{9a}); 4.27 (m, 2H, J=6.8 Hz, H_{9b}); 4.71 (d, 2H, H_7); 6.66 (s, 4H, H_2,H_6)
 **Vanillin**	3.91 (s, 3H, OMe); 6.94 (d, 1H, J=8.8 Hz, H_5); 7.42 (d, 1H, J=6.8 Hz, H_6); 7.43 (s, 1H, H_2); 9.74 (s,1H, COH).
 **Divanillin**	3.95 (s, 6H, OMe); 7.42 (d, 2H, J=1.84 Hz, H_2); 7.55 (d, 2H, J=1.84 Hz, H_6); 9.86 (s, 2H, CHO).
 **4-O-5 divanillin**	3.93 (s, 3H, OMe'); 3.94 (s, 3H, OMe); 6.78 (d, 1H, J=8.0 Hz, $H_{5'}$); 7.19 (d, 1H, H_6); 7.40* (d, 2H, H_2); 7.48 (dd, 1H, J=2.0, 8.0 Hz, $H_{6'}$); 7.54 (d, 1H, J=2.0 Hz, $H_{2'}$); 9.74 (s, 1H, CHO); 9.78 (s, 1H, CHO'). * signal under vanillin signals

[a)] The following abbreviations are used in reporting the NMR data: s (singlet), d (doublet), t (triplet), and m (multiplet)). Signals are reported as follows: chemical shift (δ, ppm), multiplicity, integration, and coupling constants (Hz).

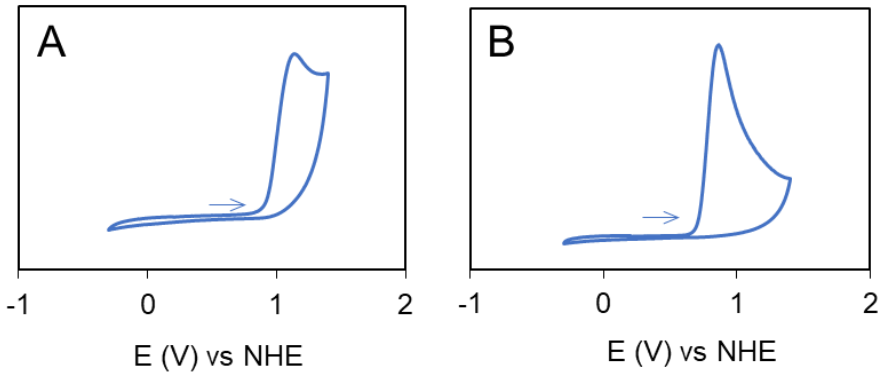


Figure SI. 2.1. Cyclic voltammograms of vanillin at pH 4 (A) and pH 8 (B) showing the irreversibility of the oxidation process (potential window: -0.3 to 1.4 V; v = 100 $mV.s^{-1}$). The arrow indicates the initial scan direction. (E_{pa} – anodic peak potential)

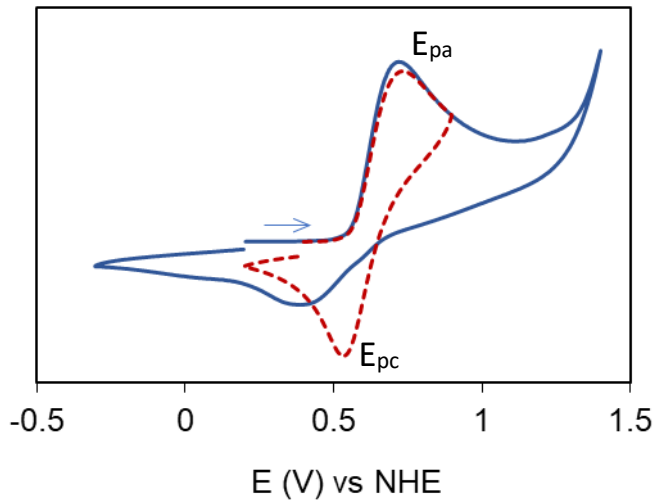


Figure SI. 2.2. Cyclic voltammograms of methyl syringate at pH 8. The dashed line represents both oxidation and reduction waves in the isolated process (potentials window: -0.3 to 1.4 V; v = 100 mVs^{-1}). (E_{pa} – anodic peak potential; E_{pc}- cathodic peak potential).

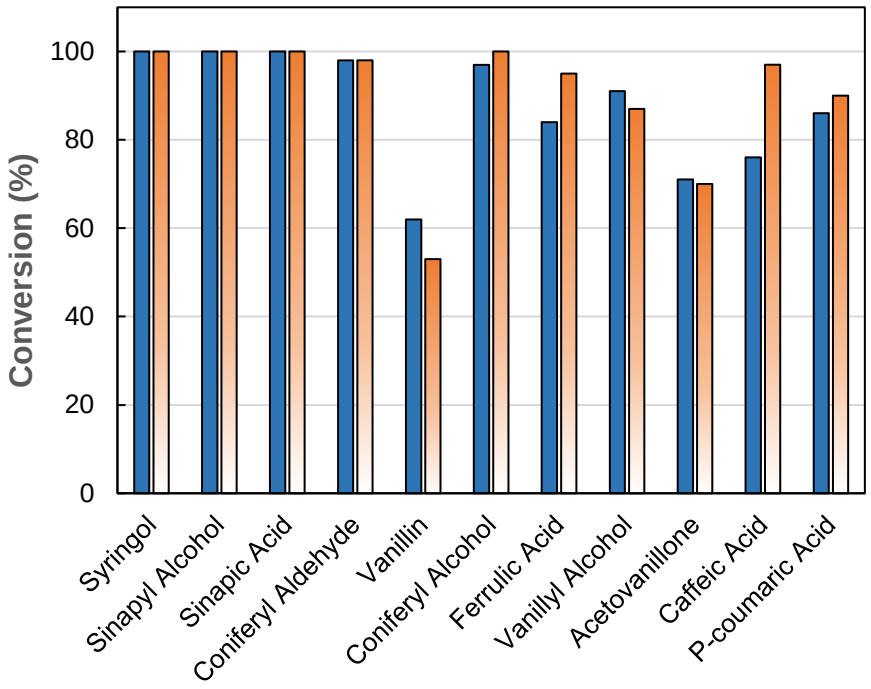


Figure SI. 2.3. Substrate conversion yields measured by HPLC after 1 h (blue) and 24 h (red) of reaction with 6E10 variant, using 0.5 mM of phenolics at pH 8. Reactions started by adding 0.5 mM H_2O_2.

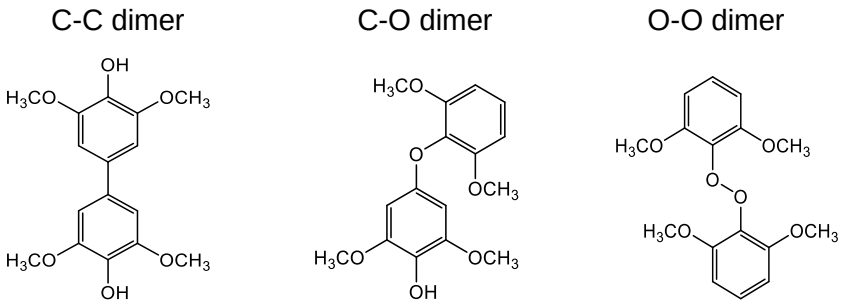


Figure SI. 2.4. Possible dimeric structures (m/z = 307) from 6E10 oxidation of syringol ($C_{16}H_{18}O_6$; MW: 306.314)

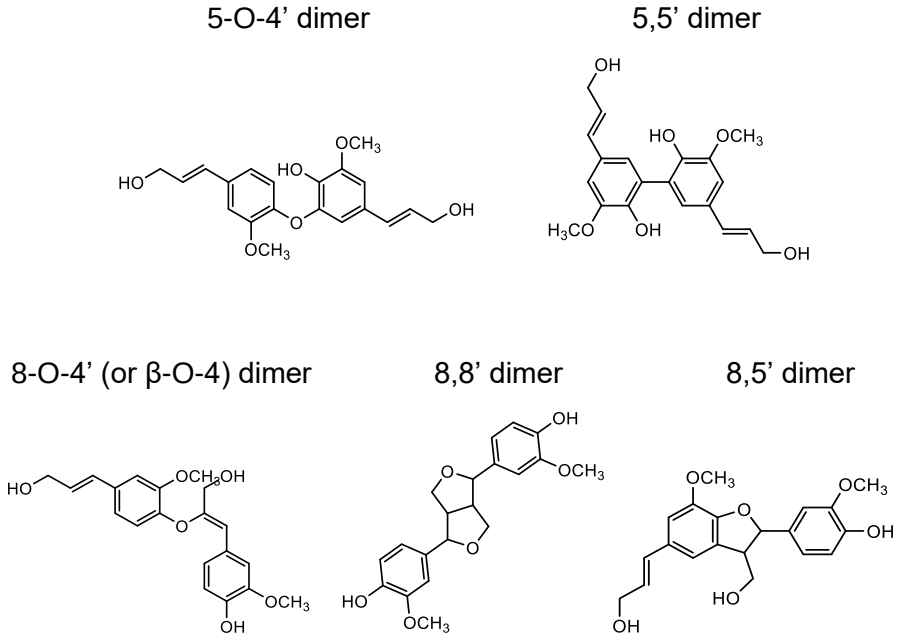


Figure SI. 2.5. Possible dimeric lignan and neolignan products (m/z = 359) from 6E10 oxidation of coniferyl alcohol ($C_{20}H_{22}O_6$; MW: 358.39)

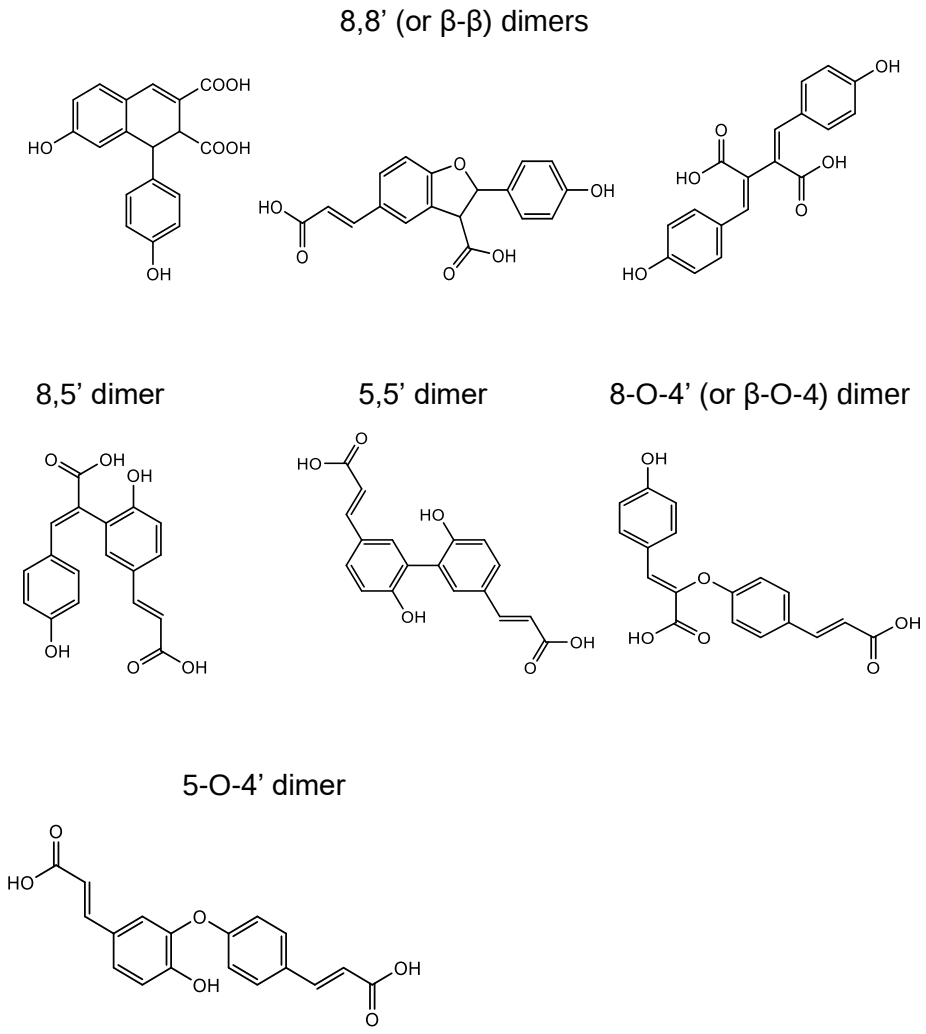


Figure SI. 2.6. Possible dimeric lignan and neolignan products (m/z = 327) from 6E10 oxidation of coumaric acid ($C_{18}H_{14}O_6$; MW: 326.304)

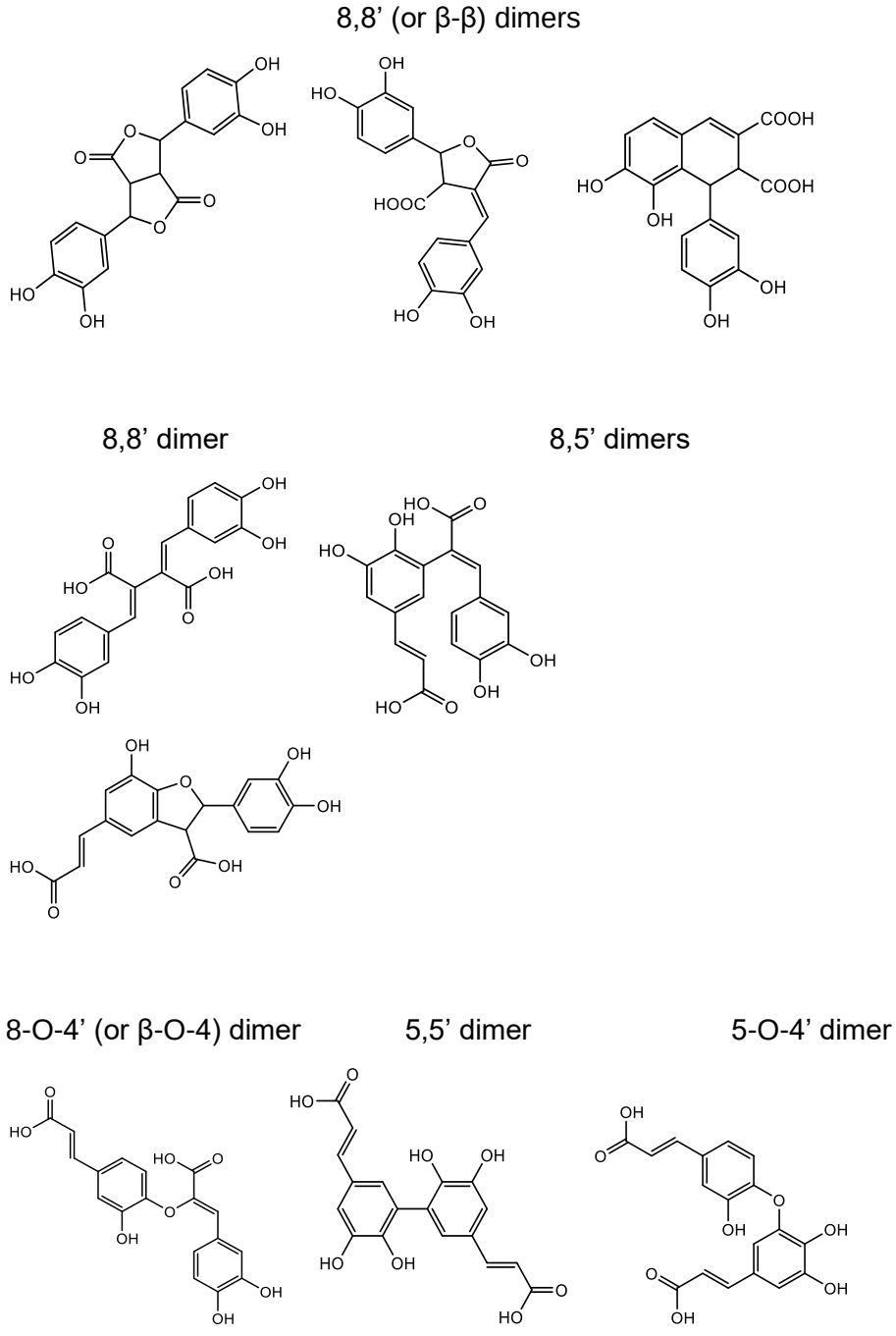


Figure SI. 2.7. Possible dimeric lignan and neolignan products (m/z = 359) from 6E10 oxidation of caffeic acid ($C_{18}H_{14}O_8$; MW: 358.302)

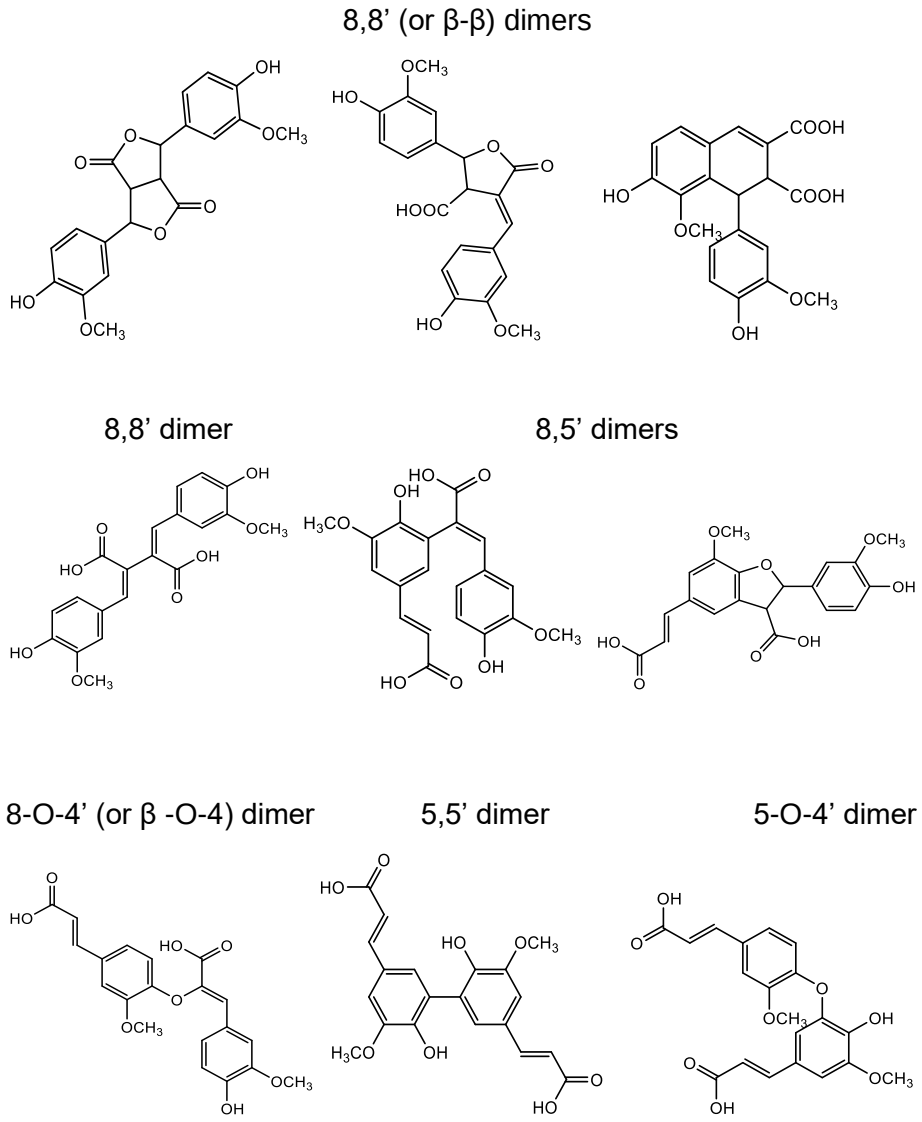


Figure SI. 2.8. Possible dimeric lignan and neolignan products (m/z = 387) from 6E10 oxidation of ferulic acid ($C_{20}H_{18}O_8$; MW: 386.100)

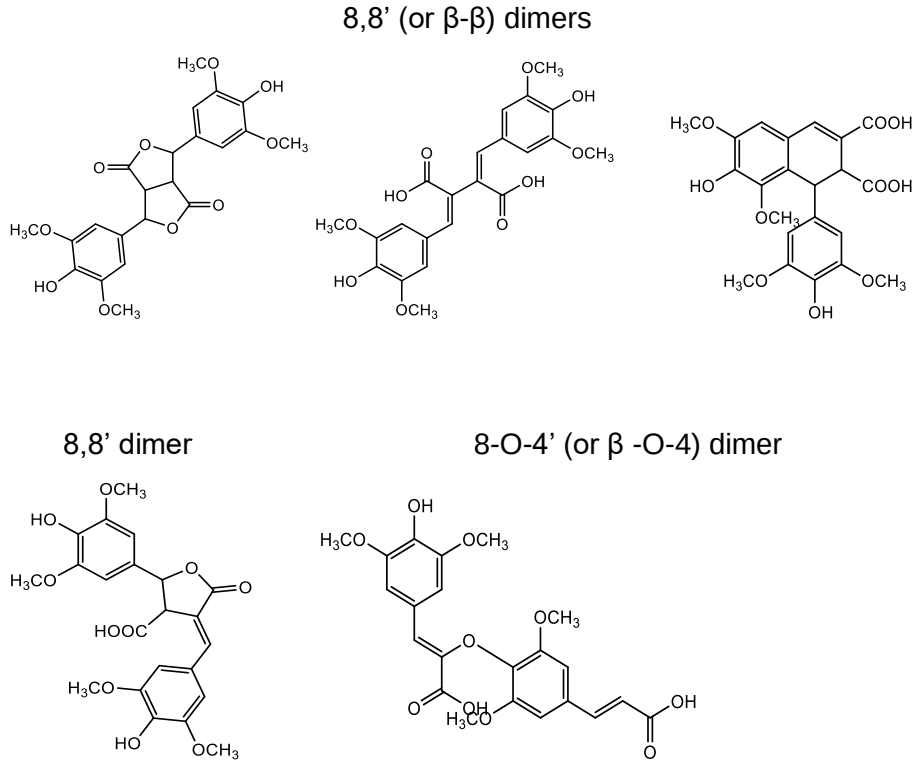


Figure SI. 2.9. Possible dimeric lignan and neolignan products (m/z = 447) from 6E10 oxidation of sinapic acid ($C_{22}H_{22}O_{10}$; MW: 446.408)

Annex II

Table SI. 3.1. 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	592150 (68900)	73015 (12699)	110187 (14483)
Unique reflections	81145 (12914)	14987 (2546)	54918 (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)
$\langle I/\sigma(I)\rangle$	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 (Karplus and Diederichs, 2012).
[b] $R_{sym} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i (hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations from symmetry-related reflections (Arndt et al., 1968).
[c] $R_{meas} = \Sigma_{hkl} [N/(N(hkl) -1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i (hkl)$, where N(hkl) is the data multiplicity, $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations from symmetry-related reflections. It is an indicator of the agreement between symmetry related observations (Diederichs and Karplus, 1997).
[d] $R_{p.i.m.} = \Sigma_{hkl} [1/(N(hkl) -1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i (hkl)$, where N(hkl) is the data multiplicity, $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ 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 (Weiss, 2001).

[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.

Table SI. 3.2. Properties of the tunnels and cavities surrounding the heme molecular access in *Pp*DyP wild type, 29E4 and 6E10.

	Tunnel 1				Tunnel 2				Cavity			
	Residues	Length (Å)	Radius (Å)	Polarity	Residues	Length (Å)	Radius (Å)	Polarity	Residues	Area ($Å^2$)	Volume ($Å^3$)	Depth (Å)	Polarity
Wild type	L120, G124, D126, Y130, E131, **D132**, G133, T134, E135, N136, I179, R181, N186, H197, V198, T201, V213, **R214**, R215, S216, V217, L227, F229, V230, L257	14.03	1.34	0.52	H125, G129, Y130, E131, **D132**, G133, T134, E135, I179, G180, R181, R182, L183, S184, D185, N186, E187, H197, V198, T201, **R214**, L257	11.54	1.33	0.64	E135, R181, E188, V198, A202, Q203, E204, F211	275.1	106.5	7.5	0.57
29E4	L120, G124, Y125, **D132**, G133, N136, H197, T201, V213, **R214**, R215, S216, V217, L227, F229, V230, L257	9.24	1.40	0.41	Y125, D126, G129, Y130, E131, **D132**, G133, T134, E135, I179, G180, R181, R182, L183, S184, D185, N186, E187, H197, V198, T201, **R214**, L227, F229, L257	15.88	1.27	0.56	E135, K188, V198, A202, Q203, E204	299.0	118.7	8.4	0.67
6E10	D126, **D132**, G133, N136, H197, T201, M212, **R214**, R215, S216, V217, L227, F229, V230, L257	9.64	1.33	0.53	Y125, G129, Y130, E131, **D132**, G133, T134, E135, I179, G180, R181, R182, L183, S184, D185, N186, E187, H197, V198, T201, **R214**, L227, F229, L257	13.89	1.32	0.54	E135, R181, K188, V198, A202, Q203, E204, F211	301.3	118.5	9.1	0.63

Table SI. 3.3. Accessible surface area (ASA, %) of residues are part of loops 1 (L1), 2 (L2), and 3 (L3) of *Pp*DyP wild type, 29E4 and 6E10.

L1	ASA (%)			L2	ASA (%)			L3	ASA (%)		
	WT	29E4	6E10		WT	29E4	6E10		WT	29E4	6E10
L120	17	15	23	I179	0	1	0	F206	7	10	4
H121	3	2	26	G180	2	2	0	E207	92	90	90
R122	52	50	73	R181	5	4	7	P208	79	76	87
G123	69	67	68	R182	31	43	42	E209	59	57	37
G124	12	19	17	L183	30	31	31	A210	2	5	53
H/Y125	15	31	39	S184	76	72	71	F211	27	30	34
D126	0	0	0	D185	53	54	47	M212	1	1	1
L127	3	1	23	N186	22	17	23	V213	0	1	0
T128	4	1	1	E187	48	59	63	R214	4	4	7
G129	34	39	35	E188	38	33	26	R215	16	18	15
Y130	1	2	2	L189	13	11	23	S216	11	15	12
E131	36	32	31	D190	76	76	69	V217	6	6	5
D132	14	20	16	D191	83	85	96	S218	1	1	1
G133	35	35	33	A192	12	11	10	W219	0	0	0
T134	33	37	32	P193	41	41	43	A220	19	18	49
E135	43	38	40	A194	41	41	38	D221	35	35	31
N136	14	16	22	S195	30	30	29	Q222	61	67	87
P137	13	19	9	A196	0	0	0	R223	29	29	30
T138	65	62	54	H197	0	0	0	G224	33	30	35
D139	68	70	77	V198	1	5	3	A225	9	8	19
				K199	31	34	27	G226	0	0	0
				R200	6	6	3				
				T201	0	0	0				
				A202	25	13	19				
				Q203	2	1	2				
				E204	72	58	75				
				S205	46	41	42				

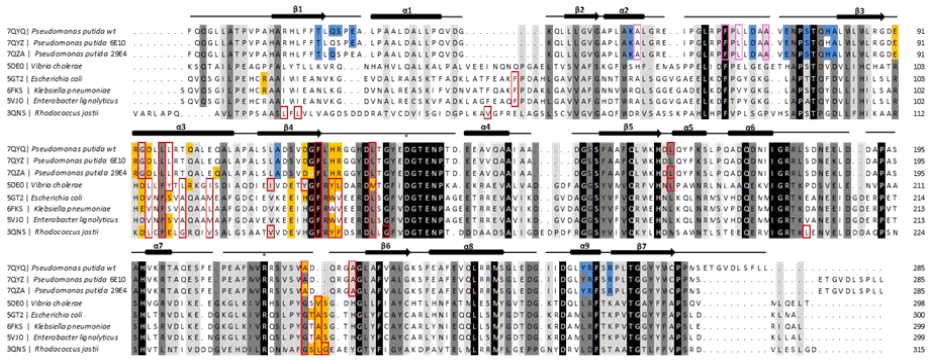


Figure SI. 3.1. Amino acid sequence alignment based on the known X-ray crystal structures of DyPs belonging to class P. Three-dimensional multi-body superposition of poly-peptide chains was performed with MODELLER (Webb and Sali, 2014). The *Pp*DyP secondary structure is shown above the alignment. The 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.

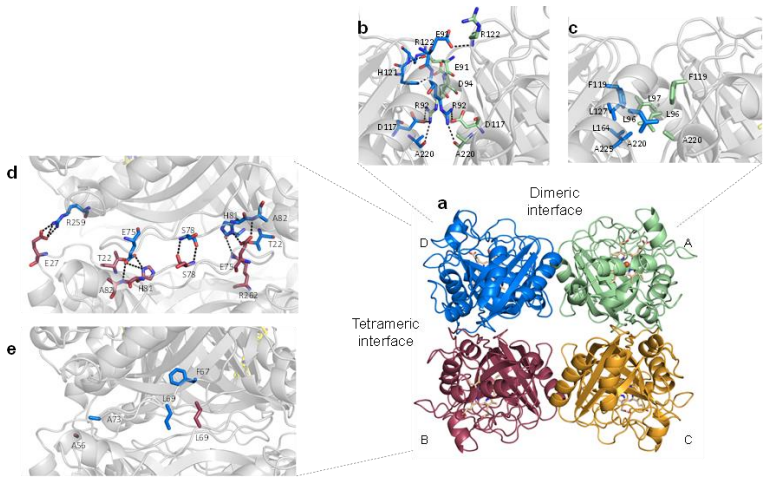


Figure SI. 3.2. Overall structure and heme access pathways in *Pp*DyP. **(a)** Dimeric and tetrameric interface interactions are involved in *Pp*DyP oligomerization. Analysis of the *Pp*DyP interfaces: the subunits AD/BC and BD/AC are the dimeric and tetrameric interfaces, respectively. Polar **(b)** 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

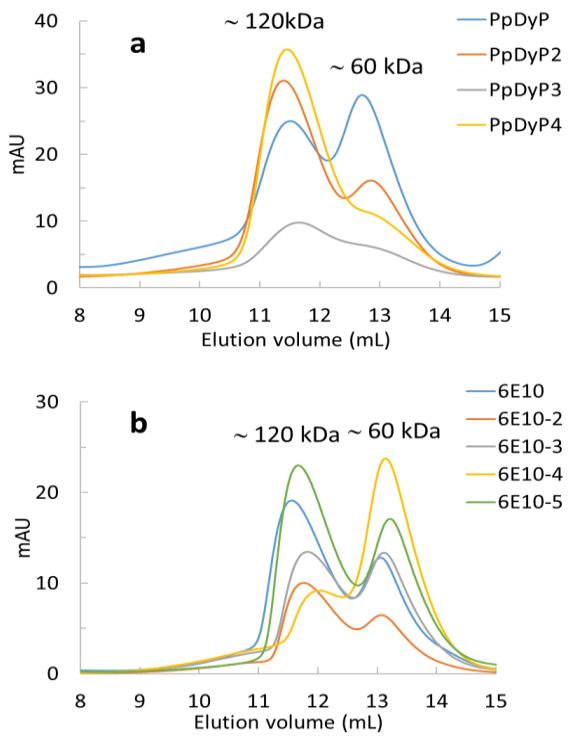


Figure SI. 3.3. Size-exclusion chromatography of (**a**) *Pp*DyP and (**b**) 6E10 equivalent enzyme preparations purified from different purification batches exhibiting random ratios of the tetrameric (128 kDa) and dimeric (64 kDa) forms.

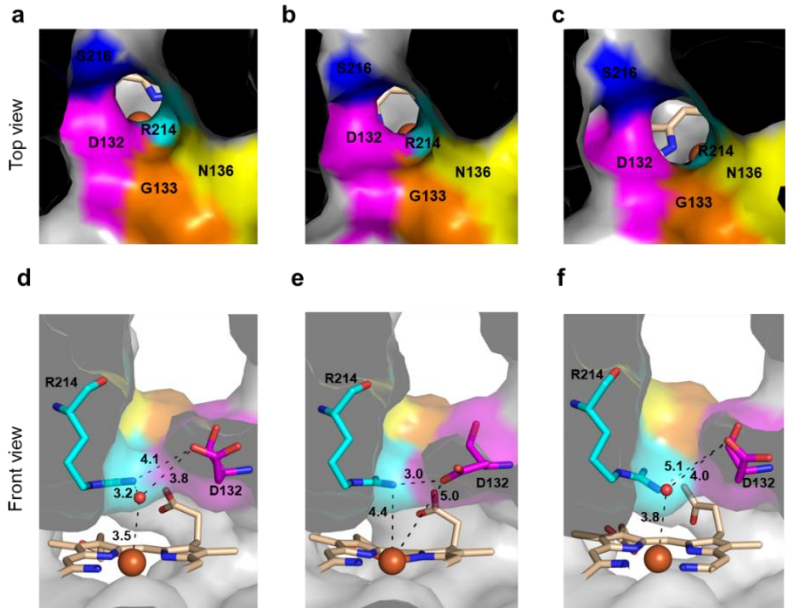


Figure SI. 3.4. Representation of the residues limiting tunnel 1 (T1) in *Pp*DyP wild type (a, d) and variants 29E4 (b, e) and 6E10 (c, f). Top: Representation of the solvent-accessible surface area of residues limiting the entrance of T1 in wild type, 29E4, and 6E10 variants, respectively. Bottom: Cut-view of the protein surface (light grey) highlighting the T1 region and catalytic residues, D132 and R214, with carbon atoms in pink and cyan, respectively. 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. The iron is shown as an orange sphere. The water molecules are shown as red spheres. The interatomic distances (Å) are shown as dashed lines.

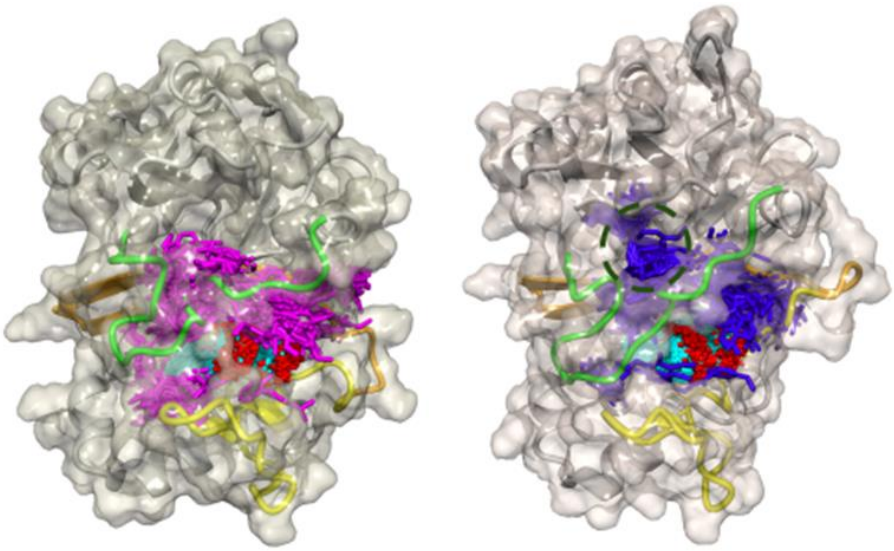


Figure SI. 3.5. Molecular representation of the results from docking DMP to *Pp*DyP (left) and to 6E10 variant (right). 664 protein structures coming from the MD simulations at pH 8 were used in each case. All binding modes with binding energy lower than -5 kcal/mol are represented. DMP is shown in magenta (wild type) and dark blue (6E10) sticks and heme cofactor in cyan-carbon sticks. L1, L2, and L3 loop regions are depicted in green, yellow, and orange, respectively. The largest entrance to T1 in 6E10 is highlighted with a dashed dark green circle.

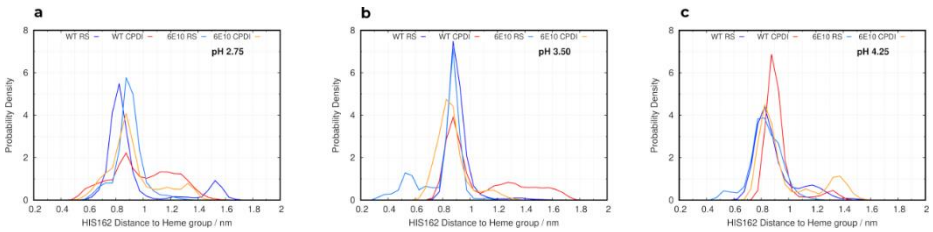


Figure SI. 3.6. Probability density distributions of the distances between His162 and the Fe atom of the heme group performed at pH 2.75 (a), 3.5 (b), and 4.25 (c) of RS and Cpd I of wild type and variant 6E10.

Annex III

Table SI. 4.1. List of mutations suggested by PROSS for the monomeric and dimeric structures of *Pp*DyP.

PROSS	
Monomer (Chain B)	**Dimer (Chain B-D)**
	L7I
P11G	
R17M	R17V
A28D	A28D
A35R	A35R
	V49I
L69Q	
E91D	E91D
	L96F
L97H	L97H
	T99A
	E103I
P108D	
L110F	L110F
S111E	S111V
	S115E
L120R	L120R
H121Y	H121Y
G123D	G123D
T138Q	T138Q
D139G	D139G
E140D	E140D
Q144E	Q144E
A149D	A149D
L159Q	
Q165D	Q165D
S169A	
D174E	D174E
A194E	A194E
	V217M
Q222G	Q222G
L232C	
K234H	K234H
E237D	E237D
V241A	
S275G	
E276D	
T277G	
V279I	
	F283A

Table SI. 4.2. List of mutations suggested by FireProt for the monomeric and dimeric structures of *Pp*DyP.

FireProt		
Monomer (Chain B)	Dimer	
	Chain B	Chain D
A9L	A9L	A9L
T22I	T22F	
S25F	S25F	S25F
A35L		
A54V	A54V	A54V
L69A		
A73G		
E91D	E91D	E91D
D94E	D94E	D94E
T99M		
Q100R	Q100R	Q100R
A105M		
L110F	L110F	L110F
S111V		
G118A	G118A	G118A
L120R		L120R
H125R	H125R	H125R
D139G	D139G	D139G
Q144W	Q144A	
A149L	A149L	A149L
S152C		
A155V	A155V	A155V
S169A	S169A	
D174E	D174E	D174E
N177M		
S184L		
V217M	V217M	V217M
S128P	S218P	S218P
	Q222G	
L232F	L232F	
E237D	E237D	E237D
	S247A	
G256A	G256A	G256A
Y258F	Y258F	Y258F
T277G		
V279L	V279L	V279L