



The  
University  
Of  
Sheffield.

**Protein Engineering of a Dye Decolorizing Peroxidase from  
*Pleurotus ostreatus* For Efficient Lignocellulose Degradation**

**Abdulrahman Hirab Ali Alessa**

A thesis submitted in partial fulfilment of the requirements for the degree of  
Doctor of Philosophy

The University of Sheffield  
Faculty of Engineering  
Department of Chemical and Biological Engineering

September 2018

## **ACKNOWLEDGEMENTS**

Firstly, I would like to express my profound gratitude to my parents, my wife, my sisters and brothers, for their continuous support and their unconditional love, without whom this would not be achieved.

My thanks go to Tabuk University for sponsoring my PhD project.

I would like to express my profound gratitude to Dr Wong for giving me the chance to undertake and complete my PhD project in his lab. Thank you for the continuous support and guidance throughout the past four years. I would also like to thank Dr Tee for invaluable scientific discussions and technical advices.

Special thanks go to the former and current students in Wong's research group without whom these four years would not be so special and exciting, Dr Pawel; Dr Hossam; Dr Zaki; Dr David Gonzales; Dr Inas,; Dr Yomi, Dr Miriam; Jose; Valeriane, Melvin, and Robert.

## SUMMARY

Dye decolorizing peroxidases (DyPs) have received extensive attention due to their biotechnological importance and potential use in the biological treatment of lignocellulosic biomass. DyPs are haem-containing peroxidases which utilize hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) to catalyse the oxidation of a wide range of substrates. Similar to naturally occurring peroxidases, DyPs are not optimized for industrial utilization owing to their inactivation induced by excess amounts of H<sub>2</sub>O<sub>2</sub>. Furthermore, DyPs are active only under acidic conditions and typically lose activity at neutral or alkaline pH. A dye decolorizing peroxidase from the *Pleurotus ostreatus* (Pleos-DyP4) was identified recently as a first fungal DyP oxidizing Mn<sup>2+</sup> to Mn<sup>3+</sup> similar to other fungal peroxidases. However, despite its unique pH and thermal stability, similar to other DyPs, it is not suited for industrial applications.

Protein engineering methods are widely used to enhance the stability and catalytic efficiency of biocatalysts to render them suitable for industrial purposes. Different directed evolution approaches (namely, error-prone PCR and saturation mutagenesis) were used to construct mutant libraries of DyP4. For protein expression studies, the mutant enzymes were co-expressed with OsmY protein (a novel secretion-enhancing protein) in order to secrete intracellular protein into the media and hence facilitate the screening of mutants. ABTS assay was used to screen for mutants with improved activities in 96-well microtiter plates. Four rounds of error-prone PCR (*ep*PCR) and saturation mutagenesis led to the identification of a mutant with an approximately 10-fold improvement in total activity and resistance to H<sub>2</sub>O<sub>2</sub> inactivation in comparison with the wild type (WT). This study showcases the usefulness of the OsmY-based secretion mechanism of protein in *E. coli* as a tool in facilitating the screening of DyP4 mutants, and potentially of other heterologous protein variants in *E. coli* – the preferred host for expression and directed evolution studies.

Chapter (6) is written for a manuscript and will be submitted for publication.

Abdulrahman HA Alessa, David Gonzalez Perez, Hossam EM Omar Ali, Alex Trevaskis, Kang Lan Tee, Xu, Jun and Tuck Seng Wong. 2019. Engineering of DyP4 using an OsmY-based secretion mechanism to facilitate the directed evolution approach in bacteria. *Bioresources and Bioprocessing*.

# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
SUMMARY .....	iii
TABLE OF CONTENTS.....	v
LIST OF FIGURES .....	x
LIST OF TABLES .....	xvi
Appendix.....	xviii
Abbreviation .....	xix
CHAPTER 1 Introduction and literature review .....	1
1.1 Introduction .....	2
1.2 Lignocellulose as a source of energy .....	3
1.2.1 Lignocellulose composition .....	3
1.2.2 Pretreatment of lignocellulose .....	8
1.2.3 White-rot and brown-rot fungi.....	13
1.3 Lignocellulose-degrading enzymes.....	13
1.3.1 Laccases .....	14
1.3.2 Peroxidases .....	14
1.4 Potential application of peroxidases.....	26
1.4.1 Bioremediation.....	26
1.4.2 Odour pollution.....	29
1.4.3 Bioremediation of azo dyes .....	29
1.4.4 Limitations of DyP4 for industrial-scale utilization .....	30
1.5 Protein engineering .....	32
1.5.1 Directed evolution.....	32
1.5.2 Rational design.....	35
1.5.3 Screening.....	36
CHAPTER 2 Materials and methods .....	39
2.1 Materials.....	40
2.2 Media preparation .....	41
2.2.1 2× TY media .....	41
2.2.2 TY AIM media.....	41
2.2.3 Tryptone yeast extract agar plates.....	41

2.3	Molecular cloning methods .....	41
2.3.1	DNA plasmid isolation .....	41
2.3.2	DNA gel extraction (PCR purification) .....	42
2.3.3	PCR clean-up .....	43
2.3.4	DNA gel electrophoresis .....	43
2.4	Bacterial transformation methods .....	44
2.4.1	Calcium chloride method .....	44
2.4.2	Electroporation method .....	45
2.5	Molecular cloning of DyP4 .....	46
2.5.1	Cloning of DyP4-tag into pET24a .....	46
2.5.2	Cloning of pET24a-DyP4-without tag .....	46
2.6	Protein expression .....	47
2.6.1	Expression of pET24a-DyP4 .....	47
2.6.2	Optimization of DyP4 secretion using a pET24a-OsmY-DyP4 vector .....	48
2.7	Analysis of protein expression .....	48
2.7.1	SDS-PAGE .....	48
2.7.2	Preparation of protein samples .....	49
2.7.3	Activity assays for DyP4 .....	50
2.8	Purification .....	51
2.8.1	Ion-exchange chromatography .....	51
CHAPTER 3 Encapsulation of DyP4 into the encapsulin nanocompartment using a dual expression system. ....		54
3.1	Introduction .....	55
3.1.1	Encapsulin .....	55
3.1.2	DyP4 .....	56
3.1.3	Encapsulation .....	56
3.1.4	Aim of the study .....	58
3.2	Methodology .....	59
3.2.1	Molecular cloning of pET24a-Encap .....	59
3.2.2	Molecular cloning of pET24a-DyP4-tag .....	60
3.2.3	Transformation in BL21 (DE3) using the CaCl <sub>2</sub> method .....	60
3.2.4	Protein expression of pET24a-Encapsulin and pET24a-DyP4-tag .....	60
3.2.5	Analysis of protein expression .....	61
3.2.6	Large-scale recombinant production of DyP4 and encapsulin .....	61




3.2.7	Purification of ReEncapsulin .....	61
3.2.8	Purification of DyP4-tag .....	64
3.2.9	Molecular cloning of ReEncapsulin and DyP4-tag in Due expression systems	64
3.3	Results .....	70
3.3.1	Molecular cloning of pET24a-Encap and pET24a-DyP4-tag.....	70
3.3.2	Recombinant protein production in BL21 (DE3) .....	70
3.3.3	Purification of encapsulin .....	71
3.3.4	Purification of DyP4-tag .....	73
3.3.5	Purification of encapsulated DyP4 into <i>ReEncapsulin</i> .....	78
3.3.6	UV-spectral analysis of purified DyP4 and encapsulated-DyP4.....	79
3.3.7	Enzymatic activity of encapsulated DyP4-tag .....	80
3.4	Discussion .....	81
3.5	Conclusion.....	82
CHAPTER 4 Substitution of methionine residues to enhance DyP4 stability against increased hydrogen peroxide concentration.....		83
4.1	Introduction .....	84
4.1.1	Mechanism of inactivation by excess H <sub>2</sub> O <sub>2</sub> .....	85
4.1.2	Aim of the study.....	85
4.2	Material and methods .....	86
4.2.1	Materials .....	86
4.2.2	Protein modelling.....	86
4.2.3	Construction of mutants using OneClick program .....	86
4.2.4	CaCl <sub>2</sub> heat shock transformation in DH5 $\alpha$ .....	87
4.2.5	Sequencing.....	88
4.2.6	CaCl <sub>2</sub> transformation in BL21 (DE3) .....	88
4.2.7	Protein expression of the WT and mutants .....	88
4.2.8	Cell harvesting and protein purification .....	88
4.2.9	Spectra and OD measurement/quantification .....	88
4.2.10	Enzymatic activities of the WT and mutants .....	89
4.3	Results .....	90
4.3.1	Engineering the methionine residues of DyP4.....	90
4.3.2	Construction of mutants .....	91
4.3.3	Sequencing results .....	92

4.3.4	Expression and purification of DyP4 variants .....	92
4.4	Discussion .....	95
	Conclusion.....	97
CHAPTER 5 Oxidation of a non-phenolic lignin model substrate, three S-type lignin units, and synthetic dye by DyP4.....		
		98
5.1	Introduction .....	99
5.1.1	Veratryl alcohol .....	100
5.1.2	Reactive Blue 19 (RB19) .....	100
5.1.3	S-type lignin units .....	100
5.1.4	Aim and objectives .....	102
5.2	Methodology .....	103
5.2.1	DyP4 expression and purification.....	103
5.2.2	Oxidation assays .....	103
5.3	Results and discussion.....	106
5.3.1	Protein expression and purification of DyP4 .....	106
5.3.2	Enzymatic oxidation of VA by DyP4 .....	107
5.3.3	Decolourization activity of RB19 by DyP4 .....	109
5.3.4	Effect of pH on Pleos-DyP4 activity with acetosyringone .....	110
5.3.5	Enzymatic oxidation of S-type lignin models.....	111
5.4	Conclusion.....	117
CHAPTER 6 Engineering of DyP4 using an OsmY-based secretion mechanism to facilitate the directed evolution approach in bacteria .....		
		118
6.1	Introduction .....	119
6.1.1	Protein engineering .....	119
6.1.2	Dye-decolorizing peroxidases (DyPs) .....	120
6.1.3	<i>E. coli</i> .....	120
6.1.4	Secretion of proteins in <i>E. coli</i> .....	121
6.1.5	Aim .....	123
6.2	Methodology .....	124
6.2.1	Cloning of OsmY and DyP4 in pET24a .....	124
6.2.2	Optimization of DyP4 screening using the OsmY mechanism .....	124
6.2.3	Validation and optimization of the ABTS assay as a screening method for DyP4 mutants .....	125
6.2.4	Screening of the WT plate .....	125



6.2.5	Directed evolution.....	126
6.2.6	Site-directed mutagenesis .....	129
6.2.7	Site-saturation mutagenesis .....	130
6.2.8	Large-scale recombinant production, purification and characterisation of DyP4 variants .....	140
6.3	Results .....	141
6.3.1	Validation of ABTS HTS assay .....	141
6.3.2	Molecular cloning .....	143
6.3.3	Error-prone PCR .....	144
6.3.4	First round of evolution .....	149
6.3.5	Second round of evolution .....	153
6.3.6	Third round of evolution .....	157
6.3.7	Fourth round of evolution .....	161
6.3.8	Saturation mutagenesis .....	162
6.3.9	Protein expression and purification of DyP4 variants .....	168
6.4	Discussion .....	171
6.5	Conclusion.....	174
CHAPTER 7	Summary and future work.....	175
7.1	Summary .....	175
7.2	Prospective .....	178
References	.....	181
Appendix	.....	191

## LIST OF FIGURES

<b>Figure 1-1:</b> Major components of lignocellulose.....	5
<b>Figure 1-2:</b> Tree, to a plant cell, to a cell wall giving a detailed chemical structure of the main components of lignocellulose. (  , Cellulose), subunits of D-glucose linked by 1,4-glycosidic bonds forming disaccharide cellobiose. Cellobiose forms the long-chain elemental fibril, (  , Hemicellulose) consist of 5- and 6-carbon monosaccharide units such as hexoses (D-glucose, D-galactose, D-mannose), pentoses (D-arabinose, D-xylose), and sugar acids, (  ; Lignin), formed by the oxidative coupling of three different phenylpropane building blocks: monolignols coniferyl alcohol, sinapyl alcohol, and <i>p</i> -coumaryl alcohol. This figure was generated using ChemDraw software. ....	7
<b>Figure 1-3:</b> Outline of the process of bioethanol production from lignocellulose. This figure was adapted from (Achinas and Euverink, 2016). ....	8
<b>Figure 1-4:</b> Scheme to illustrate lignocellulose structure and the goal of lignocellulose pretreatment. This figure was adapted from (Kumar et al., 2009).....	9
<b>Figure 1-5:</b> Schematic representation of the classification of haem peroxidases. A detailed classification of Pleos DyP4 (the enzyme investigated in this project) is given, along with other lignin-degrading peroxidases such as MnPs, LiPs, and VPs. ....	16
<b>Figure 1-6:</b> Resting state of heme peroxidases (prosthetic group with low-spin Fe <sup>III</sup> ). L constitute a histidine ligand in the majority of heme containing peroxidases and a cysteine ligand in some other heme containing peroxidases. This figure was adapted from (Hollmann and Arends, 2012).....	18
<b>Figure 1-7:</b> Overview of the general reaction cycle of peroxidases. (Reducing substrate, In-H) (In, radical specie). This figure was adapted from (Hollmann and Arends, 2012). 19	19
<b>Figure 1-8:</b> 3D model structure of ReEncapsulin showing (A) the whole carton structure of the encapsulin oligomer (60 monomers), (B) the interaction of five monomers of encapsulin, and (C) one type of opening that exists in encapsulin nanocompartment (5-6 Å).This figure was generated with Pymol and using (PDB 3dkt.1) as a template model.24	24
<b>Figure 1-9:</b> Overview of directed evolution cycles in <i>E. coli</i> . Gene of interest undergoes cycles of gene mutation using a suitable mutagenesis method (1), these genes are ligated into an expression vector (2), then transformed into a suitable expression host e.g <i>E. coli</i> (3), in the next step, mutants are expressed (4), and screened with a suitable colorimetric assay (5), and the desired mutants selected to parent the next generation (6), the cycle is repeated until the desired property is obtained. ....	33
<b>Figure 1-10:</b> An overview of rational design process in <i>E. coli</i> . ....	35
<b>Figure 2-1:</b> Reaction scheme for ABTS assay oxidation by DyP4.....	51
<b>Figure 3-1:</b> A scheme to illustrate the design of (1) <i>Re</i> -Encapsulin-His-tag and (2) Pleos-DyP4-tag. (1) ReEncapsulin was tagged with a his-tag at its C-terminal to facilitate the purification process, (2) Pleos-DyP4-tag; DyP4 was tagged with a signal peptide sequence (37 aa) in its C-terminal that is responsible for directing the enzyme into the encapsulin nanocompartment.....	57
<b>Figure 3-2:</b> Plasmid map of pET24a- <i>Re</i> Encapsulin (created with SnapGene). ReEncapsulin was inserted into pET24a vector using restriction sites <i>EcoRI</i> and <i>NdeI</i> ..	59

**Figure 3-3:** Plasmid map of pET24a-(+)-DyP4 (created with SnapGene). DyP4 was inserted into pET24a vector using restriction sites *EcoRI* and *NdeI*. ..... 60

**Figure 3-4:** Plasmid map of pRSFDuet-Re-Encap-DyP4-tag vector (created with SnapGene). ReEncapsulin was inserted into pET24a vector using restriction sites *EcoRI* and *NcoI* while DyP4 was inserted using *NdeI* and *XhoI* restriction sites. .... 67

**Figure 3-5:** Plasmid map of pACYCDuet-Re-Encapsulin-DyP4-tag vector (created with SnapGene). ReEncapsulin was inserted into pET24a vector using restriction sites *EcoRI* and *NcoI* while DyP4 was inserted using *NdeI* and *XhoI* restriction sites. .... 67

**Figure 3-6:** Comparison between the expression of DyP4-tag and encapsulin proteins induced by 0.1 mM or 1 mM of IPTG. Lane 1, molecular weight markers; lanes 2 and 3, soluble fraction of DyP4-tag induced with 0.1 mM or 1 mM of IPTG, respectively; lanes 4 and 5, soluble fractions of encapsulin induced with 0.1 mM or 1 mM of IPTG, respectively ..... 71

**Figure 3-7:** Purification of encapsulin by affinity chromatography. (A) Lane 1, molecular markers; lanes 2–8, eluted fraction of encapsulin and size exclusion chromatography. (B) Lane 1, molecular markers; lanes 2–3, eluted fractions from the first peak; lanes 4–5, eluted fractions from the second peak. .... 72

**Figure 3-8:** SDS-PAGE for the three purification steps of DyP4-tag. Q5 column (A) lane 1; molecular ladder, lanes 2–4; crude extract, flow-through, and wash, respectively, lanes 5–6; whole soluble fractions of DyP4. SP column (B) lane 1; molecular ladder, lanes 2–4; crude extract, flow-through, and wash, respectively, lanes 5–7; whole soluble fractions of DyP4. Superdex 75 column (C), lane 1; molecular ladder, lanes 2–7; whole soluble fractions of DyP4. .... 74

**Figure 3-9:** PCR amplification of encapsulin and DyP4-tag genes. DNA gel electrophoresis 1% was used to visualize the products. The expected PCR product was 0.8 kb for encapsulin (A) and 1.6 kb for DyP4-tag (B). .... 75

**Figure 3-10:** Restrictive digestion analysis of pACYCDuet-Encap (A), and pRSFDuet-Encap (B) at restriction sites *EcoRI* and *NcoI*. The expected digested products are approximately 4 kb for pACYCDuet, 3.8 kb for pRSFDuet, and 0.8 kb for encapsulin... 76

**Figure 3-11:** Restrictive digestion analysis of pRSFDuet-Encap-DyP4-tag (A) and pACYCDuet-Encap-DyP4-tag (B) at restriction sites *NdeI* and *XhoI*. The expected digested products are approximately 4.8 kb for pACYCDuet-Encap, 4.6 kb for pRSFDuet-Encap, and 1.6 kb for DyP4-tag. .... 77

**Figure 3-12:** SDS-PAGE of the coexpression of encapsulin and DyP4. (A) Lane 1, molecular weight marker; 2, crude extract; 3, flow-through; 4, wash; 5–7, purified fractions of ReEncapsulin coexpressed with DyP4-tag in pACYCDuet. (B) Lane 1, molecular weight marker; 2, crude extract; 3, flow-through; 4–6, purified fractions of ReEncapsulin coexpressed with DyP4-tag in pRSFDuet vector..... 78

**Figure 3-13:** UV-vis spectrum showing the characteristic absorption pattern of the purified fractions of DyP4-tag, encapsulated DyP4, and Encapsulin. (A) Wavelength scan measurements between 300 nm and 800 nm. (B) Wavelength scan measurements between 350 nm and 500 nm for eluted fractions of encapsulin only (black), and three eluted fractions of the encapsulated DyP4..... 79

**Figure 3-14:** Analysis of the enzymatic activity of encapsulated DyP4 into ReEncapsulin. Activity of DyP4 peroxidase using ABTS assay. The reaction was started by the addition of hydrogen peroxide and activity was measured as the change in absorbance at 405 nm. Two negative controls (red line with a buffer, brown line with encapsulin) were used; the green line indicates the packaged DyP4 (fraction 24) .....80

**Figure 4-1:** 3D model structure of Pleos DyP4 using the relative template. (A) The three methionine residues are shown in sticks and purple colour. This model was visualized with Pymol using (PDB: 1avf) as a template model. (B) Identified Met residues for mutagenesis in Pleos-DyP4, M253, M43, and M77 with 17 Å, 27.3 Å, and 30.2 Å from the heme iron respectively. ....90

**Figure 4-2:** DNA gel for PCR in site-directed mutagenesis using the OneClick program. (A) Amplified products (6.8 kb) for mutants M43L, M253L, and M253F using Q5 DNA polymerase. (B) Successfully amplified product (6.8 kb) for M77L using (PfuUltra high-fidelity DNA polymerase. ....91

**Figure 4-3:** Wavelength scan measurement for DyP4 variants taken for wavelengths 700 nm to 300 nm. Equilibration buffer contained the elution buffer of the variants (1 ml), 1 ml of the eluted fractions of the variants after the second step of chromatography was used to estimate the concentration of the protein. ....92

**Figure 4-4:** Hydrogen peroxide tolerance of DyP4 WT and Met-replaced mutants M43L, M77L, M253L, and M253F. Reaction mixtures (a total of 200 µL) contained 10 mM (185 µL) ABTS, 5 µL of diluted enzyme, and 10 µL of H<sub>2</sub>O<sub>2</sub> (4 mM, 8 mM, 12 mM, 16 mM, and 20 mM).....93

**Figure 5-1:** Purification of DyP4 using three steps of chromatography. (A) Q5 column (1<sup>st</sup> step) (A) lane 1; molecular ladder, lanes 2–5; crude extract of DyP4-tag, crude extract of DyP4, flow-through, and wash, respectively, lanes 6–8; whole soluble fractions of DyP4. SP column (2<sup>nd</sup> step) (B) lane 1; molecular ladder, lanes 2–4; crude extract, flow-through, and wash, respectively, lanes 5–8; whole soluble fractions of DyP4. Superdex 75 column (3<sup>rd</sup> step) (C), lane 1; molecular ladder, lane 2; crude extract, lanes 3–6; whole soluble fractions of DyP4. (D), (E), and (F) shows the chromatogram of the eluted fractions in the first, second, and third step of DyP4 purification respectively..... 106

**Figure 5-2:** Increase in absorbance at 310 nm due to the oxidation of VA by DyP4. A total of 1 ml reaction mixture contained 2.0 mM of VA and 25 µl or 50 µl of diluted purified enzyme (DyP4), and 4 mM of H<sub>2</sub>O<sub>2</sub> (50 µL). Different volumes of DyP4 (50 µL, blue line and 25 µL, green line. Negative control, red line. Each reaction was carried out in at least three replicates and the average is presented. .... 108

**Figure 5-3:** Decrease in absorbance at 595 nm due to the decolourization of RB19 by DyP4. .... 109

**Figure 5-4:** pH activity profile of DyP4 with acetosyringone. .... 110

**Figure 5-5:** Oxidation spectra of acetosyringone. (A) Oxidation of acetosyringone by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 520 nm over 60 min time. .... 112

<b>Figure 5-6:</b> Oxidation spectra of syringaldehyde. (A) Oxidation of syringaldehyde by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 520 nm over 60 min time. ....	113
<b>Figure 5-7:</b> Oxidation spectra of sinapic cid. (A) Oxidation of sinapic acid by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 500 nm over 60 min time. ....	113
<b>Figure 5-8:</b> Oxidation spectra of violuric acid. (A) Oxidation of violuric acid by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 521 nm over 60 min time. ....	114
<b>Figure 6-1:</b> Host expression systems used routinely in directed evolution studies. This figure was adapted from (Pourmir and Johannes, 2012). ....	121
<b>Figure 6-2:</b> Plasmid map of pET24a-OsmY-DyP4 (created with SnapGene).....	124
<b>Figure 6-3:</b> (A) Scheme of the HTS protocol for WT of DyP4 with ABTS assay, and (B) validated HTS with 11.5 % coefficient of variation (CV) based on the oxidation of ABTS by WT-DyP4. The activities of DyP4 from different replicates of the same clone are plotted in descending order (Black dots), and the average is plotted (Red dots).....	142
<b>Figure 6-4:</b> Restrictive digestion for pET24a-OsmY-Dyp4 with <i>BamHI</i> and <i>EcoRI</i> restriction enzymes. The restrictive digestion mixture prepared in a total of 100 $\mu$ L containing 1 $\times$ of CutSmart buffer, 3000 ng of pET24a-OsmY-Dyp4, 1 U of <i>BamHI</i> and 1 U of <i>EcoRI</i> enzymes, and 59 $\mu$ L deionized water, mixed and incubated overnight at 37°C. 20 $\mu$ L of 6X DNA Loading Dye added and the 120 $\mu$ L was loaded onto 1% agarose and 5.9 kb band were gel extracted and purified. The cut vector (5.9 kb) ◀, and the cut insert (DyP4) is 1.5 kb ◀.....	143
<b>Figure 6-5:</b> Analysis of PCR products on a 1% (w/v) DNA agarose gel showing products of different mutagenic reactions of <i>epPCR</i> ; A) ◀ three different conditions; B) ◀ medium and high mutagenic conditions. Side products ◀. Low mutagenic condition contained 1.5 mM of MgCl <sub>2</sub> , 0.01 mM of MnCl <sub>2</sub> , 0.3 mM of dNTP mix, 4.5 pmol of forward and reverse primers, 3.5 ng/ $\mu$ L of DNA template, and 1.25 U of <i>Taq</i> DNA polymerase. Medium mutagenic condition contained 7 mM of MgCl <sub>2</sub> , 0.2 mM of dATP, 0.2 mM of dGTP, 1 mM of dTTP, 1 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/ $\mu$ L of DNA template, and 1.25 U of <i>Taq</i> DNA polymerase. High mutagenic condition contained 7 mM of MgCl <sub>2</sub> , 0.05 mM of MnCl <sub>2</sub> , 0.2 mM of dATP, 0.2 mM of dGTP, 0.2 mM of dTTP, 0.2 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/ $\mu$ L of DNA template, and 1.25 U of <i>Taq</i> DNA polymerase. ....	144
<b>Figure 6-6:</b> Analysis of PCR products on a 1% (w/v) DNA agarose gel showing the products of saturation mutagenesis using the 4-primer method for position V56. (A) Amplification of two fragments (fragment 1: 168 bp and fragment 2: 1347 bp). (B) Full-length PCR for the two fragments (product at 1.5 kbp).....	145
<b>Figure 6-7:</b> Analysis of PCR products on a 1% (w/v) DNA agarose gel showing the PCR product of the saturation mutagenesis for position N312 using the NEBase Changer method in the second lane.....	146
<b>Figure 6-8:</b> Analysis of PCR products on a 1% (w/v) DNA agarose gel showing the products of saturation mutagenesis using the 4-primer method for positions (A306: F1 918	

bp and F2 597 bp), (R109: F1 327 bp and F2 1188 bp), (N227: F1 681 bp and F2 834 bp), (H374: F1 1122bp and F2 393 bp)..... 147

**Figure 6-9:** Combination of two fragments for positions A306, R109, N227, and H374 using full-length PCR. Reaction products were loaded into 1% agarose gel and the 1.5 kb band was cut and purified for each position to be used for ligation and transformation. 148

**Figure 6-10:** HTS for mutants in the first round of evolution using WT-DYP4 as a template for *ep*PCR. (A) Low mutagenic condition, (B) Medium mutagenic condition, (C) High mutagenic condition. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 150  $\mu$ L of ABTS (10 mM), 20  $\mu$ L of secreted enzyme, and 4 mM of H<sub>2</sub>O<sub>2</sub> (50  $\mu$ L). ..... 150

**Figure 6-11:** Normalized activity to OD for mutants A306V (Blue line) and N312S (Brown line), and the WT (Orange line). (A) Activity measured 2 minutes after the addition of hydrogen peroxide to initiate the oxidation of ABTS; (B) and (C) after 10 minutes and 20 minutes of adding hydrogen peroxide, respectively. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. .... 151

**Figure 6-12:** (A) The 3D protein model structure of DyP4 was generated with Pymol and using (PDB: 1afv) as a template model showing the position of A306V and N312S mutations. (B) detailed view of the Heme in DyP4 is given and position of the mutations shown in Blue stick. .... 152

**Figure 6-13:** Total activity profiles for mutants generated using N312S as the parental strain. (A) Low mutagenic condition, (B) medium mutagenic condition, (C) high mutagenic condition. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 150  $\mu$ L of ABTS (10 mM), 20  $\mu$ L of secreted enzyme, and 4 mM of H<sub>2</sub>O<sub>2</sub> (50  $\mu$ L). ..... 154

**Figure 6-14:** Normalized activity to OD for three significantly improved mutants (I56V, Brown line), (H347R, Green line), and (N227D, Red line), in the second round of evolution using mutant (N312S, Blue line) as a parental strain. Measurements were taken at 405 nm after 2 min (A), 10 min (B), 20 min (C), and 30 min (D) of initiating oxidation of ABTS by addition of hydrogen peroxide. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. The reaction mixture (total of 220  $\mu$ L) contained 10 mM of ABTS (150  $\mu$ L), 20  $\mu$ L of media (secreted enzyme), and 4 mM of hydrogen peroxide (50  $\mu$ L). ..... 155

**Figure 6-15:** HTS screening results for mutants screened in the third round of evolution using I56V mutant as a parental strain. A and B, low mutagenic condition; C and D, medium mutagenic condition; and E and F, high mutagenic condition. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 150  $\mu$ L of ABTS (10 mM), 20  $\mu$ L of secreted enzyme, and 4 mM of H<sub>2</sub>O<sub>2</sub> (50  $\mu$ L). ... 158

**Figure 6-16:** Normalized activity to OD for the most-improved mutant (F6, R109K, Brown line) with the parental strain (A5, I56V, Blue line) and the WT (Orange line). Measurements taken at wavelength 405 nm after 2 min (A), 10 min (B), 20 min (C), and 30 min (D) of initiating oxidation of ABTS by addition of hydrogen peroxide. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. The reaction mixture (total of 220  $\mu$ L) contained 10

mM of ABTS (150  $\mu$ L), 20  $\mu$ L of media (secreted enzyme), and 4 mM of hydrogen peroxide (50  $\mu$ L). ..... 159

**Figure 6-17:** UV spectra for ABTS oxidation by DyP4 to determine the optimal wavelength. Reaction mixture was prepared in a 1 ml cuvette containing 925  $\mu$ L of ABTS (10 mM) pH 3.4, 25  $\mu$ L of purified diluted DyP4, and 50  $\mu$ L of H<sub>2</sub>O<sub>2</sub> (4mM). Wavelength scan measurements recorded 15 seconds after addition of H<sub>2</sub>O<sub>2</sub> to initiate the activity.. 161

**Figure 6-18:** Total activity profiles for mutants generated by site saturation mutagenesis for positions N312 (A), V56 (B), A306 (C), R109 (D), N227 (E), and H374 (F) using F6 mutant as a parental strain. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 10 mM of ABTS (150  $\mu$ L), 10  $\mu$ L of secreted enzyme, and 4 mM of H<sub>2</sub>O<sub>2</sub> (50  $\mu$ L). ..... 163

**Figure 6-19:** Normalized activity to OD for the most-improved mutant (D4,N227S)) with (P.S, F6,R109K)) and the WT. Measurements taken at wavelength 405 nm after 2 min (A), 10 min (B), 20 min (C), and 30 min (D) of initiating oxidation of ABTS by addition of hydrogen peroxide. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. The reaction mixture (total of 210  $\mu$ L) contained 10 mM of ABTS (150  $\mu$ L), 10  $\mu$ L of media (secreted enzyme), and 4 mM of hydrogen peroxide (50  $\mu$ L). ..... 164

## LIST OF TABLES

<b>Table 1-1:</b> Lignocellulosic biomass composition of common plants .....	4
<b>Table 1-2:</b> Host organisms commonly used for recombinant expression of lignin degrading enzymes that are hydrogen peroxide-dependant.....	17
<b>Table 1-3:</b> Significant factors to consider for a biocatalyst in industrial applications. Information in this table is adapted from (Lorenz and Eck, 2005).....	31
<b>Table 2-1:</b> Primers used to amplify DyP4 without a tag.....	46
<b>Table 2-2:</b> Composition of the buffer used in the first step of chromatography for DyP4-tag purification (Ion exchange chromatography- Q column). .....	52
<b>Table 2-3:</b> Composition of the buffer used in the second step of chromatography for DyP4-tag purification (Ion exchange chromatography - SP column).....	53
<b>Table 2-4:</b> Composition of the buffer used in the third step of chromatography for DyP4-tag purification (size exclusion chromatography - Gel filtration). .....	53
<b>Table 3-1:</b> Composition of the buffer used in the first step of chromatography for ReEncapsulin purification (Affinity chromatography).....	63
<b>Table 3-2:</b> Composition of the buffer used in the second step of chromatography for ReEncapsulin purification (Size exclusion chromatography).....	63
<b>Table 3-3:</b> List of primers used to amplify Encap and DyP4.....	65
<b>Table 3-4:</b> PCR mixture for Encapsulin amplification .....	66
<b>Table 3-5:</b> PCR mixture for DyP4-tag amplification.....	66
<b>Table 3-6:</b> PCR program to amplify Encapsulin.....	66
<b>Table 3-7:</b> PCR program to amplify DyP4 .....	67
<b>Table 4-1:</b> Designed primers for the construction of mutants .....	87
<b>Table 4-2:</b> Optical density of DyP4 variants.....	93
<b>Table 5-1:</b> Substrates used in this study.....	115
<b>Table 6-1:</b> Overview of <i>ep</i> PCR programs used for library construction .....	127
<b>Table 6-2:</b> Reaction mixtures of ligation in the first round of evolution .....	127
<b>Table 6-3:</b> Mutagenic primers used for site saturation mutagenesis at position N312..	131
<b>Table 6-4:</b> PCR mixtures for site saturation mutagenesis at position N312 .....	131
<b>Table 6-5:</b> PCR program for site saturation mutagenesis at position N312.....	132
<b>Table 6-6:</b> Primers used for five amino acid positions using the 4-primer method.....	133
<b>Table 6-7:</b> PCR program to amplify fragment (1) for position V56.....	134
<b>Table 6-8:</b> PCR program to amplify fragment (2) for position V56.....	134
<b>Table 6-9:</b> PCR program to amplify fragment (1) for position A306.....	134
<b>Table 6-10:</b> PCR program to amplify fragment (2) for position A306.....	135
<b>Table 6-11:</b> PCR program to amplify fragment (1) for position R109.....	135
<b>Table 6-12:</b> PCR program to amplify fragment (2) for position R109.....	135
<b>Table 6-13:</b> PCR program to amplify fragment (1) for position N227 .....	136
<b>Table 6-14:</b> PCR program to amplify fragment (2) for position N227 .....	136
<b>Table 6-15:</b> PCR program to amplify fragment (1) for position H374.....	136
<b>Table 6-16:</b> PCR program to amplify fragment (2) for position H374.....	137
<b>Table 6-17:</b> Mixtures for full-length PCR performed for the 5 amino acid positions ...	138



<b>Table 6-18:</b> Two-stage full-length PCR program using two fragments for each position .....	138
<b>Table 6-19:</b> List of potentially improved mutants and mutation types in the first round of evolution .....	152
<b>Table 6-20 :</b> List of potentially improved mutants and mutation types in the second round of evolution .....	156
<b>Table 6-21:</b> List of potentially improved mutants and mutation types that caused the improvement in the third round of evolution .....	160
<b>Table 6-22:</b> List of potentially improved mutants and mutation types that caused the improvement in the fourth round of evolution using saturation mutagenesis .....	165

## Appendix

<b>Appendix (1):</b> Analysis of PCR products on a 1% (w/v) DNA agarose gel showing products of different mutagenic reactions of <i>ep</i> PCR in the second round of evolution using N312S mutant as a template for <i>ep</i> PCR. ◀ Main bands for Low, Medium and High mutagenic conditions (1.5 kb); Side products ◀ .....	192
<b>Appendix (2):</b> List of the sequencing sequences used to sequence mutants of Pleos-DyP.....	192
<b>Appendix (3):</b> DNA and protein sequences for Pleos-DyP4-tag.....	193
<b>Appendix (4):</b> DNA and protein sequences for ReEncapsulin.....	197
<b>Appendix (5):</b> DNA and protein sequences for DyP4-WT sequence.....	199
<b>Appendix (6):</b> DNA and protein sequence for DyP4 mutant (F6) N312S.....	203

## Abbreviation

ABTS	2,2'-Azino-bis(3-ethylbenzthiazoline-6-sulfonic acid)
APO	Aromatic peroxygenase
bp	Base pair
CaCl <sub>2</sub>	Calcium chloride
CV	Coefficient of variation
°C	Degree Celsius
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DyPs	Dye Decolorizing Peroxidases
DyP4	Dye Decolorizing Peroxidase 4
<i>E. coli</i>	<i>Escherichia coli</i>
<i>ep</i> PCR	Error prone PCR
H <sub>2</sub> O	Water
H <sub>2</sub> O <sub>2</sub>	Hydrogen peroxide
Hrs	Hour
kDa	KiloDalton
kbp	Kilobase pair
LiP	Lignin peroxidase
L	Liter
HST	High-throughput screening
pI	Isoelectric point
μL	Microliter
μg	Microgram
mM	Millimolar
mg	Milligram
ml	Millilitre

Min	Minute
MnP	Manganese peroxidase
nm	Nanometr
ng	Nanogram
OD <sub>600</sub>	Optical density OD <sub>600</sub>
<i>P. pastoris</i>	<i>Pichia pastoris</i>
<i>P. ostreatus</i>	<i>Pleurotus ostreatus</i>
“P.S”	Parental strain
rpm	Revolutions per minute
Sec	Seconds
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel
UV	Ultraviolet
VP	Versatile peroxidase
WT	Wild type
UV-Vis	Ultraviolet-visible

# **CHAPTER 1 Introduction and literature review**

## 1.1 Introduction

## 1.2 Lignocellulose as a source of energy

### 1.2.1 Lignocellulose composition

### 1.2.2 Pretreatment of lignocellulose

### 1.2.3 White-rot and brown-rot fungi

## 1.3 Lignocellulose degrading enzymes

### 1.3.1 Laccases

### 1.3.2 Peroxidases

## 1.4 Potential applications of peroxidases

### 1.4.1 Bioremediation

### 1.4.2 Odour pollution

### 1.4.3 Bioremediation of azo dyes

## 1.5 Protein engineering

### 1.5.1 Directed evolution

### 1.5.2 Rational design

### 1.5.3 Bioremediation

## **1.1 Introduction**

Today, there is an increasing demand for a sustainable source of energy as the global economy is largely dependent on fossil energy sources such as, natural gas, coal and oil. These sources are used in the production of electricity, fuels, chemicals, and other goods. As the continued utilization of these resources is not considered to be sustainable, alternative sources must be found and used (Uihlein and Schebek, 2009). Another reason to search for alternative sources of energy is the increased levels of pollution and the associated consequences such as global warming. Renewable sources of energy such as water, wind, geothermal, solar, and biomass can serve as an excellent alternative for the energy industry. However, biomass also represent a potential alternative source of energy for the chemical industry and for the production of fuel (Sarkar et al., 2012). Bioethanol is produced from biomass-derived sugars by microbial fermentation. The use of bioethanol as an alternative to fossil fuel is less harmful to the environment as it does not contribute to greenhouse gas emission and has other health and environmental benefits. In the first-generation feedstock, bioethanol was produced from sugar crops such as sugar beet and sugar cane, and from agriculture cereal such as maize, wheat, and barley. The major disadvantages of the first-generation feedstock is the impact on food security and the use of lands that should be designated to the cultivation of crops for human and animal consumption (Walker, 2011). Furthermore, energy-yielding crops require large areas of land for increased production, which, in contrast, will reduce and limit the availability of land for the cultivation of crop foods. However, this can be solved by the application of second-generation feedstock, that is, the production of bioethanol from lignocellulose. This is considered to be the most promising option to replace fossil fuel and to overcome the limitations of the first-generation feedstock (Moreno et al., 2015, Walker, 2011).

## **1.2 Lignocellulose as a source of energy**

Over the past 20 years, the production of ethanol from lignocellulose has received extensive attention due to its notable advantages. Lignocellulose is an abundant, renewable, and cheap feedstock for bioethanol production. This approach is also environmentally friendly, as waste materials such as wheat straw, corn straw, rice straw, and sugarcane bagasse are used for the production of bioethanol and other useful products (Sarkar et al., 2012). By 2022, the use of biofuels in the United States is expected to reach 36 billion gallons, of which at least 21 billion gallons will be obtained from lignocellulosic materials (Schnepf and Yacobucci, 2010).

### **1.2.1 Lignocellulose composition**

Lignocellulosic biomass mainly comprises cellulose, hemicellulose, and lignin, with a small amount of pectin. The percentage of these components vary from plant to plant (Table 1-1). Bioethanol production requires the conversion of cellulose and hemicellulose into smaller units (glucose), followed by fermentation to convert these units into bioethanol (Balat, 2011).

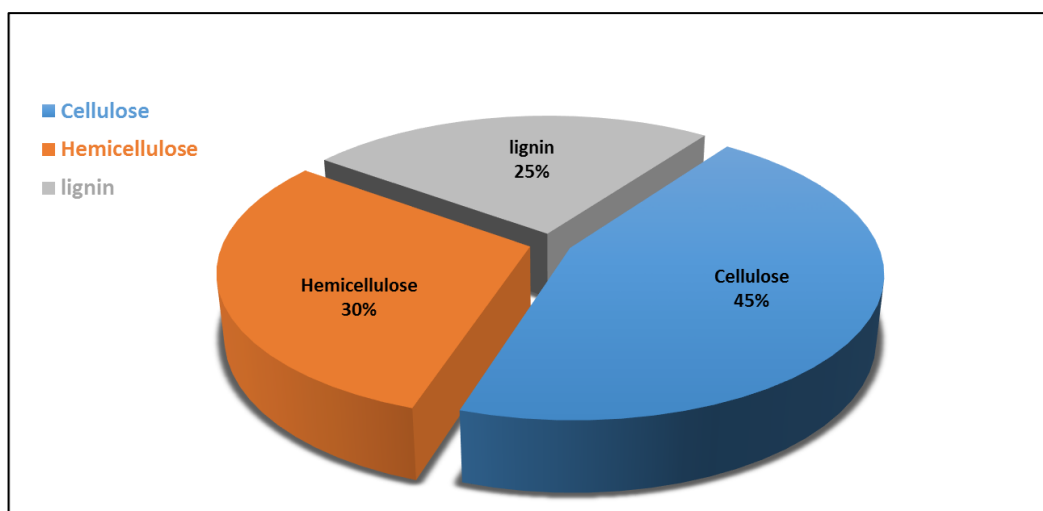
**Table 1-1:** Lignocellulosic biomass composition of common plants

<b>Feedstock</b>	<b>Cellulose (%)</b>	<b>Hemicellulose (%)</b>	<b>Lignin (%)</b>	<b>Reference</b>
Alfalfa	21.8	12.4	9.7	(Dijkerman et al., 1997)
Bamboo	49-50	18-20	23	(Menon and Rao, 2012b)
Banana waste	13.2	14.8	14	(Shahzadi et al., 2014)
Barley hull	34	36	19	(Kim et al., 2008)
Barley straw	36-43	24-33	6.3-9.8	(Menon and Rao, 2012a)
Blue agave	31-55	8-17	7-12	(Sorek et al., 2014)
Bagasse	38	27	20	(Walker, 2011)
Coconut fibre	17.7	2.2	34	(Dijkerman et al., 1997)
Coffee pulp	24	8.9	19.4	(Dijkerman et al., 1997)
Corn stover	42	28	26	(Biswas et al., 2015)
Eucalyptus	50	13	28	(Walker, 2011)
Grasses	25-40	25-50	10-30	(Shahzadi et al., 2014)
Hardwood	40-55	24-40	18-25	(Shahzadi et al., 2014)
Softwood	45-50	25-35	25-35	(Shahzadi et al., 2014)
Maple	44.9	17.3	20.7	(Zhao et al., 2012)
Miscanthus	52	26	13	(Brosse et al., 2012)
Douglas fir	44	19.2	30	(Xu and Huang, 2014)
Olive tree	25	15.8	16.6	(Cara et al., 2008)
Pine	43.3	20.5	28.3	(Xu and Huang, 2014)
Poplar	44.7	18.5	26.4	(Sorek et al., 2014, Xu and Huang, 2014)
Perennial grass	37-45	19-25	17-21	(Haffner et al., 2013)
Red oak	43.4	18.9	25.8	(Zhao et al., 2012)
Reed	39.5	29.8	24	(Li et al., 2009)
Rice husk	24	27	13	(Walker, 2011)
Switch grass	31-45	20-30	12-18	(Walker, 2011)
Salix	43	22	26	(Walker, 2011)
Wheat straw	44.1	23.8	20.5	(Xu and Huang, 2014)
White oak	43.6	18	23.2	(Zhao et al., 2012)
Willow	49.3	14.1	20	(Bridgeman et al., 2008)



## Cellulose

Cellulose is a major component of lignocellulosic biomass, and is regarded as the most abundant organic polymer on earth, comprising approximately 35%–40% of the dry weight of wood. Being a polymer, Cellulose is made up of D-glucose subunits linked by 1,4-glycosidic bonds forming disaccharide cellobiose. Cellobiose in turn forms the long-chain elemental fibril (Figure 1-2). The glucose subunits are linked tightly with extensive intramolecular and intermolecular hydrogen bonds, making breakdown of the polymer more difficult (Balat, 2011, Isikgor and Becer, 2015). The conversion of cellulose using cellulases into glucose is of significant importance, taking into account that cellulose likely represents about half of the organic carbon in the biosphere (Zhou et al., 2012).



**Figure 1-1:** Major components of lignocellulose

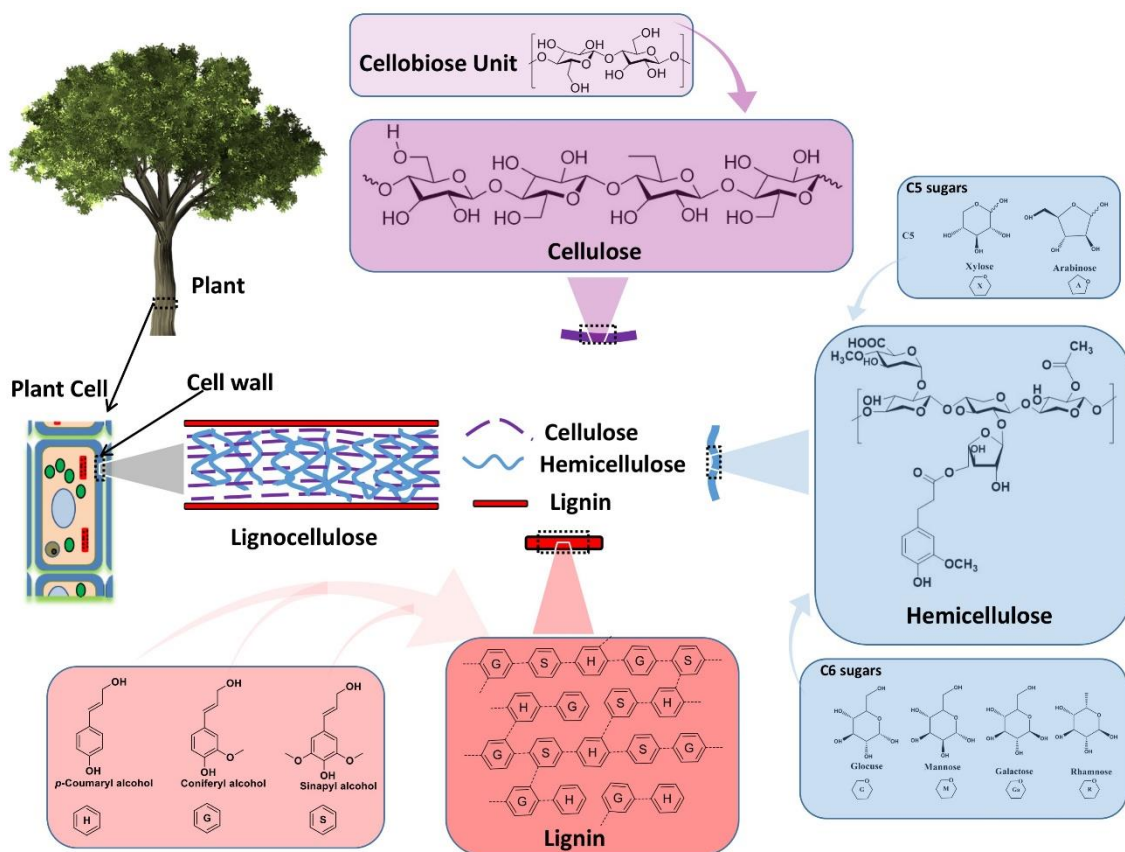
## Hemicellulose




Hemicellulose is another major constituent of lignocellulose, comprising around 25%–30% of the total dry weight of wood. In contrast to cellulose, the structure of hemicellulose is amorphous and features lateral or branched chains, which are composed of heteropolymers such as galactomannan, xylan, arabinoxylan, xyloglucan, glucuronoxylan, and glucomannan. These heteropolymers are composed of various types of 5- and 6-carbon monosaccharide units such as hexoses (D-glucose, D-galactose, D-mannose), pentoses (D-arabinose, D-xylose), and sugar acids (Figure 1-2). Hemicellulose

strengthens the plant cell wall by tethering cellulose microfibrils. The composition of hemicellulose differs among plant species. Softwood hemicelluloses, for instance, are composed mainly of glucomanna, whereas hardwood hemicelluloses are composed mostly of glucuronoxylan (Scheller and Ulvskov, 2010, Isikgor and Becer, 2015, Kumar et al., 2008).

### **Lignin**

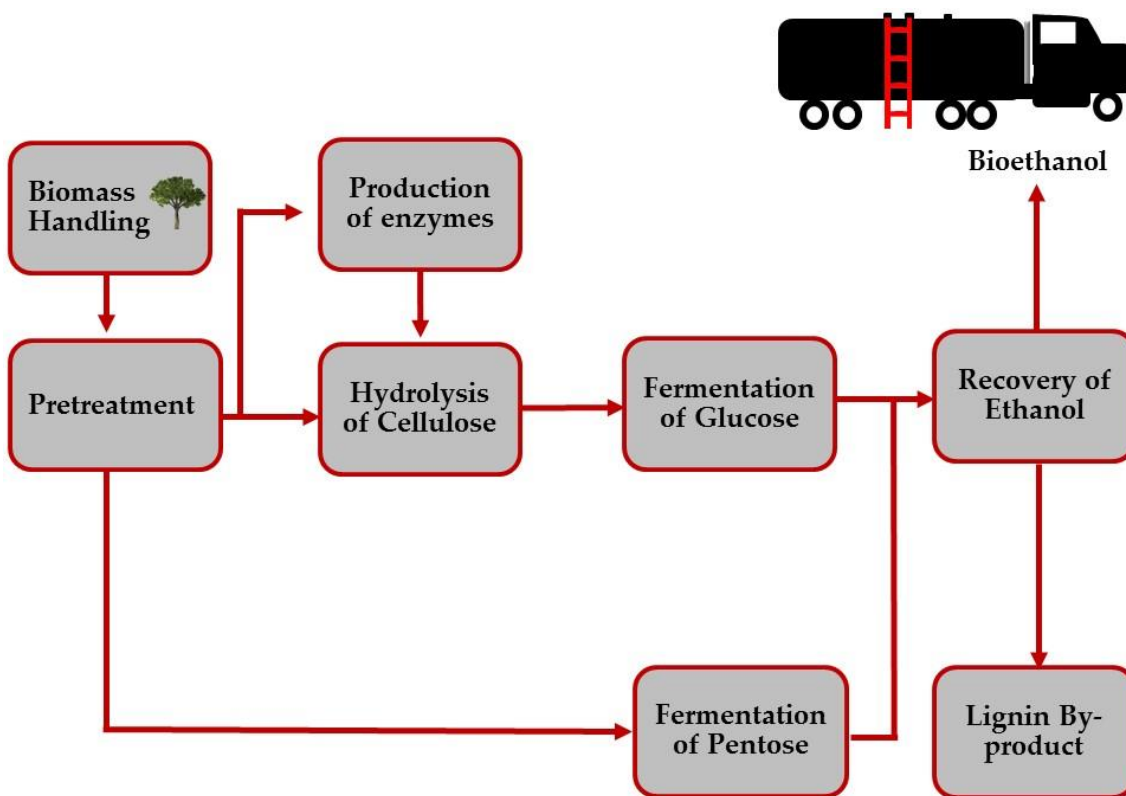
Lignin is the third most prevalent constituent of lignocellulose and, just after cellulose, regarded as the second most abundant organic polymer on earth. Around 25% of the total dry weight of lignocellulose is made up of lignin. The structure of lignin is formed by the oxidative coupling of three different phenylpropane building blocks: the monolignols coniferyl alcohol, sinapyl alcohol, and *p*-coumaryl alcohol (Figure 1-2). These three building blocks of lignin vary between softwood and hardwood lignins, and these units are connected by linkages such as aryl-glycerol  $\beta$ -arylether, C-C, and arylether. Lignin acts to strengthen the cell walls of plants and to provide protection against microbial attack and diseases. Unlike cellulose and hemicellulose, lignin does not contain fermentable sugars. The presence of lignin in the cell wall prevents the enzymatic hydrolysis of cellulose and hemicellulose and prevents the access to these fermentable sugars. The breakdown of lignin is therefore a critical step in accessing fermentable sugars, covered and protected by lignin polymer for fermentation to produce useful products. Furthermore, the breakdown of lignin is beneficial for the synthesis of complex chemicals and other useful fine chemicals (Isikgor and Becer, 2015, Agbor et al., 2011, Zeng et al., 2014, Sanchez, 2009, Pérez et al., 2002).



**Figure 1-2:** Tree, to a plant cell, to a cell wall giving a detailed chemical structure of the main components of lignocellulose. (  , Cellulose), subunits of D-glucose linked by 1,4-glycosidic bonds forming disaccharide cellobiose. Cellobiose forms the long-chain elemental fibril, (  , Hemicellulose) consist of 5- and 6-carbon monosaccharide units such as hexoses (D-glucose, D-galactose, D-mannose), pentoses (D-arabinose, D-xylose), and sugar acids, (  ; Lignin), formed by the oxidative coupling of three different phenylpropane building blocks: monolignols coniferyl alcohol, sinapyl alcohol, and *p*-coumaryl alcohol. This figure was generated using ChemDraw software.

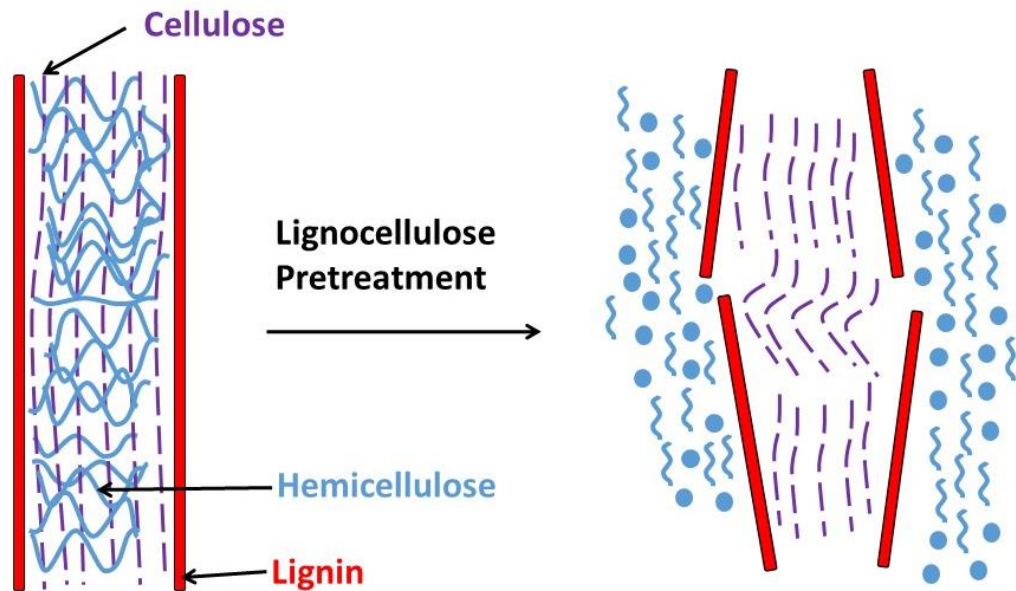
### 1.2.2 Pretreatment of lignocellulose

The process of producing bioethanol from lignocellulose feedstock includes the following steps: (i) breakdown of lignin to enable access to fermentable sugars, (ii) enzymatic hydrolysis of cellulose and hemicellulose, (iii) fermentation of sugars, and (iv) purification and recovery of ethanol (Figure 1-3) (Achinas and Euverink, 2016).



**Figure 1-3:** Outline of the process of bioethanol production from lignocellulose. This figure was adapted from (Achinas and Euverink, 2016).

The pretreatment of lignocellulose is a critical step for the production of cellulosic ethanol and has a direct effect on the overall process. The cost of this step is around 30% of the overall cost of the bioconversion process. The ultimate objective in this step is to enhance enzymatic access to cellulose during the hydrolysis step by (i) disrupting the crystalline structure of cellulose, and (ii) breaking down the structure of lignin, which acts as a physical barrier of cellulose (Figure 1-4). Pretreatment approaches are generally classified into four main categories: physical treatment, chemical treatment, combined physio-chemical treatment, and biological treatment (Mosier et al., 2005, Asgher et al., 2014, Sarkar et al., 2012).



**Figure 1-4:** Scheme to illustrate lignocellulose structure and the goal of lignocellulose pretreatment. This figure was adapted from (Kumar et al., 2009).

### **1.2.2.1 Physical treatment**

Physical pretreatment involves the comminution of the lignocellulose biomass. In this process, chipping, grinding, milling, and shredding can be used to decrease the degree of polymerization and to increase the surface area for enzymatic activity. Appropriate milling methods can be used to produce ethanol and methane, as no inhibitor is formed during this process. However, milling may not be economically feasible, owing to its high energy requirements. Extrusion is another method used in the physical treatment of lignocellulose, and although it is similar to milling, it requires high energy and may not be appropriate for large-scale applications. Gamma rays may also be used in this type of treatment to increase the surface area and to lower crystallinity; however, this method is not without its drawbacks, including its high cost, especially for large quantities of lignocellulose, and its potential environmental effects and safety risks (Agbor et al., 2011, Kumar and Sharma, 2017).

### **1.2.2.2 Chemical treatment**

Organic solvents, alkali, acids, and ionic liquids are used to alter the native structure of lignocellulose and thus facilitate its subsequent degradation.

#### **Alkali treatment**

Pretreatment with alkaline solutions such as potassium hydroxide, sodium hydroxide, hydrazine, and anhydrous ammonia exposes the internal surface of the lignocellulose to reduce cellulose crystallinity and the degree of polymerization. It also causes disruption of the structure of lignin and breaks the bonds between lignin and other carbohydrates, increasing enzyme accessibility to fermentable sugars. Alkaline treatment is more efficient for lignocellulosic material with low lignin content than for lignin-rich material (Agbor et al., 2011).

### **Acid pretreatment**

Acid pretreatment is mainly used to solubilize the hemicellulose fraction and thus increase the accessibility of enzymes to cellulose. Diluted acid is preferable to concentrated acid, which can form inhibitors and cause the degradation of cellulose (Agbor et al., 2011). However, pretreatment with concentrated acid yields higher amounts of degradation product compared with diluted acid (Alvira et al., 2010).

### **Organic solvent**

In this pretreatment method, organic solvent or aqueous solution mixtures are used to remove lignin and facilitate the enzymatic hydrolysis of cellulose. Some types of organic solvents used in this pretreatment include methanol, ethanol, ethylene glycol, acetone, and tetrahydrofurfuryl alcohol. Despite the fact that this method of pretreatment results in the production of lignin as a pure by-product, it might not be economically feasible due to the high price of organic solvent (Alvira et al., 2010).

#### **1.2.2.3 Combined physical and chemical treatment**

Steam explosion, liquid hot water (LHW), and ammonia fibre explosion (AFEX) are some common methods that are being used in a combined physical and chemical treatment approach. Steam explosion treatment involves the exposure of the biomass to high-pressure saturated steam at high temperatures (160–260°C). This treatment causes the separation of individual fibres and the disruption of the structure of the cell wall. It also causes hemicellulose solubilization and alters the lignin structure to allow enzyme to access to cellulose. The formation of inhibitors and a low efficiency for biomass from softwood compared with agricultural biomass are some of the drawbacks of this method (Kumar and Sharma, 2017, Xu and Huang, 2014). In LHW pretreatment, water in the form of liquid is used at a high temperature (160–240°C). The main advantages of this method are (i) reduced formation of inhibitors, (ii) high recovery rate of sugars, and (iii) cost effectiveness, with no requirement for the addition of chemicals and inhibitors. Its

main disadvantage is that it is not economically feasible for industrial use due to its high energy and water requirements compared with the previous method (Xu and Huang, 2014). In ammonia fibre explosion, which is similar to steam explosion, the lignocellulose biomass is treated with high-pressure liquid ammonia at high temperatures (60–100°C) for 30–60 min. This causes the swelling and breakdown of the lignocellulose structure. Although this method has advantages such as (i) the use of ammonia, which can be recycled; (ii) a high recovery rate of sugars; and (iii) an absence of inhibitor formation, it is more efficient for non-woody biomass than woody biomass (Kim, 2018).

#### **1.2.2.4 Biological treatment**

Biological treatment methods involve the use of whole microorganisms such as fungi or the enzymes produced by these microorganisms. Brown-rot fungi, white, and soft-rot fungi produce enzymes that are involved in the degradation of cellulose and hemicellulose and can alter only lignin, while other fungi produce enzymes that can efficiently degrade lignin, hemicellulose, and, to some extent, cellulose (Wan and Li, 2012). The approach of utilizing fungal treatment for the production of cellulosic ethanol has received significant attention due to its economic and environmental advantages. These advantages include (i) lower energy requirements compared with other methods, (ii) lower production of fermentation inhibitors, (iii) lower cost, and (iv) operability under mild environmental conditions (Wan and Li, 2012). Despite its remarkable advantages, biological treatment continues to suffer from certain limitations, mainly its insufficient hydrolysis rate for industrial utilization in comparison with other methods (Agbor et al., 2011, Wan and Li, 2012).

The following sections focus on biological treatment and its current state, as well as its advantages and limitations, and how to overcome these limitations.



### **1.2.3 White-rot and brown-rot fungi**

Wood-rotting fungi are classified into white- and brown-rot fungi according to the effect caused to wood during decay (Rytioja et al., 2014). White-rot fungi are the only known organisms that can efficiently decompose lignin and polysaccharides (cellulose and hemicellulose) in the plant biomass. As a result of degradation by white-rot fungi, the residual wood becomes white in colour (Hatakka and Hammel, 2010). White-rot fungi represent around 90% of all identified wood-rotting basidiomycetes and are primarily found naturally occurring on angiosperm wood species (Rytioja et al., 2014). Normally, the (S) type lignin units are degraded to a greater extent than the other lignin type (G), as the latter is more resistant to degradation (Hatakka and Hammel, 2010). In contrast to white-rot fungi, brown-rot fungi occur mainly on gymnosperm wood species and represent around 7% of all identified species of wood decay basidiomycetes. Brown-rot fungi can efficiently degrade cellulose and hemicellulose, but can only alter the structure of lignin. Thus, the residual product after the decomposition of wood has a brownish colour (Hatakka and Hammel, 2010, Rytioja et al., 2014).

### **1.3 Lignocellulose-degrading enzymes**

The best-known lignin-degrading enzymes are manganese peroxidase (MnP), lignin peroxidase (LiP), versatile peroxidase (VP), laccase, and dye decolorizing peroxidase (DyP) (Chen and Wan, 2017). Ligninolytic peroxidase enzymes such as MnP, LiP, and VP, along with DyP enzymes, utilize hydrogen peroxide to catalyse the oxidation of aromatic units, whereas laccase enzymes utilize oxygen to catalyse the oxidation of aromatic units. Over the past several years, there has been growing interest in the development of fungal biocatalysts for the large-scale degradation of lignin. However, due to the challenges of expressing proteins from fungi, this interest has shifted to the identification of bacterial biocatalysts to degrade lignin (Bugg and Rahmanpour, 2015).

### 1.3.1 Laccases

Laccases are copper-containing oxidases that utilize the oxygen molecule from air to catalyse the oxidation of a wide range of phenolic and aromatic compounds (Kunamneni et al., 2008). For substrates that are not oxidized by laccases, another mechanism can be used by these enzymes, which involves the oxidation of redox mediators to form radicals that are able to oxidize substrates not naturally oxidized by the enzyme. Laccases are produced by the majority of white-rot fungi, and are also produced in other fungi, bacteria, and in higher plants. The best-characterized laccases are those of fungal origin. Given their potential application in lignin degradation, laccases have been studied extensively since they were first discovered in 1883 (Kunamneni et al., 2008). Main applications of laccases are dye decolorization, biopulping, bioremediation, synthetic chemistry, cosmetics, food processing, and biosensors. Laccases have several physiological functions, including their involvement in (i) the biosynthesis or degradation of lignin, (ii) the production of spore pigments in fungi, and (iii) the metabolism of iron, as well as their role as virulence factors. Laccases of bacterial origin have been successfully expressed in *E. coli*, whereas laccases from fungal origin have been expressed in *Saccharomyces cerevisiae*, *Aspergillus oryzae*, *Aspergillus niger*, *Aspergillus sojae*, *Trichoderma reesei*, *Pichia pastoris*, *Pichia methalonica*, *Yarrowia lipolytica*, *Kluyveromyces lactis*, tobacco, and maize (Kunamneni et al., 2008, Christopher et al., 2014).

### 1.3.2 Peroxidases

Peroxidases are hydrogen peroxide utilizing enzymes that catalyse the oxidation of different organic and inorganic compounds. Peroxidases are found in different organisms such as plants, bacteria, and fungi. Most peroxidases contain haem as a prosthetic group (Conesa et al., 2002). Based on structural similarity, peroxidases are classified into two main superfamilies: the cyclooxygenase superfamily, which includes peroxidases from

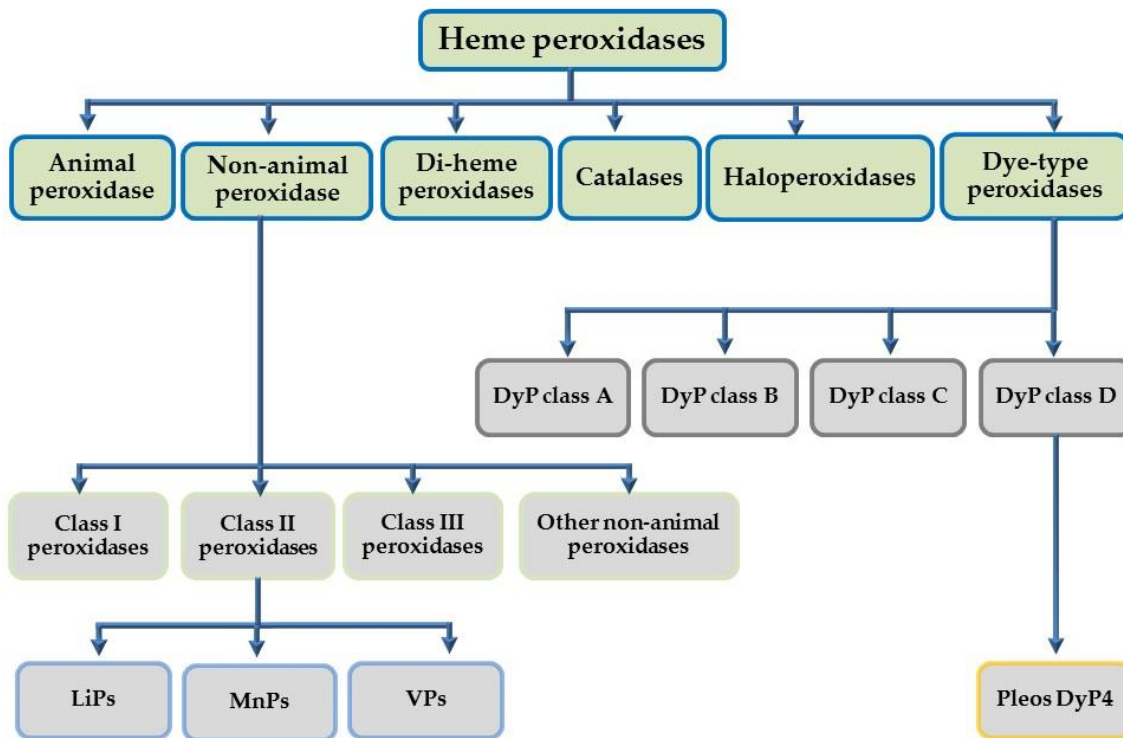
vertebrates, and the peroxidases of plant, bacterial, and fungal origin (Battistuzzi et al., 2010). The vast majority of already-identified peroxidases are haem-containing peroxidases (Zamocky and Obinger, 2010), and are grouped into three main superfamilies and two families: (1) the CDE superfamily, (2) the peroxidase-cyclooxygenase superfamily, and (3) the peroxidase-catalase superfamily (Zamocky et al., 2014). A fourth superfamily of peroxidase-peroxygenases has been described in a recent review (Zamocky et al., 2015).

Haem-containing peroxidases have several potential applications in waste water treatment, bioremediation, synthesis of aromatic chemicals, biological diagnostics, medical biosensors, lignin degradation, and dye decolorizing (Bansal and Kanwar, 2013).

### **Classification**

The CDE superfamily is also known as the peroxidase-chlorite dismutase superfamily. Two families are known to exist within this superfamily, the chlorite dismutases (CIDs) and dye-decolorizing peroxidases (DyPs) (Zamocky et al., 2015). Members of the peroxidase-cyclooxygenase superfamily have previously been known as animal peroxidases, which can be misleading as it continues to be used in public databases. Members of this superfamily exist in all domains of life, and have the Pfam accession number PF03098. The peroxide-catalase superfamily is also known as the plant, fungal, and bacterial peroxidases. This superfamily is the most abundant superfamily in the protein databases, and has an accession number of PF00141 in the Pfam database. Members of this superfamily have been found to exist in the Archaea and Bacteria domains; the Plantae, Fungi, and Metazoan kingdoms; and among many species of Chromista and Protozoa (Zamocky et al., 2015). The peroxide-peroxygenase superfamily is the smallest of the known peroxidase superfamilies due to the fact that its full phylogenomic extent has not yet been established. Unlike other superfamilies, this superfamily shows peroxidase as well as peroxygenase activity, and differs also in that

cysteine is proximal haem ligand rather than histidine. Chloroperoxidases (CPOs) and aromatic peroxygenases (APOs) are the best-characterized members within this superfamily. Other members remain to be analysed and assigned to this superfamily in the Pfam database under accession number PF01328 (Zamocky et al., 2015).



**Figure 1-5:** Schematic representation of the classification of haem peroxidases. A detailed classification of Pleos DyP4 (the enzyme investigated in this project) is given, along with other lignin-degrading peroxidases such as MnPs, LiPs, and VPs.

**Table 1-2:** Host organisms commonly used for recombinant expression of lignin degrading enzymes that are hydrogen peroxide-dependant

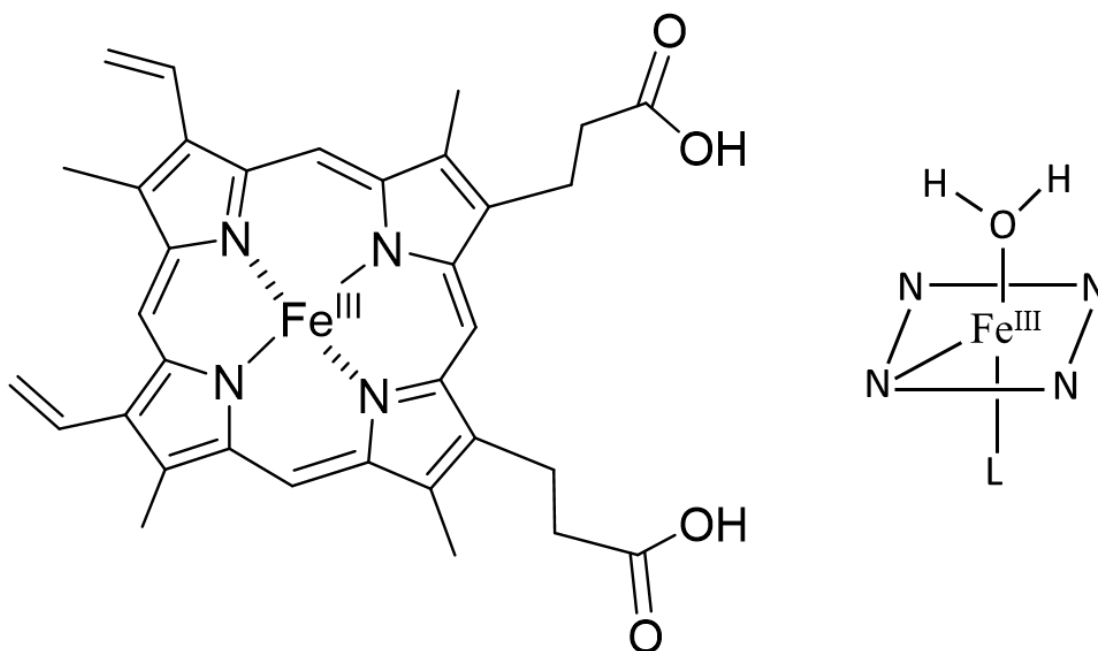
Enzyme	<i>E. coli</i>	<i>S. cerevisiae</i>	<i>A. niger</i>	<i>P. pastoris</i>	<i>A. oryzae</i>
Aromatic peroxygenases (APOs)		+		+	
Laccases		+	+	+	
Dye decolorizing peroxidases (DyPs)	+				+
Manganese peroxidases (MnPs)				+	+
Lignin peroxidases (LiPs)	+		+		
Versatile peroxidases (VPs)		+	+		

### 1.3.2.1 Peroxidases active site structure and general reaction mechanism

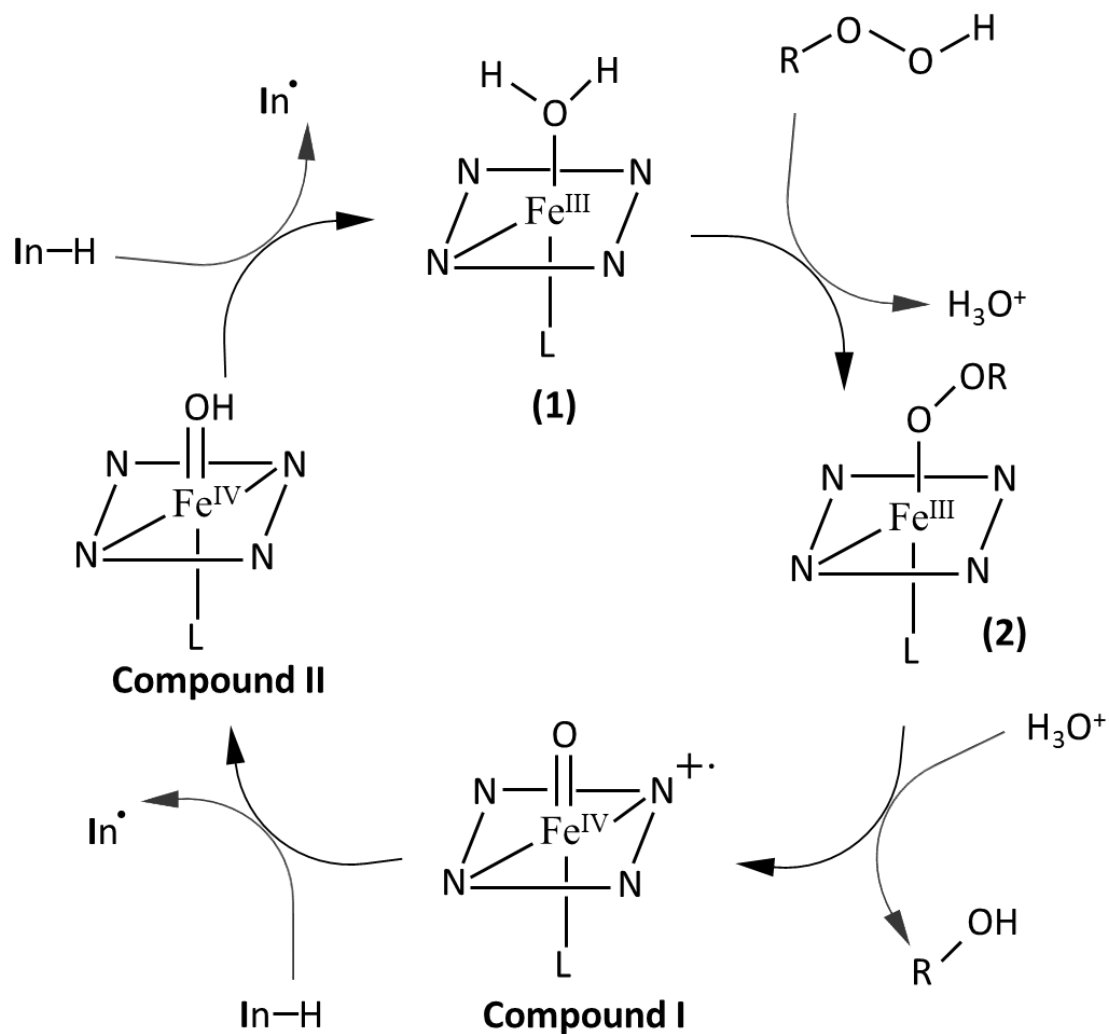
Peroxidases have some common features, including similar overall protein folding and a similar haem pocket architecture. The catalytic cycle of peroxidases is also similar in that they oxidize several substrates to radicals by using H<sub>2</sub>O<sub>2</sub> as a final electron acceptor. This results in the formation of compound I and compound II intermediates (Ogola et al., 2015).

The active site of heme containing peroxidases contains, in its resting state, contains a ferriprotophyrin IX prosthetic group (Figure 1-6 ) (van Rantwijk and Sheldon, 2000). In the majority of heme containing peroxidases, a histidine is the fifth (proximal) ligand of the iron atom and in some other peroxidases it is a cysteine (Hollmann and Arends, 2012).

Water is the sixth ligand and during the reaction mechanism of peroxidases, this ligand is replaced by hydrogen peroxide leading to formation of a peroxo complex (intermediate 2). Compound I is formed - following the O-O bond cleavage - which contains an oxyferryl ( $\text{Fe(IV) = O}$ ) centre and a cation radical located on the heme, which undergoes a one-electron reduction, and the first substrate molecule is oxidized and Compound II is formed. Compound II returns to the resting ground state after oxidation of a second reducing substrate and the release of a second water molecule (figure 1-7) (Hollmann and Arends, 2012, Battistuzzi et al., 2010).



**Figure 1-6:** Resting state of heme peroxidases (prosthetic group with low-spin  $\text{Fe}^{\text{III}}$ ). L constitute a histidine ligand in the majority of heme containing peroxidases and a cysteine ligand in some other heme containing peroxidases. This figure was adapted from (Hollmann and Arends, 2012).



**Figure 1-7:** Overview of the general reaction cycle of peroxidases. (Reducing substrate, In-H) ( $\text{In}^\bullet$ , radical specie). This figure was adapted from (Hollmann and Arends, 2012).

### 1.3.2.2 Lignin peroxidases

Lignin peroxidases (LiPs) are extracellular haem proteins that were firstly described in 1983 in the basidiomycete *Phanerochaete chrysosporium* Burdsall. LiP has been found to exist in a number of species of white-rot basidiomycetes as well as in actinomycetes (Maciel et al., 2010). The catalytic cycle of LiP is similar to that of HRP and of some other peroxidases, in that it utilizes hydrogen peroxide to catalyse the oxidation of different substrates. However, LiP has unique features, including a high redox potential and very low pH optima (Bugg et al., 2011). LiPs are therefore of importance in various industrial processes, given their broad substrate range and high redox potential (Maciel et al., 2010).

### 1.3.2.3 Manganese peroxidases

Manganese peroxidases (MnPs) are haem-containing glycoproteins, the first of which was described in the mid-1980s in *Phanerochaete chrysosporium* fungus. MnP is typically produced in several forms, and in the fungal strain *Ceriporiopsis subvermispora*, around 11 different isoforms have been characterized, all with different isoelectric points (Bugg et al., 2011). MnPs are found in all white-rot fungi, in which it is more prevalent than LiP (Maciel et al., 2010). The main role of MnP is to oxidize  $Mn^{2+}$  to  $Mn^{3+}$ , which in turn oxidizes a number of phenolic substrates (Li et al., 1999). MnP shares 43% of sequence identity with LiP, has a molecular weight ranging from 38 to 62.5 kDa, and is around 350 amino acids residues in length (Plácido and Capareda, 2015). MnP has a similar catalytic cycle to LiP and HRP whereby the native ferric enzyme reacts with hydrogen peroxide, leading to the formation of Compound I, an oxo-ferryl intermediate, followed by the interaction of this compound with  $Mn^{2+}$  to form Compound II and  $Mn^{3+}$  (Bugg et al., 2011). MnP has potential applications in bioremediation, dye decolorization, pulp bleaching, biochemical pulping, and the production of beneficial chemicals from residual lignin in biorefineries and pulp and paper side-streams. However, certain factors



limit the applications of MnP, including those concerning the native production of the enzyme such as low productivity and slow growth. The lack of an efficient recombinant production process is another such factor (Jarvinen et al., 2012).

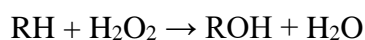
#### **1.3.2.4 Versatile peroxidases**

Versatile peroxidases (VPs), the third type of ligninolytic peroxidase, are glycoproteins that combine the typical properties of lignin and manganese peroxidases and microbial/plant peroxidases. Interestingly, VPs oxidize Mn(II) as well as phenolic and nonphenolic aromatic compounds. VP was discovered in 1999 in members of the *Pleurotus* genus, such as *P. ostreatus*, and *P. eryngii*. Other VPs were subsequently found to exist in basidiomycetes such as *B. adusta*, *Bjerkandera* sp. strain BOS55, *B. fumosa*, and *Bjerkandera* sp. B33/3 (Dashtban et al., 2010). VPs from *Pleurotus* were first described as MnP enzymes due to their Mn-oxidizing activity, but later designated as a new type of peroxidase. Similar to MnP, VP oxidizes Mn<sup>2+</sup>, and, similarly to LiP, also oxidizes high redox potential aromatic compounds (Maciel et al., 2010).

Among peroxidases from basidiomycetes, much attention has been paid to VPs because of their catalytic versatility and unique properties that distinguish them from other peroxidases. The main disadvantage of VPs in commercial applications is similar to that for MnP, in that it is not available in large quantities (Dashtban et al., 2010). This issue can be resolved, however, by the use of DNA recombinant technology. Attempts to produce VPs using heterologous expression systems have also been successful, as have the engineering of LiP or MnP to generate novel peroxidases with features similar to those of naturally occurring VPs (Dashtban et al., 2010). Over the past few years, VPs have been engineered in *S. cerevisiae* using direct evolution methods, resulting in improved functional expression, activity at alkaline pH, and oxidative stability (Gonzalez-Perez and Alcalde, 2018).

### 1.3.2.5 Aromatic peroxygenases

More than a decade ago, novel peroxide-consuming enzymes termed aromatic peroxygenases (APOs) were discovered in *Agrocybe aegerita*. APOs are enzymes that, when fuelled by catalytic amounts of hydrogen peroxide, catalyse the oxidation of aromatic substrates (Molina-Espeja et al., 2014). The characteristic spectroscopic properties of this novel enzyme indicate that it is a haem-thiolate protein (Pecyna et al., 2009), which has been found to be a true peroxygenase that can effectively transfer oxygen from peroxide to a number of different organic substrates such as aromatic molecules (Gutierrez et al., 2011). The following equation shows how this enzyme catalyses the oxidation of its substrates:



The biological function of APOs remains unclear; however, it may be involved in the detoxification of plant components as well as in the transformation of lignin and humus (Molina-Espeja et al., 2015). APO has a similar function to that of classic haem peroxidase, acting as catalyst to oxidize ABTS and phenolic substrates. It also has a similar function to cytochrome P450 enzymes and chloroperoxidase (CPO), acting as a selective monooxygenase for a number of substrates (Poraj-Kobielska et al., 2012). In comparison with cytochrome P450s, APOs have a significant advantage over cytochrome P450s, including their independence from expensive compounds such as NAD(P)H as an electron donor. For APOs to function, only H<sub>2</sub>O<sub>2</sub> is required (Piontek et al., 2010). The sequence identity between AaeAPO and CPO is only 30%, but this proportion is larger than that of other haem enzymes such as cytochrome P450 and LiP (Piontek et al., 2013).

The utilization of aromatic peroxygenases on an industrial scale remains limited due to a lack of suitable expression hosts (Molina-Espeja et al., 2014). Protein engineering of peroxygenases was performed in *Saccharomyces cerevisiae* and *Pichia pastoris*, and

resulted in increased expression and enhanced characteristics of the enzyme (Molina-Espeja et al., 2014).

### **1.3.2.6 Dye decolorizing peroxidases**

Dye decolorizing peroxidases (DyPs) are newly discovered haem peroxidases first found in fungi. DyPs are named after their mechanism of degradation of different dyes. Later, further members were found to exist in other fungi, as well as in bacteria, indicating a natural occurrence similar to that of peroxidases. DyPs are mostly found in bacteria, with only few DyPs exist in fungi and higher eukaryotes (Colpa et al., 2014). DyPs have potential biotechnological applications since they can catalyse the transformation of different substrates. DyPs effectively decolorize industrial dyes, and can degrade lignin by oxidizing  $Mn^{2+}$  or phenolic and non-phenolic structures. Unlike other classical haem peroxidases, DyPs retain their activity over different temperatures and pH ranges. This characteristic, alongside their ease of expression in *E. coli* and other bacteria, may favour the utilization of DyPs over other peroxidases (Singh and Eltis, 2015).

#### **Dye decolorizing peroxidase 4 (DyP4) from *Pleurotus ostreatus***

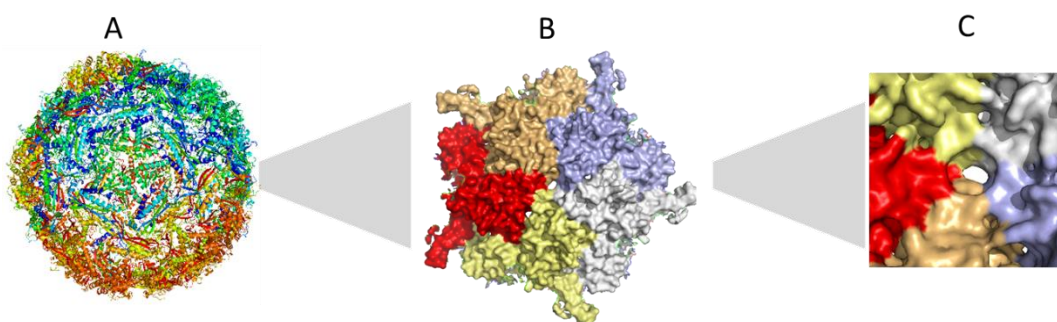
DyP4 is a recently discovered enzyme in *Pleurotus ostreatus*; the second edible mushroom in the world . DyP4 belongs to the DyPs superfamily and shows interesting features as a fungal DyP. Among the unique characteristics of DyP4 are its capability to oxidize  $Mn^{2+}$  to  $Mn^{3+}$ , a characteristic that was assigned to MnP and VP for peroxidases from fungi. Moreover, DyP4 is among the most thermostable DyP described to date, and it also has high pH stability (Fernandez-Fueyo et al., 2015).

Due to its unique characteristics regarding pH and thermal stability, as well as its ability to catalyse a number of lignin model substrates, mediators, and some synthetic industrial dyes, Pleos DyP4 was selected for investigation in this project. Attempts to engineer

DyP4 using protein engineering methods for improved characteristics are described in the experimental chapters (Chapter 4 and Chapter 6).

### Encapsulins

Encapsulins are a type of protein nanocompartment which have been recently discovered in bacteria and archaea (Giessen, 2016), and have a potential applications as a novel delivery mechanism. Studies on the structure of encapsulin revealed that there are multiple openings in its thin shell (Figure 1-7C). These pores might serve to control the exchange of small molecules between the encapsulin lumen and cytosol. It has been suggested that encapsulins that can encapsulate other cargo proteins might have the same openings based on sequence similarity between the encapsulin shell proteins. There are three main types of openings in all encapsulins, some containing positively charged amino acids while some form a negatively charged channel. The diameter of the holes of all encapsulins ranges from 5 to 6 Å and the chemical structure differs between the holes. The holes might act to allow small molecules to enter the shell of the encapsulin nanocompartment while blocking larger molecules (Nichols et al., 2017, Sutter et al., 2008).



**Figure 1-8:** 3D model structure of ReEncapsulin showing (A) the whole carton structure of the encapsulin oligomer (60 monomers), (B) the interaction of five monomers of encapsulin, and (C) one type of opening that exists in encapsulin nanocompartment (5-6 Å). This figure was generated with Pymol and using (PDB 3dkt.1) as a template model.

Enzyme packaging into the shell of the encapsulin nanocompartment could protect the enzyme under extreme conditions such as high temperature or low/high pH values. Encapsulated enzymes are still able to function as small molecule substrates and can enter through the pores of the encapsulin nanocompartment (Sutter et al., 2008).

Cargo proteins have been found to have a C-terminal sequence motif that is responsible for directing proteins into the encapsulin nonocompartment. Examples of cargo proteins that have a conserved C-terminal include DyP, FLP, ruberythrin, and hemerythrin. This C-terminal was later named the cargo loading peptide (CLP), and sufficient evidence exists to show its involvement in the loading of cargo protein into the encapsulin nanocompartment. Tagging the C-terminal of heterologous proteins such as luciferase or fluorescent protein was found to lead to encapsulation. In contrast, the deletion of CLP from a cargo protein prevents encapsulation (Nichols et al., 2017).

## **1.4 Potential application of peroxidases**

### **1.4.1 Bioremediation**

The release of pollutants into the environment is a serious global issue. Air, soil, and water are being contaminated mainly due to the use of toxic chemicals such as pesticides in agriculture or as a result of industrialization (Adenipekun et al., 2012). Due to the nature and characteristics of these xenobiotics, degradation by indigenous flora and fauna is not easily achieved (Adenipekun et al., 2012). Some of the best-known chemical pollutants that have serious toxic effects on human health as well as in the environment include polycyclic aromatic hydrocarbons (PAHs), polychlorinated dibenzo-*p*-dioxin (PCDDs), and coplanar polychlorinated biphenyle (PCBs) (Sakaki et al., 2013, Adenipekun et al., 2012).

#### **Polycyclic aromatic hydrocarbons**

PAHs are major pollutants present in soil, air, and sediments. The major concern regarding these pollutants is their potential mutagenicity, carcinogenicity, and toxicity. Examples of PAHs include pyrene, chrysene, naphthalene, acenaphthalene, phenanthrene, fluoranthene, and anthracene (Mrozik et al., 2003).

#### **Polychlorinated dibenzodioxins**

PCDDs are recognized as environmental contaminants having extreme toxicity. The biodegradation of PCDDs occurs naturally via microorganisms such as bacteria; however, the degradation of 2,3,7,8-tetraCDD, the most toxic dioxin, tends to be inadequate for practical applications. For bioremediation purposes, the discovery or evolution of toxic metabolizing enzymes is a promising approach (Kasai et al., 2010).

#### **Coplanar polychlorinated biphenyle**

PCBs are xenobiotic chlorinated aromatic compounds used in industrial applications such as dielectric fluids, lubricants, and plasticizers. PCBs were widely synthesized in the

1920s due to their excellent thermal and electrical properties. In 1979, PCBs became restricted in many countries due to their toxicity and persistence in the environment (Aken et al., 2010). Currently, the most widely effective practice used for their remediation is incineration. Given the high cost of this practice in terms of money or energy requirements, however, the development of alternative and cost-effective remediation methods is essential (Adenipekun et al., 2012). In comparison to traditional methods used to eliminate pollutants, bioremediation is a generally safer, more cost effective, and less disruptive treatment (Alcalde et al., 2006).

**Bioremediation** can be described as the use of plants or microorganisms such as fungi, bacteria, and algae or their degradative enzymes to remove pollutants from the environment. In the of process bioremediation, enzymes from these microorganisms are used to convert pollutants to innocuous products (Karigar and Rao, 2011). In microbial biodegradation, whole microorganisms are used to clean up pollution (Alcalde et al., 2006). In this approach, fungi, yeast, or bacteria are stimulated to grow in a particular polluted area to perform the desired activities and clean up contamination (Singh et al., 2014). One of the major drawbacks of microbial bioremediation is that microbes must be able to grow in severe adverse conditions that differ significantly from those in the laboratory. Another drawback is that the introduction of genetically modified organisms (GMOs) into a given ecosystem might cause alterations in the ecosystem, and is therefore strictly regulated (Alcalde et al., 2006). These factors, among others, limit the use of microbial bioremediation. Enzymatic bioremediation has therefore become an alternative and interesting approach. In contrast to microbial bioremediation systems, enzymatic bioremediation is less complex (Alcalde et al., 2006), and has many advantages from an environmental point of view compared with the use of microorganisms or chemicals. For example, enzymes can be produced on a high scale with improved activity and stability at a lower cost due to the use of recombinant DNA technology. Furthermore, chemical

and microbiological processes can generate products with toxic side effects, while enzymes are typically digested after treatment by their producers (Alcalde et al., 2006).

### **Oxidation of PAHs**

Enzymes of the CYP superfamily that metabolize PAHs include CYP101, CYP102, CYP1A1, CYP1A2, and CYP1B1 (Kumar, 2010). Peroxidases such as lignin peroxidase and manganese peroxidase can also oxidize PAHs (Bansal and Kanwar, 2013). Manganese peroxidase MnP from *Bjerkandera* sp. have been found to degrade three PAHs, namely pyrene, anthracene, and dibenzothiophene (Yadav and Yadav, 2015). Aromatic peroxygenases can alter the chemical characteristics of PAHs through oxyfunctionalization and thus eliminate their physiological and ecotoxicological effects (Piontek et al., 2013). APO of *Agrocybe aegerita* was found to perform the following activities: *N*-oxidation of pyridine, hydroxylation of the ring and side-chain of toluene, and hydroxylation and sulfoxidation of dibenzothiophene and, selectively, epoxidate naphthalene (Aranda et al., 2010).

### **Oxidation of PCDDs**

The insertion of a single oxygen atom into PCDDs results in the formation of epoxide, and is recognized as the initial reaction in the metabolism of PCDDs by CYP. Two CYP enzymes, CYP1A1 and CYP1A2, play a significant role in the metabolism of PCDDs (Inui et al., 2014). Since several species of white-rot fungi are shown to degrade PCDD, it is suggested that LiP and MnP are involved in this degradation process (Bansal and Kanwar, 2013). The extracellular lignin peroxidase from *P. chrysosporium* is capable of oxidizing CPDDs (Sakaki and Munetsuna, 2010).

### **Oxidation of PCBs**

Peroxidases that degrade PCBs include extracellular LiP and MnP secreted by fungi such as *Pleurotus ostreatus*, *Coriolopsis polyzona*, *Trametes veriscolor*, and *Phanerochoete chrysosporium*.



**Chlorophenols** are another highly distributed pollutant having a significant impact on human health. Pentachlorophenol (PCP) and tetrachlorophenol (TCP) are commonly used as insecticides, herbicides, and fungicides, and in the synthesis of other pesticides. PCP can cause DNA adducts and, potentially, carcinogenesis. CPO is the only peroxidase found to transform halogenated alinins, and HRP, VP, LiP, along with CPO, have been found to transform halogenated phenols (Longoria et al., 2008).

#### **1.4.2 Odour pollution**

Odour pollution is a major environmental concern arising from the production of large-scale livestock. Current methods applied to the elimination of odour from livestock manure are neither efficient nor cost effective (Yan et al., 2016). Peroxidases have a potential application in the deodorization of swine manure. HRP, for example, has been shown to effectively deodorize manure (Ye et al., 2009), offering an attractive alternative to other expensive approaches such as dietary management and intense aeration (Hamid and Khalil-ur-Rehman, 2009). Recently, a study of LiP combined with one of three peroxides was conducted to examine its efficiency in manure deodorization. Compared with a similar study of HRP, LiP offered a simple, low-cost, and feasible method to combat odour pollution, and its electron acceptors are environmentally safe and inexpensive (Yan et al., 2016).

#### **1.4.3 Bioremediation of azo dyes**

Azo dyes represent about 50% of all dyes synthesized for utilization in the cosmetic, paper, printing, food, leather, and textile industries. Each year, approximately 50,000 tons of textile dyes are released into the environment, causing a direct threat of toxicity and mutagenicity to living organisms. The treatment and degradation of azo dyes is preferably attributed to fungal systems (Singh et al., 2015). One peroxidase shown to degrade azo

dyes is CPO from *Caldariomyces fumago* (Zhang et al., 2012). Horseradish peroxidase (HRP) is also known to degrade azo dyes (de Souza et al., 2007).

Anthraquinone dyes are another important textile dye. Anthraquinone and Azo dyes have a variety of colours, and typically used in dyeing polyamide fibres, cellulosic fabric, and wool. Triphenylmethane dyes are another group of dyes that includes crystal violet, malachite green, and pararosaniline. The presence of chromogens is a characteristic of these dyes (Ogola et al., 2015). Triphenylmethane dyes are mostly used to dye nylon, wool, polyacrylonitrile-modified nylon, cotton, and silk. A major limitation of these dyes is their accumulation as recalcitrant compounds, since the majority of them are stable against biodegradation, temperature, and light. DyPs are potential biocatalysts and have received significant attention owing to their effective oxidation of AQ dyes, a feature not observed among other peroxidases such as HRP (Ogola et al., 2015).

#### **1.4.4 Limitations of DyP4 for industrial-scale utilization**

Similar to other haem peroxidases and laccases, DyPs are generally not well-suited or applicable to large-scale industrial applications. Despite the potential of DyPs to be used as biocatalysts in different fields, their applications are limited by a number of factors. Among these is their inactivation in the presence of excess amounts of hydrogen peroxide, similar to other haem peroxidases. The inactivation of haem peroxidases occurs due to the formation of compound III species that causes haem bleaching and irreversible inactivation. Moreover, inactivation might be due to the modification of a protein through the oxidation of specific residues such as methionine, tryptophan, cysteine, histidine, or tyrosine. Using protein engineering approaches, residues which are not stable against oxidation by hydrogen peroxide are replaced with those that are more resistant, and this approach has been successfully employed to increase the stability of different haem-peroxidases against hydrogen peroxide (Ogola et al., 2010).

DyPs are generally only active at low pH and this has limited their potential applications. It is strongly recommended to alter their optima pH using protein engineering methods in order to broaden their applicability. This can be achieved by the construction of DyP variants with an active site is similar to that of plant peroxidases known to be highly active at a neutral pH (Colpa et al., 2014). While this approach did not succeed in a recent study using oligonucleotide-directed mutagenesis, another method that used a directed evolution method was successful in increasing the pH optima from pH 4.5 to pH 8.5 (Brissos et al., 2017, Singh et al., 2012).

Protein engineering methods such as directed evolution and/or rational design have been used to effectively tailor haem-containing peroxidases and enhance their properties (Table 1-2), including stability, thermostability, pH stability, and catalytic efficiency, since naturally occurring enzymes are not optimized for industrial use (Lorenz and Eck, 2005).

**Table 1-3:** Significant factors to consider for a biocatalyst in industrial applications. Information in this table is adapted from (Lorenz and Eck, 2005).

Activity	Efficiency	Stability	Specificity
Turnover frequency (kcat)	Space-time yield	pH	Substrate range
pH profile	Product inhibition	Temperature	Substrate specificity ( $K_m$ and $K_{cat}/K_m$ )
Temperature profile	Production yield	Solvent	Regioselectivity
Specific activity	By-product inhibition	By-product	Conversion yield

## **1.5 Protein engineering**

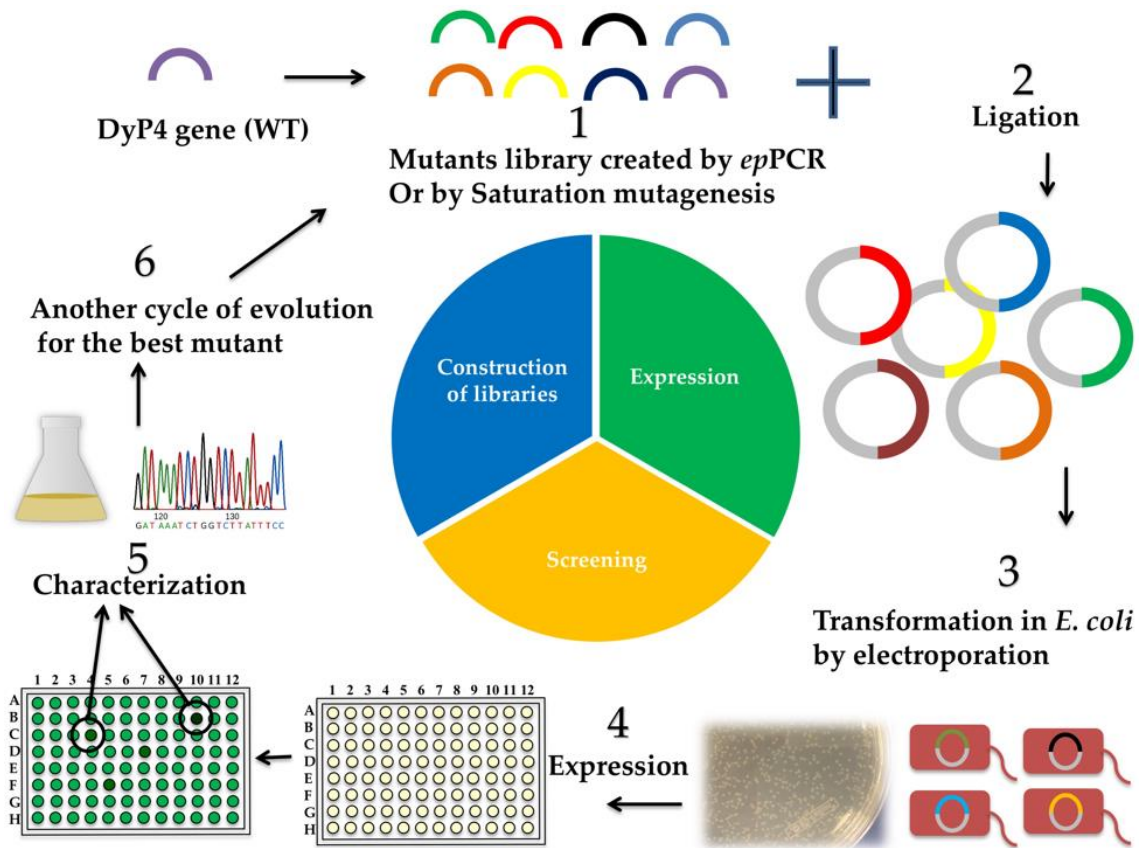
Protein engineering can be described as the art of modifying existing proteins or the design of novel proteins to obtain enhanced proteins suitable for industrial utilization. Due to advancements in recombinant DNA technology and techniques in molecular biology, the alteration of protein properties, structure, and sequence is now possible. Modification of the amino acid sequence of a protein of interest is done via nucleotide insertion, substitution, or deletion in the coding gene (Galzie, 1991). Initially, when protein engineering methods were first used to alter proteins, the focus was on a rational design approach. This restricted protein engineering to proteins whose structural information was well characterized. Later, a second method was developed in which large libraries were constructed and screened (Woodley, 2013).

Two main methods are routinely used in protein engineering studies, rational design and directed evolution. In the rational design approach, the availability of data on protein structure or function is essential to modify a specific amino acid of a protein of interest. In contrast, the directed evolution approach does not require prior knowledge about the structure or function of the protein. Here, the gene coding a protein of interest undergoes several rounds of mutations followed by screening and selection for enhanced phenotypes of the protein (Chen, 2001).

### **1.5.1 Directed evolution**

Genetic variation in DNA can occur naturally from several causes such as reactive chemicals or ultraviolet light, which can damage DNA. Genetic variation can also occur during genome duplication when errors can be introduced. It is possible, however, to create random genetic variation in modern laboratories using polymerase chain reaction (PCR) methods. Methods of directed evolution include error-prone PCR (*epPCR*),

saturation mutagenesis, and DNA-shuffling. Of all of these methods, *ep*PCR is the most widely used in directed evolution studies due to its simplicity (Leemhuis et al., 2009).



**Figure 1-9:** Overview of directed evolution cycles in *E. coli*. Gene of interest undergoes cycles of gene mutation using a suitable mutagenesis method (1), these genes are ligated into an expression vector (2), then transformed into a suitable expression host e.g *E. coli* (3), in the next step, mutants are expressed (4), and screened with a suitable colorimetric assay (5), and the desired mutants selected to parent the next generation (6), the cycle is repeated until the desired property is obtained.

### **Error-prone PCR**

In *ep*PCR, a DNA polymerase amplifies the target gene in an error-prone manner to allow the random incorporation of mutations. The fidelity of DNA amplification can be decreased by changing the normal optimal conditions used routinely in PCR amplification. Some of the methods used to decrease the fidelity of DNA amplification include (i) the use of *Taq* DNA polymerase, (ii) the use of unequal concentrations of dNTPs, and (iii) the addition of chemicals such as  $MnCl_2$  (Tee and Wong, 2014). High concentrations of  $MgCl_2$  ions are also added to the reaction in *ep*PCR to lower the fidelity of *Taq* DNA polymerase during DNA synthesis. The addition of  $MnCl_2$  to the reaction also causes the incorporation of errors (Mohan et al., 2011).

### **Saturation mutagenesis**

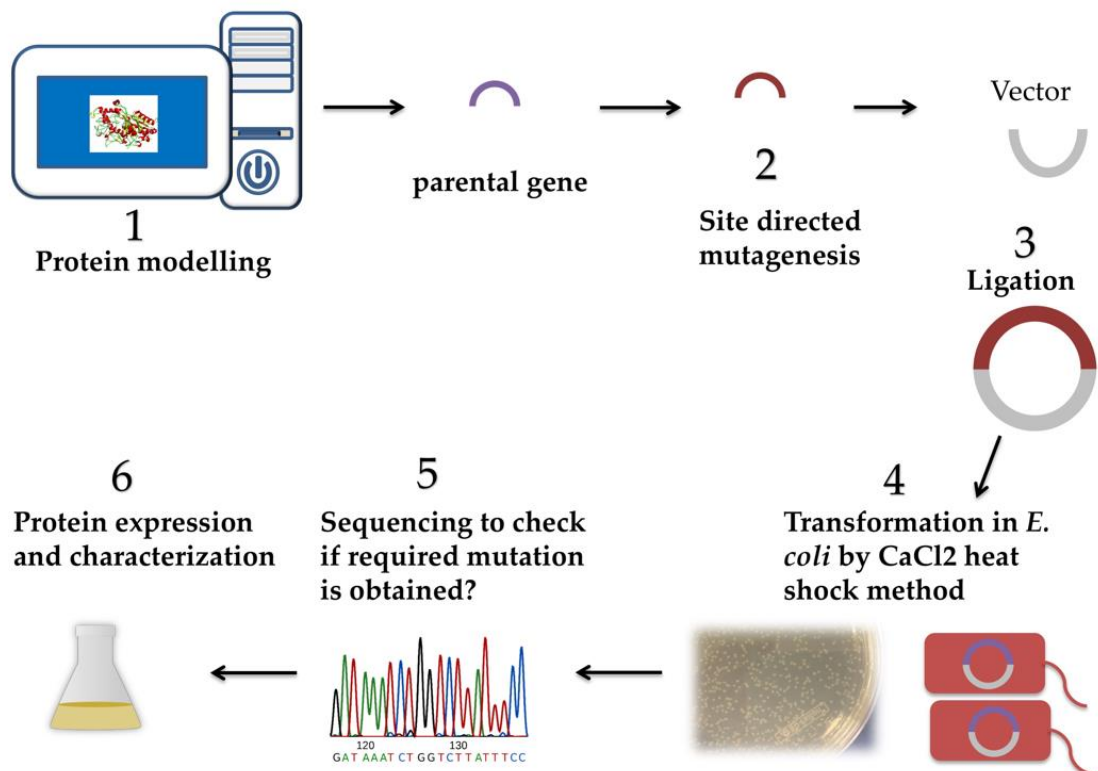
Another technique used in directed evolution is saturation mutagenesis. In this method, knowledge is required about the important functional part of the protein. Variants generated by saturation mutagenesis are superior to those generated by *ep*PCR (Leemhuis et al., 2009). One advantage of saturation mutagenesis is the reduction of library size compared with *ep*PCR techniques (Nannemann et al., 2011). However, this method depends heavily on structural knowledge to determine the specific region in enzymes for targeted saturation mutagenesis (Leemhuis et al., 2009).

### **Gene shuffling**

Gene shuffling is a commonly used technique in the directed evolution approach. In this method, which was originally described by Stemmer, a gene is fragmented with DNase and the random fragments prime each other in a PCR reaction with no external addition of primers. Monticello and colleagues have developed a related method in which they used DNase-mediated fragmentation, although the reassembly method is different (Packer and Liu, 2015). Stemmer's seminal process may have inspired many other researchers to develop protocols for gene recombination (Nannemann et al., 2011).

### 1.5.2 Rational design

In contrast to the directed evolution approach, detailed structural data are essential in rational design. Analysis of the biochemical data and 3D structure of the protein of interest is conducted to identify the target amino acid to mutate. There are a number of online programs which can be used to design primers for PCR to introduce a single mutation to the target amino acid (Steiner and Schwab, 2012). Rational design is not as laborious as directed evolution since only one or a few mutants are screened after the rational construction of mutant(s). The design of a high-throughput screening assay is not required in the rational design approach, another advantage over directed evolution. The main limitation of this approach is the unavailability of detailed structural knowledge for most of identified proteins (Steiner and Schwab, 2012).



**Figure 1-10:** An overview of rational design process in *E. coli*.

### **1.5.3 Screening**

The development of an enzymatic assay is a key step in protein engineering, regardless of the approach used. The assay must be cost- and time-effective, technically simple, reflect the activity of the enzyme under the desired reaction conditions, and maintain accuracy and sensitivity. These assays are usually based on optical detection (fluorescence, luminescence, and colour).

The screening of libraries to identify the desired mutants may be the most important step in directed evolution. Enzyme activity is usually determined by measuring optical absorbance or fluorescence. For the screening process, colonies are picked and transferred individually to microtiter plates to grow cells in a suitable cultivation media and express the protein. Cell disruption is required to release protein expressed intracellularly for enzymatic activity using an optimized and validated screening assay. Mutants can be screened in microtiter plates using a microtiter plate-reader spectrophotometer. Multiple experiments are then conducted to develop and validate a reliable screening assay that is sensitive enough to detect the improved mutants. The screening method must have low variability of the WT or parental strains to eliminate false positives or negatives. High variability would affect the capture and identification of desired mutants as it can be associated with a high rate of false negatives or positives. It is important to measure the Coefficient of variation (CV) for the WT or parental strain, expressed and screened from a single colony in a microtiter plate, and to ensure that the CV is as low as possible by optimizing the screening conditions (Arnold and Georgiou, 2003).

#### **ABTS assay for peroxidase screening**

The activity of peroxidases and laccases is screened using an ABTS (2,2'-azinobis (3-ethylbenzothiazoline-6-sulfonic acid)) assay in the presence of hydrogen peroxide. Spectrophotometric observation is used to determine the production of green radical cation. The ABTS assay works well as a 96-well plate screening method due to its high



extinction coefficient and reproducibility as well as its high solubility. ABTS has other advantages such as stability at high temperature and low toxicity, and that the green dye produced is relatively stable. It is a highly flexible assay that is suitable for functional selection and screening (Reymond, 2006).

The ABTS assay has been used widely in protein engineering methods to improve the expression or other features of peroxidases. For example, the ABTS assay has been used to improve the functional expression of APO in *S. cerevisiae* (Molina-Espeja et al., 2014) and in *P. pastoris* (Molina-Espeja et al., 2015), and to improve the total activity of laccases expressed in *S. cerevisiae* (Camarero et al., 2012). This assay has also been used to improve the catalytic efficiency of PpDyP from *Pseudomonas putida* MET94 expressed in *E. coli* (Brissos et al., 2017).

## **Project objectives and thesis organization**

- i. **Chapter (2)** Detailed description of all materials and methods used in the project.
- ii. **Chapter (3)** Encapsulation of DyP4 into the encapsulin nanocompartment using a due expression system. In this experiment, DyP4 was tagged with a C-terminal sequence which acts to direct the encapsulation of DyP4 into the encapsulin nanocompartment.
- iii. **Chapter (4)** Enhancement of DyP4 stability against hydrogen peroxide by replacing susceptible residues to oxidation with residues that are more resistant to oxidation. In this experiment, all methionine residues were replaced with leucine or phenylalanine.
- iv. **Chapter (5)** Exploration of the potential of DyP4 to oxidize lignin model substrates such as VA or natural mediators such as acytociryngone, sinapic acid, syringaldehyde, and violuric acid. Another aim of these experiments was to validate a colorimetric HTS assay for the screening of DyP4 mutants created by directed evolution methods.
- v. **Chapter (6)** Engineering of DyP4 using an OsmY-based secretion mechanism to facilitate the directed evolution approach in bacteria. In this experiment, mutants were created by *ep*PCR for three generations, and for a fourth generation using site saturation mutagenesis.
- vi. **Chapter (7)** Conclusion and future work.

## **CHAPTER 2 Materials and methods**

2.1 Materials

2.2 Media preparation

2.3 Molecular cloning methods

2.4 Bacterial transformation

2.5 Molecular cloning of DyP4

2.6 Protein expression

2.7 Analysis of protein expression

2.8 Purification

## 2.1 Materials

All kits used for PCR purification, gel extraction, and plasmid mini prep were obtained from QIAGEN, Omega Bio-Tek, and (NucleoSpin® Gel and PCR Clean-up Kit, Germany), respectively. Restriction enzymes such as *Dpn1*, *EcoRI*, *Nde1*, *Nco1*, *XhoI*, *BamHI*, and T4 DNA ligase were obtained from New England Biolabs. *Taq* DNA Polymerase and Q5 High-Fidelity DNA Polymerase were obtained from New England Biolabs. Bacteria strains of *E. coli* BL21 (DE3), C41 (DE3), and DH5 $\alpha$  were obtained from Novagen. SeaKem LE agarose and ethidium bromide were obtained from Lonza and Merck, respectively.

2,2'-Azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), Remazol Brilliant Blue R (RB19), and hydrogen peroxide solution 30% were obtained from Sigma. Acetosyringone (3',5'-dimethoxy-4'-hydroxyacetophenone), syringaldehyde (3,5-dimethoxy-4-hydroxybenzaldehyde), sinapic acid (3,5-dimethoxy-4-hydroxycinnamic acid), violuric acid monohydrate (2,4,5,6(1H,3H)-pyrimidinetetrone 5-oxime), and veratryl alcohol (3,4-dimethoxybenzyl alcohol) were obtained from Sigma Aldrich.

Reagents for SDS gel electrophoresis including tetramethylethylenediamine (TEMED) and acrylamide bis-acrylamide 30% were obtained from AppliChem and Severn Biotech Ltd., respectively. Tris base, glycine, and SDS were obtained from Fisher, Sigma, and VWR, respectively.

## **2.2 Media preparation**

### **2.2.1 2× TY media**

For routine propagation and cultivation of *E. coli* strains BL21 (DE3), C41 (DE3) and DH5 $\alpha$ , 2× TY media was used. The following components were mixed to prepare 2× TY media: 16 g of tryptone, 10 g of yeast extract, and 5 g of NaCl, and the final volume was adjusted to 1 L with ddH<sub>2</sub>O and autoclaving was performed to sterilize the media.

### **2.2.2 TY AIM media**

TY AIM is used widely to express proteins in lac-derived expression systems in *E. coli*. TY AIM was prepared by dissolving 16 g of tryptone 10 g of yeast extract, 3.3 g of (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 6.8 g of KH<sub>2</sub>PO<sub>4</sub>, 7.1 g of Na<sub>2</sub>HPO<sub>4</sub>, 0.5 g of glucose, 2.0 g of  $\alpha$ -lactose, and 0.15 g of MgSO<sub>4</sub> in deionized water, adjusted to a volume of 1 L and autoclaved.

### **2.2.3 Tryptone yeast extract agar plates**

TYE (TYE) agar plates were prepared by dissolving 10 g of tryptone, 5 g of yeast extract, 4 g of NaCl, and 15 g of agar in deionized water to 1L and autoclaving. After cooling to approximately 50°C, 50  $\mu$ g/ml of kanamycin was added and the media was poured into sterile plates, left at room temperature to solidify, and stored at 4°C until use.

## **2.3 Molecular cloning methods**

### **2.3.1 DNA plasmid isolation**

For the routine propagation and isolation of plasmid constructs, the *E. coli* strain DH5 $\alpha$  was used. One of two kits was routinely used for this purpose, the QIAprep Spin Miniprep Kit (QIAGEN) or the Omega Bio-Tek kit. DH5 $\alpha$  containing the plasmid of interest was grown overnight in 5 ml of 2× TY media containing kanamycin or another appropriate antibiotic. Cells were harvested by centrifugation of 3–5 mL of overnight culture at 10,000 rpm for 2 min. Media was removed and the cell pellets were suspended in 250  $\mu$ L of solution 1 or P1 buffer. Next, 250  $\mu$ L of P2 buffer or solution 11 was added to lyse

cells by gently inverting the tube, with lysis time lasting for no more than 4 minutes. Then, 350  $\mu$ L of N3 or solution 11 was added, the tube was inverted several times, and centrifugation at maximum speed was performed for 10 minutes.

A transfer with caution of the supernatant into the spin column was performed, and followed by centrifugation at 10,000 rpm for two min for the DNA to bind to the column. The solution of the flow-through was removed, and the bound DNA was washed with 750  $\mu$ L of HBC or PE buffer. Two centrifugation steps at 10,000 rpm for 2 min were performed, the flow-through was again removed, and the empty tube was centrifuged for 2 min at 10,000 rpm to remove residual ethanol followed by incubation at 50°C for 3 minutes. Elution buffer (50  $\mu$ L) was added to the column in a 1.5 mL centrifuge tube, and the tube was incubated at room temperature for 5 minutes for high DNA recovery. DNA was finally eluted by centrifugation for 2 minutes at maximum speed, and quantified using a NanoDrop 1000 spectrophotometer.

### **2.3.2 DNA gel extraction (PCR purification)**

QIAquick Gel Extraction Kit (QIAGEN) or a NucleoSpin® Gel and PCR Clean-up Kit (Germany) were used routinely to carry out the DNA gel extraction. The target DNA sample was loaded onto a DNA gel and run for 90 minutes at 100 V. The DNA gel was visualized under a transilluminator and the desired band was excised and transferred to a 50 ml Falcon tube. The required volume of TN1 buffer (200  $\mu$ L of NT1 buffer for each 100 mg agarose gel) was added to the piece of gel, which was then incubated at 50°C for 10 minutes and gently vortexed for 2–3 minutes to allow the gel to dissolve completely. Next, 700  $\mu$ L of the dissolved gel was pipetted to a spin column and centrifuged for 30 seconds at 10,000 g for the DNA to bind to the column. The flow-through was removed, and the remaining dissolved gel was transferred to the column. The column was washed by the addition of 700  $\mu$ L of NT3 buffer and centrifugation at 10,000 g for 60 seconds. This step was performed twice before discarding the flow-through, and centrifugation for

the empty tube was performed at maximum speed for 2 minutes to remove residual ethanol. The column was placed in a 1.5-mL centrifuge tube, and incubation at 70°C was performed for few minutes to ensure removal of residual ethanol. Elution with 30 µL of elution Buffer NE was performed in the following step and the tube was incubated for 5 minutes at room temperature for high recovery of DNA, and then centrifuged at maximum speed for 2 minutes. NanoDrop (Thermo Fisher Scientific) was used to measure the concentration of the eluted DNA.

### **2.3.3 PCR clean-up**

PCR purification was routinely carried out using either a QIAquick PCR Purification Kit (QIAGEN) or a NucleoSpin® Gel and PCR Clean-up Kit (Germany). In the process of PCR purification, 200 µL of NT1 buffer was added to each 100 µL of the PCR mixture, and the solution was mixed by pipetting and transferred to the column to bind DNA. The remainder of the procedure was conducted as for the DNA gel extraction procedure.

### **2.3.4 DNA gel electrophoresis**

Agarose gel was prepared routinely for the analysis of DNA using gel electrophoresis by the dissolve of 0.5 g of agarose in 50 of 1× TBE buffer to prepare 1% of agarose gel. The mixture was heated by microwaving for 1 minute at maximum temperature to dissolve the agarose in the buffer. Heating for few seconds was also performed before allowing the hot mixture to cool at room temperature. After the mixture was cooled, 2 µL of ethidium bromide was added, the agarose was poured into a gel tray with the required well comb in place, and the gel was left at room temperature for 30–40 minutes to solidify. The required volumes of samples were loaded into the wells, and gel electrophoresis was run for 90 minutes. The DNA gel was visualized using a GenoSmart VWR gel documentation system.

## **2.4 Bacterial transformation methods**

### **2.4.1 Calcium chloride method**

#### **Principle**

Calcium chloride ( $\text{CaCl}_2$ ) is a commonly used chemical transformation method to incorporate plasmid DNA into bacterial cells. It is based on the use of positively charged calcium ions ( $\text{Ca}^{+2}$ ) to render bacterial cells permeable to DNA uptake. Positively charged calcium ions bind to negatively charged lipopolysaccharides (LPS) and DNA. Treatment with  $\text{CaCl}_2$  causes the formation of small pores in the cell membrane, allowing DNA to pass into bacterial cell after heat shock.

#### **Protocol**

Bacteria (DH5 $\alpha$ , C41 (DE3), or BL21 (DE3)) were cultured for 16 hrs in 5 ml 2 $\times$  TY media at 37°C with shaking at 250 rpm. Next, 5 ml of 2 $\times$  TY media was inoculated with 100  $\mu\text{L}$  of the overnight culture and grown at 37°C with shaking at 250 rpm for 1 hr and 15 minutes until the desired optical density at  $\text{OD}_{600}$  was reached (0.5–0.6). For each single transformation, a total of 1 ml of the grown culture was used. Centrifugation was applied for 2 minutes at 2800 rpm, followed by pipetting to remove supernatants. Then, suspension of cells in 500  $\mu\text{L}$  of pre-chilled 50 mM  $\text{CaCl}_2$  was performed and followed by centrifugation at maximum speed for two minutes. Supernatants were again removed by a pipette and the cells were suspended in 500  $\mu\text{L}$  of the pre-chilled 50 mM  $\text{CaCl}_2$ , followed by incubation on ice for 10 minutes for transformation with intact plasmid or for 30 minutes for transformation with a ligation mixture. One  $\mu\text{L}$  of isolated plasmid or 5  $\mu\text{L}$  of ligation mixture was added, followed by incubation on ice for 10 minutes or 30 minutes, respectively. To heat shock cells, incubation at 42°C for 1 minute was performed, followed by incubation on ice for an additional 3 minutes. Next, 800  $\mu\text{L}$  of 2 $\times$  TY media (pre-warmed at 37°C) was added and cells were grown for 1 hr at 37°C with shaking at 250 rpm. TYE agar plates supplemented with kanamycin were pre-warmed at 37°C for 1 hr. For transformation with purified plasmids, 100  $\mu\text{L}$  of the



outgrowth was plated on an agar plate, while for transformation with ligation mixtures and for higher efficiency transformations, 1100  $\mu\text{L}$  was removed by centrifugation at 2800 rpm for 2 minutes, and the remaining 100–200  $\mu\text{L}$  was suspended and plated on the pre-warmed agar plates and incubated for 14–16 hrs at 37°C to allow growth of the transformants and formation of bacterial colonies.

#### **2.4.2 Electroporation method**

##### **Principle**

Electroporation is another commonly used transformation method in molecular biology to incorporate plasmid DNA into bacterial cells. In this method, a specific voltage of electrical pulse is applied to a cell suspension for a few microseconds. This leads to the formation of pores in the cell membrane and allows the DNA molecules to pass into cells. Given its significantly higher efficiency than the  $\text{CaCl}_2$  method, electroporation is used when a very high transformation efficiency is needed, for example when creating mutant libraries (Chapter 6).

##### **Protocol**

An overnight culture of BL21 (DE3) was prepared in 5 mL 2 $\times$  TY media and grown at 37°C with shaking at 250 rpm for 15–16 hrs. Then, 50 mL of 2 $\times$  TY media in a 250-mL flask was inoculated with 500  $\mu\text{L}$  of the overnight culture and allowed to grow at 37°C with shaking at 250 rpm for 2 hrs until the desired  $\text{OD}_{600}$  was reached (0.5–0.6). Next, the cell suspension was transferred to a 50-mL Falcon tube and centrifuged at 5000 rpm and 4°C for 10 minutes. Supernatants were removed by pipetting, and cells were re-suspended in sterile pre-chilled Milli-Q water and centrifuged at 8000 rpm and 4°C for 15 minutes. For higher efficiency, this step was performed three times. Cells were concentrated and re-suspended in the appropriate concentration. For each transformation, 40  $\mu\text{L}$  of the concentrated cell suspension was used, and 1  $\mu\text{L}$  of ligation mixture was added, mixed, and electroporated using the appropriate program (P1; 1700 V and 5 ms).

One ml of SOC media was added immediately after electroporation, mixed, and transferred to a sterile 1.5-mL microcentrifuge tube, and then placed at 37°C with shaking at 250 rpm for 1 hr. TYE agar plates were pre-warmed at 37°C for 60 minutes. Centrifugation was performed for the outgrowth for 2 minutes at 2800 rpm, and 800 µL of the supernatant was removed. The remaining supernatant was mixed and plated onto the agar plates supplemented with kanamycin and placed at 37°C to allow growth of the transformants and formation of bacterial colonies.

## 2.5 Molecular cloning of DyP4

### 2.5.1 Cloning of DyP4-tag into pET24a

The GeneScript synthesis service was used for the synthesis of DyP4-tag (DNA sequence, Appendix 3). The target signal peptide sequence (FLDDPPDAPTRLV-PEATFTAPISDSLIGIGSLKRSAQQ) was tagged to the DyP4 C-terminus to direct the encapsulation of DyP4 into the ReEncapsulin nanocompartment (Chapter 3). *EcoRI* and *NdeI* restriction sites were used to insert the DyP4-tag into pET24a for recombinant protein expression in *E. coli*. Sequencing was performed by Eurofins to confirm the insertion of DyP4 into the pET24a vector.

### 2.5.2 Cloning of pET24a-DyP4-without tag

For experiments other than the encapsulation of DyP4 into encapsulin (Chapter 3), DyP4 without tag was used (DNA sequence, Appendix 5). To remove the tag from DyP4, two primers (Table 2-1) were designed to amplify DyP4 without tag from the plasmid (pET24a-DyP4-tag).

**Table 2-1:** Primers used to amplify DyP4 without a tag

Primer	Oligonucleotide sequences (5'->3')
1	Fwd: 5' TATACATATGATGACCACCCGGCGCCGCGCTG 3' Rev: 5' ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC 3'

PCR to amplify DyP4-without tag was performed in a 50  $\mu$ L reaction mixture containing 1 $\times$  Q5 buffer, 0.2 mM of dNTPs, 100 ng of template (pET24a-DyP4-tag), 1  $\mu$ M of forward and reverse primers, and 1 U of Q5 high-fidelity DNA polymerase. Distilled water was added to adjust the final volume to 50  $\mu$ L. PCR with 30 cycles of 98°C for 10 sec, 72°C for 30 sec, and 72°C for 1 min was performed and followed by final extension for 2 min at 72°C before incubation at 8°C until gel electrophoresis. Following PCR and DNA gel visualization to confirm the success of PCR amplification, template digestion with *DpnI* was performed overnight at 37°C followed by PCR clean up, restrictive digestion with *EcoRI* and *NdeI* restriction enzymes overnight, and ligation into a linearized pET24a vector. Restriction analysis was performed to confirm the ligation and insertion of DyP4 into the pET24a vector. Analysis using SDS-PAGE and an ABTS assay was carried out to confirm the expression of DyP4.

## **2.6 Protein expression**

### **2.6.1 Expression of pET24a-DyP4**

Protein expression of DyP4 was carried out in the *E. coli* BL21 (DE3) strain. Transformation with the  $\text{CaCl}_2$  method was performed in BL21 (DE3) using pET24a-DyP4. A single colony from the TYE agar plate was picked in 5 mL 2 $\times$  TY media with 50  $\mu$ M kanamycin and grown overnight at 37°C with shaking at 250 rpm for 15–16 hrs. In a 1-L flask, 400 mL of 2 $\times$  TY media containing 50  $\mu$ M kanamycin was inoculated with 2 mL of overnight culture and placed at 37°C with shaking at 250 rpm for 2 hrs until the  $\text{OD}_{600}$  reached approximately 0.6. Next, induction of IPTG with a final concentration of 1 mM was performed, and expression was initiated at 25°C for 24 hrs. After expression, cells were harvested by centrifugation at 8000 rpm for 10 minutes at 4°C. Cell pellets were stored at  $-20^\circ\text{C}$  until purification or other analysis.

## **2.6.2 Optimization of DyP4 secretion using a pET24a-OsmY-DyP4 vector**

To determine the optimal conditions for the secretion of DyP4 using the OsmY mechanism (Chapter 6), two bacterial *E. coli* strains were used to transform pET24a-OsmY-DyP4: C41 (DE3) and BL21 (DE3). Expression was carried out in 5 mL 2× TY TY AIM with 50 µM kanamycin at two different temperatures, 30°C and 37°C, and protein samples were screened by ABTS assay after 12, 18, and 24 hrs of expression. Analysis by ABTS assay was carried out to determine the optimal conditions for secretion (e.g. strain, temperature, time).

## **2.7 Analysis of protein expression**

### **2.7.1 SDS-PAGE**

#### **Preparation of polyacrylamide gels**

Acrylamide SDS gels (10%) were prepared in two parts, a stacking gel and a resolving gel. For the preparation of resolving gels, 4.1 mL DDI H<sub>2</sub>O, 3.3 mL 30% acrylamide, 2.5 mL 0.5 M Tris-HCl (pH 8.8), and 0.1 mL SDS (10% w/v) were mixed. The addition of 5 µL TEMED and 50 µL 10% APS was performed prior to pouring the gel. A thick layer of DDI H<sub>2</sub>O was added on the top of the resolving gel, and the gel was allowed to solidify for around 40 minutes. The DDI H<sub>2</sub>O layer was removed, the stacking gel was poured on top, and the comb was placed into the assembled gel cassette. The stacking gel was prepared using 4.1 mL DDI H<sub>2</sub>O, 3.3 mL 30% acrylamide, 2.5 mL 0.5 M Tris-HCl (pH 6.8), and 0.1 mL SDS (10% w/v). Prior to pouring the gel, 10 µL TEMED and 50 µL of 10% APS were added.

## **2.7.2 Preparation of protein samples**

### **Soluble protein**

A soluble fraction was prepared by lysis of cell pellets and disruption by mixing with lysomix, followed by incubation on ice for 1 hr and centrifugation at a maximum speed of 10 minutes. Supernatants were transferred to a new 1.5-mL microcentrifuge tube, and cell pellets were retained for the analysis of insoluble protein as described in the following section. An equivalent volume of protein sample (supernatant) was mixed with sample buffer (2× SDS reducing buffer supplemented with  $\beta$ -mercaptoethanol), boiled at 94°C for 5 min to denature the protein, and centrifuged at maximum speed for 2 min. Volumes of 5–10  $\mu$ L) were used for SDS-PAGE analysis.

### **Insoluble protein**

For the determination of insoluble protein, cell pellets were re-suspended in 1× SDS reducing buffer supplemented with  $\beta$ -mercaptoethanol, boiled at 94°C for 5 min, and centrifuged at maximum speed for 2 min. Volumes of 5–10  $\mu$ L were subjected to SDS-PAGE.

### **Electrophoresis**

The electrophoresis chamber was filled with 1× running buffer, and the required amount of protein sample was loaded (5, 10, or 20  $\mu$ L) according to the experiment and type of protein sample. The required amount of the molecular weight marker was loaded into the first lane, and electrophoresis was performed at 200 V for 35 min or until the blue dye reached the bottom of the gel.

### 2.7.3 Activity assays for DyP4

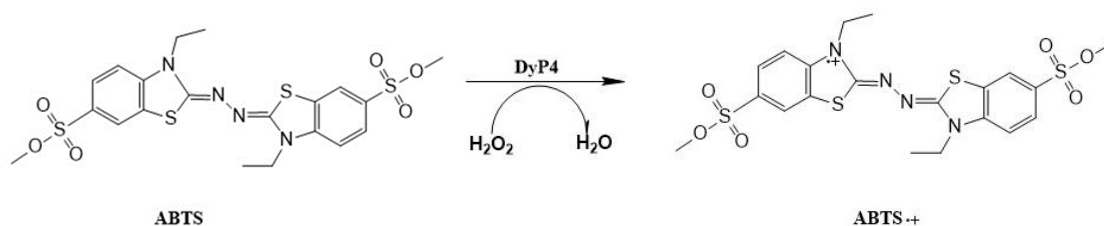
#### ABTS assay

The ABTS assay was used in this project to (i) detect and confirm the expression and activity of DyP4, (ii) test the enzymatic activity of encapsulated DyP4 into encapsulin (Chapter 3), and (iii) screen mutants created by site-directed mutagenesis or by *ep*PCR and saturation mutagenesis for the improvement of stability against higher concentrations of hydrogen peroxide (Chapter 4) or to increase total activity (Chapter 6).

DyP4 activity can be screened colorimetrically with the ABTS assay based on the oxidation of ABTS, 2,2'-azinobis (3-ethylbenzothiazoline-6-sulfonic acid) by the enzyme in the presence of H<sub>2</sub>O<sub>2</sub> (Figure 2-1). Spectrophotometric observation is conducted for the production of green radical cation. ABTS assay is suitable for 96-well plate screening due to its high extinction coefficient and reproducibility as well as its high solubility. ABTS has other advantages such as its stability at high temperature, low toxicity, and relative stability of the green dye produced (Reymond, 2006). It is also applicable for use with peroxidases and oxidases as well as laccases. It is thus a highly adaptable assay that works well for functional selection and screening (Reymond, 2006).

#### Citric acid- buffer-Na<sub>2</sub>HPO<sub>4</sub> buffer

ABTS assay was conducted using a in 0.1 M-citric acid-0.2-Na<sub>2</sub>HPO<sub>4</sub> buffer 3.4 pH (optimal pH for the activity of DyP4 by ABTS assay is approximately pH 3.5, as reported by (Fernandez-Fueyo et al., 2015). The required volume of the buffer was prepared as described in the <https://www.sigmaaldrich.com/life-science/core-bioreagents/biological-buffers/learning-center/buffer-reference-center.html>.



**Figure 2-1:** Reaction scheme for ABTS assay oxidation by DyP4

## 2.8 Purification

Cell pellets were lysed by sonication (10 seconds on, 20 seconds off for 15 minutes, for a total of 5 minutes on; 70% amplitude). Purification of DyP4 was carried out in three steps: two ion-exchange chromatography (IEX) steps using a HiTrap Q column 5 mL, and HiTrap SP column 5 mL followed by size exclusion chromatography using HiLoad 26/600 Superdex 75 pg. All purification steps were performed using an ÄKTA Pure system (GE Healthcare Life Sciences) Germany.

### 2.8.1 Ion-exchange chromatography

IEX has been used widely to purify and separate molecules based on their charge. IEX has been applied successfully to the purification and separation of proteins, and is probably the most widely used chromatography technique for this purpose, mainly because it is simple, easy to control, and has a high capacity and resolving power (Khan, 2012). Anion-exchange and cation-exchange chromatography are two approaches of IEX. In anion-exchange chromatography, a positively charged matrix allows the binding of negatively charged molecules, while in cation-exchange chromatography, a negatively charged matrix allows the binding of positively charged molecules. The first step to consider in the selection of buffers for purification is pH, as this will determine the charge of the protein of interest (Figure 2-3). The isoelectric point (pI) of proteins is initially calculated, and is the pH at which the protein carries no net charge. Online tools such as

ExPASy (<http://ca.expasy.org/tools/protparam.html>) can be used to determine the pI of a protein. Components used to prepare the buffer should not bind to the beads in IEX; for example, acetate can be used to prepare buffer for cation-exchange chromatography but not for anion-exchange chromatography (Duong-Ly and Gabelli, 2014).

#### **HiTrap Q column (anion-exchange chromatography)**

Purification of DyP4 using a HiTrap Q column was the first step of purification. Pellets were lysed by sonication for 5 min using the following program: 10 seconds on, 20 seconds off, 70% amplitude. The lysis, equilibration, and washing buffers consisted of 50 mM Tris-HCl (pH 8.5) and 1 mM EDTA. Protease inhibitors, DNase, RNase, and lysomix were added to the lysis buffer (Table 2-2). The elution buffer consisted of 50 mM Tris-HCl (pH 8.5), 1 mM EDTA, and 1 M NaCl. Crude extracts were centrifuged at 4°C for 40 min at 8000 rpm, filtrated using a Whatman® Puradisc (pore size 0.45 µm) filter, and loaded onto a Q column connected to the AKTA Pure system. Two fractions of the purified protein were combined for the next step of purification.

**Table 2-2:** Composition of the buffer used in the first step of chromatography for DyP4-tag purification (Ion exchange chromatography- Q column).

<b>Buffer: (Lysis, equilibration and</b>		<b>First step</b>
<b>Component</b>	<b>Concentration</b>	<b>Column</b>
Tris-HCl (pH 8.5)	50 mM	Q column
EDTA	1 mM	
<b>Buffer: Elution</b>		
<b>Component</b>	<b>Concentration</b>	
Tris-HCl (pH 8.5)	50 mM	
EDTA	1 mM	
NaCl	1 M	



### HiTrap SP column (cation-exchange chromatography)

Two fractions from the first step (4 mL in total) were combined and mixed in a lysis buffer (Table 2-3), containing 25 mM acetate (pH 4.0), centrifuged, filtrated, and loaded onto the AKTA Pure system for purification. The elution buffer contained 25 mM acetate (pH 4.0) and 1 M NaCl. Two fractions of the purified protein were combined for the next step of purification.

**Table 2-3:** Composition of the buffer used in the second step of chromatography for DyP4-tag purification (Ion exchange chromatography - SP column).

<b>Buffer: (Lysis, equilibration and</b>		Second step
<b>Component</b>	<b>Concentration</b>	<b>Column</b>
acetate (pH 4.0)	25 mM	SP column
<b>Buffer: Elution</b>		
<b>Component</b>	<b>Concentration</b>	
acetate (pH 4.0)	25 mM	
NaCl	1 M	

### Size-exclusion chromatography (gel filtration)

The third and final step of the purification of DyP4 was carried out using a size-exclusion chromatography method with a HiLoad 26/600 Superdex 75 pg column. An equilibration buffer at pH 4 was used for the pre-equilibration of the column with 1.5 column volume, and consisted of 50 mM of  $C_6H_8O_7 \cdot H_2O$ , 50 mM of  $Na_2HPO_4$ , 100 mM NaCl, and 10% glycerol (Table 2-4). Eluted fractions (2 mL) were collected and analysed using SDS-PAGE and ABTS assay. The desired purified fractions were combined, frozen in liquid nitrogen, and stored at  $-80^\circ C$  until use.

**Table 2-4:** Composition of the buffer used in the third step of chromatography for DyP4-tag purification (size exclusion chromatography - Gel filtration).

<b>Buffer: (Lysis, equilibration and</b>		Third step
<b>Component</b>	<b>Concentration</b>	<b>Column</b>
$C_6H_8O_7 \cdot H_2O$	50 mM	Gel filtration
$Na_2HPO_4$	50 mM	
NaCl	100 mM	
Glycerol	10%	

## **CHAPTER 3    Encapsulation of DyP4 into the encapsulin nanocompartment using a dual expression system.**

3.1 Introduction

3.2 Methodology

3.3 Results

3.4 Discussion

3.5 Summary

### **3.1 Introduction**

The recently identified class of protein nanocompartments known as encapsulins are found in bacteria and archaea where they encapsulate ‘cargo proteins’ (Giessen, 2016). Investigations have shown that DyP-type peroxidases or ferritin-like proteins are naturally encapsulated in the encapsulin nanocompartment (Contreras et al., 2014, Giessen and Silver, 2017, He et al., 2016, Rahmanpour and Bugg, 2013). Packaging of enzymes in the encapsulin nanocompartment occurs naturally and is believed to have a specific biological function. Some cargo proteins have a specific C-terminal peptide sequence that is thought to direct their encapsulation into the encapsulin nanocompartment, while others have alternative mechanisms to regulate their encapsulation. What makes encapsulin a highly interesting platform is the possibility of applying this encapsulation mechanism to other non-native proteins. This can be achieved by fusion or by tagging the target protein with a C-terminal peptide sequence. Moreover, the ability of encapsulin nanocompartments to display proteins inside and release them under certain conditions, such treatment at low pH, has tremendous potential for nano-biotechnology applications. Finally, in the deconstruction of lignocellulose, the targeting of a peroxidase into the nanocompartment might also have a specific application (Giessen, 2016, Rahmanpour and Bugg, 2013).

#### **3.1.1 Encapsulin**

Encapsulin is a nanocompartment that exists inside bacterial cells and has potential applications in drug discovery and nanobiotechnology (Tamura et al., 2015). The crystal structure of encapsulin from the hyperthermophil *Thermotoga maritima* has been described (Sutter et al., 2008). Compared with bacterial microcompartments (BMCs), encapsulin nanocompartments have potential feasibility to be used as a nanomaterial since they are significantly smaller and less complex than BMCs. The spherical nanocompartment comprises 60 monomers and has a molecular weight of 31 kDa and

diameter of 21–24 nm (Tamura et al., 2015).

### **Encapsulin from *Rhodococcus erythropolis* N771 (ReEncapsulin)**

In *Rhodococcus erythropolis*, the native system of encapsulin is made up of a capsid protein and a cargo protein (DyP). This DyP has a C-terminal targeting sequence of 37 amino acids that is responsible for directing the peroxidase into the encapsulin nanocompartment. The C-terminal sequence has been used in a study to package two guest proteins into ReEncapsulin. In this study, the C-terminal sequence was genetically fused to the C-terminus of firefly luciferase (Luc) and enhanced green fluorescent protein (eGFP) proteins. Both proteins were successfully encapsulated and remained active (Giessen, 2016).

### **3.1.2 DyP4**

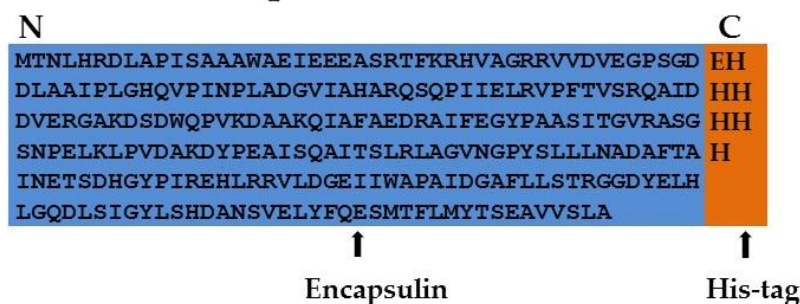
Dye-decolorizing peroxidase (DyP4) from *Pleurotus ostreatus* mushroom was described in 2015 as a DyP that can catalyse the oxidation of a number of lignin model substrates as well as some other important industrial dyes. DyP4 has been found to be one of the most thermostable peroxidases identified to date, and also has high pH stability (Fernandez-Fueyo et al., 2015).

### **3.1.3 Encapsulation**

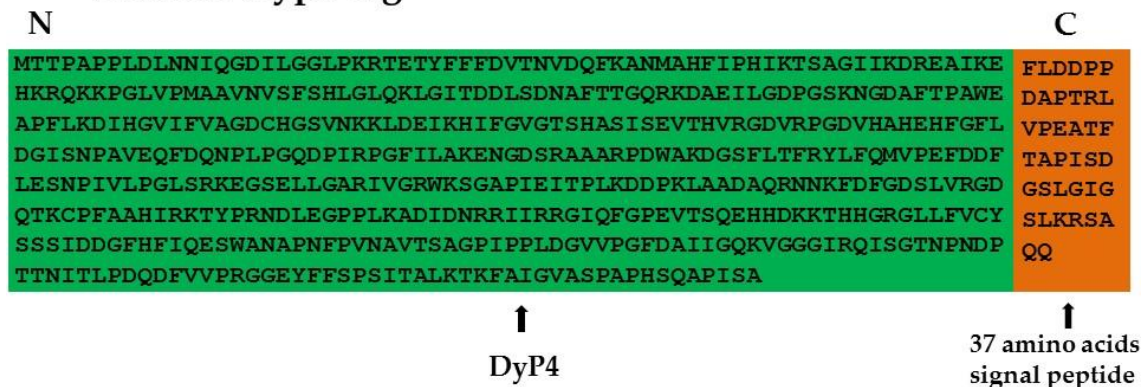
Two mechanisms in which a guest protein can be packaged into the encapsulin nanocompartment have been reported. The first approach is through the disassembly and assembly of encapsulin and a DyP by incubating both proteins at lower pH for nanocompartment disassembly, then increasing the pH to assemble the nanocompartment and form the specie. The second approach is fusion of the C-terminal of the guest protein with the amino acids of signal peptide responsible for directing the encapsulation of the target protein into the encapsulin nanocompartment. Herein, to package the DyP4 enzyme into the ReEncapsulin nanocompartment, the C-terminus of DyP4 was fused with a C-

terminal signal peptide of *R. erythropolis* N771 DypB peroxidase, which has previously been reported to direct the encapsulation of two guest proteins into the ReEncapsulin nanocompartment.

## 1. ReEncapsulin



## 2. Pleos-Dyp4-tag



**Figure 3-1:** A scheme to illustrate the design of (1) *Re-Encapsulin-His-tag* and (2) *Pleos-Dyp4-tag*. (1) *Re-Encapsulin* was tagged with a his-tag at its C-terminal to facilitate the purification process, (2) *Pleos-Dyp4-tag*; *Dyp4* was tagged with a signal peptide sequence (37 aa) in its C-terminal that is responsible for directing the enzyme into the encapsulin nanocompartment.

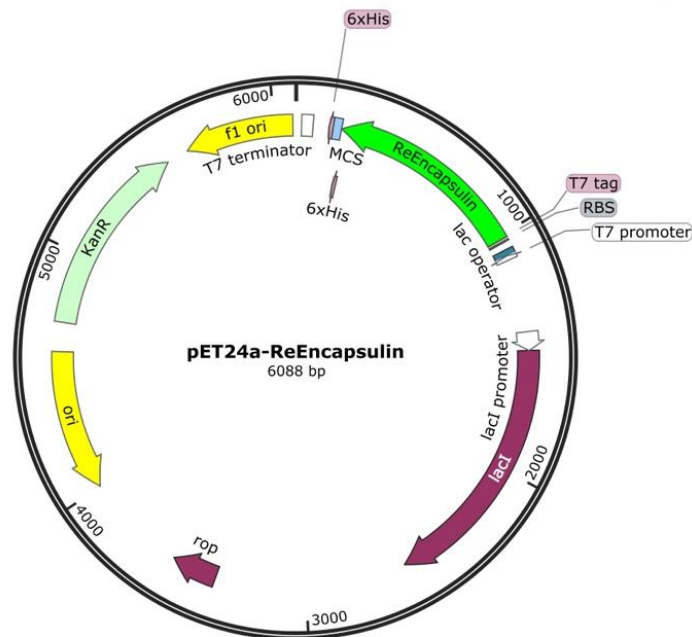
### **3.1.4 Aim of the study**

In this study, the expression of ReEncapsulin and DyP4-tag in BL21 (DE3) is described, along with mechanism of encapsulating DyP4 into the encapsulin nanocompartment. Investigations were also carried out to determine whether (i) fusing DyP4 with a signal peptide on its C-terminal – which would act to direct the encapsulation of DyP4 into the ReEncapsulin nanocompartment – would affect the expression or activity of DyP4, and (ii) the encapsulated DyP4 remains active upon encapsulation, and if so, whether this activity is reduced, increased, or similar to that of the purified DyP4.

## 3.2 Methodology

### 3.2.1 Molecular cloning of pET24a-Encap

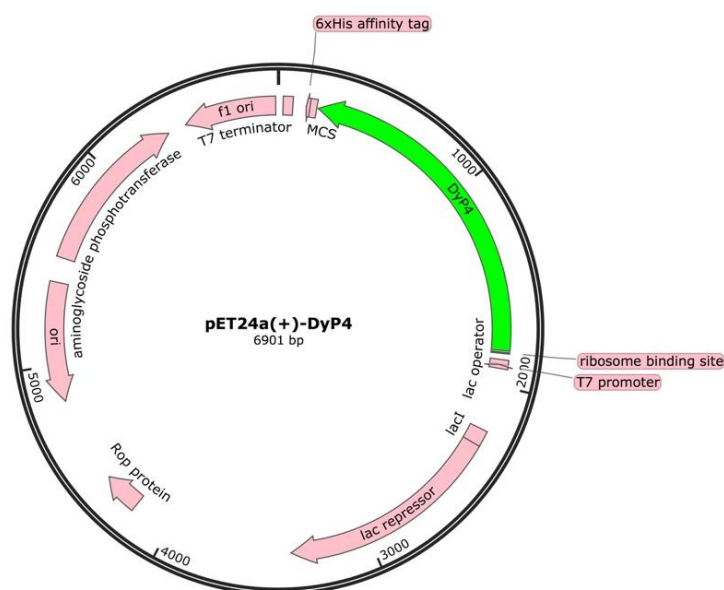
The GeneScript synthesis service was used for the synthesis of ReEncapsulin (DNA sequence, Appendix 4). Basic molecular cloning methods (restrictive digestion and ligation) were used to clone the encapsulin gene from its carrier vector into the expression vector pET24a-Encapsulin. Briefly, double digestion (for carrier vector of encapsulin) with *EcoRI* and *NdeI* restriction enzymes was performed. Gel extraction of the product was performed, followed by ligation with pET24a that had already been double digested with the same restriction enzymes, *EcoRI* and *NdeI*. The insertion of Encap into the pET24a vector was confirmed by sequencing (Eurofins sequencing services).



**Figure 3-2:** Plasmid map of pET24a-*ReEncapsulin* (created with SnapGene). *ReEncapsulin* was inserted into pET24a vector using restriction sites *EcoRI* and *NdeI*.

### 3.2.2 Molecular cloning of pET24a-DyP4-tag

Molecular cloning of DyP4-tag into pET24a is described in Chapter 2 (Materials and methods).



**Figure 3-3:** Plasmid map of pET24a(+)-DyP4 (created with SnapGene). DyP4 was inserted into pET24a vector using restriction sites *EcoRI* and *NdeI*.

### 3.2.3 Transformation in BL21 (DE3) using the CaCl<sub>2</sub> method

Chemical transformations were carried out using a pET24a plasmid harbouring either the DyP4 or encapsulin gene and were performed using the CaCl<sub>2</sub> heat shock method. The procedure is detailed in Chapter 2.

### 3.2.4 Protein expression of pET24a-Encapsulin and pET24a-DyP4-tag

A singly colony from an agar plate of transformed BL21 (DE3) was grown at 37°C for 15 hrs, and 50 µL of overnight culture was then added to 50 mL of 2× TY media containing kanamycin in 250-mL conical flasks and left to grow for 2 hrs until the desired OD was reached (0.5–0.6). Next, 0.1 mM and 1 mM of IPTG was added for pET24a-Encapsulin and pET24a-DyP4-tag, respectively, and protein expression was allowed to proceed at 25°C and 200 rpm for 24 hrs. Two IPTG concentrations were used for both proteins, and the optimal concentration of each was used for all subsequent experiments.



### **3.2.5 Analysis of protein expression**

SDS-PAGE analysis was performed to confirm the expression of DyP4 and encapsulin using 10% polyacrylamide gels. Pellets were suspended in lysomix, incubated on ice, and centrifuged at maximum speed for 10 minutes. For one part of the supernatant (soluble), one part of SDS reducing buffer supplemented with 5%  $\beta$ -mercaptoethanol was added, mixed, boiled, and loaded onto SDS-PAGE gels for analysis. For DyP4, an ABTS assay was also performed to confirm expression as the formation of an intense green colour indicating the oxidation of ABTS by DyP4. Alternatively, expression was detected by UV spectrophotometer to monitor the increase in absorbance at a wavelength of 420 nm.

### **3.2.6 Large-scale recombinant production of DyP4 and encapsulin**

Using the optimal concentration of IPTG, large-scale expression of DyP4 and encapsulin was carried out in BL21 (DE3) cells in 400 mL of 2 $\times$  TY media with kanamycin (50 mg/L) in 1 L conical flasks. Cells were harvested by centrifugation at 6000 rpm and 4°C for 5 minutes, media was removed, and the cell pellets were stored at -20 °C until use.

### **3.2.7 Purification of ReEncapsulin**

The purification of ReEncapsulin required two steps, and after the first step, highly pure protein was obtained, as described in the results section. The first step was performed by Affinity chromatography using a HisTrap column 5 mL (GE Healthcare Life Sciences). In the second step, size-exclusion chromatography using a Hiload 26/600 Superdex 200 pg gel filtration column (GE Healthcare Life Sciences) was performed. All purification steps were performed using an AKTA Pure system (GE Healthcare Life Sciences).

## **Affinity chromatography**

### **Principle**

The fusion of a protein of interest with a tag has been used widely to facilitate the purification of recombinant proteins. A His-tag constituting of six or more histidine residues is the most widely used fusion tag in protein expression and purification. Immobilized metal ion affinity chromatography (IMAC) is widely used to purify His-tagged proteins since histidine residues have affinity to immobilized metal ions such as  $Zn^{+2}$ ,  $Cu^{+2}$ ,  $Ni^{+2}$ , and  $Co^{+2}$ . The specificity and affinity to polyhistidine tag vary for these metal ions. One step of IMAC purification using the His-tag typically results in fractions of the eluted protein of interest with high purity, sufficient for downstream applications. This is due to the high affinity and specificity of the His-tag (Block et al., 2009, Spriestersbach et al., 2015).

### **Procedure**

Pellets recovered from protein expression in 400 mL of expression media, were mixed with 50 mL of lysis buffer (Table 3-1) and protease inhibitor was added to inhibit the activity of proteolytic and phospholytic enzymes released after cell lysis, which would cause protein inactivation or degradation. Lysomix, DNase, and RNase were added to the mixture, and pellets were lysed by sonication for 5 min using the following program: 10 seconds on, 20 seconds off, 70% amplitude. Lysis, equilibration, and washing buffers were prepared using 50 mM Tris-HCL (pH 8.0), 500 mM NaCl, 25 mM  $MgCl_2$ , and 0.5 mL of 2-mercaptoethanol. The elution buffer consisted of 50 mM Tris-HCL (pH 8.0), 500 mM NaCl, 25 mM  $MgCl_2$ , 0.5 mL of  $\beta$ -mercaptoethanol, and 500 mM imidazole. Crude extracts were centrifuged at 4°C for 40 min at 8000 rpm, filtrated using a Whatman® Puradisc (pore size 0.45  $\mu$ m) filter, and loaded onto a HisTrap HP 5 mL column connected to the AKTA Pure system. Encapsulin fractions were eluted in a 0–500 mM imidazole linear gradient (total volume 60 mL). Seven fractions of the purified

protein (14 mL) were concentrated to about 4 mL and subjected to gel filtration in the next step of purification.

**Table 3-1:** Composition of the buffer used in the first step of chromatography for ReEncapsulin purification (Affinity chromatography).

<b>Buffer: (Lysis, equilibration and washing)</b>		<b>First step</b>
<b>Component</b>	<b>Concentration</b>	<b>Column</b>
Tris-HCl (pH 8)	50 mM	HisTrap column
NaCl	500 mM	
MgCl <sub>2</sub>	25 mM	
β-Mercaptoethanol	0.5 ml	
<b>Buffer: Elution</b>		
<b>Component</b>	<b>Concentration</b>	
Tris-HCl (pH 8)	50 mM	
NaCl	500 mM	
MgCl <sub>2</sub>	25 mM	
β-Mercaptoethanol	0.25 ml	
imidazole	500 mM	

#### Size exclusion (gel filtration)

A total of 4 mL of concentrated protein from the first step was injected into the Hiload 26/600 Superdex 200 pg gel filtration column (GE Healthcare Life Sciences). The equilibration buffer contained 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 25 mM MgCl<sub>2</sub>, 1 mM DTT, and 10% glycerol (Table 3-2). Eluted fractions were analysed by SDS-PAGE, and the pure eluted fractions were combined and stored in 10% glycerol at -80°C until use.

**Table 3-2:** Composition of the buffer used in the second step of chromatography for ReEncapsulin purification (Size exclusion chromatography).

<b>Buffer: (Lysis, equilibration and washing and Elution)</b>		<b>Second step</b>
<b>Component</b>	<b>Concentration</b>	<b>Column</b>
Tris-HCl (pH 7.5)	50 mM	Size exclusion
mM MgCl <sub>2</sub>	25 mM	
NaCl	500 mM	
DTT	1 mM	
glycerol	10%	

### **3.2.8 Purification of DyP4-tag**

The purification of DyP4-tag was carried out in three chromatography steps, similar to the purification of DyP4, including two ion-exchange chromatography steps using a Q column and SP column. A gel filtration step was the third step of purification, using a HiLoad 26/600 Superdex 75 pg column according to the procedure described in Chapter 2 (section 2.8).

### **Enzymatic activity for DyP4-tag**

Two colorimetric assays were used to verify whether tagging DyP4 with a C-terminal signal peptide might affect or inhibit activity via the ABTS and RB19 assays. The advantage of using these assays is that the reaction occurs fast (either formation of green colour in case of ABTS assay, or the disappearance of the blue dye colour in case of RB19 assay) as a result of oxidation by the enzyme (DyP4-tag).

### **3.2.9 Molecular cloning of ReEncapsulin and DyP4-tag in Due expression systems**

#### **Molecular cloning of pACYCDuet and pRSFDuet vectors**

Double digestion for both vectors was performed using *EcoRI* and *NcoI* restriction sites to insert encapsulin, while *NdeI* and *XhoI* restriction sites were used to insert the DyP4-tag. Double-digested vectors were gel extracted, purified, ligated with encapsulin, and transformed into DH5 $\alpha$  using the CaCl<sub>2</sub> method. Next, restriction analysis was performed to confirm the insertion of encapsulin into both vectors before the insertion of the DyP4-tag in a process similar to the insertion of encapsulin but at different restriction sites.

#### **Molecular cloning of Re-Encap-DyP4-tag-pACYCDuet and Re-Encap-DyP4-tag-pRSFDuet**

*EcoRI* and *NcoI* restriction sites were used to insert the encapsulin gene into a pACYCDuet vector. ReEncapsulin was PCR-amplified from pET24a-encapsulin using the primers shown in Table 3-1. *DpnI* was added to digest the parental template, followed

by double digestion with *EcoRI* and *NcoI* restriction enzymes. Next, PCR purification was performed and the purified products were ligated into pACYCDuet. The ligated mixtures were used for CaCl<sub>2</sub> heat shock transformation in DH5 $\alpha$  followed by a mini prep step and restriction analysis to verify the insertion of encapsulin into both vectors. Following the successful insertion of encapsulin, DyP4-tag was PCR-amplified from pET24a-DyP4 using the primers shown in Table 3-3. *DpnI* was added to digest the parental template, followed by double digestion with *NdeI* and *XhoI* restriction enzymes. Purification was performed and followed by ligation in both vectors using T4 DNA ligase. Transformation in DH5 $\alpha$  was performed and followed by mini prep and restriction analysis to confirm the insertion of DyP4-tag in both vectors.

**Table 3-3:** List of primers used to amplify Encap and DyP4

Primer	Oligonucleotide sequences (5'->3')
Encapsulin	Fwd: 5' GAGCTCGAATTCATTAGTGGTGGTGGTGG 3'
	Rev: 5' GATATACCATGGGCACCAACCTGCACCGTG 3'
DyP4-tag	Fwd: 5' TATACATATGATGACCACCCCGGCGCCGCGCTG 3'
	Rev: 5' ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC 3'

## PCR mixture

**Table 3-4:** PCR mixture for Encapsulin amplification

Component	Stock Concentration	Volume ( $\mu\text{L}$ )
Distilled water	–	33
Q5 buffer	5 $\times$	10
dNTPs	10 mM each	1
pET24a-Encapsulin	100 ng/ $\mu\text{L}$	0.5
Primer 1	20 $\mu\text{M}$	2.5
Primer 2	20 $\mu\text{M}$	2.5
Q5 high-fidelity DNA polymerase	2 U/ $\mu\text{L}$	0.5
Total reaction volume		50

**Table 3-5:** PCR mixture for DyP4-tag amplification

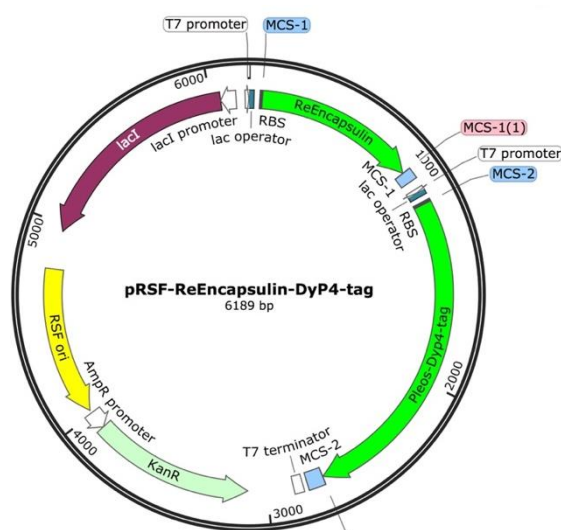
Component	Stock Concentration	Volume ( $\mu\text{L}$ )
Distilled water	–	33
Q5 buffer	5 $\times$	10
dNTPs	10 mM each	1
pET24a-DyP4-tag	100 ng/ $\mu\text{L}$	0.5
Primer 1	20 $\mu\text{M}$	2.5
Primer 2	20 $\mu\text{M}$	2.5
Q5 high-fidelity DNA polymerase	2 U/ $\mu\text{L}$	0.5
Total reaction volume		50

**Table 3-6:** PCR program to amplify Encapsulin

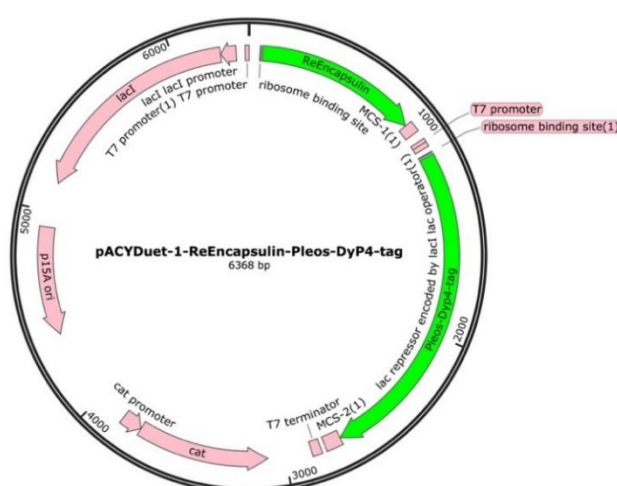
Step	Temperature	Time
1	98°C	30 sec
2	98°C	10 sec
3	72°C	30 sec
4	72°C	1 min
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

**Table 3-7:** PCR program to amplify DyP4

Step	Temperature	Time
1	98°C	30 sec
2	98°C	10 sec
3	72°C	30 sec
4	72°C	1 min, 45 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold



**Figure 3-4:** Plasmid map of pRSFDuet-Re-Encap-DyP4-tag vector (created with SnapGene). ReEncapsulin was inserted into pET24a vector using restriction sites *EcoRI* and *NcoI* while DyP4 was inserted using *NdeI* and *XhoI* restriction sites.



**Figure 3-5:** Plasmid map of pACYCDuet-Re-Encapsulin-DyP4-tag vector (created with SnapGene). ReEncapsulin was inserted into pET24a vector using restriction sites *EcoRI* and *NcoI* while DyP4 was inserted using *NdeI* and *XhoI* restriction sites.

### **Transformation and Due expression in BL21 (DE3) and purification**

Chemical transformation using the CaCl<sub>2</sub> method was carried out in BL21 (DE3) using pACYCDuet-Re-Encapsulin-DyP4-tag and pRSFDuet-Re-Encap-DyP4-tag vectors. The detailed procedure of this transformation method is given in Chapter 2 (section 2.4.1).

### **Due expression in BL21 (DE3)**

Expression in BL21(DE3) was performed using the optimized conditions described in section 3.2.4 for the expression of encapsulin. It should be noted here that large-scale expression (400 mL of expression media resulted in a high concentration of expressed encapsulin and encapsulated DyP4, and significant aggregation of the protein was thus reported. The expression scale was therefore reduced to 100 mL of the expression media (2× TY media) induced with 0.1 mM of IPTG.

### **Purification of encapsulated DyP4-tag**

The purification of encapsulated DyP4 was performed in a single step using a HisTrap HP 5 mL column. To avoid protein aggregation, which occurs when the concentration of purified encapsulin is high, protein expression and purification were performed on a smaller scale and immediately followed by SDS-PAGE and spectrometric and ABTS assay analysis.

### **Analysis of encapsulated DyP4 in ReEncapsulin**

The eluted fractions from the affinity chromatography column were analyzed using SDS-PAGE and ABTS assay. SDS-PAGE analysis was performed using 10% polyacrylamide gels. To one part of protein solution, one part of SDS reducing buffer supplemented with β-mercaptoethanol was added, mixed, heated, and loaded onto the gel for electrophoresis. Staining with Coomassie Brilliant Blue was performed, followed by destaining and visualization of the protein bands.

Since the spectral measurement of haem-containing proteins gives a peak at a specific wavelength, spectral analysis was also performed in 1 mL of the eluted fraction using a



UV spectrophotometer.

ReEncapsulin and DyP4 were cloned into the due expression vectors pRSFDuet and pACYCDuet. Purification of the expressed Encap-DyP4 was performed with affinity chromatography in the first step of purification. Thus, as the encapsulin protein has a C-terminal His-tag, it will bind to the affinity column (HisTrap column).

### **Enzymatic activity of encapsulated DyP4 into ReEncapsulin**

An ABTS assay was used to compare the activity of encapsulated DyP4 eluted from three fractions. The reaction mixture contained 10 mM of ABTS (185  $\mu$ L) and 5  $\mu$ L of the eluted fraction of encapsulated DyP4, and 4 mM of hydrogen peroxide (10  $\mu$ L) was added to initiate the reaction. For the negative controls, two reactions were performed where encapsulin or PBS buffer were added instead of encapsulated DyP4. The reaction mixtures were performed in a volume of 200  $\mu$ L in a 96-well plate, and the measurements were recorded at a wavelength of 405 nm for 5 minutes.

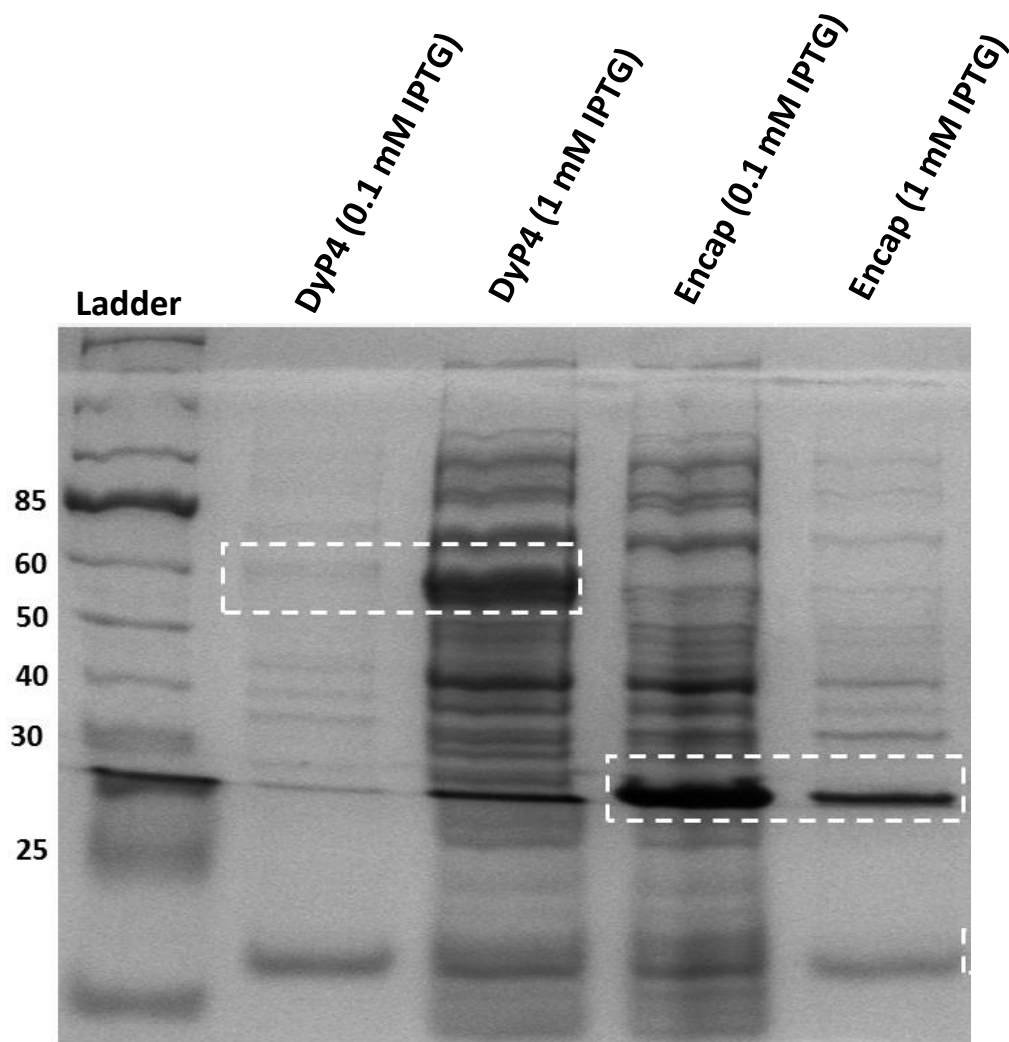
### **3.3 Results**

#### **3.3.1 Molecular cloning of pET24a-Encap and pET24a-DyP4-tag**

Sequencing results confirmed the insertion of encapsulin into pET24a and DyP4 into pET24a. To determine whether tagging DyP4 with a C-terminal peptide sequence caused any aggregation or loss of activity, ABTS and RB19 assays were used to monitor activity at 405 nm and 595 nm, respectively.

#### **3.3.2 Recombinant protein production in BL21 (DE3)**

The expression of encapsulin and DyP4-tag was carried out successfully in the BL21 (DE3) *E. coli* strain as confirmed by SDS-PAGE (Figure 3-6). Expression in the C41 (DE3) strain was not detected (data not shown). Two concentration of IPTG were evaluated for both proteins, and 0.1 mM was found to be optimal for encapsulin while 1 mM of IPTG was optimal for DyP4.

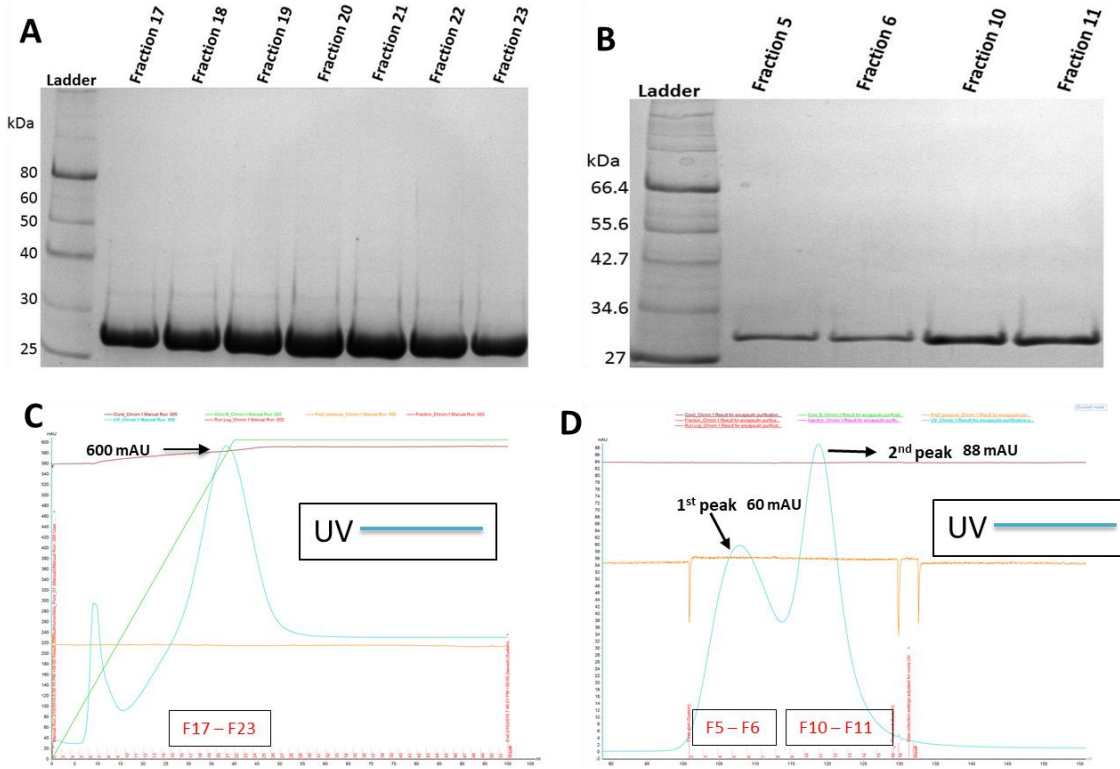


**Figure 3-6:** Comparison between the expression of DyP4-tag and encapsulin proteins induced by 0.1 mM or 1 mM of IPTG. Lane 1, molecular weight markers; lanes 2 and 3, soluble fraction of DyP4-tag induced with 0.1 mM or 1 mM of IPTG, respectively; lanes 4 and 5, soluble fractions of encapsulin induced with 0.1 mM or 1 mM of IPTG, respectively

### 3.3.3 Purification of encapsulin

Eluted fractions after the first step of purification (Fractions 17-23), (Figure 3-7A) were collected and analysed by SDS-PAGE and found to be already highly pure after the first step and only a few thin bands (around 35 kDa) were seen which were removed after purification by gel filtration (Figure 3-7B). Fractions 17 – 23 were combined, concentrated and then loaded in the size exclusion column, and the resulting appropriate fractions (fractions 5 and 6 from the first peak, and fractions 10 and 11 from the second

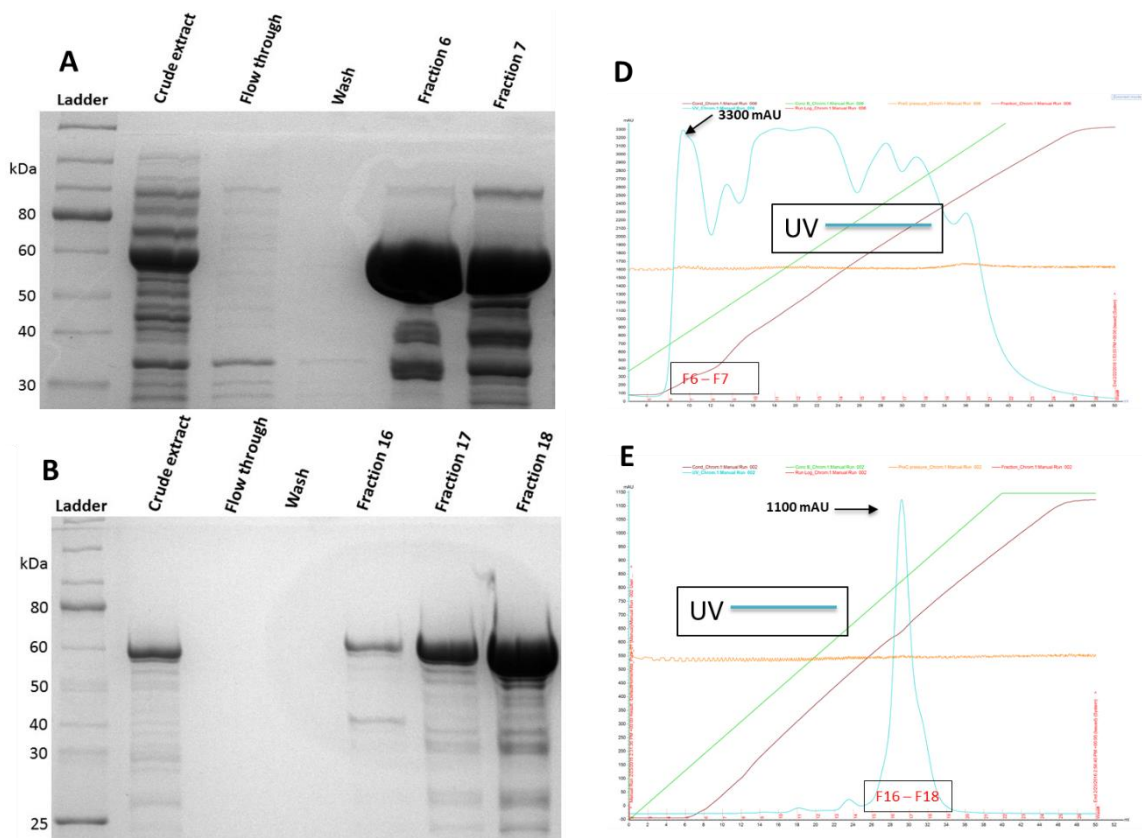
peak) (Figure 3-7B) were combined and stored at  $-80^{\circ}\text{C}$  until use.

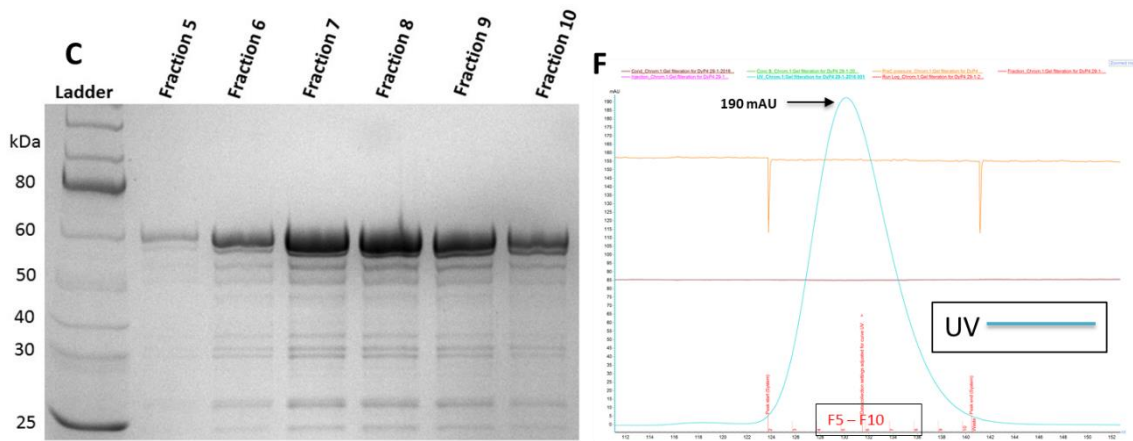


**Figure 3-7:** Purification of encapsulin by affinity chromatography. (A) Lane 1, molecular markers; lanes 2–8, eluted fraction of encapsulin and size exclusion chromatography. (B) Lane 1, molecular markers; lanes 2–3, eluted fractions from the first peak; lanes 4–5, eluted fractions from the second peak.

### 3.3.4 Purification of DyP4-tag

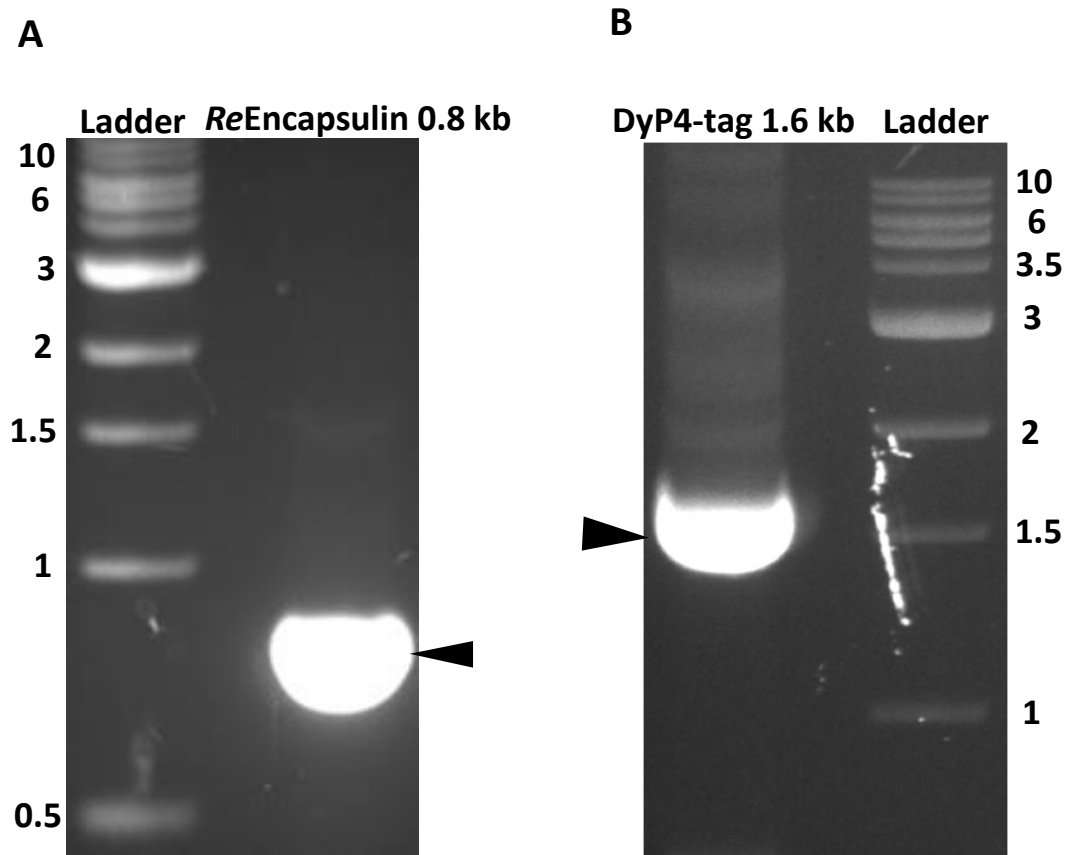
Large-scale production (400 mL) of *Pleos*-DyP4 as the active protein was achieved by culturing BL21 (DE3) transformants at 25°C and 200 rpm for 24 hrs. Cells were disrupted and lysed, centrifuged, and the supernatant was filtrated and loaded into the appropriate column for purification using the AKTA Pure system. Three chromatographic steps were performed for the purification of DyP4-tag. The first step was ion-exchange chromatography with a Q5 column. SDS-PAGE (Figure 3-8A) showed the thick bands of the soluble fraction of DyP4-tag (59 kDa) in lanes 5 and 6; other bands of impurities can also be seen. The second step of purification was carried out using an SP column, and, as shown in Figure 3-8B, significant amounts of impurities were removed. Most of these contaminants were removed after the final step using a HiLoad 26/600 Superdex 75 pg column (Figure 3-8A-C), though a small amount of impurities remained.





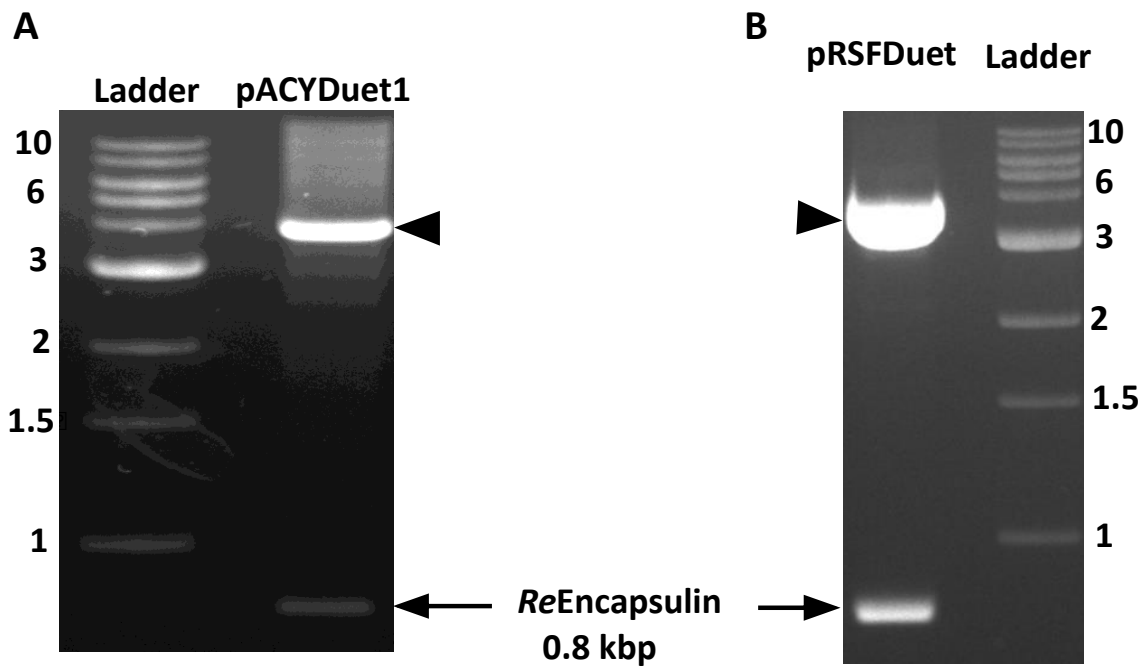
**Figure 3-8:** SDS-PAGE for the three purification steps of DyP4-tag. Q5 column (A) lane 1; molecular ladder, lanes 2–4; crude extract, flow-through, and wash, respectively, lanes 5–6; whole soluble fractions of DyP4. SP column (B) lane 1; molecular ladder, lanes 2–4; crude extract, flow-through, and wash, respectively, lanes 5–7; whole soluble fractions of DyP4. Superdex 75 column (C), lane 1; molecular ladder, lanes 2–7; whole soluble fractions of DyP4.

### 3.3.5 Molecular cloning of Re-Encap-DyP4-tag-pACYCDuet and Re-Encap-DyP4-tag-pRSFDuet



**Figure 3-9:** PCR amplification of encapsulin and DyP4-tag genes. DNA gel electrophoresis 1% was used to visualize the products. The expected PCR product was 0.8 kb for encapsulin (A) and 1.6 kb for DyP4-tag (B).

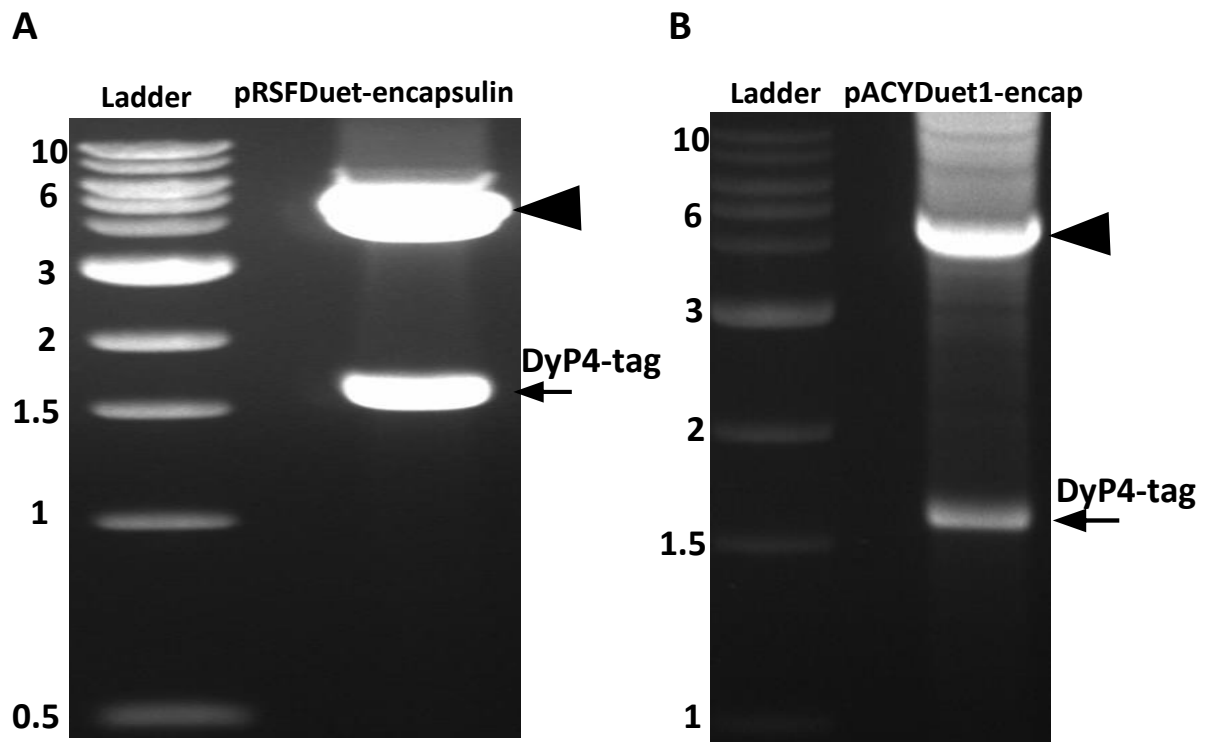
To insert encapsulin and DyP4-tag into the due expression vectors, PCR amplification of both genes was carried out (Figure 3-9, A and B). The encapsulin gene was inserted first by ligating the double-digested PCR products with *EcoRI* and *NcoI* with pACYCDuet and pRSFDuet vectors, which were previously double digested with the same restriction enzymes.



**Figure 3-10:** Restrictive digestion analysis of pACYCDuet-Encap (A), and pRSFDuet-Encap (B) at restriction sites *EcoRI* and *NcoI*. The expected digested products are approximately 4 kb for pACYCDuet, 3.8 kb for pRSFDuet, and 0.8 kb for encapsulin.

Ligated mixtures were used for chemical transformation using the  $\text{CaCl}_2$  method in DH5 $\alpha$ , followed by extraction of plasmids, restrictive digestion, and DNA gel analysis. The results confirmed the successful insertion of the encapsulin gene into both vectors, pACYCDuet and pRSFDuet (Figure 3-10, A and B).



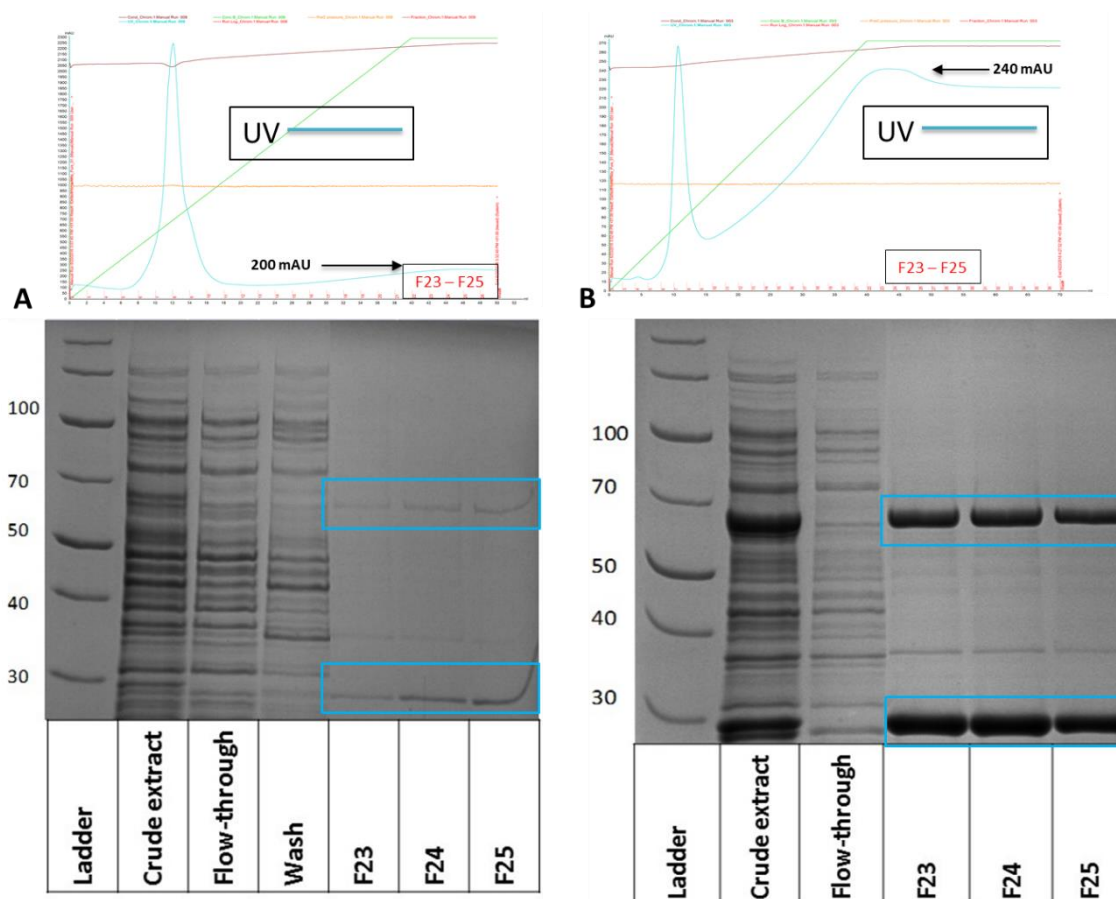


**Figure 3-11:** Restrictive digestion analysis of pRSFDuet-Encap-DyP4-tag (A) and pACYCDuet-Encap-DyP4-tag (B) at restriction sites *Nde*I and *Xho*I. The expected digested products are approximately 4.8 kb for pACYCDuet-Encap, 4.6 kb for pRSFDuet-Encap, and 1.6 kb for DyP4-tag.

Double-digested PCR products of the amplified DyP4-tag were ligated with pACYCDuet-Encap and pRSFDuet-Encap that were previously double-digested with the same restriction enzymes. Ligated mixtures were used for transformation into DH5 $\alpha$ , followed by the extraction of plasmids from DH5 $\alpha$  and restriction analysis. The successful insertion of DyP4-tag insert into pACYCDuet-Encap and pRSFDuet-Encap was achieved as shown by restriction analysis (Figure 3-11, A and B).

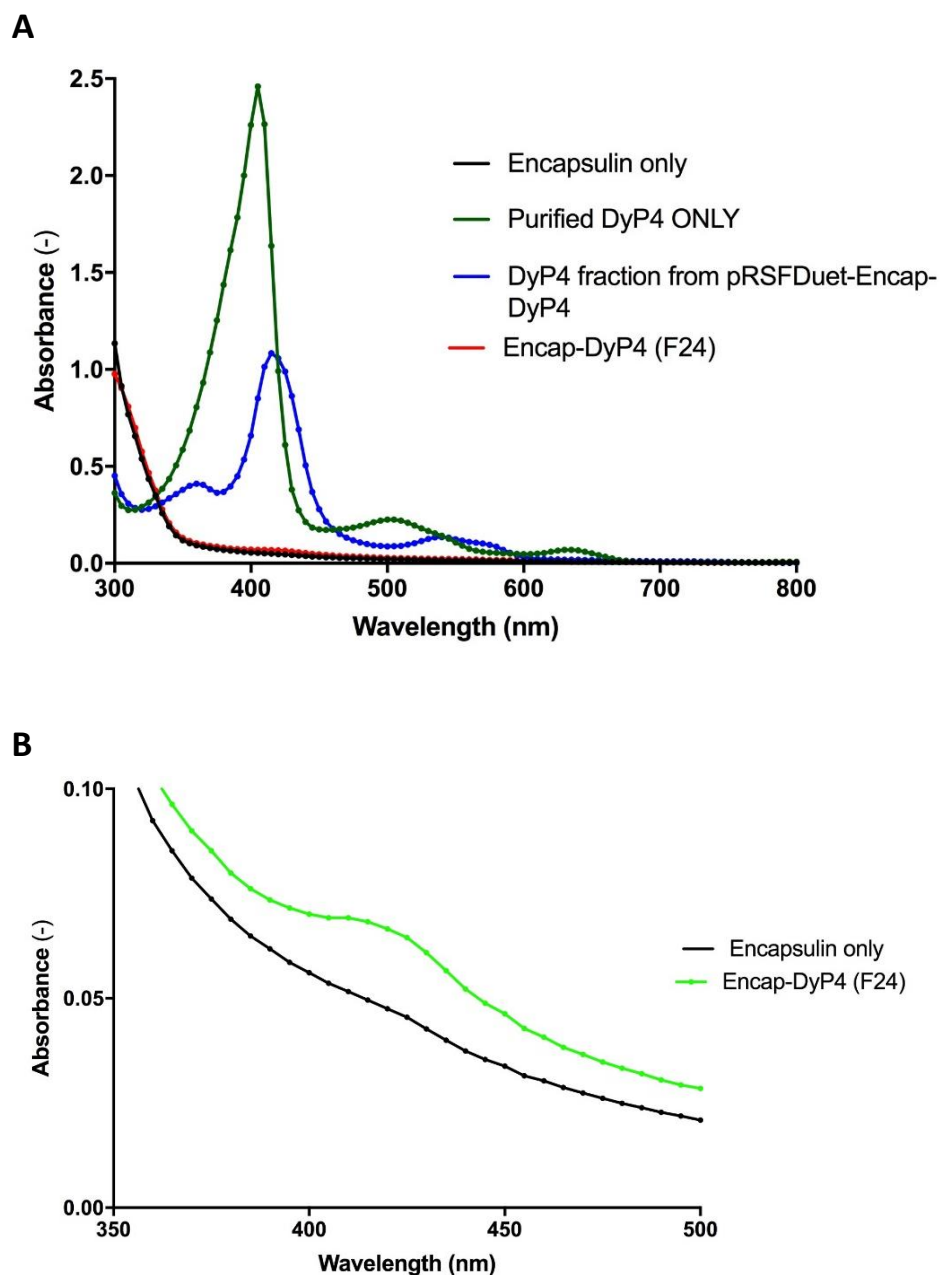
### 3.3.5 Purification of encapsulated DyP4 into ReEncapsulin

ReEncapsulin was coexpressed with DyP4-tag in 100 mL of expression culture in the BL21 (DE3) strain. Pellets were lysed and centrifuged, and the supernatants were loaded into a HisTrap 5 mL column for affinity chromatography purification. Eluted fractions were analysed by SDS-PAGE, and the results showed the presence of fractions of the coexpression of ReEncapsulin and DyP4-tag (Figure 3-13; A, lanes 5–7 for expression in a pACYCDuet vector; and B, lanes 4–6 for expression in a pRSFDuet vector). The SDS-PAGE results (figure 3-12) showed the presence of two bands, suggesting formation of the Encap-DyP4 protein complex.



**Figure 3-12:** SDS-PAGE of the coexpression of encapsulin and DyP4. (A) Lane 1, molecular weight marker; 2, crude extract; 3, flow-through; 4, wash; 5–7, purified fractions of ReEncapsulin coexpressed with DyP4-tag in pACYCDuet. (B) Lane 1, molecular weight marker; 2, crude extract; 3, flow-through; 4–6, purified fractions of ReEncapsulin coexpressed with DyP4-tag in pRSFDuet vector.

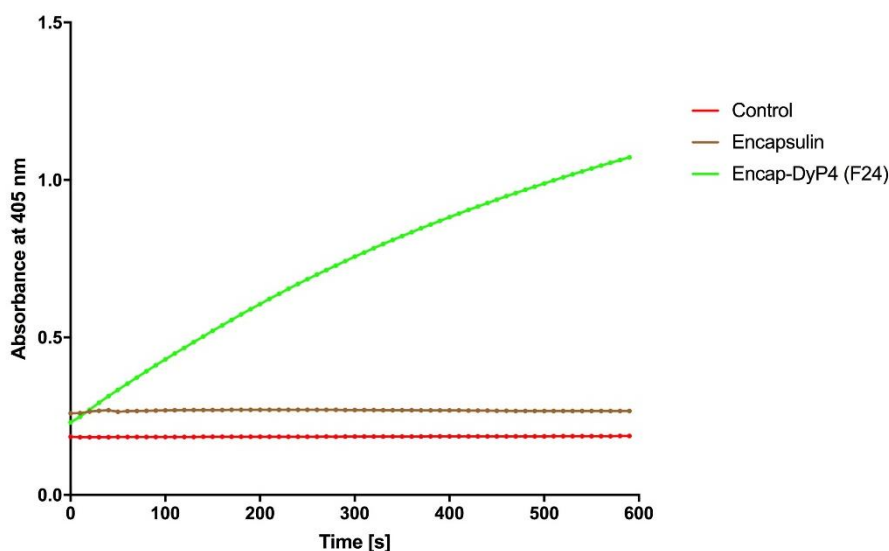
### 3.3.6 UV-spectral analysis of purified DyP4 and encapsulated-DyP4



**Figure 3-13:** UV-vis spectrum showing the characteristic absorption pattern of the purified fractions of DyP4-tag, encapsulated DyP4, and Encapsulin. (A) Wavelength scan measurements between 300 nm and 800 nm. (B) Wavelength scan measurements between 350 nm and 500 nm for eluted fractions of encapsulin only (black), and three eluted fractions of the encapsulated DyP4

### 3.3.7 Enzymatic activity of encapsulated DyP4-tag

Encapsulated DyP4 was assayed with ABTS to determine the activity of the packaged enzyme. ABTS is used to test the activity of peroxidases, including Pleos DyP4, and is assayed calorimetrically at a wavelength of 420 nm. The results of the ABTS assay showed that the encapsulated DyP4 retained its activity, with a possible reduction in activity, however, which could be due to the encapsulation reported in previous studies. Figure 3-14 shows the increase in absorbance at 405 nm due to oxidation of ABTS substrate by the packaged DyP4 from the eluted fraction of the encapsulated DyP4.



**Figure 3-14:** Analysis of the enzymatic activity of encapsulated DyP4 into ReEncapsulin. Activity of DyP4 peroxidase using ABTS assay. The reaction was started by the addition of hydrogen peroxide and activity was measured as the change in absorbance at 405 nm. Two negative controls (red line with a buffer, brown line with encapsulin) were used; the green line indicates the packaged DyP4 (fraction 24)

### **3.4 Discussion**

The interesting family of nanocompartments known as encapsulins have recently been discovered in prokaryotes by Sutter et al (Sutter et al., 2008). Encapsulated proteins have a C-terminal sequence that is responsible for targeting their encapsulation. The present study demonstrated that Pleos-DyP4 can be encapsulated into a ReEncapsulin nanocompartment by the fusion a C-terminal sequence that can act as a cargo-leading peptide to direct the encapsulation of DyP4. This provides evidence that the encapsulin system is also applicable to non-native enzymes. This finding is similar to that reported in previous studies where non-native proteins were loaded into encapsulin nanocompartments by fusion of the C-terminal sequence to the guest C-terminus sequence. There is strong evidence in this work to claim that the encapsulation of DyP4 has been successful. The eluted fractions from peaks from a nickel column known to bind strongly to His-tagged proteins (encapsulin in this study) showed the presence of a band representing DyP4, along with another band for encapsulin. The eluted fractions from the peaks where the encapsulin band was found, along with DyP4, were used to test peroxidase activity in the ABTS assay. The assayed fractions of the encapsulated DyP4 showed activity in the ABTS assay, although the activity seemed to be reduced. This might provide further evidence to support the successful encapsulation of what might be attributable to the very small size pores of the nanocompartment, thus affecting the permeability. UV-vis scan measurement of the encapsulated DyP4 showed a haem-type peak, similar to the purified DyP4 (3-14B) but with a reduced peak value which, again, could be due to encapsulation.

### **3.5 Conclusion**

The packaging of Pleos DyP4 was performed in the present work by fusing the C-terminal sequence (37 aa) of ReDypB to the C-terminus of DyP4. It seemed that the packaged DyP4 remained active when assayed with ABTS, but with possible reduced activity due to encapsulation. More studies are needed to investigate this further. The present work provides further evidence that non-native enzymes can be encapsulated in encapsulin using CLP, as described here. Overall, this mechanism of packaging enzymes into the encapsulin nanocompartment is of potential biotechnological value since (i) being a stable shell, encapsulin protects the enzyme from degradation by the cell or might protect the cell from toxic enzymes, and (ii) encapsulated enzymes can be released at low pH, which is suitable for DyP4 with its high activity at low pH (Nichols et al., 2017). Taken together, the specific targeting sequence can be used to package non-native proteins into the encapsulin nanocompartment, which has tremendous potential for use in nanobiotechnology applications using this delivery system, and could also be used in the destruction of lignocellulose (Rahmanpour and Bugg, 2013).

## **CHAPTER 4 Substitution of methionine residues to enhance DyP4 stability against increased hydrogen peroxide concentration**

4.1 Introduction

4.2 Materials and Methods

4.3 Results

4.4 Discussion

4.5 Summary

## 4.1 Introduction

Peroxidases are a significant group of enzymes that are widely distributed in nature and use a peroxide to catalyse the oxidation of a broad range of substrates. They can be classified into haem-containing non-haem-containing proteins (Fawal et al., 2013). Haem peroxidases are found in animals, plants, bacteria, and fungi and utilize hydrogen peroxide ( $H_2O_2$ ) to oxidize a broad range of organic and inorganic substrates. Oxidation of substrates by haem peroxidase occurs in three steps, which require the formation of Compound I and Compound II. Peroxidases have potential applications in bioremediation such as for contaminated waste water; in the paper industry, such as for biopulping and biobleaching; and for diagnostic and analytical applications in medicine (Hamid and Khalil-ur-Rehman, 2009). Dye-decolorizing peroxidases (DyPs) are a novel family of haem peroxidases which have received significant attention due to their role in the decolourization of anthraquinone dyes and the degradation of lignin (Shrestha et al., 2017) (Singh and Eltis, 2015). Dye-decolorizing peroxidase 4 (DyP4) is a newly identified peroxidase from the edible mushroom *Pleurotus ostreatus* (Pleos-DyP4) which has been found to efficiently oxidize low and high redox potential dyes such as Reactive Blue 19 and Reactive Black 5. Interestingly, DyP4 has also been found to oxidize  $Mn^{2+}$  to  $Mn^{3+}$ . Moreover, DyP4 has high thermal and pH stability (Fernandez-Fueyo et al., 2015).



#### **4.1.1 Mechanism of inactivation by excess H<sub>2</sub>O<sub>2</sub>**

One of the main limitations in utilizing haem peroxidases in industrial applications is their inactivation by hydrogen peroxide. The mechanism is complicated, poorly understood, and referred to as ‘suicide inactivation’. The formation of Compound III species causes the inactivation of haem peroxidases through haem bleaching and irreversible inactivation. The oxidation of amino acids that are not stable against oxidation such as methionine (Met), tryptophan (Trp), tyrosine (Tyr), cysteine (Cys), and histidine (His) by hydrogen peroxide leads to protein modification and, as a result, protein inactivation (Ayala et al., 2011, Valderrama et al., 2002). The conventional method of increasing hydrogen peroxide stability of enzymes is to rationally substitute the residues that are easily oxidized with residues that are more resistant to oxidation by hydrogen peroxide. Increasing the hydrogen peroxide stability of haem peroxidases has been reported in several studies using different strategies of protein engineering such as site-directed mutagenesis (rational design) or random mutagenesis (directed evolution), or a combination of both (Brissos et al., 2017, Cherry and Lamsa, 2004, Cherry et al., 1999, Gonzalez-Perez et al., 2014, Miyazaki and Takahashi, 2001, Miyazaki-Imamura et al., 2003, Morawski et al., 2001, Ogola et al., 2010).

#### **4.1.2 Aim of the study**

This study aimed to investigate the effect of replacing susceptible residues to hydrogen peroxide oxidation with residues that are more resistant to oxidation. In the present study, three methionine residues were replaced with leucine and/or phenylalanine. A semi-rational protein engineering method using site-directed mutagenesis was used to replace the three methionine residues of the *Pleurotus ostreatus* DyP4 with residues that are more stable against oxidation, namely leucine and phenylalanine.

## **4.2 Material and methods**

### **4.2.1 Materials**

All materials used in this study were obtained as described in Chapter 2, and cultivation media were also prepared as described in Chapter 2.

### **4.2.2 Protein modelling**

Homology modelling is used widely in site-directed mutagenesis studies to build a 3D structure of a protein of interest using established 3D structures for related proteins. The SWISS-MODEL workspace is an online tool used widely for this purpose, and the models are generated by identifying a related protein with a known structure and using it as a template model to build a 3D model of the protein of interest (Bordoli et al., 2009). In this experiment, a protein model was created using the most structurally similar template proteins identified in SWISS-MODEL, (PDB: 1avf). The 3D-model structure was visualized using Pymol. The crystal structure of Pleos-DyP4 has recently been solved at 1.56 Å resolution (PDB: 6fsk) but it is still not released into the protein data bank (Fernandez-Fueyo et al., 2018).

### **4.2.3 Construction of mutants using OneClick program**

The OneClick program (<http://tucksengwong.staff.shef.ac.uk/OneClick/>) was used to design primers to substitute methionine with leucine or phenylalanine Table 4-1 (Warburton et al., 2015). OneClick is an online program, established to facilitate the design of mutagenic primers in site-directed mutagenesis and saturation mutagenesis. Design of mutagenic primers using OneClick takes only a few steps, where the user can choose the position of the amino acid to mutate and the amino acid to replace. A wide range of the commercially available DNA polymerases and plasmids being used in site-directed mutagenesis or saturation mutagenesis have been incorporated into the OneClick program. In addition, OneClick can provide information on the Polymerase Chain

Reaction (PCR) mixture, PCR cycling condition, estimated size of the product of PCR, and the agar plate to use after the transformation (Warburton et al., 2015).

**Table 4-1:** Designed primers for the construction of mutants

Mutant	Oligonucleotide sequences
<b>M43L</b>	F: 5' AAAGCGAACCTGGCGCACTTCATCCCGCACATTAAGACCAGCGCGG 3'
	R: 5' GAAGTGCGCCAGGTTTCGCTTTAAATTGATCAACGTTGGTCACGTCG 3'
<b>M77L</b>	F: 5' CTGGTGCCGCTGGCGGCGGTGAACGTTAGCTTTAGCCACCTGGGCC 3'
	R: 5' CACCGCCGCCAGCGGCACCAGACCCGGTTTCTTCTGACGTTTGTGT 3'
<b>M253L</b>	F: 5' CTGTTCCAACCTGGTGCCGGAGTTTGACGATTTCTGGAAAGCAACC 3'
	R: 5' CTCCGGCACCAGTTGGAACAGGTAACGGAAGGTCAGAAAGCTACCA 3'
<b>M253F</b>	F: 5' CTGTTCCAATTTGTGCCGGAGTTTGACGATTTCTGGAAAGCAACC 3'
	R: 5' CTCCGGCACAAATTGGAACAGGTAACGGAAGGTCAGAAAGCTACCA 3'

Mutants were constructed using pET24a-DyP4 as a template. PCR reactions mixtures were prepared in a total volume of 50  $\mu$ L containing 1 $\times$  Q5 Reaction Buffer, 50 ng DNA template plasmid (pET24a-DyP4), 200  $\mu$ M of each of the dNTPs, 20  $\mu$ M of each primer, and 1 U of Q5 high-fidelity DNA polymerase. The PCR reaction cycle was 98°C for 30 sec, followed by 9 cycles of 98°C for 8 sec, 55°C for 20 sec, and 72°C for 2 min 51 sec. The reaction was then paused, and both tubes were mixed and distributed equally between the two tubes, followed by 19 cycles of 98°C for 8 sec, 55°C for 20 sec, and 72°C for 2 min 51 sec. The final step was 2 min at 72°C before cooling at 8°C. Reactions were digested with 1  $\mu$ L of high-fidelity *DpnI* at 37°C overnight to digest the parental methylated DNA.

#### 4.2.4 CaCl<sub>2</sub> heat shock transformation in DH5 $\alpha$

The *DpnI* digested mixture (5  $\mu$ L) was used for chemical transformation using the CaCl<sub>2</sub> heat shock method in DH5 $\alpha$  cells. The detailed protocol for this method is given in Chapter 2 (section 2.4.1). A single colony from each plate containing the mutants was

picked and grown overnight in 2× TY media containing kanamycin for plasmid isolation to confirm that the desired mutations were successfully obtained.

#### **4.2.5 Sequencing**

Extracted plasmids were used to sequence the obtained mutants to determine whether the desired mutations were obtained. DNA sequencing was performed by Eurofins using the T7 promoter sequence 5' TAATACGACTCACTATAG ' 3 for mutants M34L and M77L, and 5' CAGCGAAGTGACCCATGTTC' 3 for mutants M253L and M253L.

#### **4.2.6 CaCl<sub>2</sub> transformation in BL21 (DE3)**

The resulting mutant plasmids were used for transformation in BL21 (DE3) as described in the materials and methods section (2.4.1) of Chapter 2.

#### **4.2.7 Protein expression of the WT and mutants**

Expression of the WT and the four mutants was carried out in *E. coli* BL21 (DE3). Cells were grown in 50 mL of 2× TY at 37°C until an OD<sub>600</sub> of approximately 0.6 was reached, and then induced with 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) followed by cultivation for 24 hrs at 25°C and 200 rpm.

#### **4.2.8 Cell harvesting and protein purification**

Cells were harvested by centrifugation at 4°C at 6000 rpm for 5 min and pellets were stored at -20°C until required. Purification of the WT and mutants was performed using the ÄKTA Pure system with two ion-exchange chromatography steps. In the first step, a HiTrap Q column 5 mL was used, and two fractions of the purified protein (total of 4 mL) were used for the second step using a HiTrap SP column 5 mL. The procedure for both ion-exchange chromatography steps was as described in Chapter 2 (section 2.8).

#### **4.2.9 Spectra and OD measurement/quantification**

To overcome the issues occurred during the purification of mutants and WT, it was

necessary to use the exact concentration of variants to validate the results of stability of DyP4 against hydrogen peroxide. Optical density measurement at A206, A280, and A340 using a spectrophotometer and wavelength scan measurement was performed in a 1 mL volume from each fraction of the purified mutants and the WT to determine the concentration of protein.

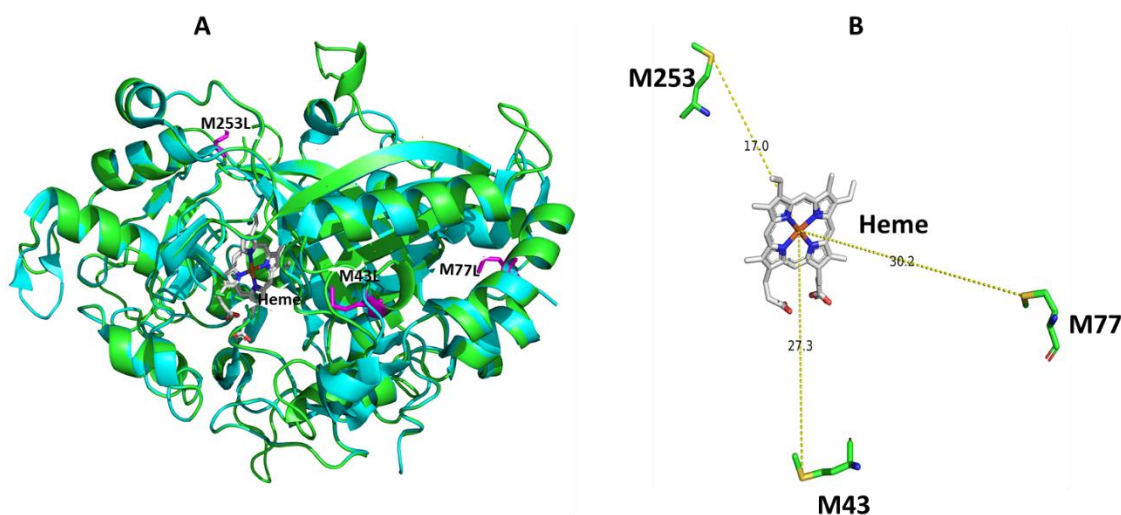
#### **4.2.10 Enzymatic activities of the WT and mutants**

To investigate the hydrogen peroxide stability of the WT and mutants, five different concentrations of hydrogen peroxide were used in an ABTS assay: 4 mM initially, followed by 8, 12, 16, and 20 mM of hydrogen peroxide. The reaction mixture contained 10 mM of ABTS (185  $\mu$ L) and 5  $\mu$ L of diluted purified enzyme, and the reaction was started with hydrogen peroxide (10  $\mu$ L). A multiplate reader was used to monitor the increase in absorbance at a wavelength of 405 nm. At least three replicates were performed for each reaction.

## 4.3 Results

### 4.3.1 Engineering the methionine residues of DyP4

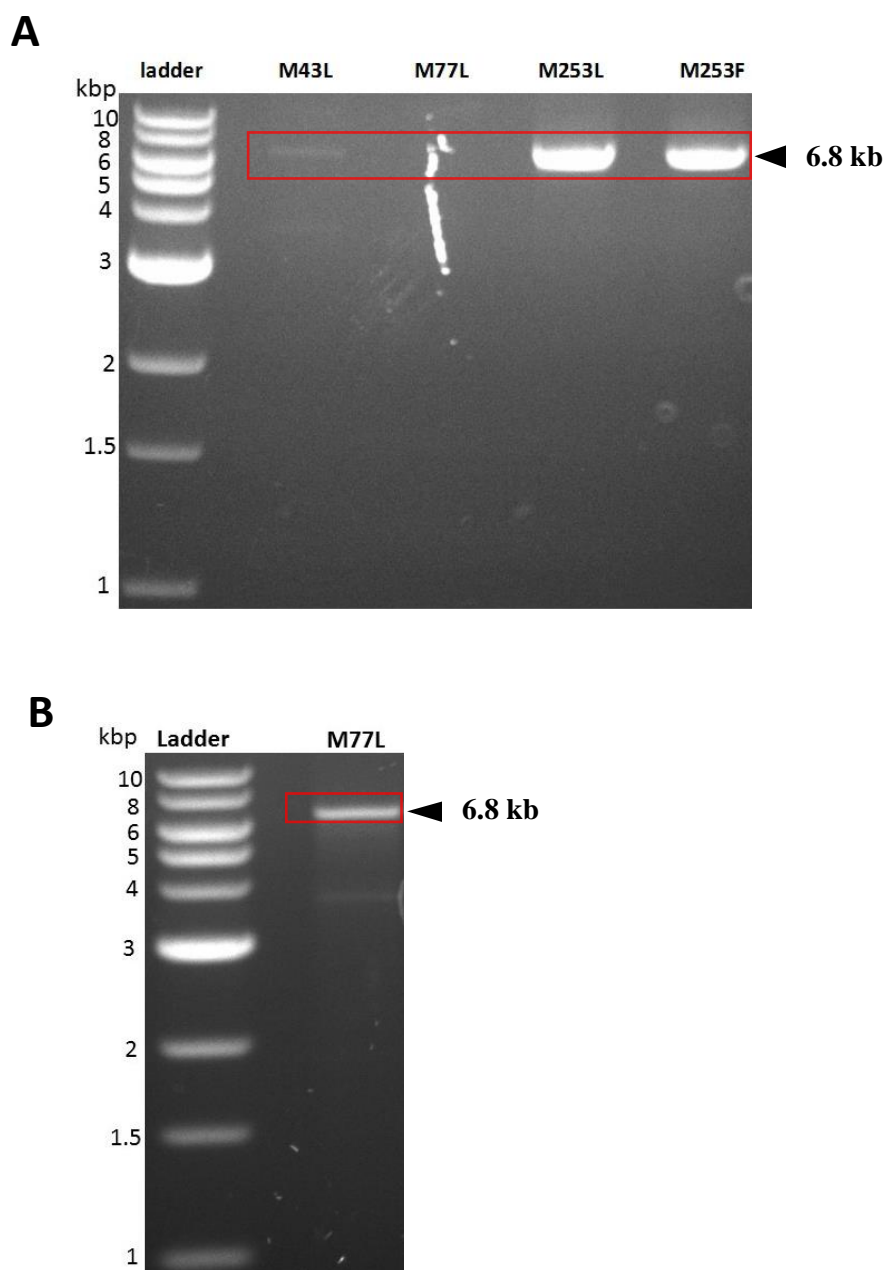
Methionine residues are the most susceptible residues to oxidation by high concentrations of hydrogen peroxide in peroxidase enzymes. There are only three methionine amino acids in the DyP4 protein. Using the SWISS-MODEL tool, a 3D structure model was built for DyP4 using the most similar template. Figure 4-1A shows the positions of the engineered methionine of the structural model of DyP4, constructed using the applicable template. None of the three methionine amino acids in the DyP4 protein are close to the haem active site, as shown by the 3D model (Figure 4-1B).



**Figure 4-1:** 3D model structure of Pleos DyP4 using the relative template. (A) The three methionine residues are shown in sticks and purple colour. This model was visualized with Pymol using (PDB: 1avf) as a template model. (B) Identified Met restudies for mutagenesis in Pleos-DyP4, M253, M43, and M77 with 17 Å, 27.3 Å, and 30.2 Å from the heme iron respectively.

### 4.3.2 Construction of mutants

To introduce a site mutation replacing methionine with leucine or phenylalanine, the OneClick program was used to design partially overlapping primers using the pET24a-DyP4 vector.



**Figure 4-2:** DNA gel for PCR in site-directed mutagenesis using the OneClick program. (A) Amplified products (6.8 kb) for mutants M43L, M253L, and M253F using Q5 DNA polymerase. (B) Successfully amplified product (6.8 kb) for M77L using (PfuUltra high-fidelity DNA polymerase.

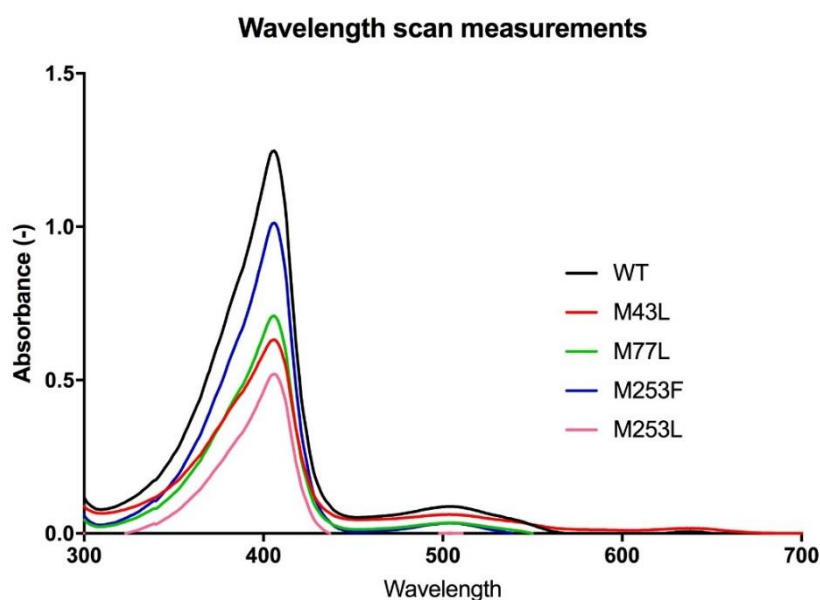
Using one-stage PCR, a clear band was obtained for the required product for positions M253L and M253F, and a thin band was observed for position M43L (Figure 4-2, A). However, for mutant M77L, no band was obtained with one-stage or two-stage PCR performed using Q5 high-fidelity DNA polymerase. Another polymerase (PfuUltra high-fidelity DNA polymerase AD) was used in the PCR reaction for mutant M77L, and this time there was a clear band for the product using one-stage PCR (Figure 4-2, B).

### 4.3.3 Sequencing results

Sequencing confirmed the replacement of Met with Leu (ATG to CTG) and Met to Phe (ATG to TTT) for positions 43M, 77M, and 253M.

### 4.3.4 Expression and purification of DyP4 variants

From the AKTA Pure chromatogram for the purification of WT and mutants, it was clear that there was a significant variation in the amount of purified protein for each variant of DyP4. This was confirmed by the results of the spectra (Figure 4-3) and by optical density measurements (Table 4-2).

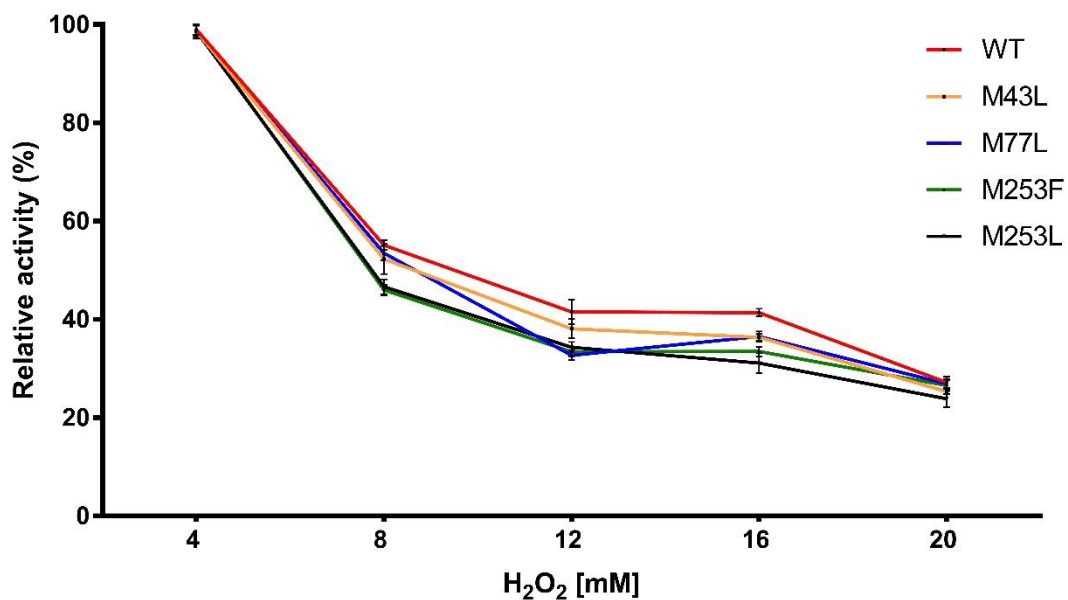


**Figure 4-3:** Wavelength scan measurement for DyP4 variants taken for wavelengths 700 nm to 300 nm. Equilibration buffer contained the elution buffer of the variants (1 ml), 1 ml of the eluted fractions of the variants after the second step of chromatography was used to estimate the concentration of the protein.



**Table 4-2:** Optical density of DyP4 variants

Variant	Optical density			Concentration (mg/mL)
	A260	A280	A340	
WT	0.435	0.500	0.227	0.448
M43L	0.261	0.301	0.138	0.259
M77L	0.268	0.300	0.141	0.270
M253L	0.186	0.212	0.099	0.187
M253F	0.360	0.410	0.186	0.380



**Figure 4-4:** Hydrogen peroxide tolerance of DyP4 WT and Met-replaced mutants M43L, M77L, M253L, and M253F. Reaction mixtures (a total of 200  $\mu$ L) contained 10 mM (185  $\mu$ L) ABTS, 5  $\mu$ L of diluted enzyme, and 10  $\mu$ L of H<sub>2</sub>O<sub>2</sub> (4 mM, 8 mM, 12 mM, 16 mM, and 20 mM).

Overall, the Met-replaced mutants behaved similarly to the WT of DyP4 in the presence of increased concentrations of hydrogen peroxide (Figure 4-4). The relative activity for WT and mutants using 4 mM of hydrogen peroxide was set to 100%. The values are presented as the average of at least three replicates. When 8 mM of hydrogen peroxide was used in the ABTS activity assay, the WT retained around 53% activity while the less stable mutants (M253L and M253F) retained around 47% activity. Higher concentrations of hydrogen peroxide had an effect on the activity of all variants of DyP4. At 20 mM of hydrogen peroxide (the highest concentration used in this experiment), all variants retained less than 27% of activity, with no significant variation among them.

## 4.4 Discussion

Inactivation by hydrogen peroxide is a major limitation of the utilization of haem peroxidase for large-scale industrial applications (Martinez et al., 2009). The oxidation stability against hydrogen peroxide has been increased for certain peroxidases using site-directed mutagenesis, namely MnP (Miyazaki and Takahashi, 2001), VP (Bao et al., 2014), and DyP (Ogola et al., 2010), in which the substitution, deletion, or insertion of amino acids more resistant to oxidation is typically performed. Here, all three methionine residues were replaced with residues that are more resistant to oxidation by hydrogen peroxide to investigate its effect and whether this approach can improve the stability of DyP4 towards hydrogen peroxide. However, the results obtained in this experiment showed that, the substitution of Met residues with those conferring higher resistant to oxidation did not lead to an increase or improvement in hydrogen peroxide stability (Figure 4-4). Specifically, the stability of the WT was virtually the same as that of all of four mutants in which the methionine residues were replaced.

It remains unknown whether the location of the three methionine residues is close to the haem and active site, since the template model used to build the 3D structure of DyP4 shares less than 50% of its identity. At the same time, and it can be seen in Figure 4-2, the methionine residues displayed on the 3D model might not be located in the haem area, particularly M43 and M77, which tend to be located on the surface of the enzyme. In some previously reported studies, the replacement of methionine successfully improved the stability of the peroxidase against hydrogen peroxide, while in other studies; replacement did not lead to an increase in the stability of the peroxidase against increased concentrations of hydrogen peroxide. An obvious effect of the location of the substituted methionine can be observed, as the closer the methionine residue is to the active site and the heme, the higher the probability that the stability of the enzyme will improve after replacing susceptible residues using site-directed mutagenesis (Ogola et al., 2010, Bao et

al., 2014).

Using site directed-mutagenesis, other amino acids that are susceptible to oxidation by hydrogen peroxide can be replaced with amino acids that are more resistant to oxidation. Cys389, Tyr390, and Trp405 are located in the heme area as suggested by the 3D structure model and they could be targeted for site-directed mutagenesis. Alternatively, directed evolution might be adjusted to engineer DyP4 for higher tolerance against hydrogen peroxide (See Chapter 6).

## **Conclusion**

Increasing the stability of peroxidases against high concentrations of hydrogen peroxide is a major challenge to the practical use of peroxidases in industrial applications. While directed evolution of peroxidases has led to improvements in peroxidase stability towards hydrogen peroxide in some studies, rational design involving the replacement of residues susceptible to oxidation with those that are more resistant has been the method of choice. In this work, four mutants were generated in which all three methionine residues of the WT-DyP4 were replaced with residues more resistant to oxidation by hydrogen peroxide. The results show that there was no significant variation between the WT and the four mutants with respect to resistance to oxidation by hydrogen peroxide, and this may be because of the location of the methionine residues on the enzyme. This finding was similar to those of previously reported studies where no improvement was shown for substituted methionine residues that are not close to the haem or the active site.

These mutants are located on the surface of the enzyme, as suggested by the 3D model structure, and this might explain why no improvement in DyP4 stability against hydrogen peroxide stability was observed using this approach in the present work. The use of directed evolution to increase the resistance of DyP4 against inactivation by hydrogen peroxide is a potential alternative, as well targeting other amino acids that are easily oxidised by hydrogen peroxide and replace them with amino acids that are more resistant to oxidation.

# **CHAPTER 5 Oxidation of a non-phenolic lignin model substrate, three S-type lignin units, and synthetic dye by DyP4**

5.1 Introduction

5.2 Methodology

5.3 Results and Discussion

5.4 Summary

## 5.1 Introduction

The use of biorefinery systems to produce biofuels and biochemicals from lignocellulose has been a focus of extensive research in recent years. Lignocellulose as a source of energy does not compete with agricultural food production, and is abundant in nature and generally considered a general waste. One critical step in this process is the pretreatment of lignocellulose to break down lignin and enable accessibility to fermentable sugars. In recent decades, different approaches have been investigated in an attempt to convert lignin to allow the significant conversion of lignocellulose-containing sugars. The optimal approach needs to be economically feasible to achieve the successful conversion of lignocellulose into useful products such as fuels and chemicals at low cost (Agbor et al., 2011, FitzPatrick et al., 2010, Min et al., 2015, Min et al., 2013). Lignin is second to cellulose, the most prevalent biopolymer and renewable source of carbon on earth. There is an urgent need for the complete degradation of lignin, as the majority is wasted because no effective method for lignin conversion is available (Shrestha et al., 2017). Naturally occurring enzymes found in certain fungi and bacteria depolymerize lignin efficiently. Fungal laccases and peroxidases such as lignin peroxidases (LiP), manganese peroxidases (MnP), and versatile peroxidases (VP), have been identified as lignin depolymerizers as reported in several studies. In addition to these enzymes, the newly discovered family of haem peroxidases, the dye-decolorizing peroxidases (DyPs), have been found to oxidize lignin model substrates (Janusz et al., 2017, Pollegioni et al., 2015).

### **5.1.1 Veratryl alcohol**

Due to its simplicity, the oxidation of veratryl alcohol (VA) (chemical structure; Table 5-1) by peroxidase is a widely used assay to determine the potential of biocatalysts to degrade lignin. Peroxidases catalyse the oxidation of VA to veratraldehyde, which can be detected by UV spectrophotometry at a wavelength of 310 nm. Given its simplicity, the VA assay is the method of choice for the screening of lignin degradation (Arora and Gill, 2001, Min et al., 2015). The enzymatic oxidation of VA has been studied using different microbial enzymes, either with direct oxidation by the enzyme or using mediators. For instance, the effect of mediators for MnP and laccases was studied using non-phenolic compounds which are not naturally oxidized by these enzymes. However, the addition of mediators was found to enhance oxidation reactions to non-phenolic substrates such as VA (Nousiainen et al., 2012). The oxidation of VA by fungal LiP, VP, and DyPs, along with some peroxidases from plants, occurs under acidic pH conditions; specifically, under pH 5 (Liers et al., 2010). VA oxidation testing for different peroxidases has been reported in several studies (Yang et al., 2011, van Bloois et al., 2010, Ruiz-Duenas et al., 2009, Min et al., 2015).

### **5.1.2 Reactive Blue 19 (RB19)**

Along with azo dyes, anthraquinone dyes are used in significant quantities in the textile industry. Due to their recalcitrance and resistant to biodegradation, these dyes accumulate and are an important environmental concern (Ogola et al., 2015). DyPs are capable of decolourizing anthraquinone and azo dyes along with MnP, LiP, VP, and laccases (Kornilowicz-Kowalska and Rybczynska, 2014). See Table 5-1 for the chemical structure of RB19.

### **5.1.3 S-type lignin units**

White-rot fungi decompose lignin via oxidation mechanisms in which the arylether ( $\beta$ -O-4) linkages are cleaved alongside the degradation of p-hydroxyphenyl (H), guaiacyl lignin



units (G), and syringyl (S) lignin units. As a result, some phenolic compounds such as acids, aldehydes, and ketones are released, and can act as redox mediators. Mediators can be described as small molecules that, after being oxidized by enzymes, diffuse from the oxidation site to other regions where they can oxidize other substrates. The oxidation of mediators can have tremendous advantages because (i) compounds that are not oxidized by enzymes can be oxidized in this manner by the mediators, and (ii) the mediators are small molecules that diffuse to oxidize large substrates that enzymes cannot reach (Baiocco et al., 2003).

Less than three decades ago, extensive work started to find synthetic mediators and justify their enzymatic oxidation mechanisms. Violuric acid (chemical structure; Table 5-1), is among the most effective synthetic mediators for laccases that, upon oxidation, can oxidize non-phenolic lignin structures (Canas and Camarero, 2010). Being a synthetic redox mediator of the NOH type, violuric acid has been used widely to enhance the activity of laccases (Pogni et al., 2007). The enzymatic oxidation of violuric acid produces iminoxyl radicals that are highly stable, and its purple colour is detectable in the visible spectrum (Kim et al., 2003, Pardo et al., 2013). Violuric acid has been used as a reporter assay for mutants generated for laccases by directed evolution. Interestingly, screening results have shown that improvements in mutants identified by screening with natural mediators such as syringaldehyde and acetosyringone could be detected by violuric acid assay (Pardo et al., 2013).

#### **5.1.4 Aim and objectives**

Spectrophotometric assays were used in this study based on oxidation by a Pleos-DyP4 peroxide enzyme. These assays can be used for the phenolic compounds of the S-type lignin unit, syringaldehyde, acetosyringone, and sinapic acid and can screen mutants of DyP4 created by directed evolution methods. Violuric acid substrate was selected to represent synthetic mediators and can be used in a similar way to natural mediators of S-type lignin compounds. Furthermore, Reactive Blue 19 (RB19) and veratryl alcohol (VA) substrates were selected to represent AQ dyes and non-phenolic compounds, respectively. We sought to develop a simple, rapid, and reliable assay for lignin degradation. All chemical structures for substrates used in this chapter are shown in (Table 5-1).

## **5.2 Methodology**

### **5.2.1 DyP4 expression and purification**

DyP4-tag and DyP4 with tag removed were expressed in BL21 (DE3) cells in 400 mL of 2× TY media induced with 1 mM IPTG. The collection, lysis, and purification of cell pellets was as described in Chapter 2 (section 2.8).

### **5.2.2 Oxidation assays**

#### **VA oxidation assay**

The oxidation of VA by peroxidases varies in pH requirements, with LiP and some DyPs oxidizing VA only in acidic environments of between pH 2 and 5 for LiP and pH 1.2 and 4.5 for DyPs (Hofrichter et al., 2010). VA was oxidized to veratryl aldehyde to determine the potential of lignin breakdown. The reaction mixture was prepared in a 1 mL volume containing 2.0 mM of VA and 25 µl of diluted purified enzyme (DyP4), and the reaction was started by addition of 4 mM of H<sub>2</sub>O<sub>2</sub> (50 µL). The oxidation product veratryl aldehyde ( $\epsilon_{310}$ , veratryl aldehyde = 9.3 mM<sup>-1</sup>cm<sup>-1</sup>) was detected by UV spectrophotometry at 310 nm.

#### **RB19 oxidation assay**

An RB19 substrate solution was prepared in 0.1 M citric acid-0.2 M Na<sub>2</sub>HPO<sub>4</sub> buffer with pH 3.4. The oxidation assay was performed in a 1 mL volume containing 200 µM of RB19 and 25 µl of diluted enzyme, and the reaction was initiated by the addition of 4 mM of H<sub>2</sub>O<sub>2</sub> (50 µL). Decolourization of RB19 by DyP4 was monitored as a decrease in absorbance at 595 nm by UV spectrophotometry over 5 minutes.

### **Sinapic acid oxidation assay**

To prepare a stock of 10 mM of sinapic acid, 22.4 mg of sinapic acid was dissolved in 10 mL ethanol. Changes in the UV-visible spectra of sinapic acid oxidation by DyP4 were monitored by spectrophotometry in 0.1 M citric acid-0.2 M Na<sub>2</sub>HPO<sub>4</sub> buffer at pH 4.0. The reaction mixture was prepared in a cuvette in a total volume of 1 mL, containing 250 μM of sinapic acid, and 25 μL of diluted DyP4, and 4 mM hydrogen peroxide (50 μL) was added to initiate the oxidation reaction.

### **Acetosyringone oxidation assay**

A stock concentration of 80 mM acetosyringone was prepared by dissolving 157 mg of acetosyringone in 10 mL of ethanol. Changes in the UV-Vis spectra of acetosyringone oxidation by DyP4 in 0.1 M citric acid-0.2 M Na<sub>2</sub>HPO<sub>4</sub> buffer at pH 4.0 were observed by the spectrophotometer. In a 1 mL reaction mixture, the oxidation of 2 mM acetosyringone by DyP4 (25 μL diluted enzyme) was initiated by the addition of 4 mM of hydrogen peroxide.

### **Optimization of pH for acetosyringone**

To determine the optimum pH for acetosyringone oxidation by DyP4, different values of pH were evaluated for the 0.1 M citric acid-0.2 M Na<sub>2</sub>HPO<sub>4</sub> reaction buffer. Reaction buffers with different pH values were prepared as described in the <https://www.sigmaaldrich.com/life-science/core-bioreagents/biologicalbuffers/learning-center/buffer-reference-center.html>. Oxidation testing to determine the optimum pH was performed at a wavelength of 520 nm using 2 mM of substrate.

### **Syringaldehyde oxidation assay**

A stock concentration of 80 mM syringaldehyde was prepared by dissolving 145.7 mg of syringaldehyde in 10 mL of ethanol. Changes in the UV-Vis spectra of acetosyringone oxidation by DyP4 in 0.1 M citric acid-0.2 M Na<sub>2</sub>HPO<sub>4</sub> buffer at pH 4.0 were observed by the spectrophotometer. The reaction mixture was prepared in a cuvette at a total volume of 1 mL, containing 2 mM syringaldehyde and 25 µl diluted DyP4, and 4 mM of hydrogen peroxide (50 µL) was added to start the reaction.

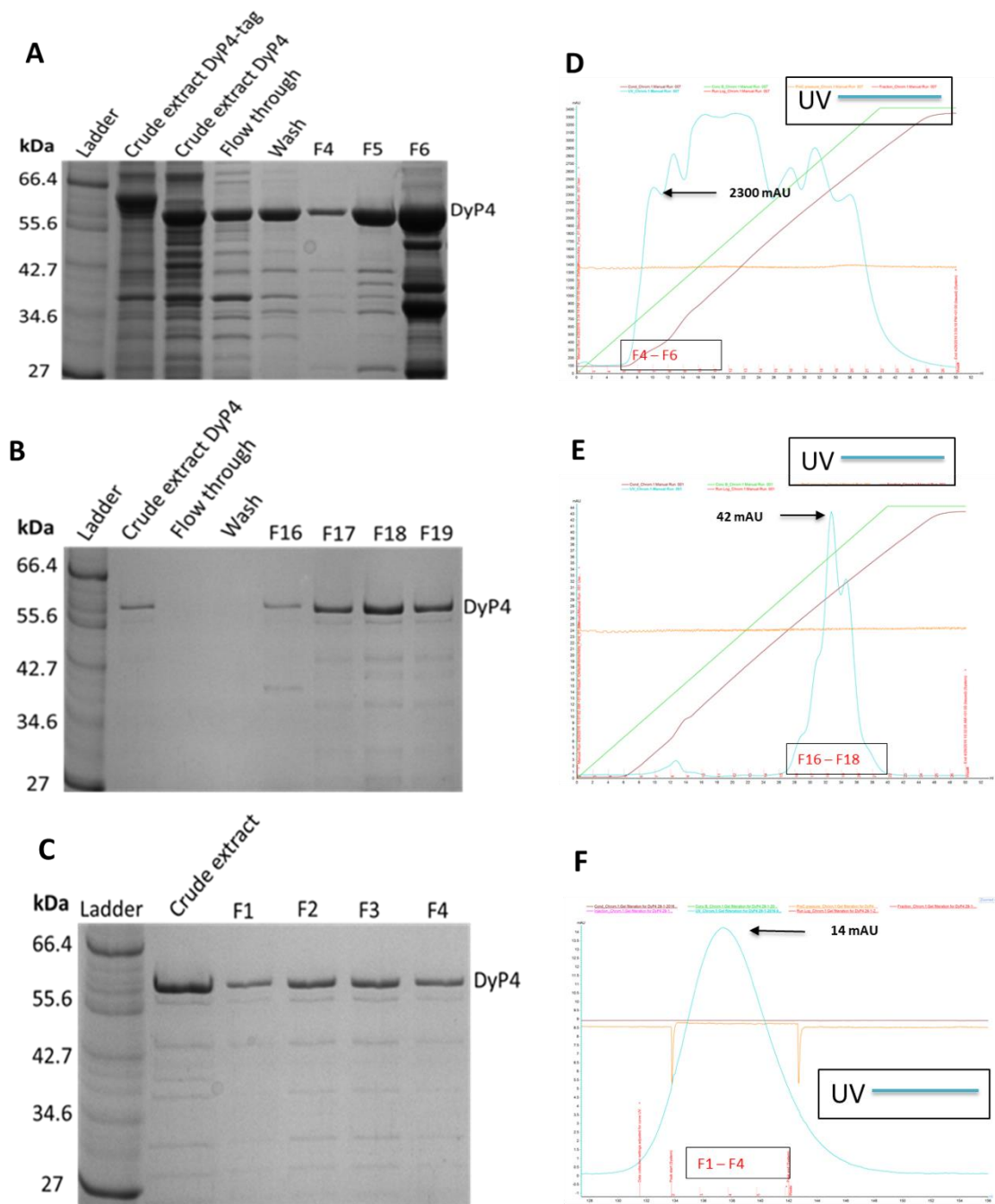
### **Violuric acid oxidation assay**

A stock solution of 800 mM violuric acid was prepared by dissolving 1.4 g of this substrate in 10 mL methanol. The reaction mixture was prepared in 0.1 M citric acid-0.2 M Na<sub>2</sub>HPO<sub>4</sub> buffer, pH 4.0. The oxidation reaction for violuric acid was conducted in a cuvette in a total volume of 1 mL, containing 20 mM of violuric acid and 25 µL of diluted enzyme, with 50 µL (4 mM) of hydrogen peroxide. Changes in the UV-Vis spectra were monitored by UV spectrophotometry.

For all four of the assays above, oxidation was carried out in a spectrophotometer for 60 minutes or more, and absorbance was measured at 0 minutes and at 10-minute intervals thereafter.

## 5.3 Results and discussion

### 5.3.1 Protein expression and purification of DyP4



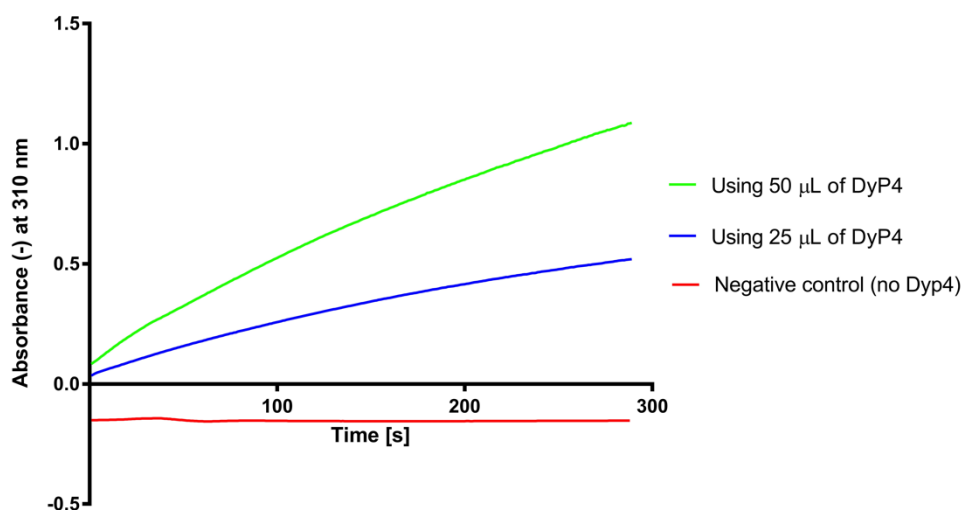
**Figure 5-1:** Purification of DyP4 using three steps of chromatography. (A) Q5 column (1<sup>st</sup> step) (A) lane 1; molecular ladder, lanes 2–5; crude extract of DyP4-tag, crude extract of DyP4, flow-through, and wash, respectively, lanes 6–8; whole soluble fractions of DyP4. SP column (2<sup>nd</sup> step) (B) lane 1; molecular ladder, lanes 2–4; crude extract, flow-through, and wash, respectively, lanes 5–8; whole soluble fractions of DyP4. Superdex 75 column (3<sup>rd</sup> step) (C), lane 1; molecular ladder, lane 2; crude extract, lanes 3–6; whole soluble fractions of DyP4. (D), (E), and (F) shows the chromatogram of the eluted fractions in the first, second, and third step of DyP4 purification respectively.

Following the successful removal of the C-terminal peptide tag from DyP4, as confirmed by SDS-PAGE and gel electrophoresis (data not shown), pET24a-DyP4 was used for small-scale expression in BL21 (DE3) in 50 mL 2× TY media containing kanamycin. An ABTS assay was used to screen activity in the crude extract, which showed normal enzymatic activity for the DyP4 enzyme. Large-scale production was then performed for DyP4; cells were grown for 2 hrs at 25°C until OD<sub>600</sub> reached 0.6 in 400 mL of 2× TY media, induced with 1 mM IPTG, and cultured for a further 24 hrs. Centrifugation at 6000 rpm for 5 min at 4°C was performed to harvest the cells. Pellets were lysed by sonication in a lysis buffer containing 50 mM Tris-HCl (pH 8.5) and 1 mM EDTA. Sonication was performed for 5 min using the following program: 10 sec on, 20 sec off, 70% amplitude. Centrifugation for 40 min at 8000 rpm and 4°C was carried out to remove cell debris. Supernatants were filtrated using a Whatman<sup>®</sup> Puradisc filter of pore size 0.45 µm and loaded onto a HisTrap Q5 column for purification using an AKTA Pure system. The process for purifying DyP4 was carried out in three steps, as described in Chapter 2 (section 2.7). Following ion-exchange chromatography using a HiTrap Q column, a concentrated band of DyP4 at 59 kDa was observed (Figure 5-1), as was the presence of impurities. After the second step of purification using a HiTrap SP column, most of impurities had been removed, although small quantity remained. The final step of purification used a HiLoad 26/600 Superdex 75 pg gel filtration column, and slight bands represented the remaining impurities.

### **5.3.2 Enzymatic oxidation of VA by DyP4**

The VA assay is the assay of choice for lignin-degrading enzymes such as LiP (Min et al., 2015). DyP4 was found to catalyse the oxidation of the non-phenolic lignin substrate (VA), as shown in Figure 5-2. An increase in absorbance at 310 nm was observed after initiating the oxidation of VA by the addition of hydrogen peroxide. Different

concentrations of purified diluted DyP4 were used and, as the concentration of DyP4 increased in the reaction, the absorbance increased accordingly. To validate the data and prevent false positive results, different reactions were carried out for the negative control; for example, a reaction without the addition of DyP4, using PBS buffer instead. Furthermore, in another control, DyP4 was added to the reaction but PBS buffer was added instead of hydrogen peroxide. For all the negative controls, no increase in absorbance was seen at a wavelength of 310 nm. For all reaction mixtures, including the negative control, at least three replicates were performed and the average is plotted (Figure 5-2).

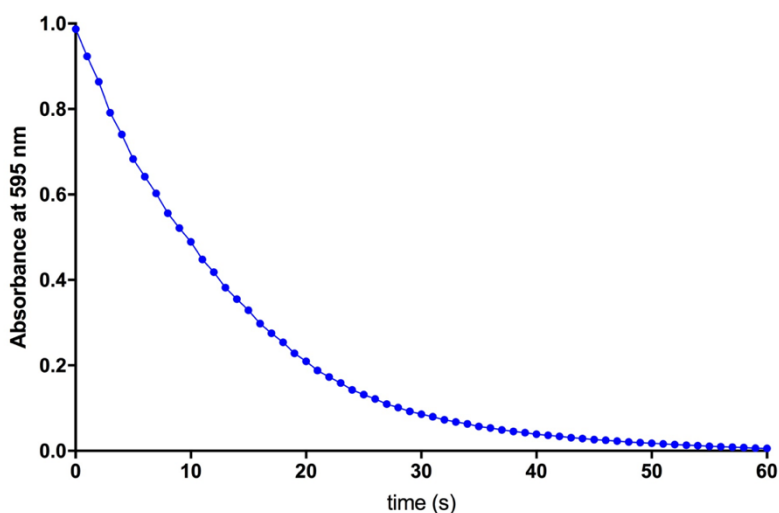


**Figure 5-2:** Increase in absorbance at 310 nm due to the oxidation of VA by DyP4. A total of 1 ml reaction mixture contained 2.0 mM of VA and 25 µl or 50 µl of diluted purified enzyme (DyP4), and 4 mM of H<sub>2</sub>O<sub>2</sub> (50 µL). Different volumes of DyP4 (50 µL, blue line and 25 µL, green line). Negative control, red line. Each reaction was carried out in at least three replicates and the average is presented.



### 5.3.3 Decolourization activity of RB19 by DyP4

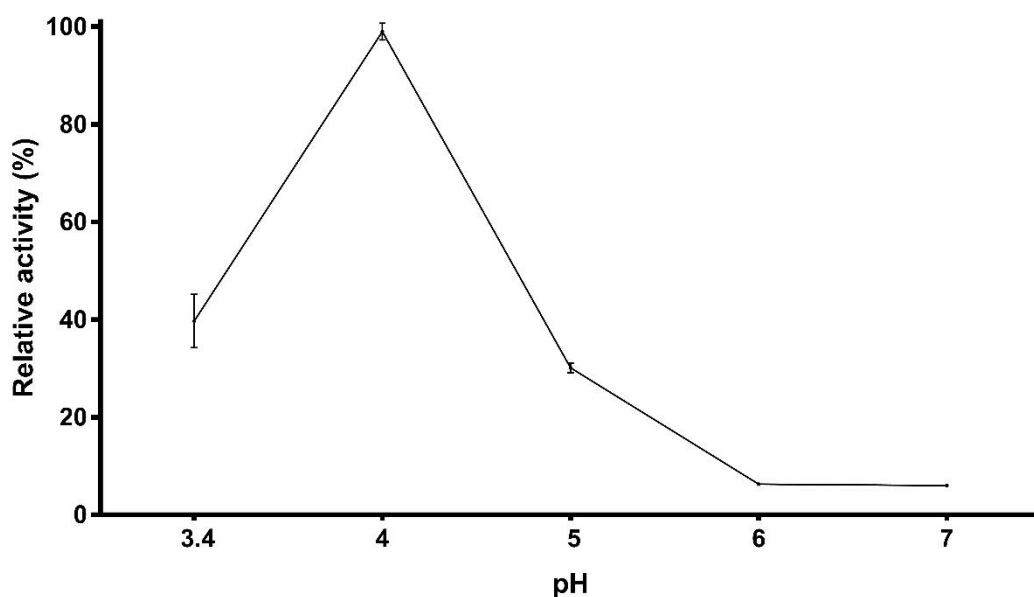
Dye decolourization of RB19 by DyP4 was monitored by UV-vis absorption spectroscopy. The reaction mixture was started by the addition of hydrogen peroxide to the diluted enzyme (DyP4) mixed with the optimal concentration of RB19 as reported elsewhere. The degradation of RB19 was monitored by measuring the decrease in absorbance at 595 nm for 5 minutes as shown in Figure 5-4. The activity of DyPs is generally affected by pH; here, the decolourization reaction was performed at pH 3.4, the optimal pH as described by Fernandez et al. (Fernandez-Fueyo et al., 2015) where the maximum decrease in absorbance was observed at around pH 3.5. This assay was also used to investigate the effect of increased concentrations of hydrogen peroxide on the activity of DyP4, along with ABTS and VA assays. The RB19 assay was also used to determine whether or not the addition of a C-terminal tag to DyP4 caused a loss of activity along with the ABTS assay (Chapter 3).



**Figure 5-3:** Decrease in absorbance at 595 nm due to the decolourization of RB19 by DyP4.

### 5.3.4 Effect of pH on Pleos-DyP4 activity with acetosyringone

To examine the effect of pH on the activity of DyP4 with acetosyringone, a range of pH values was used to carry out the reactions, namely pH 3.4, 4, 5, 6, and 7. As shown in Figure 5-4, the oxidation of acetosyringone was optimal at pH 4, and only 40% residual activity remained at pH 3.4 and 30% at pH 5. No activity for DyP4 was seen at pH 6 and pH 7 with the acetosyringone substrate.



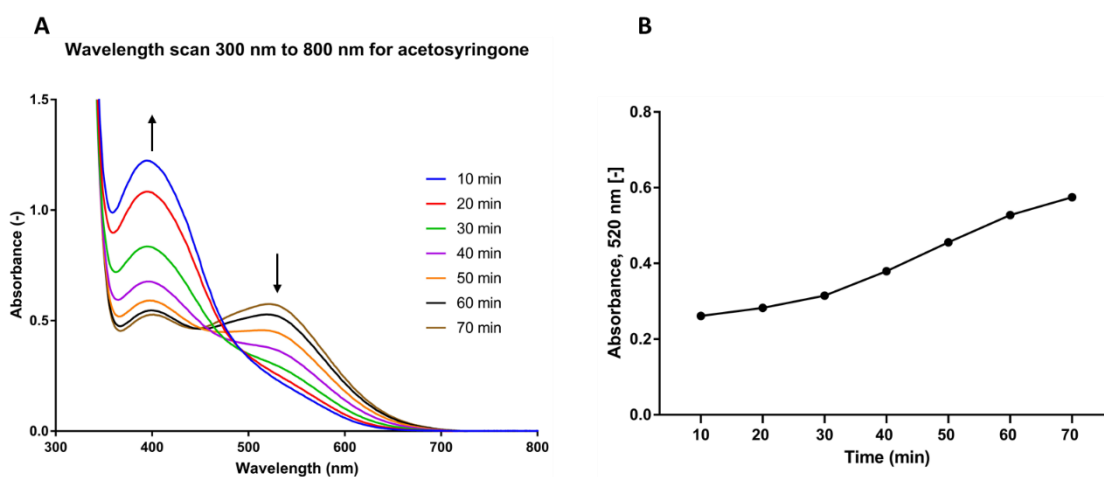
**Figure 5-4:** pH activity profile of DyP4 with acetosyringone.

### 5.3.5 Enzymatic oxidation of S-type lignin models

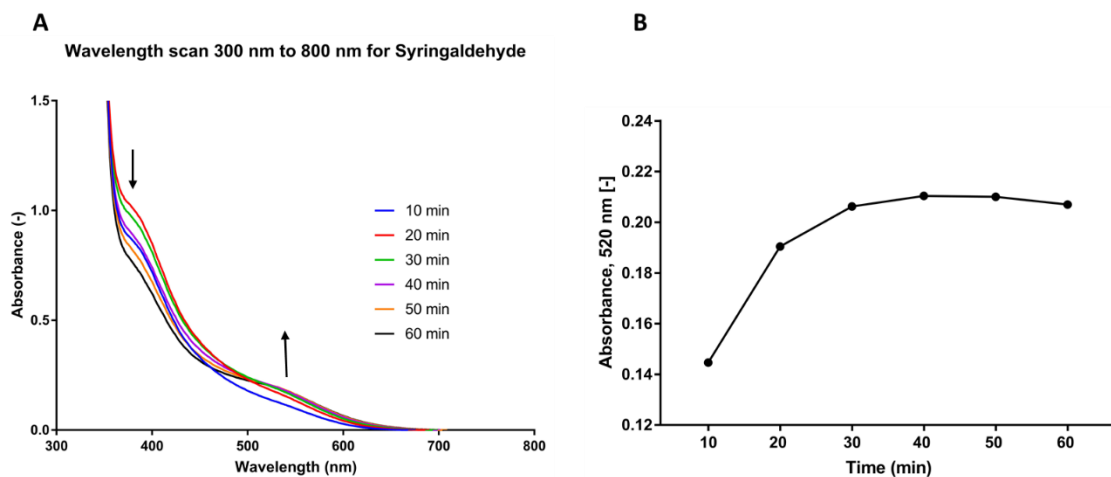
Three S-type phenolic lignin compounds, syringaldehyde, acetosyringone, and sinapic acid, were used to develop spectrophotometric assays with DyP4 peroxide enzyme. DyP4 successfully oxidized these S-type compounds, as shown in this study. The results obtained for the oxidation of these compounds were similar to those of other reported studies in the literature (Pardo et al., 2013). As can be seen in the figures 5-5 – 5-8, the oxidation of syringaldehyde, acetosyringone, and sinapic acid resulted in the appearance of different peaks of absorbance in the UV-visible spectrum. In the case of acetosyringone, for instance, upon oxidation by DyP4, the appearance of a yellow-coloured product was observed and associated with an increase in absorbance at 390 nm after 10 min of initiating the oxidation reaction. This yellow-coloured product began to turn red after 20 min and, as can be seen in Figure 5-6, the maximum absorbance shifted from 390 nm to 520 nm. The absorbance at 520 nm continued to increase over 70 min, and the red colour was maintained for several hours. In contrast to acetosyringone, syringaldehyde oxidation caused an increase in absorbance at 390 nm due to the formation of the stable yellow product (Figure 5-7). The colour of this yellowish product persisted for longer than that in the case of acetosyringone oxidation, and the slightly red-coloured product developed after 30 min. The oxidation of sinapic acid by DyP4 caused the formation of a pinkish product, and a peak at around 500 nm can be seen in Figure 5-8 with a maximum absorbance at 505 nm. Phenoxy radicals of sinapic acid are rich in  $\beta$ - $\beta'$  coupling, and following sinapic acid oxidation by the enzyme, dehydrodisinapic acid lactones are formed and are subsequently oxidized by the enzyme (Canas and Camarero, 2010). Enzymatic oxidation of this dilactone results in the formation of one or more products where the absorbance at approximately 500 nm can be observed spectrophotometrically (Camarero et al., 2008). The oxidation of violuric acid led to the formation of purple-coloured products known as iminoxyl radicals. These radicals are

highly stable, and caused the appearance of a peak as shown in figure 5-9, with a maximum absorbance at 521 nm.

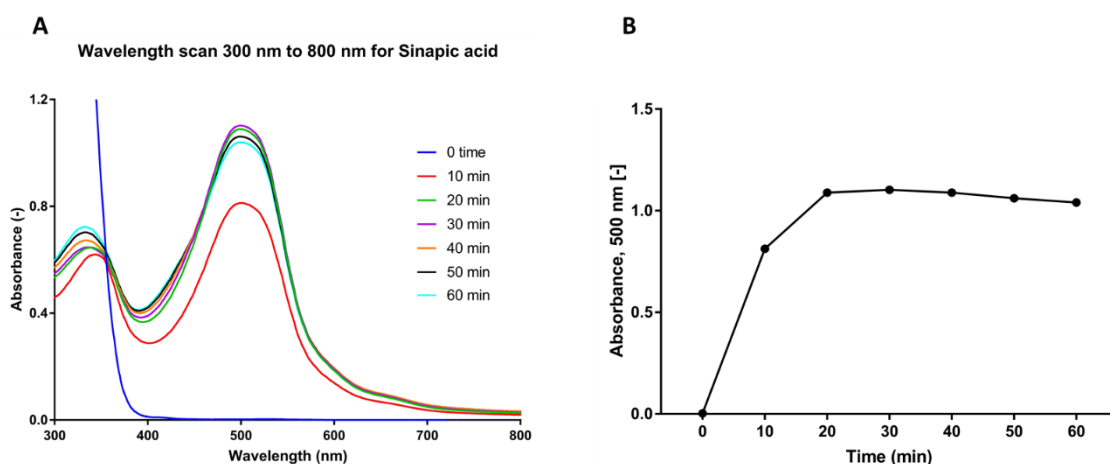
The design of enzymatic screening assays using these compounds is of significant biotechnological value. This is mainly because (i) these compounds are natural redox mediators which can oxidize substrates that cannot reach the enzymatic oxidation site due to their size, (ii) enzymatic oxidation of these redox mediators produces coloured products that can be detected in the visible spectrum (Camarero et al., 2012), and (iii) some mediators link lignin with carbohydrates in the cell wall of certain plant species. The development of enzymes with a higher efficiency of lignocellulose breakdown might reveal the means and/or lead to the evolution of enzymes with higher efficiency to decompose lignocellulose (Camarero et al., 2008, Camarero et al., 2007, Pardo et al., 2013).



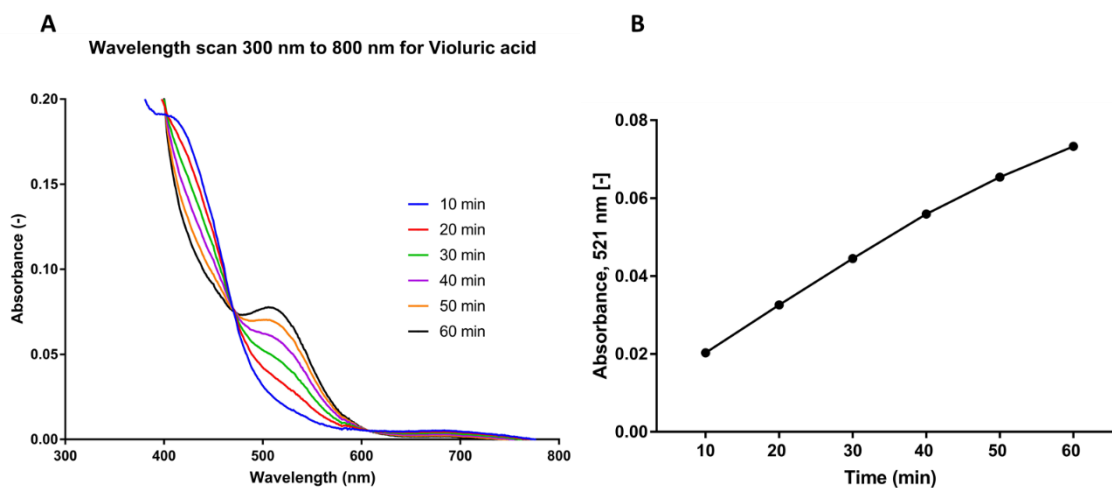
**Figure 5-5:** Oxidation spectra of acetosyringone. (A) Oxidation of acetosyringone by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 520 nm over 60 min time.



**Figure 5-6:** Oxidation spectra of syringaldehyde. (A) Oxidation of syringaldehyde by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 520 nm over 60 min time.

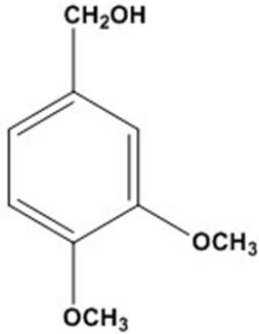
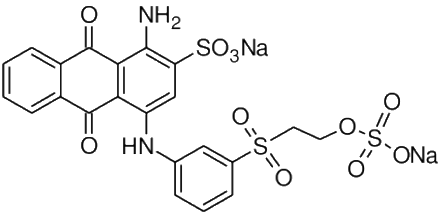
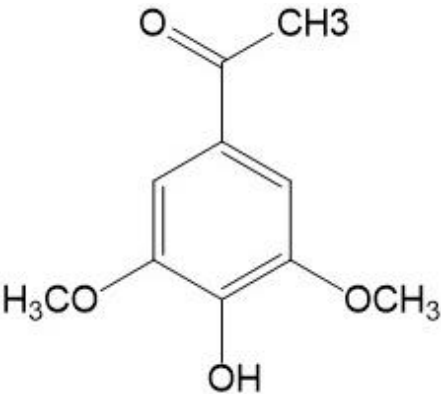


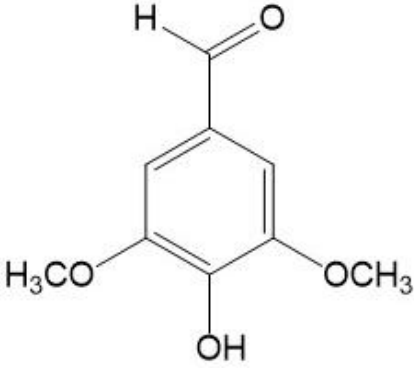
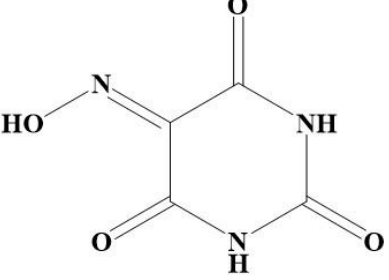
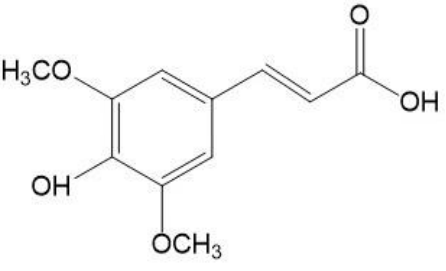
**Figure 5-7:** Oxidation spectra of sinapic acid. (A) Oxidation of sinapic acid by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 500 nm over 60 min time.



**Figure 5-8:** Oxidation spectra of violuric acid. (A) Oxidation of violuric acid by DyP4 caused changes in the UV-vis spectra. (B) Increase in the absorbance at wavelength 521 nm over 60 min time.

**Table 5-1:** Substrates used in this study

Substrate	Type	Chemical structure	$\lambda_{\text{max}}$ /nm
VA	Non-phenolic	 <chem>OCC1=CC=C(OC)C(OC)=C1</chem>	310 nm
RB19	Anthraquinone dyes	 <chem>NC1=CC=C(S(=O)(=O)[Na])C=C2C(=O)C(=O)C3=CC=CC=C3C2=O1NS(=O)(=O)CCOS(=O)(=O)[Na]</chem>	595 nm
Acetosyringone	S-type lignin units	 <chem>CC(=O)C1=CC=C(O)C(OC)=C1OC</chem>	520 nm

Syringaldehyde	S-type lignin units		370 nm
Violuric acid	Synthetic mediators		521 nm
Sinapic acid	S-type lignin units		503 nm



## **5.4 Conclusion**

The oxidation of VA by DyP4 is reported in this study, and supports previous findings for DyPs in general along with those for LiP and VP, which can directly oxidize VA, unlike Lac and MnP, which depend upon mediators for this purpose (Qin et al., 2017). These results can be used to extend the substrate spectrum of DyP4 and might help to reveal the natural role of DyPs. The oxidation of S-type lignin units generates coloured products which can be used to validate HTS assays to evolve DyP4 (Chapter 6), as well as to extend the enzymatic activity of DyP4 to reach remote substrates from catalytic sites. The enzymatic oxidation of these natural and artificial mediators can be exploited to establish and validate HTS assays for the evolution of DyP4 towards the higher catalytic conversion of plant biomass.

# **CHAPTER 6 Engineering of DyP4 using an OsmY-based secretion mechanism to facilitate the directed evolution approach in bacteria**

This chapter is written for a manuscript and will be submitted for publication.

Abdulrahman HA Alessa, David Gonzalez Perez, Hossam EM Omar Ali, Alex Trevaskis, Kang Lan Tee, Xu, Jun and Tuck Seng Wong. 2019. Engineering of DyP4 using an OsmY-based secretion mechanism to facilitate the directed evolution approach in bacteria. *Bioresources and Bioprocessing*.

## 6.1 Introduction

## 6.2 Methodology

## 6.3 Results

## 6.4 Discussion

## 6.5 Summary

## **6.1 Introduction**

Extensive studies focused on lignin-degrading enzymes from white-rot and brown-rot fungi have been performed since 1980s. Recently, bacterial peroxidases and laccases have also been investigated. Heme peroxidases such MnP, LiP, VPs, and DyPs, along with laccases, are the best-characterized lignin-degrading enzymes (Lambertz et al., 2016). Naturally occurring laccases and peroxidases are not efficient enough for industrial utilization, and usually require modification and further improvement of their characteristics. Protein engineering approaches are typically used to obtain biocatalysts with efficient properties (Martinez et al., 2009).

### **6.1.1 Protein engineering**

#### **Directed evolution**

Directed evolution is an efficient tool to modify biocatalysts and improve their characteristics to make them efficient enough to meet industrial process requirements. In this method, a gene of interest undergoes repeated cycles of (i) construction of a gene library randomly, (ii) expressing the mutants in a suitable expression system (e.g. *E. coli* or *C. cerevisiae*), and (iii) screening of mutants using an efficient assay to identify those with properties of interest (Turner, 2009).

#### **Error-prone PCR**

Error-prone PCR (*epPCR*) methods are among the most widely used approaches in protein engineering studies due to their simplicity. Using this method, it is possible to insert a mutation randomly at any position on the gene of interest. Once established, PCR can be used to amplify lengths of DNA, however it was modified in the form of *epPCR* to create libraries for a gene of interest. The DNA polymerase used in *epPCR* has reduced fidelity to introduce mutations randomly in the gene of interest during replication. Errors in the incorporation of nucleotides are enhanced by changing the normal conditions used

in PCR. In *ep*PCR, unequal concentrations of dNTPs are added to the reaction, as well as  $Mn^{2+}$ , or the addition of increased concentrations of  $Mg^{+2}$  (Porter et al., 2016). See Chapter 1, section 1.5.1 for more details about *ep*PCR.

### **Site-saturation mutagenesis**

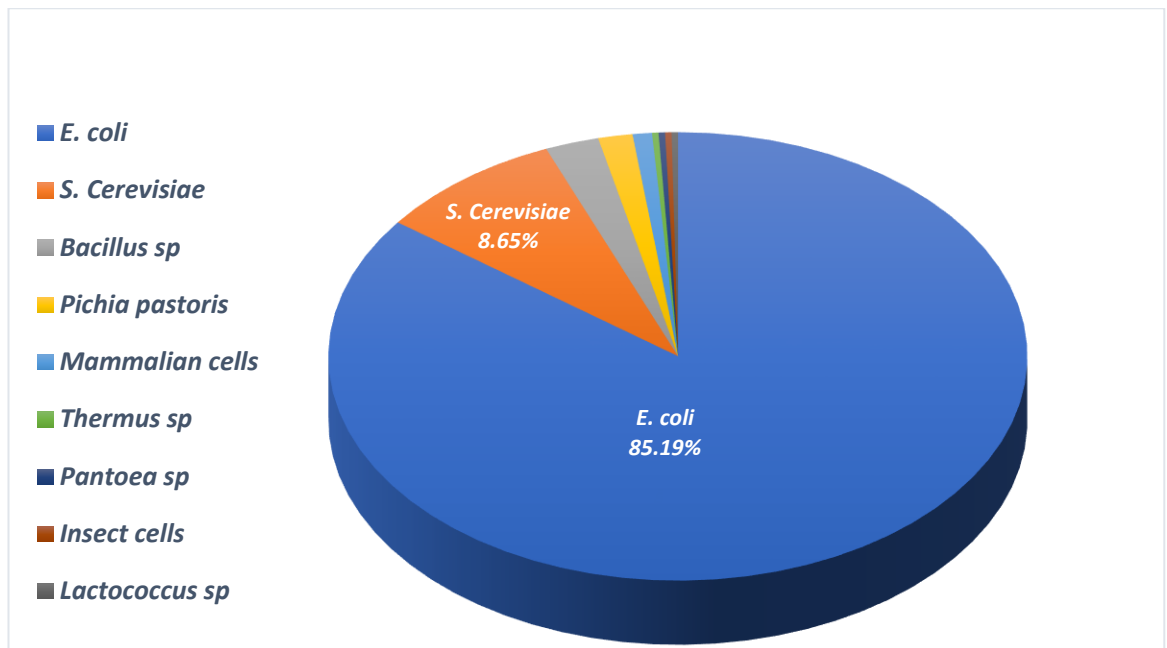
Site saturation mutagenesis (SSM) is another method that is well-established in directed evolution studies of enzymes. SSM is used to introduce a substitution at a single position in the protein of interest and allows the substitution of that amino acid at the given position to all 19 remaining amino acids. Different approaches of SSM have been generated and reported, each with its benefits and drawbacks (Li et al., 2018).

#### **6.1.2 Dye-decolorizing peroxidases (DyPs)**

DyP peroxidases are newly discovered haem-containing peroxidases that share no similarity with animal and plant peroxidases. DyPs utilize hydrogen peroxide as an electron acceptor to catalyse the oxidation of a wide range of substrates. More interestingly, DyPs have been found to play a role in lignin breakdown. DyP4 is an interesting enzyme belonging to class D of the DyP superfamily, and is one of the most thermostable peroxidases identified to date. It is the first fungal DyP found to catalyse the oxidation of  $Mn^{2+}$  to  $Mn^{3+}$  (Fernandez-Fueyo et al., 2015).

#### **6.1.3 *E. coli***

*Escherichia coli*, as a host for the functional expression of proteins, is the ‘gold standard’ from an economical and practical prospective. This is mainly because of the advantages of using bacterial systems for recombinant protein production, specifically using *E. coli*, which is fast, easy, and cheap to culture, with well-studied genetic characteristics. On the other hand, *E. coli* is not without its disadvantages as a host to express proteins from yeasts or fungi, including the lack of a mechanism to secrete proteins into the expression cultural media (Ribeiro et al., 2018).



**Figure 6-1:** Host expression systems used routinely in directed evolution studies. This figure was adapted from (Pourmir and Johannes, 2012).

#### 6.1.4 Secretion of proteins in *E. coli*

In a study conducted by Qian et al to develop a system for the extracellular production of proteins in *E. coli*, a fusion partner named OsmY was found to be the most efficient fusion partner for excretion .(Qian et al., 2008). In this study, OsmY was used as a carrier protein to secrete other proteins into the expression media. The mechanism of excretion is not fully understood; however, the signal peptide is responsible for transporting the protein to the periplasm while mature OsmY is required for secretion into the medium. While this extracellular excretion of protein might have tremendous advantages, such as simplification of the purification process, and it results in the efficient folding of the protein, we intended to use the previously reported OsmY-based secretion mechanism of protein in *E. coli* for the excretion of DyP4 in the *E. coli* strain BL21 (DE3) (Qian et al., 2008). Our aim in the present study was to apply this novel excretion system to directed evolution studies of DyP4. This mechanism could be useful to facilitate the screening process by the secretion of DyP4 into the culture media. In this case, no cell lysis step is required, as screening can be performed using an aliquot of media containing the secreted

enzyme.

Directed evolution is a highly efficient and rapid method which has been applied successfully to improve the stability and catalytic efficiency of heme peroxidases, including HRP, LiPs, VPs, and APO or UPO which have been expressed and evolved in yeast expression systems (Garcia-Ruiz et al., 2012, Gonzalez-Perez et al., 2014, Gonzalez-Perez et al., 2016, Molina-Espeja et al., 2014, Morawski et al., 2001). Recently, a bacterial DyP was evolved in a bacterial expression system with improved catalytic abilities of lignin-related phenols as well as improvements in stability for increased concentrations of hydrogen peroxide (Brissos et al., 2017).

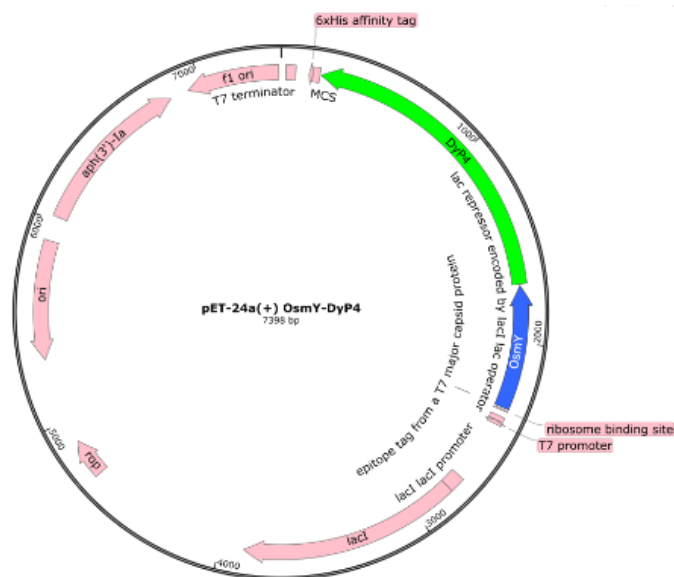
### **6.1.5 Aim**

The following work establishes a protocol for the secretion of DyP4 enzyme in BL21 (DE3) cells in a 96-well microtiter plate, constructed from libraries of DyP4 using *ep*PCR or saturation mutagenesis methods, the production and secretion of mutants in BL21 (DE3), and the screening of libraries using the most widely used assay for the screening of peroxidases, the ABTS assay. An OsmY-based secretion mechanism was used to facilitate the screening process of libraries generated by *ep*PCR or site-saturation mutagenesis. Using this mechanism, no cell lysis step is required as the enzyme is secreted into the media.

## 6.2 Methodology

### 6.2.1 Cloning of OsmY and DyP4 in pET24a

The cloned vector (pET24-a-OsmY-DyP4, Figure 6-2) for the excretion of DyP4 in *E. coli* strain BL21 (DE3) was provided by Dr Wong. Several preliminary experiments were conducted to validate and optimize the secretion of DyP4 into the media of the BL21 (DE3) strain.



**Figure 6-2:** Plasmid map of pET24a-OsmY-DyP4 (created with SnapGene).

### 6.2.2 Optimization of DyP4 screening using the OsmY mechanism

To determine the optimal conditions for the secretion of DyP using the OsmY mechanism, two bacterial *E. coli* strains, C41 (DE3) and BL21 (DE3), were used to transform pET24a-OsmY-DyP4. Expression was performed in 5 mL 2× TY AIM at 30°C or 37°C, and protein samples were screened by ABTS assay after 12, 18, and 24 hrs of expression.



### **6.2.3 Validation and optimization of the ABTS assay as a screening method for DyP4 mutants**

After the design and determination of the optimal conditions of expression in directed evolution studies, it was essential to validate the assay used to screen mutants and ensure that variations among the WT or parental strains were as low as possible. A validation experiment was conducted using the ABTS assay as a screening method. In this experiment, 100 ng/ $\mu$ L of pET24a-OsmY-DyP4 was transformed in BL21 (DE3) cells using the CaCl<sub>2</sub> heat shock method. A single colony was picked to prepare a glycerol stock and DyP4-OsmY-pET24a was pre-cultured in a 96-well microtiter plate. Protein expression and screening was performed for 24 hrs at 30°C and 1050 rpm in 2 $\times$  TY AIM. Centrifugation was performed at 4000 rpm for 10 min, and 20  $\mu$ L of media (secreted enzyme) was then used to perform activity screening to examine the variation in WT using the ABTS assay to validate this assay for directed evolution.

### **6.2.4 Screening of the WT plate**

The WT of DyP4 was screened in a 96-well microtiter plate using the optimized conditions, with 10 mM of ABTS (150  $\mu$ L), 20  $\mu$ L of secreted enzyme, and 4 mM of hydrogen peroxide (50  $\mu$ L). A multichannel pipette was used to transfer 150  $\mu$ L of ABTS solution into the 96-well plate, and 20  $\mu$ L of media (secreted enzyme) was then transferred into the 96-well microtiter plate. Mixing was performed by incubation with shaking of the microtiter plate for 2 min at 1050 rpm. Next, 50  $\mu$ L of hydrogen peroxide was added, and the solution was mixed for 2 minutes by incubation with shaking at 1050 rpm. Absorbance was immediately measured in a plate reader at 405 nm. Following the optimization of the screening assay, directed evolution was applied for DyP4.

## **6.2.5 Directed evolution**

### **First round of evolution**

#### **Digestion of pET24a-OsmY-DyP4 (vector preparation)**

Restriction digestion was achieved in pET24a-OsmY-Dyp4 using *BamHI* and *EcoRI* restriction enzymes. A reaction mixture was prepared using 1× of CutSmart buffer, 3000 ng of pET24a-OsmY-Dyp4, and 1 U of *BamHI* and *EcoRI* enzymes, and the mixture was adjusted to 100 µL using deionized water. The reaction mixture was incubated overnight at 37°C, followed by gel extraction, and an appropriate volume was used for ligation with the product of DyP4 from *epPCR*.

#### **Construction of mutants using *epPCR***

In this study, three different mutagenic conditions were used as shown below.

##### **Low mutagenic condition (1)**

In this condition, the *epPCR* mixture contained 1.5 mM of MgCl<sub>2</sub>, 0.01 mM of MnCl<sub>2</sub>, 0.3 mM of dNTP mix, 4.5 pmol of forward and reverse primers, 3.5 ng/µL of DNA template, and 1.25 U of *Taq* DNA polymerase (NEB).

##### **Medium mutagenic condition (2)**

In this condition, the *epPCR* mixture contained 7 mM of MgCl<sub>2</sub>, 0 mM of MnCl<sub>2</sub>, 0.2 mM of dATP, 0.2 mM of dGTP, 1 mM of dTTP, 1 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/µL of DNA template, and 1.25 U of *Taq* DNA polymerase (NEB).

##### **High mutagenic condition (3)**

In this condition, the *epPCR* mixture contained 7 mM of MgCl<sub>2</sub>, 0.05 mM of MnCl<sub>2</sub>, 0.2 mM of dATP, 0.2 mM of dGTP, 0.2 mM of dTTP, 0.2 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/µL of DNA template, and 1.25 U of *Taq* DNA polymerase (NEB).

**Table 6-1:** Overview of *ep*PCR programs used for library construction

Step	Temperature	Time
Initial denaturation	95°C	30 sec
30 cycles	95°C	20 sec
	68°C	30 sec
	68°C	1 min and 30 sec
Final extension	68°C	5 min
Hold	8°C	∞

**Gel extraction and PCR purification**

DNA gel electrophoresis was performed on the PCR products. For clean PCR products, *DpnI* digestion was performed overnight at 37°C, while for PCR products with side products, gel extraction was performed. For all PCR products from the three different mutagenic conditions, restriction digestion with *BamHI* and *EcoRI* was performed followed by ligation with pET24a-OsmY.

**Ligation**

A ligation mixture was prepared in a total volume of 20  $\mu$ L containing the components shown in Table 6-2.

**Table 6-2:** Reaction mixtures of ligation in the first round of evolution

Component	Volume ( $\mu$ L)			
	Low	Medium	High	Control
Water	9.7	10.1	9.8	11.5
T4 DNA ligase reaction	2	2	2	2
pET24a-OsmY (50 ng)	4.5	4.5	4.5	4.5
PCR product (X)	1.8	1.4	1.7	0
T4 DNA ligase	2	2	2	2
Total	20	20	20	20

**Transformation in BL21 (DE3)**

Transformation by the CaCl<sub>2</sub> heat shock method was performed for the WT and by the electroporation method for ligation mixtures of libraries. Electroporation was performed using 1  $\mu$ L of each of the non-purified ligation mixtures, as the transformation efficiency was more than enough to obtain the required colonies for each mutagenic condition. Thus,

no purification of ligation mixtures was performed. Detailed protocols for these transformation processes are given in Chapter 2 (section 2.4).

### **Expression of libraries**

Colonies were picked individually using sterile toothpicks and cultured in sterile 96-well microtiter plates containing 150  $\mu$ L of 2 $\times$  TY media containing kanamycin per well. Well numbers B2, E6, and H11 were inoculated with the WT or parental strain as a standard. Plates were wrapped with silicon tape and incubated at 30°C at 1050 rpm for 24 hrs. Pre-culturing was performed for all libraries using the replicator and grown at 30°C at 1050 rpm for 18 hrs, and master plates were prepared by the addition of 100  $\mu$ L of 50% sterile glycerol to each well before plates were stored at -80°C. Protein expression was initiated in 96-well microtiter plates containing 150  $\mu$ L of 2 $\times$  TY AIM with kanamycin for 24 hrs at 30°C and 1050 rpm. Microtiter plates were centrifuged at 4000 rpm for 10 min, and 20  $\mu$ L or 10  $\mu$ L of media (secreted enzyme) was then used in activity screening of mutants and the WT or parental strain at wavelength 405 nm using a spectrophotometer.

### **Large-scale expression and screening of the improved mutants and the WT/parental strain**

After each round of evolution, all possible improved mutants were characterized further by expression on a larger scale and the normalization of activity with optical density. For these mutants, plasmids were isolated from BL21 (DE3) cells and transformed into DH5 $\alpha$ , and plasmids were isolated and re-transformed in BL21 (DE3) cells using the CaCl<sub>2</sub> heat shock method. Expression was carried out in 50 mL 2 $\times$  TY AIM with kanamycin in a 250-mL flask at 30°C with shaking at 200 rpm for 24 hrs. Samples of secreted protein (1 mL) were collected after 16, 20, and 24 hrs and the optical density at OD<sub>600</sub> was measured for all samples in triplicate. Screening of mutants along with the WT was performed to identify the most-improved mutant for use in the next round of evolution. This mutant was used as the parental strain for the next round of evolution and

was also sequenced to determine the mutation that may have caused the improvement in total activity.

### **Sequencing analysis of mutants**

Mutants generated by *ep*PCR were sequenced by (Eurofins Genomics, Ebersberg Germany) to check the mutations. Each mutant was sequenced using four different primers (appendix 2), and the sequencing data were analysed using SnapGene, ApE plasmid software, and the NCBI alignment tools. The most-improved mutants in the first round of evolution were sequenced and used in the second round of evolution as the parental strain.

Four rounds of directed evolution were performed in total, although an improvement in the mutants was found only in rounds 1, 2, and 3. In the fourth round, more than 1600 mutants were screened, but no improved mutant was identified. Thus, an alternative strategy was attempted using the most improved mutant to date as the parental strain.

## **6.2.6 Site-directed mutagenesis**

### **Combination of two mutations**

In order to investigate the effect of combining two beneficial mutations acquired after the first round of evolution, site-directed mutagenesis was performed to introduce a N312S mutation into A306V mutant, and appropriate primers were designed using the OneClick program (<http://tucksengwong.staff.shef.ac.uk/OneClick/>), OneClick program described in (chapter 4- section 4.2.3) in more details.

### **Second round of evolution**

In this round, the most-improved mutant from the first round was used as a parental strain for mutagenesis. A vector was prepared as described in the first round. Libraries were prepared using the three mutagenic conditions (Appendix 1), and the expression and screening were carried out as described previously (section 6.2.5.1.6). In the second round of directed evolution, screening of 558 mutants from different mutagenic conditions was

sufficient to identify several significantly improved mutants.

### **Third round of evolution**

The most-improved mutant from the second round was used as a parental strain in the third round. Preparation of the vector (pET24a-OsmY), construction of mutants, and expression and screening were performed as in the previous rounds. A total of 1116 mutants were screened, leading to the identification of one significantly improved mutant.

### **Fourth round of evolution**

The parental strain used in this round was the most-improved mutant from round three. The construction of mutants and expression and screening were similar to previous rounds. However, due to the increase in activity as a result of the evolution in the three previous rounds, the amount of enzyme used in screening was decreased by 50%. Over 1600 mutants were screened in this round, but no improved mutant was identified. Thus, an alternative strategy was attempted.

#### **6.2.7 Site-saturation mutagenesis**

After three successful rounds of directed evolution, a fourth round was performed but was unsuccessful, as described above. Saturation mutagenesis was therefore used, and a total of six amino acid positions were chosen for site-saturation mutagenesis based on the previously generated mutants showing improvements in total activity in comparison with the WT or parental strain. Single site-directed saturation mutagenesis was performed to introduce a mutation of all 20 naturally occurring amino acids at a given position on the protein or mutant. The screening of 64 mutants for each library to increase the probability of representing all 20 amino acids was performed. Mutagenesis reactions were prepared in a total volume of 50  $\mu$ L.

### **Primers**

Using the NEBChanger method, primers were designed for the saturation mutagenesis for position 312N. Primers, reaction mixtures, and PCR programs are illustrated in Tables 6-3 to 6-5. Using the OneClick program, mutagenic primers were designed using 4-

primers methods to replace positions N56, K109, 56R, 227N, and 374H with all 20 possible amino acids (Table 6-6).

### Non-overlapping primers for position N312

**Table 6-3:** Mutagenic primers used for site saturation mutagenesis at position N312

Position	Primers
312N	Rev - 5' AAACCTTGTTMNNACGCTGCGC...CGCCAG 3' Fwd - 5' GCGCAGCGTNNKAACAAGTTT...GGCGAT 3'

**Table 6-4:** PCR mixtures for site saturation mutagenesis at position N312

Component	Stock Concentration	Amount per Reaction
Distilled water	–	35 $\mu$ L
Q5 buffer	5 $\times$	10 $\mu$ L
20 $\mu$ M mutagenic forward primer	20 $\mu$ M	1.25 $\mu$ L
20 $\mu$ M mutagenic reverse primer	20 $\mu$ M	1.25 $\mu$ L
dNTPs	10 mM each	1 $\mu$ L
DNA template	1–25 ng	1 $\mu$ L
Q5 high-fidelity DNA polymerase	2 U/ $\mu$ L	0.5 $\mu$ L
Total reaction volume		50 $\mu$ L

## PCR program

**Table 6-5:** PCR program for site saturation mutagenesis at position N312

Step	Temperature	Time
1	98°C	30 sec
2	98°C	10 sec
3	72°C	20 sec
4	72°C	3 min and 45 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

The gel images of PCR products and the following steps are explained in the saturation mutagenesis results.

### One-stage PCR (4 primer method)

Mutagenic primers were designed using the OneClick program (<http://tucksengwong.staff.shef.ac.uk/OneClick/>), and the 4 primer method (Table 6-6). More details about OneClick provided in (chapter 4- section 4.2.3). Using this method, 10 PCR reactions mixtures (Tables 6-7 to 6-16) were used to amplify two fragments of each position, followed by full-length PCR using the two fragments from each position (Table 6-18).



**Table 6-6:** Primers used for five amino acid positions using the 4-primer method

Position	Fragment	Primers
V56	F1	Rev: 5' -TTTAATMNINGCCCGCGCTGGTCTTAATGTG-3' Fwd: 5' -GATCGGATCCATGACCACCCCGGCGCCGCGCTGG-3'
	F2	Fwd: 5' -ACCAGCGCGGGCENNKATTAAGACCGTGAG-3' Rev: 5' - ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC-3'
A306	F1	Rev: 5' -ATCCGCMNNCAGTTTCGGATCGTCCTTCAG-3' Fwd: 5' -GATCGGATCCATGACCACCCCGGCGCCGCGCTGG-3'
	F2	Fwd: 5' GATCCGAAACTGNKGC GGATGCGCAGCGT-3' Rev: 5' - ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC-3'
R109	F1	Rev: 5' -CGCGTCMNNACGCTGGCCGGTGGTGAACGC-3' Fwd: 5' -GATCGGATCCATGACCACCCCGGCGCCGCGCTGG-3'
	F2	Fwd: 5' -ACCGCCAGCGTNNKGACGCGGAGATTCTG-3' Rev: 5' - ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC-3'
N227	F1	Rev: 5' -GTCACCMNNCTCCTTCGCCAGAATGAAACC-3' Fwd: 5' -GATCGGATCCATGACCACCCCGGCGCCGCGCTGG-3'
	F2	Fwd: 5' -CTGGCGAAGGAGNNKGGTGACAGCCGTGCG-3' Rev: 5' - ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC-3'
H374	F1	Rev: 5' -GTCGTGMNNTTCTTGGCTGGTCACTTCCGG-3' Fwd: 5' -GATCGGATCCATGACCACCCCGGCGCCGCGCTGG-3'
	F2	Fwd: 5' -ACCAGCCAAGAANNKACGACAAGAAAACC-3' Rev: 5' - ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC-3'

## PCR program

### Position V56

**Table 6-7:** PCR program to amplify fragment (1) for position V56

Step	Temperature	Time
1	98°C	30 sec
2	98°C	8 sec
3	55°C	20 sec
4	72°C	5 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

**Table 6-8:** PCR program to amplify fragment (2) for position V56

Step	Temperature	Time
1	98°C	2 min
2	98°C	20 sec
3	55°C	20 sec
4	72°C	34 sec
5	Repeat from step 2, 29 times	
6	72°C	3 min
7	8°C	Hold

### Position A306

**Table 6-9:** PCR program to amplify fragment (1) for position A306

Step	Temperature	Time
1	98°C	30 sec
2	98°C	8 sec
3	55°C	20 sec
4	72°C	23 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

**Table 6-10:** PCR program to amplify fragment (2) for position A306

Step	Temperature	Time
1	98°C	2 min
2	98°C	20 sec
3	55°C	20 sec
4	72°C	15 sec
5	Repeat from step 2, 29 times	
6	72°C	3 min
7	8°C	Hold

**Position R109****Table 6-11:** PCR program to amplify fragment (1) for position R109

Step	Temperature	Time
1	98°C	30 sec
2	98°C	8 sec
3	55°C	20 sec
4	72°C	23 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

**Table 6-12:** PCR program to amplify fragment (2) for position R109

Step	Temperature	Time
1	98°C	2 min
2	98°C	20 sec
3	55°C	20 sec
4	72°C	30 sec
5	Repeat from step 2, 29 times	
6	72°C	3 min
7	8°C	Hold

**Position N227****Table 6-13:** PCR program to amplify fragment (1) for position N227

Step	Temperature	Time
1	98°C	30 sec
2	98°C	8 sec
3	55°C	20 sec
4	72°C	18 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

**Table 6-14:** PCR program to amplify fragment (2) for position N227

Step	Temperature	Time
1	98°C	2 min
2	98°C	20 sec
3	55°C	20 sec
4	72°C	21 sec
5	Repeat from step 2, 29 times	
6	72°C	3 min
7	8°C	Hold

**Position H374****Table 6-15:** PCR program to amplify fragment (1) for position H374

Step	Temperature	Time
1	98°C	30 sec
2	98°C	8 sec
3	55°C	20 sec
4	72°C	29 sec
5	Repeat from step 2, 29 times	
6	72°C	2 min
7	8°C	Hold

**Table 6-16:** PCR program to amplify fragment (2) for position H374

Step	Temperature	Time
1	98°C	2 min
2	98°C	20 sec
3	55°C	20 sec
4	72°C	10 sec
5	Repeat from step 2, 29 times	
6	72°C	3 min
7	8°C	Hold

Amplified fragments were analysed by DNA gel electrophoresis and gel extraction was performed in the presence of side product bands, while for amplified fragments with no presence of other bands, *DpnI* digestion followed by purification was performed.

### Full-length PCR

Full length PCR were performed for all positions using the two fragments amplified for each position. In the first stage, only fragments added to the reaction (length and concentration of fragment (1) and (2) was considered before addition to the reaction mixture (Table 6-17)). In the second stage, 1 µL of each of the following primers was added: Rev: 5'-ATGCGAATTCTTACGCGCTGATCGGCGCTTGGCTGTGC-3', Fwd: 5'-GATCGGATCCATGACCACCCCGGCGCCGCGCTGG-3', the solution was mixed, and the reaction resumed (Table 6-18).

**Table 6-17:** Mixtures for full-length PCR performed for the 5 amino acid positions

Component	Stock Concentration	Amount per Reaction
Distilled water	–	38.5 $\mu$ L
Q5 buffer	5 $\times$	10 $\mu$ L
dNTPs	10 mM each	1 $\mu$ L
Fragment (1)	12 ng/ $\mu$ L	0.5 $\mu$ L
Fragment (2)	100 ng/ $\mu$ L	0.5 $\mu$ L
Q5 high-fidelity DNA polymerase	2 U/ $\mu$ L	0.5 $\mu$ L
Total reaction volume		50 $\mu$ L

**Table 6-18:** Two-stage full-length PCR program using two fragments for each position

Step	Temperature	Time
1	98°C	30 sec
2	98°C	8 sec
3	72°C	20 sec
4	72°C	1 min, 30 sec
	Repeat from step 2, 9 times	
5	Hold and pause at 72°C	
6	98°C	8 sec
7	72°C	20 sec
8	72°C	1 min, 30 sec
	Repeat from step 6, 19 times	
9	72°C	2 min
10	8°C	Hold

DNA gel electrophoresis was performed for the resulting mixtures to analyse the products and determine the purity or the presence of side-band products for the next step. According to the purity of products, either gel extraction was performed in the presence of side products, or *DpnI* digestion where there were no side bands.

### Ligation and transformation in BL21 (DE3)

Restrictive digestion of the purified PCR products was performed using *BamHI* and *EcoRI* restriction enzymes followed by ligation with T4 DNA ligase into pET24-a-OsmY. Ligation of these mutants was performed in a total volume of 20  $\mu$ L using T4 DNA ligase and the process was similar to that performed for *epPCR* in the four rounds of evolution

described previously in this chapter. Transformation by electroporation in BL21 (DE3) was performed using 1  $\mu$ L of the ligation mixture as described previously. Briefly, 1  $\mu$ L of purified DNA or ligation mixture was used for the electroporation of 40  $\mu$ L of concentrated BL21 (DE3), 1 ml of SOC media was added, and the mixture was incubated for 1 hr with shaking at 37°C. Transformants were plated onto TYE agar plates supplemented with kanamycin (50 mg/mL) and grown at 37°C for 15 hrs.

#### **Expression of libraries (high-throughput screening assay for secreted DyP4 (total activity))**

Colonies were picked individually and cultured in sterile 96-well microtiter plates containing 150  $\mu$ L of 2 $\times$  TY media with kanamycin per well. Well numbers B2, E6, and H11 were inoculated with the WT/parental strain as a standard. Plates were wrapped with silicon tape and incubated at 30°C and 1050 rpm for 24 hrs. Pre-culturing was performed for all libraries using the replicator with growth at 30°C and 1050 rpm for 18 hrs, and master plates were prepared by the addition of 100  $\mu$ L of 50% sterile glycerol to each well before storage at -80°C. Protein expression was started in 150  $\mu$ L of 2 $\times$  TY AIM with kanamycin for 24 hrs at 30°C and 1050 rpm.

#### **Second expression and screening of improved mutants and the WT**

Significantly improved mutants were further investigated by expression on a larger scale (50 mL 2 $\times$  TY AIM media), and the activity was normalized using OD values. The most-improved mutant was sequenced to identify the mutation that caused to the activity to improve, and was used as a parental strain for the next round of evolution.

### **6.2.8 Large-scale recombinant production, purification and characterisation of DyP4 variants**

BL21 (DE3) harbouring pET24a-WT, pET24a-F6, pET24a-D4, pET24a-OsmY-WT, pET24a-OsmY-F6, and pET24a-OsmY-D4 were cultured in 200 mL 2× TY AIM media containing kanamycin at 30°C for 24 hours for protein expression and secretion (see Chapter 2, section 2.6.2). For analysis with SDS-PAGE, the acetone precipitation method was used for the treatment of samples from the media, as the secretion seems to be low and almost undetectable on SDS-PAGE.

Acetone precipitation of the samples was performed in order to concentrate protein samples due to the low secretion level obtained when the OsmY-based secretion system was used. In brief, the required volume of pure acetone was placed at -80°C for 20 min, and then incubated at -20°C for 1 hour. Protein samples were placed in an acetone compatible tube, and four times of acetone was added to the protein samples, then mixed and incubated at -20°C for overnight. Centrifugation at maximum speed ( $17,000 \times g$ ), -4°C for 15 minutes was performed, supernatants were disposed of and uncapped tubes left open for 30 minutes at room temperature to ensure the evaporation of the remaining acetone. 35 µL of 1× SDS reducing buffer supplemented with  $\beta$ -mercaptoethanol) was added, boiled at 94°C for 5 min to denature the protein, and centrifuged at maximum speed for 2 min. 2 µL were loaded in the SDS-PAGE for analysis. Analysis with SDS-PAGE performed as explained in (chapter 2 section 2.7). Pellets of DyP4 variants (WT, F6, and D4) were used for three step of purification as described in (chapter 2 section 2.8).

#### **Enzymatic and H<sub>2</sub>O<sub>2</sub> stability analysis**

To compare the effect of hydrogen peroxide on the WT and the two obtained mutants, different concentrations of hydrogen peroxide were used in an ABTS assay as described in Chapter 4, section 4.2.10.

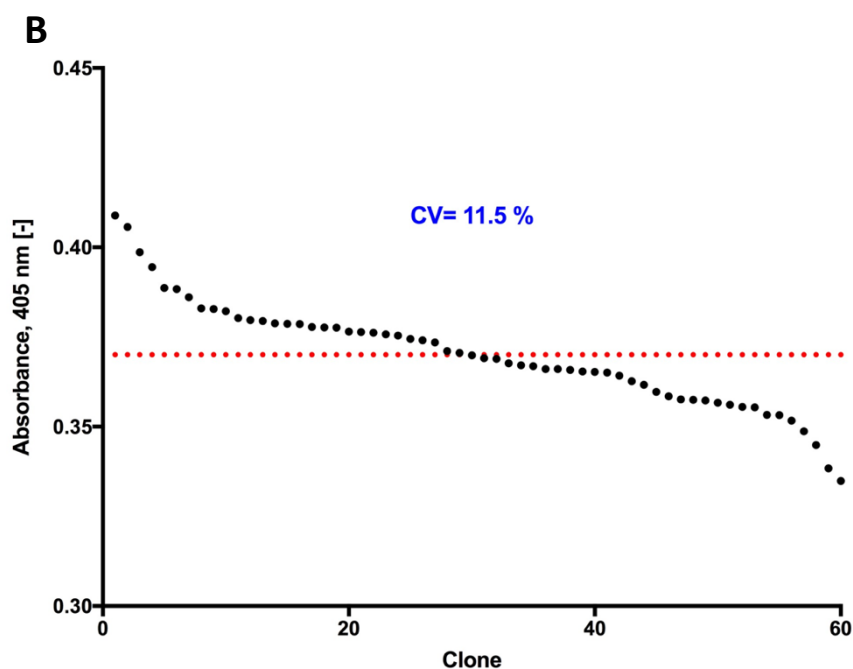
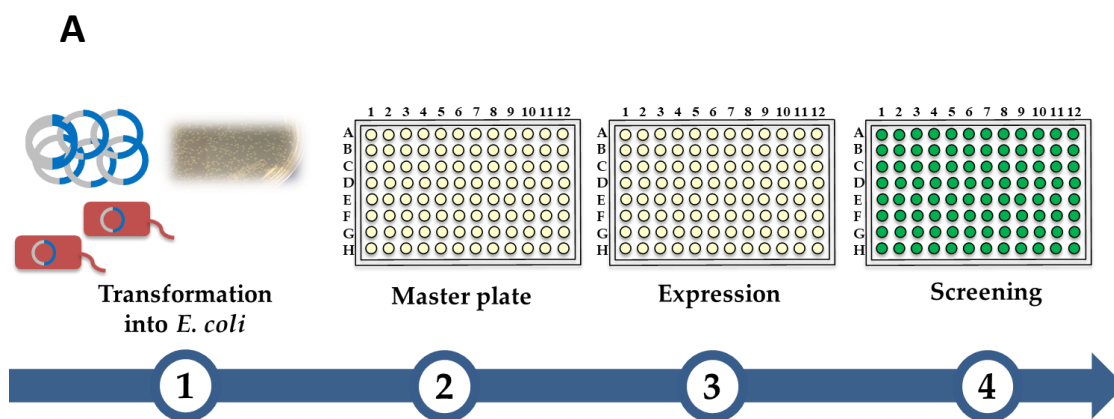


## **6.3 Results**

The OsmY-based secretion mechanism is an interesting method for the secretion of enzymes in bacteria. It can facilitate the process of directed evolution by eliminating the need for a cell lysis step. Based on experiments to determine the optimal conditions for the secretion of DyP4 using OsmY, the BL21 (DE3) strain was found to be better than C41 (DE3). In fact, no secretion was observed using the latter (data not shown). Two temperatures, 30°C and 37°C, were used to determine the optimum temperature for secretion. Increased secretion was achieved when expression was performed at 30°C, so this temperature was used for the remaining experiments for the secretion of DyP4 WT or mutants in BL21 (DE3).

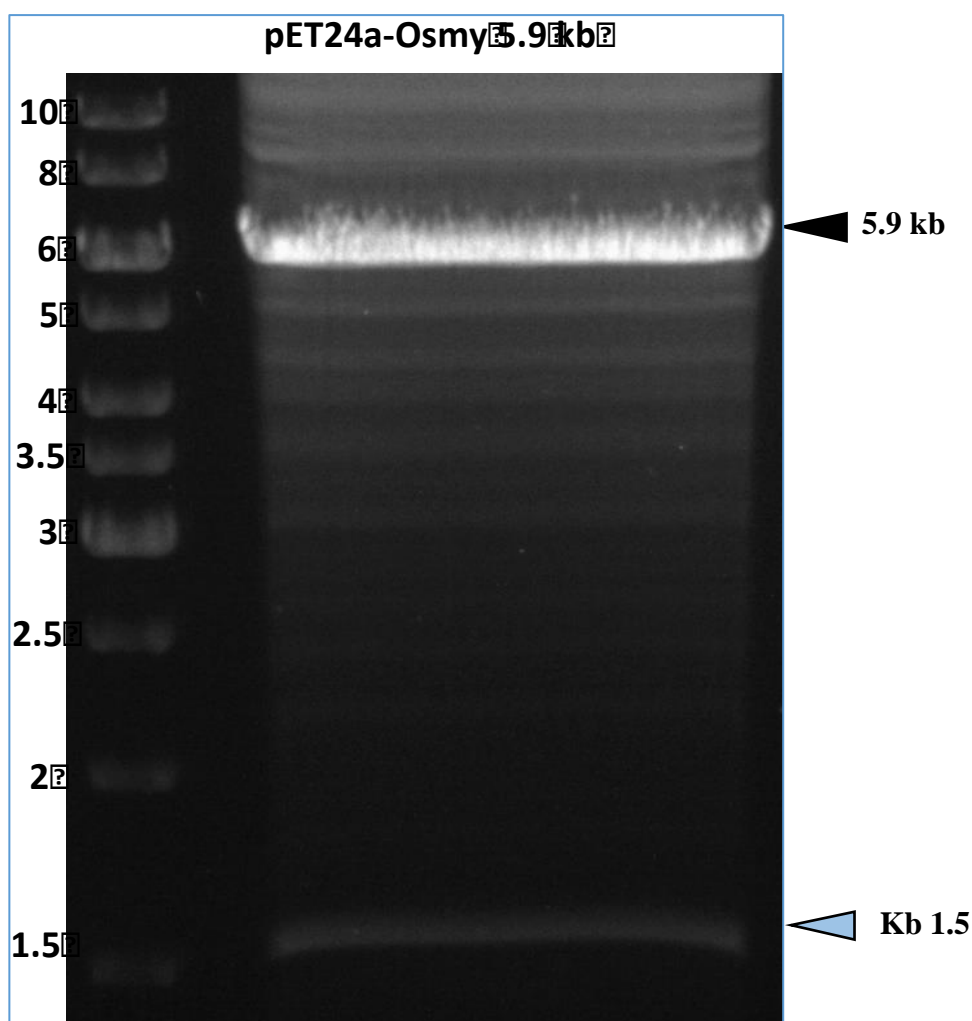
### **6.3.1 Validation of ABTS HTS assay**

To validate the ABTS assay for screening, a clone of WT was replicated in a 96-well plate and the activity was measured with ABTS. Using the ABTS assay, the CV was less than 12% and this was sufficient for the reliability and reproducibility of this assay for directed evolution (Figure 6-3). It should be noted here that evaporation was observed from the side wells, despite the wrapping of microtiter plates with silicon, and this was taken into consideration as sometimes these side wells were not included in the data presented, especially when the variation was significant.



**Figure 6-3:** (A) Scheme of the HTS protocol for WT of DyP4 with ABTS assay, and (B) validated HTS with 11.5 % coefficient of variation (CV) based on the oxidation of ABTS by WT-DyP4. The activities of DyP4 from different replicates of the same clone are plotted in descending order (Black dots), and the average is plotted (Red dots).

### 6.3.2 Molecular cloning

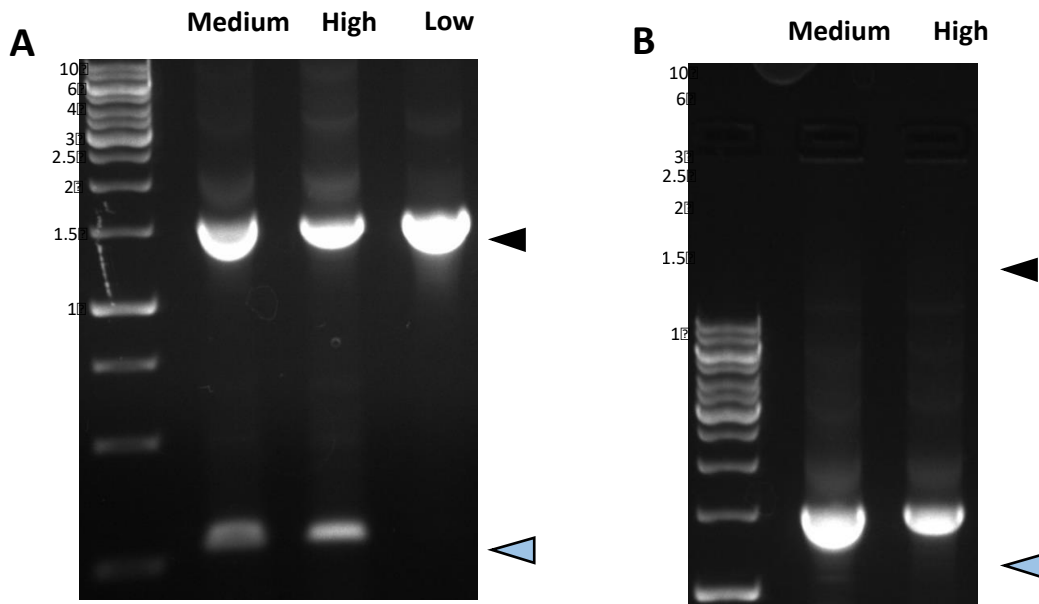


**Figure 6-4:** Restrictive digestion for pET24a-OsmY-Dyp4 with *Bam*HI and *Eco*RI restriction enzymes. The restrictive digestion mixture prepared in a total of 100  $\mu$ L containing 1 $\times$  of CutSmart buffer, 3000 ng of pET24a-OsmY-Dyp4, 1 U of *Bam*HI and 1 U of *Eco*RI enzymes, and 59  $\mu$ L deionized water, mixed and incubated overnight at 37°C. 20  $\mu$ L of 6X DNA Loading Dye added and the 120  $\mu$ L was loaded onto 1% agarose and 5.9 kb band were gel extracted and purified. The cut vector (5.9 kb)  $\blacktriangleleft$ , and the cut insert (Dyp4) is 1.5 kb  $\triangleleft$ .

To avoid the presence of non-digested parental template (pET24a-OsmY-Dyp4), 50 ng of double-digested vector was used for ligation and transformation in DH5 $\alpha$ . At least two colonies were used for plasmid extraction and to check the background for the presence of non-digested vector. This process was performed for each freshly prepared vector to eliminate the presence of the WT of Dyp4 during expression and screening, thus

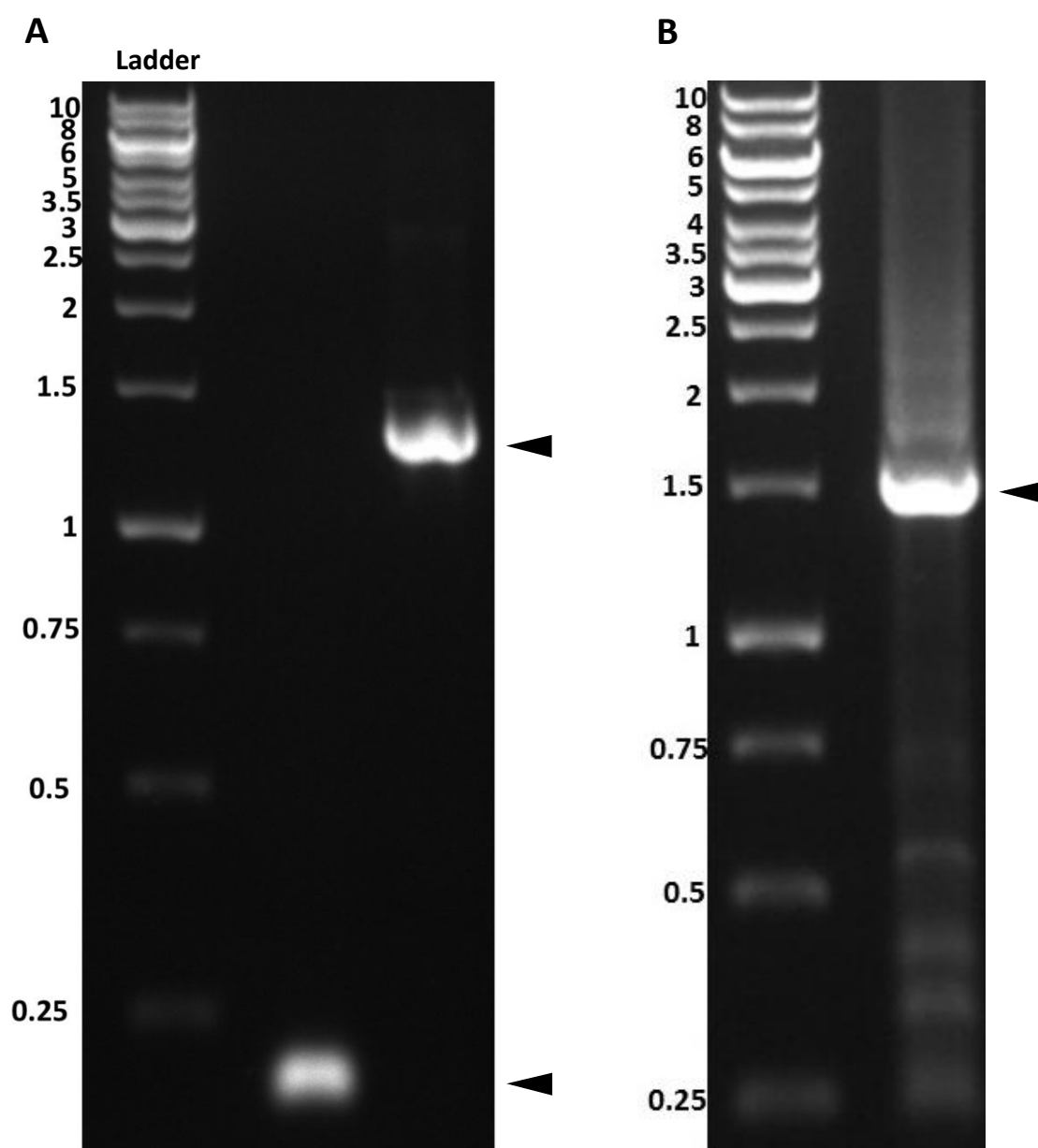
increasing the variation in mutants and the likelihood of identifying a significantly improved mutant.

### 6.3.3 Error-prone PCR



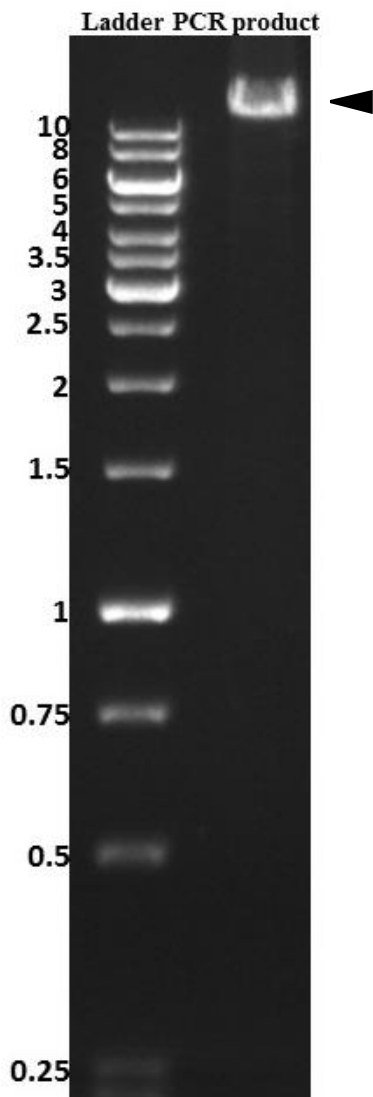
**Figure 6-5:** Analysis of PCR products on a 1% (w/v) DNA agarose gel showing products of different mutagenic reactions of *ep*PCR; A) ◀ three different conditions; B) ◀ medium and high mutagenic conditions. Side products ◀. Low mutagenic condition contained 1.5 mM of MgCl<sub>2</sub>, 0.01 mM of MnCl<sub>2</sub>, 0.3 mM of dNTP mix, 4.5 pmol of forward and reverse primers, 3.5 ng/μL of DNA template, and 1.25 U of *Taq* DNA polymerase. Medium mutagenic condition contained 7 mM of MgCl<sub>2</sub>, 0.2 mM of dATP, 0.2 mM of dGTP, 1 mM of dTTP, 1 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/μL of DNA template, and 1.25 U of *Taq* DNA polymerase. High mutagenic condition contained 7 mM of MgCl<sub>2</sub>, 0.05 mM of MnCl<sub>2</sub>, 0.2 mM of dATP, 0.2 mM of dGTP, 0.2 mM of dTTP, 0.2 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/μL of DNA template, and 1.25 U of *Taq* DNA polymerase.

For low mutagenic conditions (Figure 6-4, lane 4), a clear band was observed, as a result overnight *Dpn*I digestion was performed. However, as seen in Figure 6-5 A and B, for the two remaining mutagenic conditions, side product bands were present. As a result, these products were gel-extracted, double digested, ligated into the vector, and transformed by electroporation in BL21 (DE3).



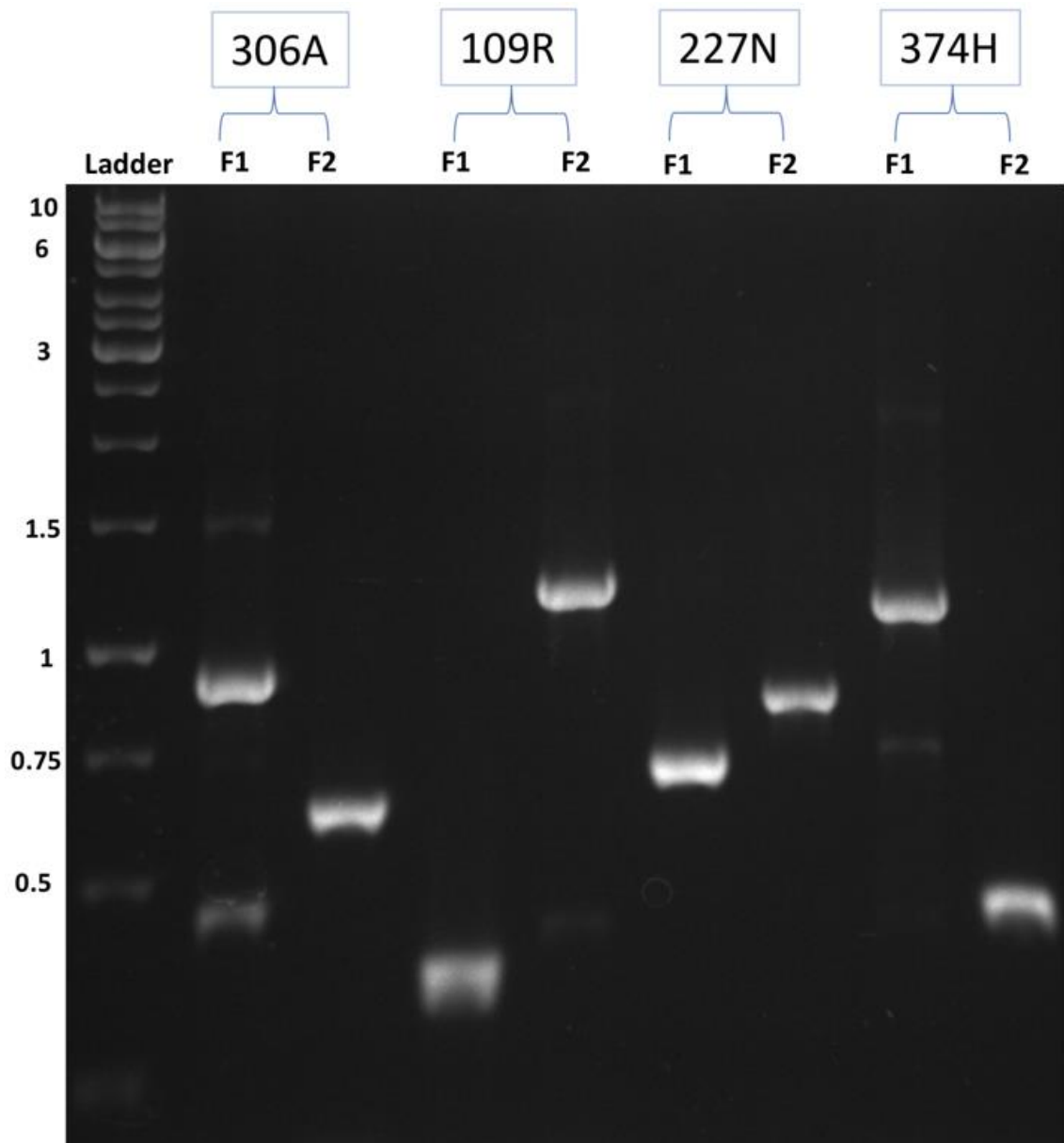
**Figure 6-6:** Analysis of PCR products on a 1% (w/v) DNA agarose gel showing the products of saturation mutagenesis using the 4-primer method for position V56. (A) Amplification of two fragments (fragment 1: 168 bp and fragment 2: 1347 bp). (B) Full-length PCR for the two fragments (product at 1.5 kbp)

Two-stage PCR amplified two fragments for position V56, fragment (1) 168 bp and fragment (2) 1347 bp (Figure 6-6). In the second stage, full-length PCR was performed using two fragments amplified in the first stage. PCR product in the second stage was at the expected size of DyP4 (1515 bp). Gel extraction was performed for the product, followed by double digestion with *Bam*HI and *Eco*RI, purification, ligation, and transformation by electroporation in BL21 (DE3).



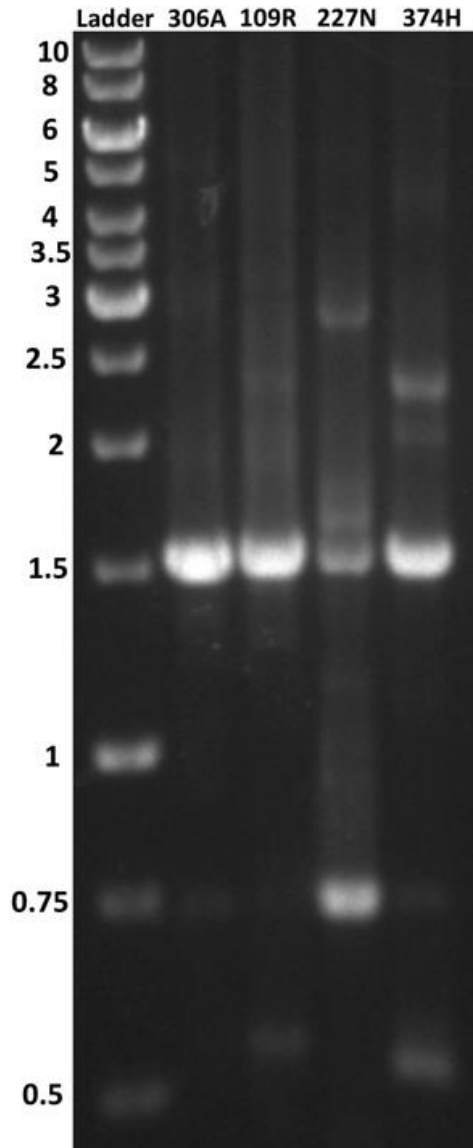
**Figure 6-7:** Analysis of PCR products on a 1% (w/v) DNA agarose gel showing the PCR product of the saturation mutagenesis for position N312 using the NEBase Changer method in the second lane

A clean product was obtained using NEBase Changer for position N312 (Figure 6-7). The parental template was digested by incubation with 1  $\mu$ L of *DpnI* enzyme at 37°C overnight. Purification was performed and, to an 8  $\mu$ L volume, 1  $\mu$ L of 10 $\times$  NEB T4 ligase buffer was added, followed by 1  $\mu$ L of 10 U/ $\mu$ L NEB T4 polynucleotide kinase (PNK) enzyme. The sample was mixed and incubated at 37°C for 1 hr, then 1  $\mu$ L of Invitrogen ligase was added and mixed, and incubation at 16°C was performed for 16 hrs. Of this ligation mixture, 1  $\mu$ L was used for transformation by electroporation in BL21 (DE3) to generate and screen mutants for the 312N position.



**Figure 6-8:** Analysis of PCR products on a 1% (w/v) DNA agarose gel showing the products of saturation mutagenesis using the 4-primer method for positions (A306: F1 918 bp and F2 597 bp), (R109: F1 327 bp and F2 1188 bp), (N227: F1 681 bp and F2 834 bp), (H374: F1 1122bp and F2 393 bp)

Two-stage PCR amplified two fragments for positions A306, R109, N227, and H374 were performed similar to position V56 as described in the previous section (Figure 6-8) and (6-9).



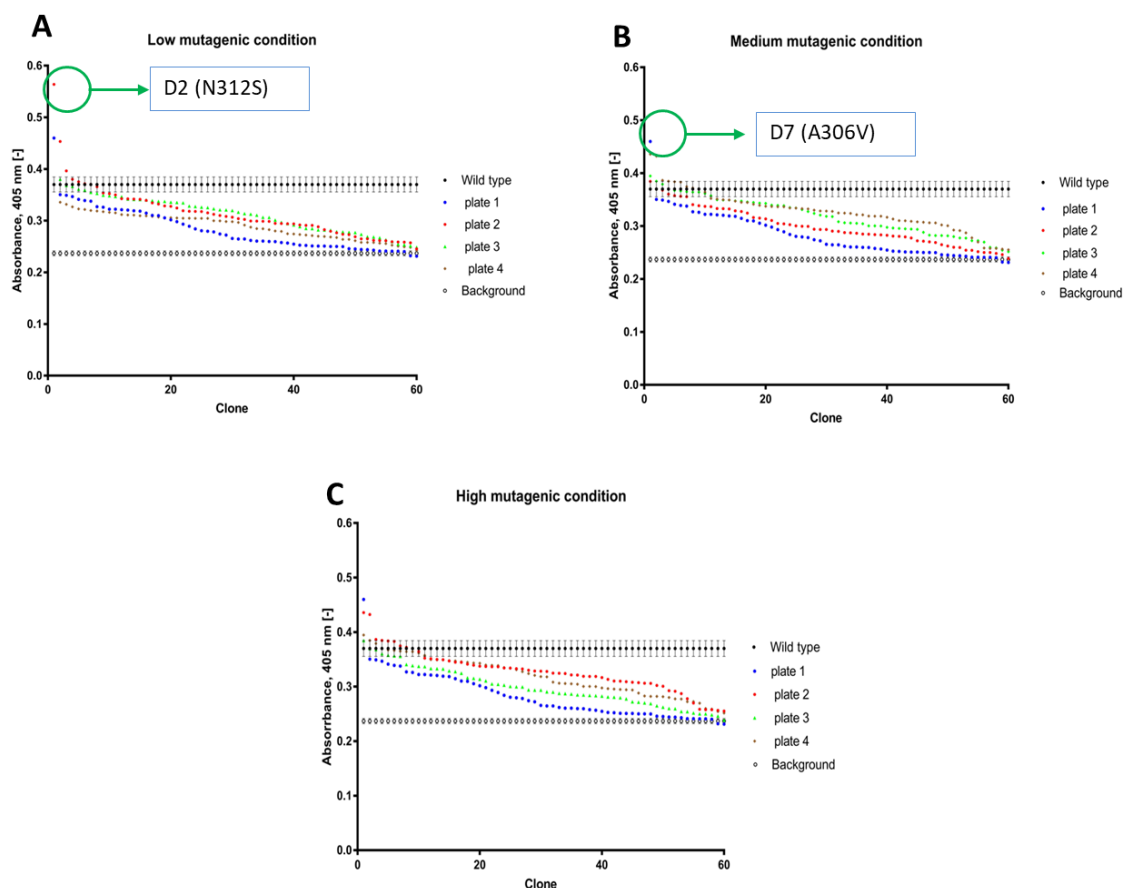
**Figure 6-9:** Combination of two fragments for positions A306, R109, N227, and H374 using full-length PCR. Reaction products were loaded into 1% agarose gel and the 1.5 kb band was cut and purified for each position to be used for ligation and transformation.



#### **6.3.4 First round of evolution**

After the construction of mutants with three different mutagenic conditions in the first round, a total of 1100 mutants were screened (Figure 6-10) and two mutants with potentially improved total activity were identified. These two mutants underwent further characterization by expression on a larger scale (50 mL), and their total activities were normalized to OD. These two mutants were found to show a significant improvement in total activity and were sequenced to detect the mutations which may have caused this improvement. Sequencing results showed that one mutant (D2) had a missense mutation and two silent mutations, while the second mutant (D7) had only one missense mutation. As shown in the 3D model structure (figure 6-12), the missense mutations of these two mutants were close to each other, and the two mutations were combined. The purpose of this investigation was to determine whether this combination could lead to any further improvement. Sequencing results confirmed the successful combination of these two mutations into a single mutant. The new combined mutant was characterized along with the WT, D2, and D7 mutants. However, screening results showed that the D2 mutant remained the most-improved mutant compared with WT, D7, and the combined mutant (data not shown). As the most improved mutant from the first round of evolution, the D2 mutant was then used as a parental strain for the second round of random mutagenesis using *ep*PCR.

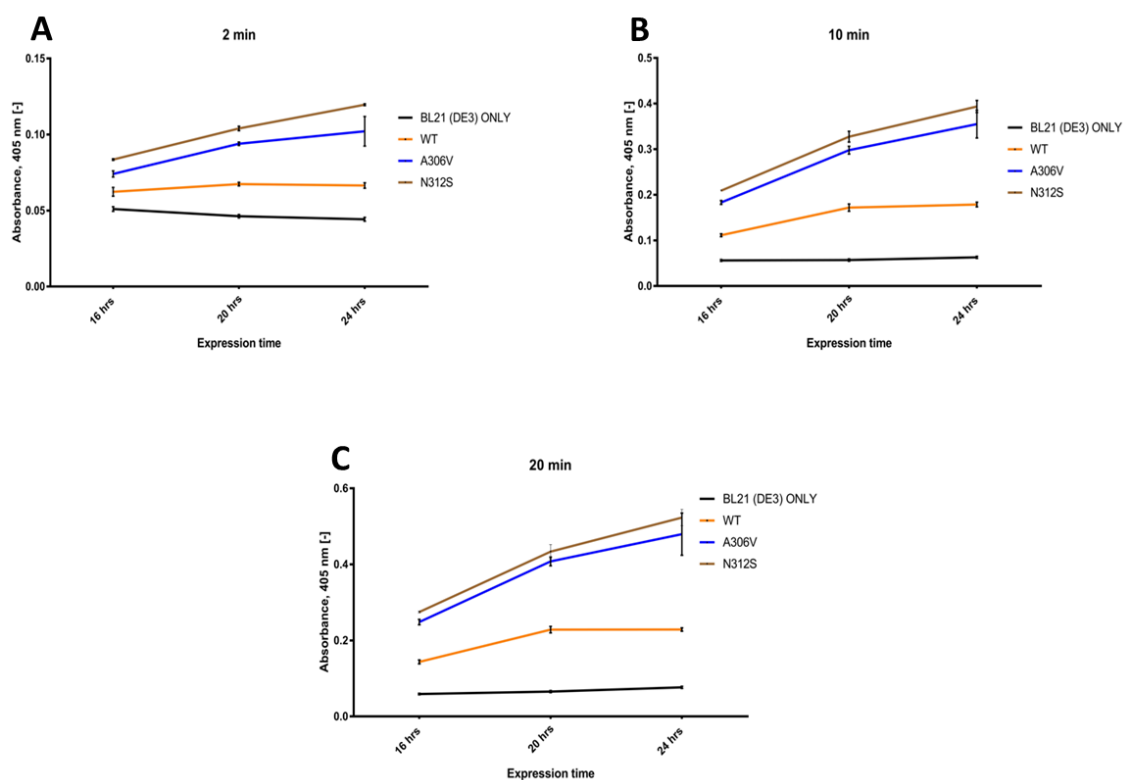
## Total activity of mutants generated in the first round of directed evolution



**Figure 6-10:** HTS for mutants in the first round of evolution using WT-DYP4 as a template for *epPCR*. (A) Low mutagenic condition, (B) Medium mutagenic condition, (C) High mutagenic condition. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 150  $\mu\text{L}$  of ABTS (10 mM), 20  $\mu\text{L}$  of secreted enzyme, and 4 mM of  $\text{H}_2\text{O}_2$  (50  $\mu\text{L}$ ).

Two mutants with a significant improvement in total activity were further characterized by expression on a larger scale. The total activities were normalized to OD and the two mutants were compared with the WT (Figure 6-11), (Table 6-19). The results shows that mutants N312S and A306V exhibited more activity than the WT with around 2-fold improvement.

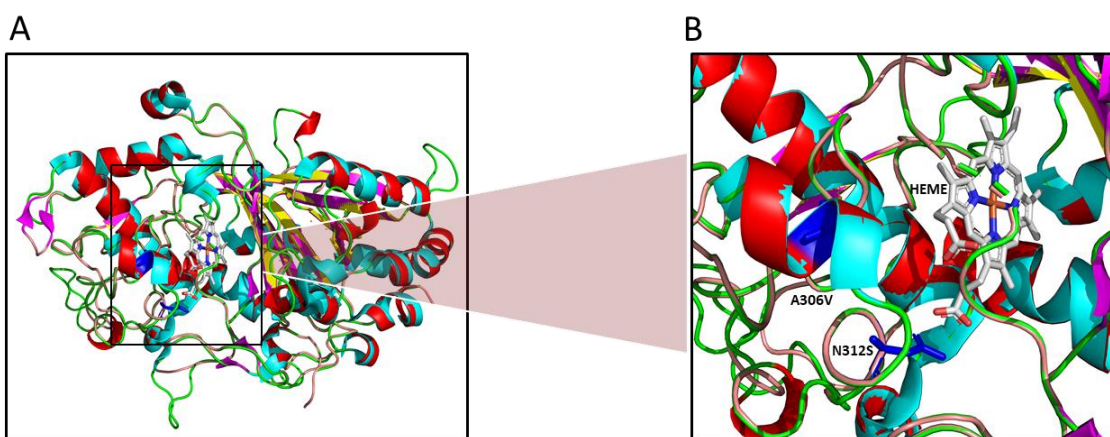
## Normalized activity



**Figure 6-11:** Normalized activity to OD for mutants A306V (Blue line) and N312S (Brown line), and the WT (Orange line). (A) Activity measured 2 minutes after the addition of hydrogen peroxide to initiate the oxidation of ABTS; (B) and (C) after 10 minutes and 20 minutes of adding hydrogen peroxide, respectively. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression.

**Table 6-19:** List of potentially improved mutants and mutation types in the first round of evolution

Mutant	Type of mutation	Original amino acid	Original codon	New amino acid	New codon	Mutant	Improvement
D2	Silent	Aspartic acid (D) 241	GAT	Aspartic acid (D) 241	GAC	-----	
D2	Missense	Asparagine (N) N312	AAC	Serine (S) 312	AGC	N312S	217%
D2	Silent	Isoleucine (I) I444	ATC	Isoleucine (I) I444	ATT	-----	
D7	Missense	Alanine (A)	GCG	Valine (V)	GTG	A306V	186%
D7	Silent	Arginine (R)	CGT	Arginine (R)	CGC	-----	
Mutants D2 (N312S) was used as a parental strain in the second round of evolution							



**Figure 6-12:** (A) The 3D protein model structure of DyP4 was generated with Pymol and using (PDB: 1afv) as a template model showing the position of A306V and N312S mutations. (B) detailed view of the Heme in DyP4 is given and position of the mutations shown in Blue stick.

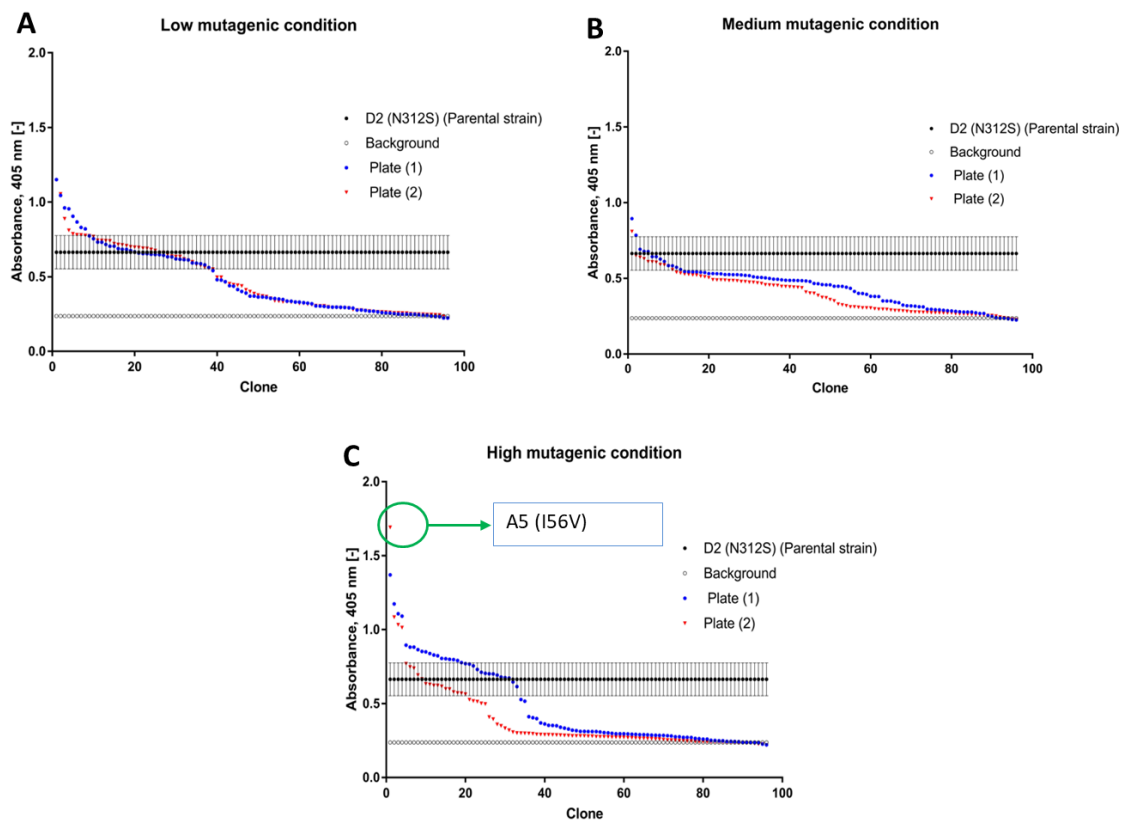
### 6.3.5 Second round of evolution

A total of 558 mutants coming from low, medium, and high mutagenic conditions were expressed in microliter plates containing 150  $\mu\text{L}$  of 2 $\times$  TY AIM and kanamycin for 24 hrs at 30°C and 1050 rpm. Three parental strain replicates (D2, N312S) coming from the same clone were included in each plate in wells (B2, E6, and G11), for comparison with the mutants.

Following centrifugation of the plates after 24 hours expression, 20  $\mu\text{L}$  of media (secreted enzyme) was used for activity screening with ABTS assay. The reaction mixture contained 150  $\mu\text{L}$  of ABTS (10 mM), 20  $\mu\text{L}$  of secreted enzyme, and the reaction initiated by adding 4 mM of  $\text{H}_2\text{O}_2$  (50  $\mu\text{L}$ ). Activities measurements were carried out using a multiplate reader (Multiskan™ FC Microplate Photometer, Thermo Scientific™) at wavelength 405 nm after 2 minutes of initiating the reaction with the addition of  $\text{H}_2\text{O}_2$ .

Few mutants showed a significant improvement in activity, and the most-improved mutant was from the high mutagenic condition (Figure 6-13A-C). Nine possibly improved mutants were further characterised by expression on a larger scale (50 ml) along with the parental strain (D2, N312S), and the WT.

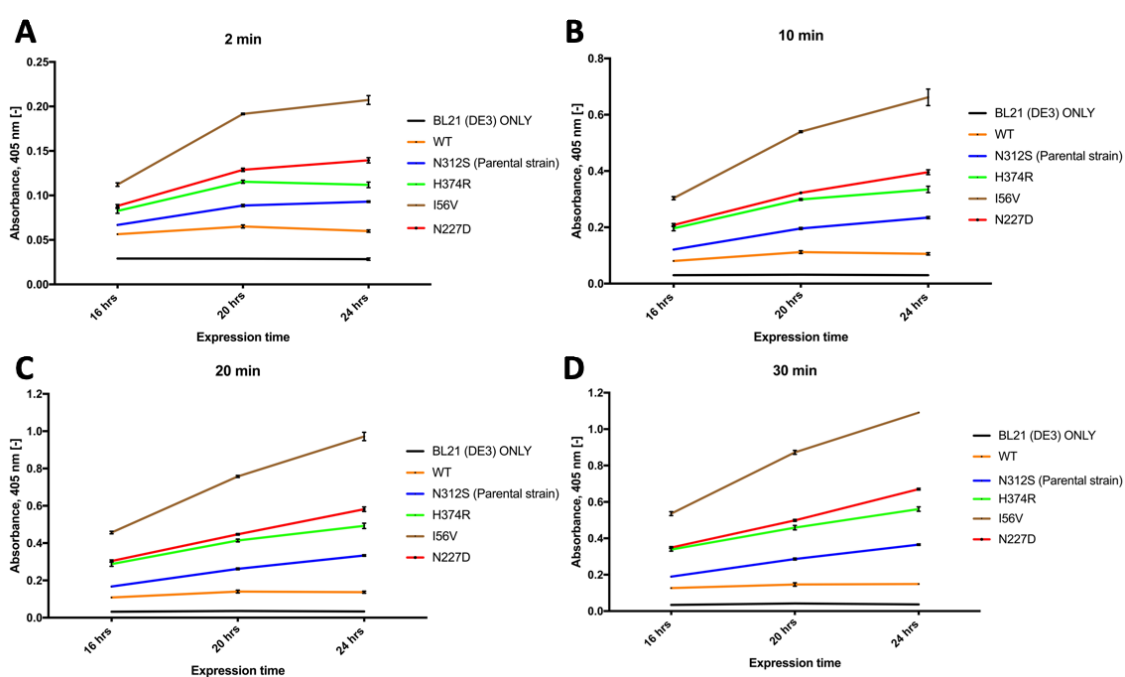
## Total activity for microtiter plates



**Figure 6-13:** Total activity profiles for mutants generated using N312S as the parental strain. (A) Low mutagenic condition, (B) medium mutagenic condition, (C) high mutagenic condition. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 150  $\mu\text{L}$  of ABTS (10 mM), 20  $\mu\text{L}$  of secreted enzyme, and 4 mM of  $\text{H}_2\text{O}_2$  (50  $\mu\text{L}$ ).

## Normalized activity

A total of nine significantly improved mutants from the three different mutagenic conditions were identified, and further investigation identified three improved mutants, of which mutants (A5, I56V) showed more than a two-fold improvement compared with the parental D2 (N312S) (Figure 6-14). Mutants (F8, N227D, and C2, H374R) showed around 2-fold improvement compared the parental strain (D2, M312S). Sequencing results showed the presence of two silent mutations in mutant (A5, I56V) while the other mutants did not have any silent mutation (Table 6-20).



**Figure 6-14:** Normalized activity to OD for three significantly improved mutants (I56V, Brown line), (H374R, Green line), and (N227D, Red line), in the second round of evolution using mutant (N312S, Blue line) as a parental strain. Measurements were taken at 405 nm after 2 min (A), 10 min (B), 20 min (C), and 30 min (D) of initiating oxidation of ABTS by addition of hydrogen peroxide. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. The reaction mixture (total of 220  $\mu$ L) contained 10 mM of ABTS (150  $\mu$ L), 20  $\mu$ L of media (secreted enzyme), and 4 mM of hydrogen peroxide (50  $\mu$ L).

**Table 6-20 :** List of potentially improved mutants and mutation types in the second round of evolution

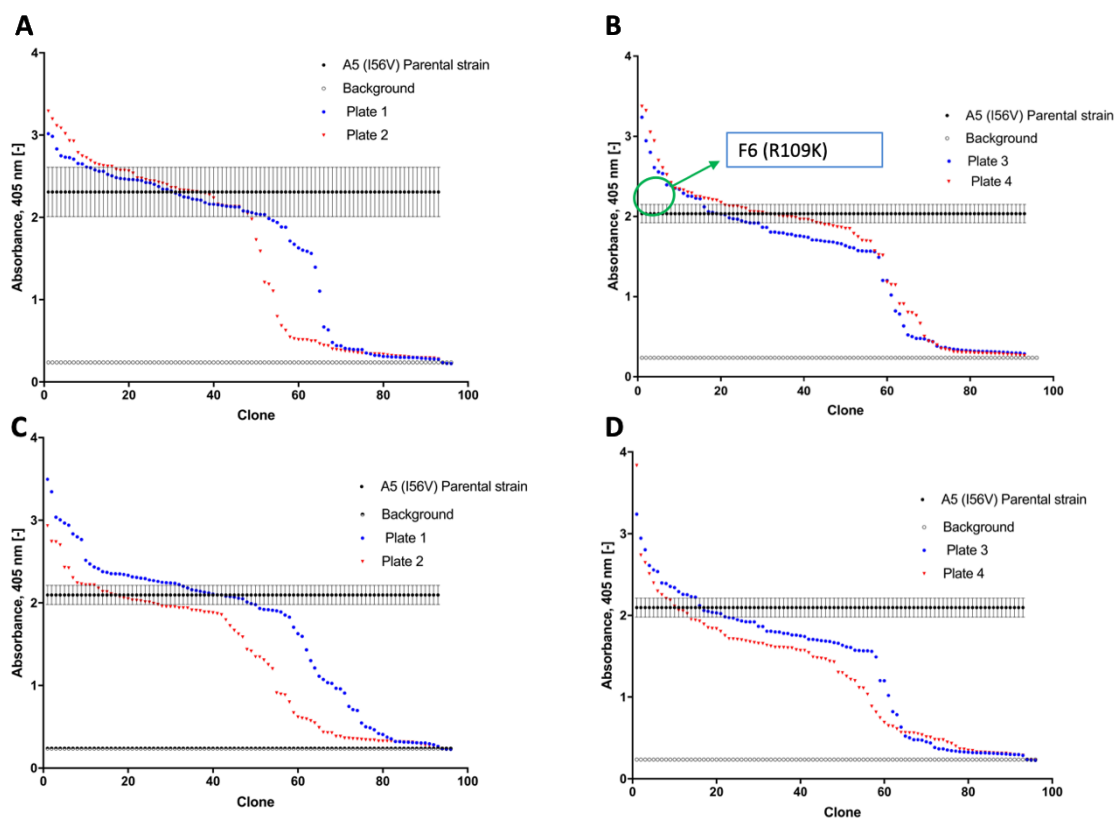
Mutant	Type of mutation	Original amino acid	Original codon	New amino acid	New codon	Mutant	Improvement
A5	Silent	Leucine (L) 245	<u>CTG</u>	Leucine (L) 245	<u>CTT</u>	-----	
A5	Missense	Isoleucine (I) 56	<u>ATC</u>	Valine (V) 56	<u>GTC</u>	I56V	302%
A5	Silent	Glycine (G)73	<u>GGT</u>	Glycine (G)73	<u>GGA</u>	-----	
C2	Missense	Histidine (H) 374	<u>CAC</u>	Arginine (R) 374	<u>CGC</u>	H374R	150%
F8	Missense	Asparagine (N) 227	<u>AAC</u>	Aspartic acid (D) 227	<u>GAC</u>	N227D	186%
Mutants A5 (I56V) was used for as the parental strain for the third round of evolution							

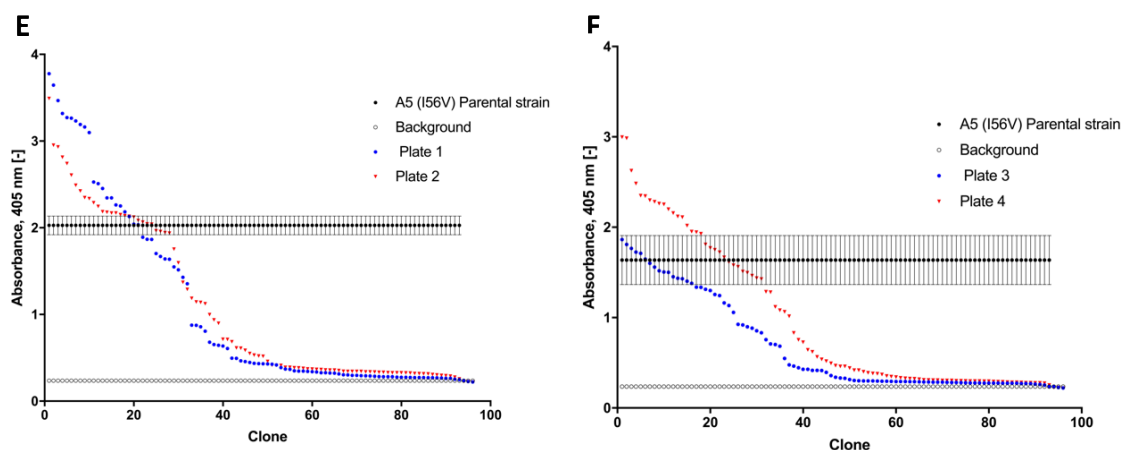


### 6.3.6 Third round of evolution

Around 1100 mutants were generated using low, medium, and high mutagenic condition and using mutant (A5, I56V) as a template for *ep*PCR. The expressing of mutants was similar to previous rounds and three parental strain replicates (A5, I56V) coming from the same clone were included in each plate in wells (B2, E6, and G11) for comparison with the generated mutants. Several mutants were shown to have a significant improvement in activity with ABTS assay compared to the parental strain (A5, I56V) (Figure 6-15A-F).

#### 6.3.6.1 Total activity in microtiter plates screening

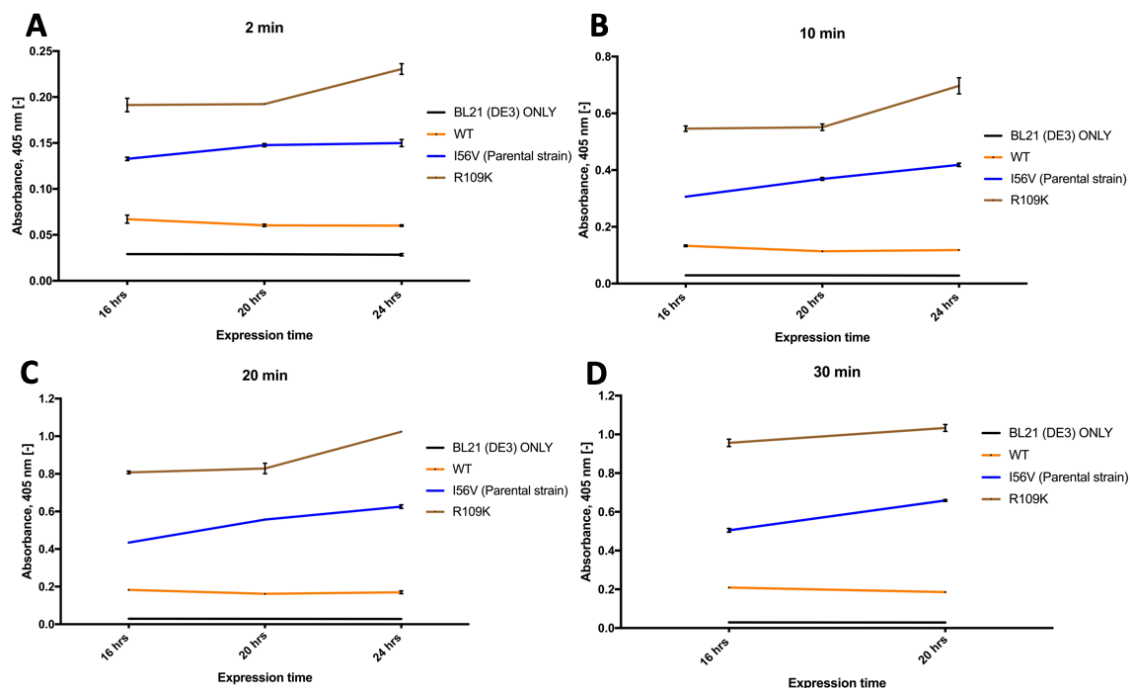




**Figure 6-15:** HTS screening results for mutants screened in the third round of evolution using I56V mutant as a parental strain. A and B, low mutagenic condition; C and D, medium mutagenic condition; and E and F, high mutagenic condition. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 150  $\mu\text{L}$  of ABTS (10 mM), 20  $\mu\text{L}$  of secreted enzyme, and 4 mM of  $\text{H}_2\text{O}_2$  (50  $\mu\text{L}$ ).

## Normalized activity

After the 1100 mutants were screened in the third round of evolution using I56V mutants as a parental strain, mutants (R109K) exhibited around two-fold improvement as confirmed by the normalised activity (Figure 6-16), (Table 6-21).



**Figure 6-16:** Normalized activity to OD for the most-improved mutant (F6, R109K, Brown line) with the parental strain (A5, I56V, Blue line) and the WT (Orange line). Measurements taken at wavelength 405 nm after 2 min (A), 10 min (B), 20 min (C), and 30 min (D) of initiating oxidation of ABTS by addition of hydrogen peroxide. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. The reaction mixture (total of 220  $\mu$ L) contained 10 mM of ABTS (150  $\mu$ L), 20  $\mu$ L of media (secreted enzyme), and 4 mM of hydrogen peroxide (50  $\mu$ L).

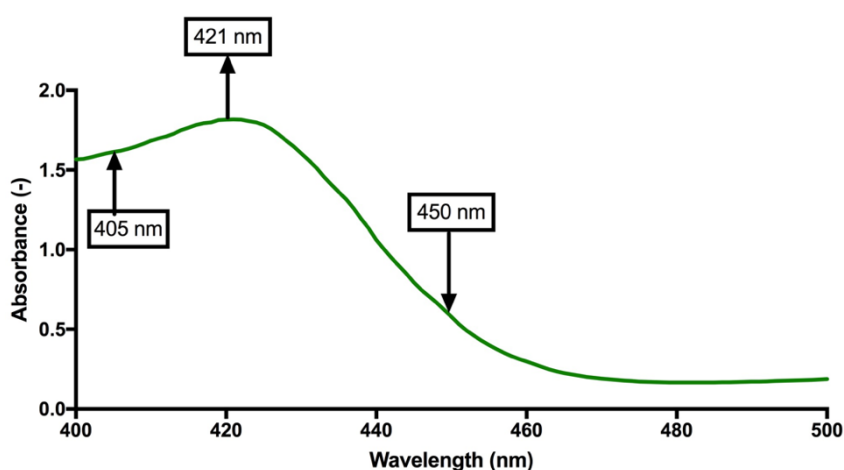
**Table 6-21:** List of potentially improved mutants and mutation types that caused the improvement in the third round of evolution

Mutant	Type of mutation	Original amino acid	Original codon	New amino acid	New codon	Mutant	Improvement
F6	Missense	Lysine (K) 109	A <u>A</u> G	Arginine (R) 109	A <u>G</u> G	K109R	160%
Mutants F6 (K109R was used as a parental strain for the fourth round of evolution and for site saturation mutagenesis)							

### 6.3.7 Fourth round of evolution

#### Rescreening at a different wavelength

Over 1600 mutants were screened in the fourth round of evolution using the F6 mutant as parental strain. However, no significantly improved mutant was obtained. Another consideration was the increase in variation among the three triplicates of the parental strain. This could have been caused by the increase in total activity, with the variation increased as a result. Screening at a different wavelength was performed for the next screening of mutants to overcome this issue. Wavelength scan measurement was conducted to determine an alternative wavelength for screening (Figure 6-17). The suggested wavelength based on the result of the UV-vis spectra for the ABTS assay was 450 nm, and the 405 nm wavelength was subsequently used only for comparison with the 450 nm wavelength and to validate the results obtained. The  $\lambda_{\text{max}}$  for measuring ABTS oxidation by DyP4 is 421 nm, which, however, was not used because the multiplate reader can measure two alternative wavelengths 405 nm or 450 nm. At the first few rounds of evolution, 405 nm was used, and after the third round, 450 nm was used as the activity became higher and as a result, the variation increased.

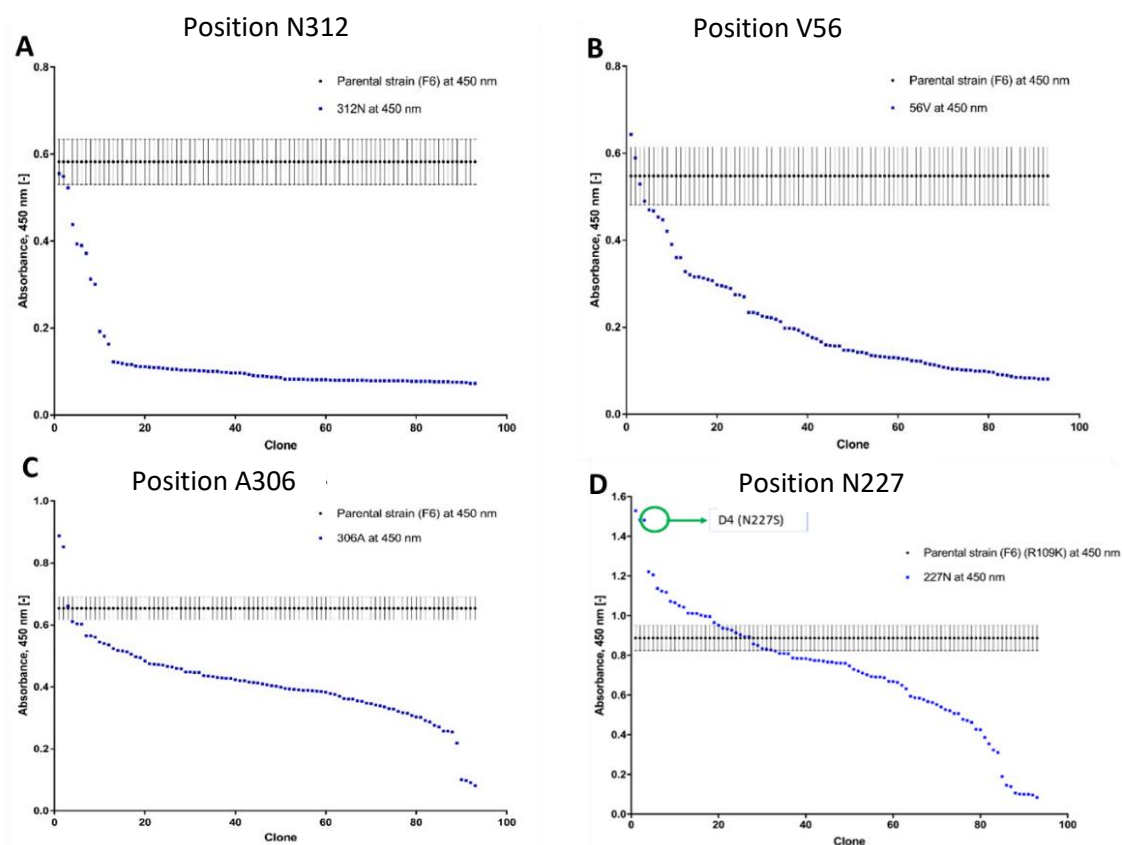


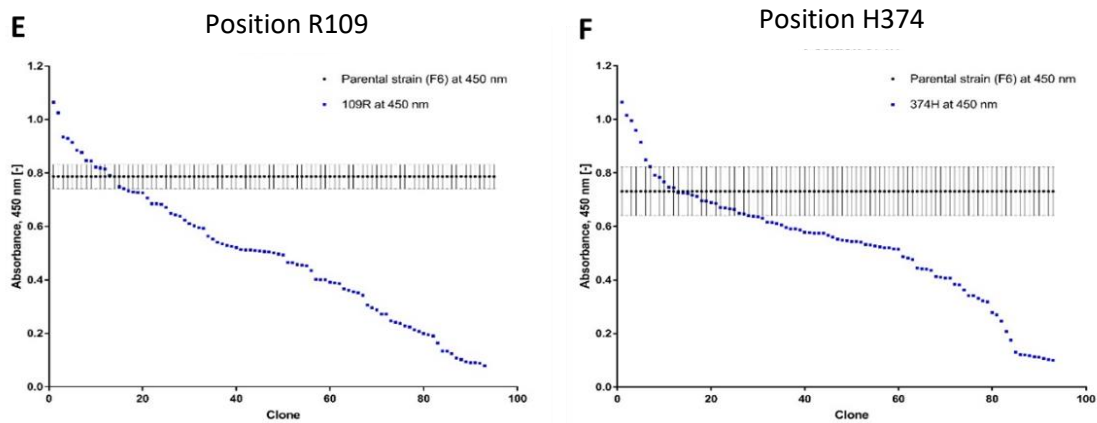
**Figure 6-17:** UV spectra for ABTS oxidation by DyP4 to determine the optimal wavelength. Reaction mixture was prepared in a 1 ml cuvette containing 925  $\mu\text{L}$  of ABTS (10 mM) pH 3.4, 25  $\mu\text{L}$  of purified diluted DyP4, and 50  $\mu\text{L}$  of  $\text{H}_2\text{O}_2$  (4mM). Wavelength scan measurements recorded 15 seconds after addition of  $\text{H}_2\text{O}_2$  to initiate the activity.

### 6.3.8 Saturation mutagenesis

#### Two strategies of saturation mutagenesis

Two approaches were attempted in site-directed mutagenesis: NEBase changer used for position 312N, and the 4-primer method using the OneClick program (<http://tucksengwong.staff.shef.ac.uk/OneClick/>) for the remaining five positions (V56, A306, R109, N227 and H374). The latter is more laborious but more efficient for saturation mutagenesis. The results of screening for libraries of site saturation mutagenesis for six positions (N312, A56, 306A, R109, N227 and H374; Figure 6-18, A-F) show that there were few mutants where a significant improvement in activity might be acquired. The most significant improvement achieved using saturation mutagenesis seems to be for position 227N (Figure 6-18,D) where the obtained improvement is approximately two-fold. Further investigation by the expression and re-screening of these potential mutants along with the parental strain confirmed the observed improvement (Figure 6-18).

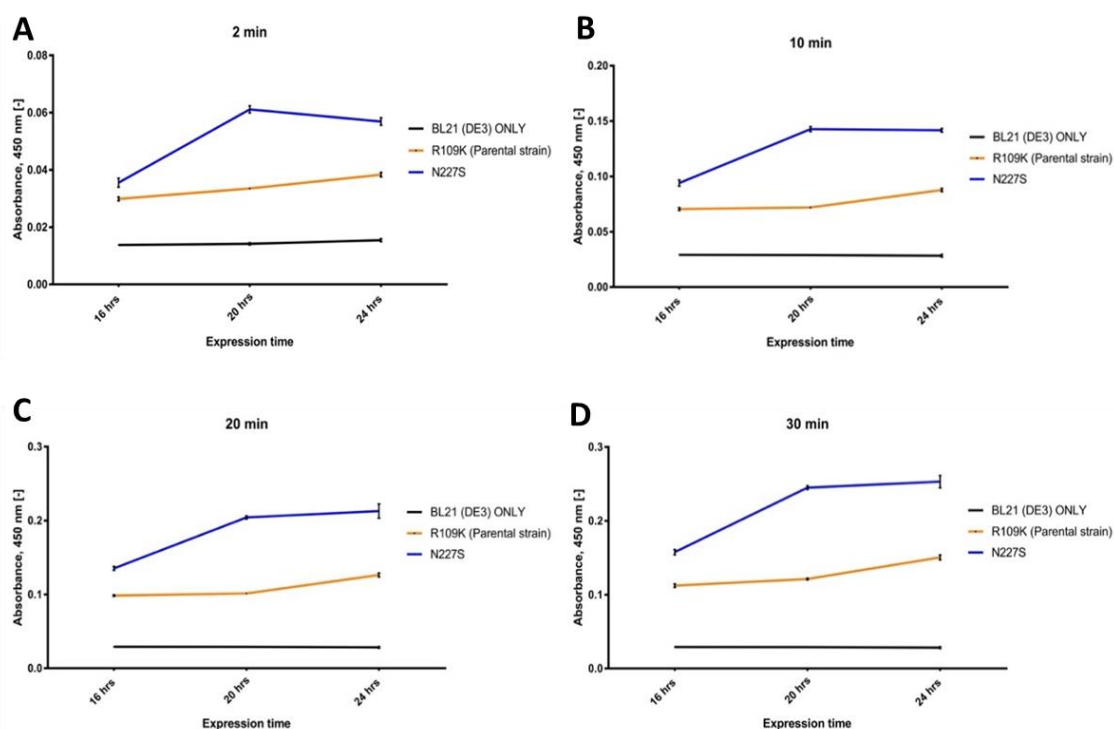




**Figure 6-18:** Total activity profiles for mutants generated by site saturation mutagenesis for positions N312 (A), V56 (B), A306 (C), R109 (D), N227 (E), and H374 (F) using F6 mutant as a parental strain. The activities with ABTS assay plotted in descending order against the clones. The reaction mixture contained 10 mM of ABTS (150  $\mu$ L), 10  $\mu$ L of secreted enzyme, and 4 mM of  $H_2O_2$  (50  $\mu$ L).

## Normalized activity

Over 90 mutants generated for each position using saturation mutagenesis were screened and the results showed several possible improved mutants, of which amino acid at position Asn227 exhibited the most improved mutants. The characterisation results confirmed this observation (Figure 6-19) with mutant N227S showing around 2-fold improvement in comparison to the parental strain.



**Figure 6-19:** Normalized activity to OD for the most-improved mutant (D4,N227S) with (P.S, F6,R109K) and the WT. Measurements taken at wavelength 405 nm after 2 min (A), 10 min (B), 20 min (C), and 30 min (D) of initiating oxidation of ABTS by addition of hydrogen peroxide. Protein expression carried out for 24 hours and samples taken for activity measurements after 16 hrs, 20 hrs, and 24 hrs of expression. The reaction mixture (total of 210  $\mu$ L) contained 10 mM of ABTS (150  $\mu$ L), 10  $\mu$ L of media (secreted enzyme), and 4 mM of hydrogen peroxide (50  $\mu$ L).

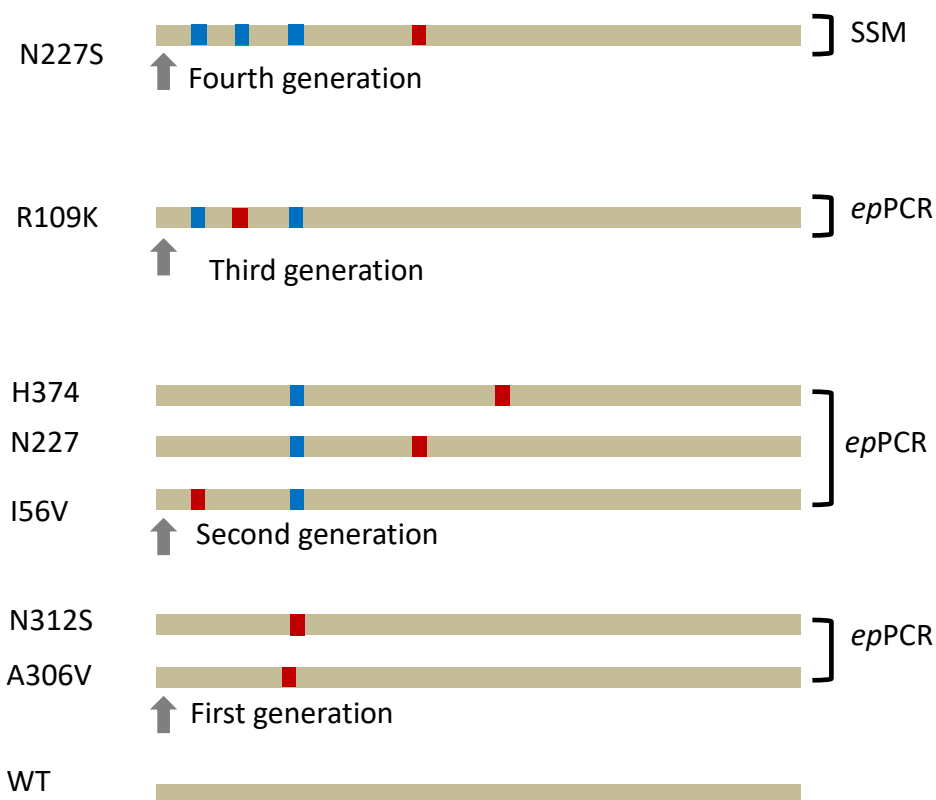


**Table 6-22:** List of potentially improved mutants and mutation types that caused the improvement in the fourth round of evolution using saturation mutagenesis

Mutant	Type of mutation	Original amino acid	Original codon	New amino acid	New codon	Mutant	Improvement
D4	Missense	Asparagine (N) 227	<u>A</u> AC	Serine (S) 227	<u>T</u> CC	N227S	167%
Mutants (D4, N227S) is most improved mutant achieved after four rounds of directed evolution							

## Characterisation of mutants and the WT

Following a successful fourth round of directed evolution using *ep*PCR and saturation mutagenesis, two mutants (F6 and D4) were characterised along with the WT. F6 mutant had three accumulative missense mutations N312S, I56V, and R109K, while the D4 mutant acquired the four missense mutations N312S, I56V, and R109K, and N227S (Figure 6-20).

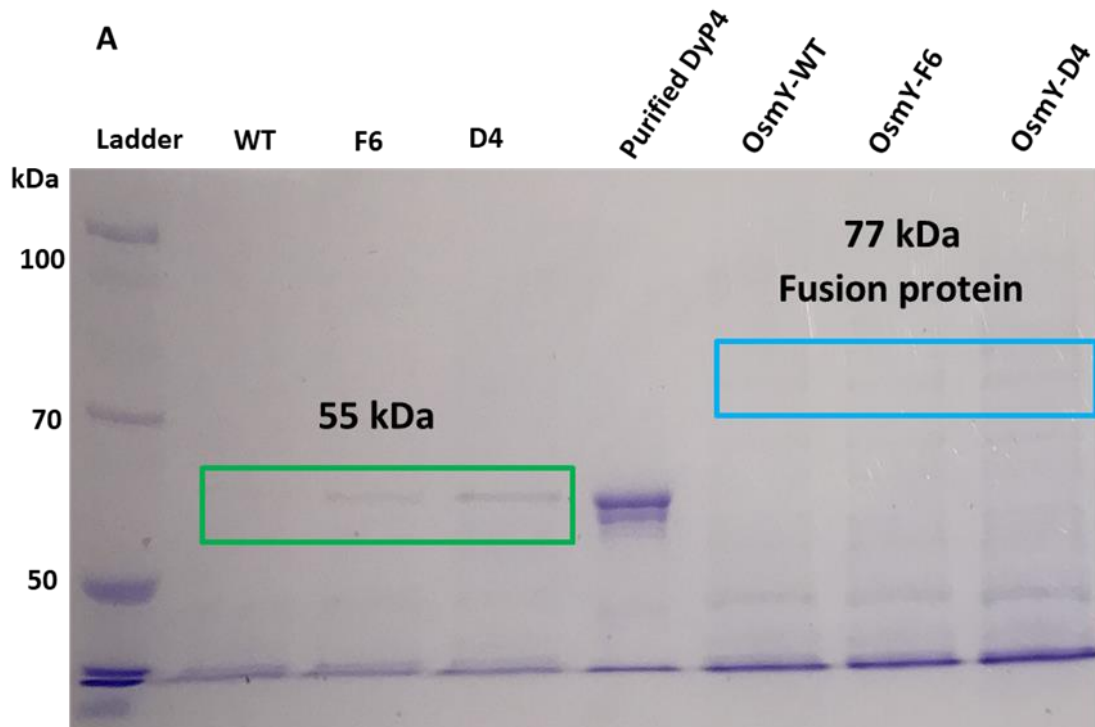


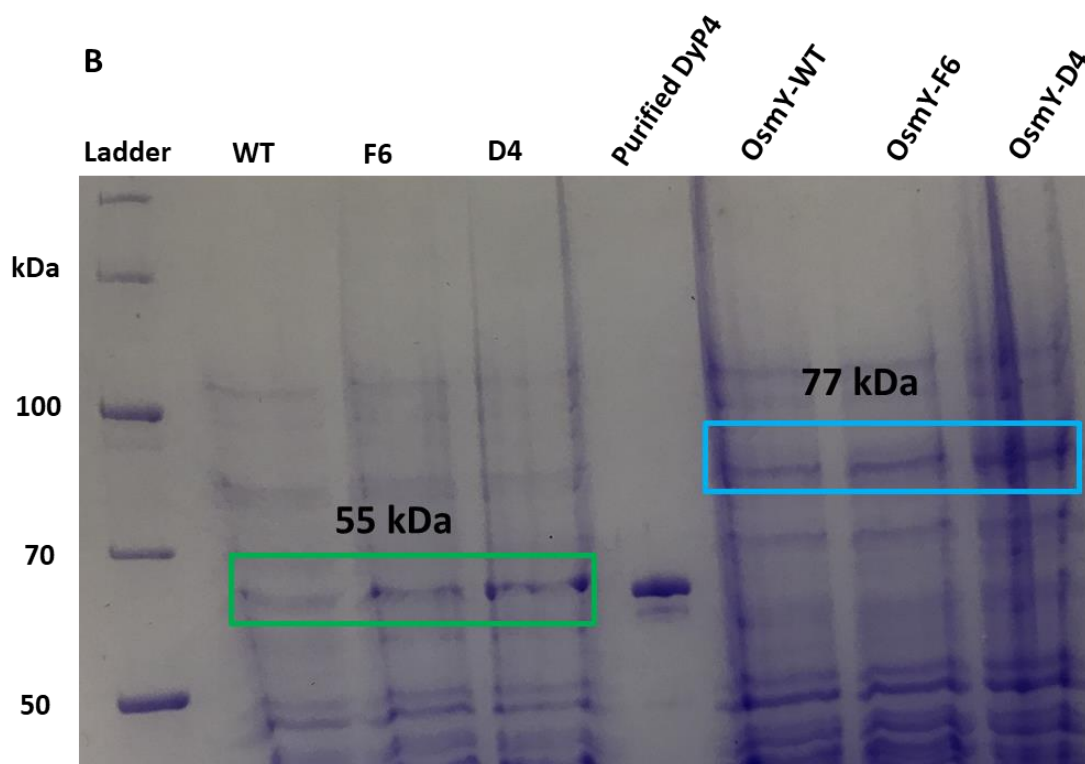
**Figure 6-20:** Four rounds of directed evolution of DyP4 using an OsmY-based secretion mechanism. (Dark Red colour, new mutation; Blue colour, accumulative mutation).

### OsmY as a Fusion Partner for Extracellular Secretion of DyP4 variants

A study was conducted to test whether the secreted proteins can be detected in SDS-PAGE. Samples from the media taken after 24 hours of expression pET24a-WT, pET24a-F6, pET24a-D4, pET24a-OsmY-WT, pET24a-OsmY-F6, and pET24a-OsmY-D4 in 200 mL 2× TY AIM media containing kanamycin at 30°C for 24 hours. Samples from the media were analysed and compared with other samples from the media, treated with acetone precipitation.

The OsmY based secretion system in BL21 (DE3) seemed to be responsible for the secretion of DyP4 into the media as confirmed by the SDS-PAGE results when DyP4 coexpressed with OsmY. This also can be seen in Figure 6-21, where a band at around 55 kDa represent the expression of DyP4 variants without OsmY, while there were bands corresponding to the fusion proteins at around 77 kDa, as suggested by the results of the SDS-PAGE with and without acetone precipitation treatment.





**Figure 6-21:** 10% SDS-PAGE results of the expression and secretion of DyP4-WT and the two mutants (F6 and D4) with and without OsmY. Protein samples (secreted protein in the media) (A) lane 1; molecular ladder, lanes 2–4; crude extract of WT, F6 , and D4 respectively, lane 5; previously purified DyP4 used here as a positive control, lanes 6–7; whole soluble fractions of WT, F6 and D4 respectively after the third step of purification using Superdex 75 column. Protein samples (secreted protein in the media) treated with acetone precipitation (B). lane 1; molecular ladder, lanes 2–4; crude extract of WT, F6 , and D4 respectively, lane 5; previously purified DyP4 used here as a positive control, lanes 6–7; whole soluble fractions of WT, F6 and D4 respectively after the third step of purification using Superdex 75 column.

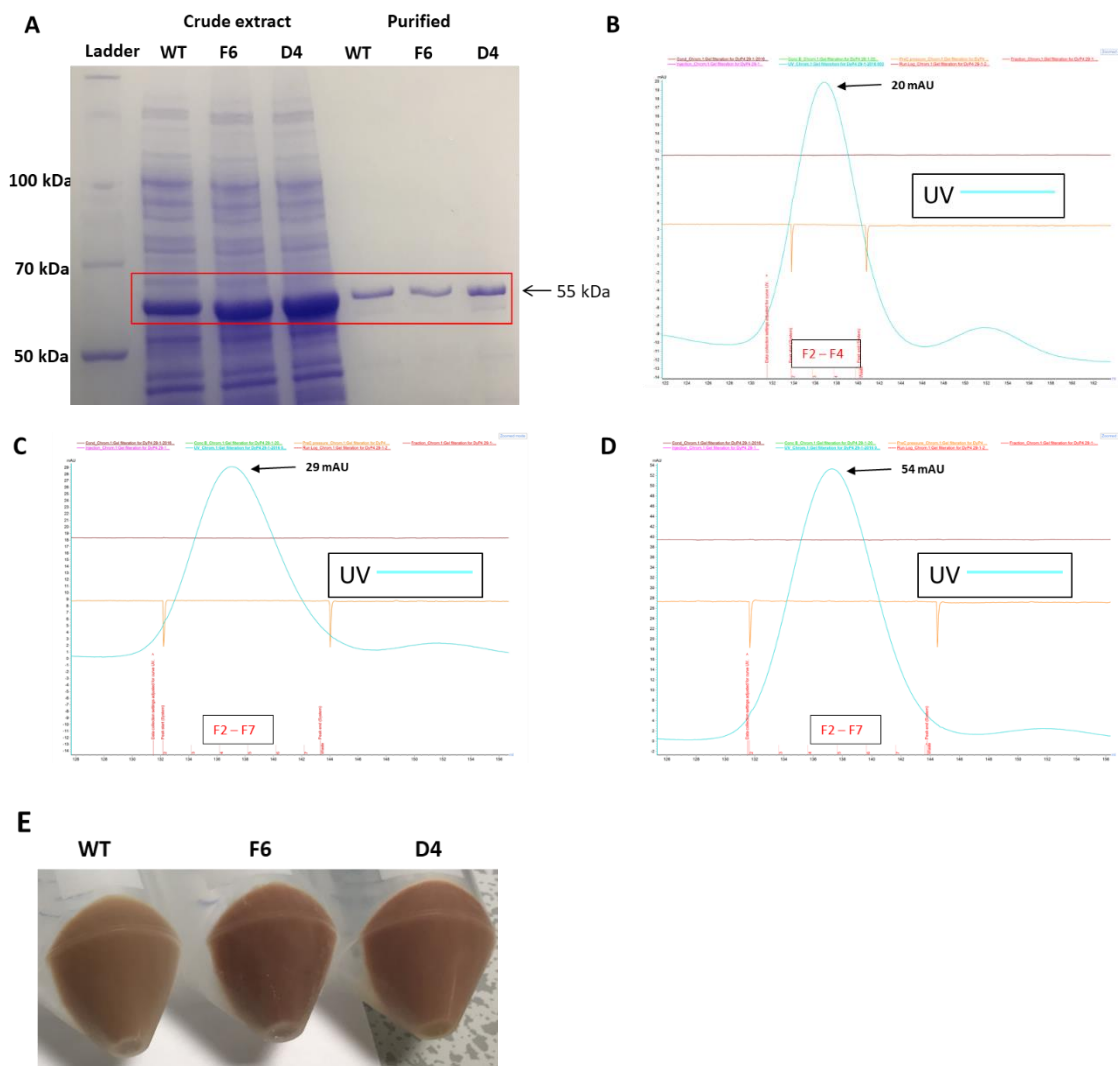
### 6.3.9 Protein expression and purification of DyP4 variants

Following the successful cloning of the two mutants in pET24a, protein expression of the WT and the two mutants (F6 and D4) was performed by culturing in BL21 (DE3) in 200 mL 2× TY AIM media containing kanamycin at 30°C for 24 hours. Centrifugation at 6000 rpm for 5 min at 4°C was performed to harvest the cells. Three steps of protein purification were carried out as described in Chapter 2, (sections 2-8). Purification results showed a significant increase of protein and heme content as shown by the SDS-PAGE results in Figure 6- 22A and as can be seen from the chromatograms of the eluted fractions during size exclusion chromatography of the variants (Figure 6-22B,C, and D) as well as

from the colour of the collected cell pellets after protein expression (Figure 6-22E). The cell pellets of F6 and D4 mutants are darker than those of the WT, indicating a higher expression level (Figure 6-20E).

#### **Enzymatic activities of the WT, and F6 and D4 mutants**

The main improvement obtained after the characterisation up to now is the increased tolerance of the mutants against inactivation by hydrogen peroxide (data not shown). Taken together, these two mutants exhibited higher yield and more resistance to oxidation by hydrogen peroxide, while no significant increase in the catalytic efficiency was obtained using ABTS assay. This might be due to the use of a high concentration of hydrogen peroxide during the evolution of mutants, and this contributes to the identification of mutants showing more resistance to oxidation by hydrogen peroxide.



**Figure 6-22:** Purification of DyP4 variants using three steps of chromatography; (A) SDS-PAGE of the purification of DyP4 variants, lane 1; molecular ladder, lanes 2–4; crude extract of WT, F6, and D4 respectively, lanes 5–7; whole soluble fraction of WT, F6, and D4 respectively after the gel filtration step of the purification. Chromatograms of the eluted fractions after gel filtration step of the purification of WT, F6, and D4 respectively (B), (C) and (D). Cell pellets of WT, F6, and D4 harvested from 200 ml of expression media (E).

## 6.4 Discussion

OsmY-based secretion mechanism was used in the present work to facilitate the screening of more than 4700 clones of DyP4 constructed in four rounds of evolution using either *ep*PCR or site saturation mutagenesis. A colorimetric assay based on the oxidation of ABTS by DyP4 was described in this work. The described HTS assay proved to be reliable, reproducible (with a CV of less than 12%), and sensitive enough to reflect and detect the improved variants obtained through three rounds of *ep*PCR and a fourth round using saturation mutagenesis. It should be noted, however, that some of the natural mediators described in the previous chapter, such as acetosyringone, syringaldehyde, sinapic acid, and violuric acid, were assayed with the secreted DyP4. Unlike for the ABTS assay however, no catalytic oxidation was seen for these assays with the secreted DyP4 using the OsmY secretion mechanism in BL21 (DE3). After the third round, the F6 mutant was used to test these assays again, but no activity was observed. Thus, the ABTS assay was the only assay used so far for the screening of DyP4 variants in this study. After screening 1100 clones in the first round, two significantly improved mutants were detected, showing approximately a two-fold improvement in ABTS activity compared with the WT. These two mutants, D2 and D7, harbouring mutations N312S and A306V, were obtained using the lower and medium mutagenic conditions, respectively. The N312S mutant showed the highest activity with ABTS, and was used to parent the second round of evolution. In the second round, over 550 clones were screened and several mutants with higher activity compared with the parental strain (N321S) were identified. Among these mutants, the A5 mutant harbouring I56V had the highest activity by ABTS assay, with a more than two-fold improvement, and was chosen for parenting the next generation of evolution. In the third round of evolution, more than 1100 clones of DyP4 using the parental strain I56V were screened by ABTS assay. This led to the detection of a mutant with around a two-fold improvement in comparison with the parental strain.

This mutant (F6), harbouring the mutation K109R, was selected to parent the next generation of evolution. In the fourth round, the most clones screened in a single round was more than 2000. However, this was not enough to identify an improved mutant, unlike in the previous rounds. The use of a new strategy became essential to proceed further with the evolution of the enzyme. Based on the positions of the amino acids where improvements were acquired in the first three rounds of evolution (figure 6-20), single site saturation mutagenesis was performed using the most improved mutant so far, the K109R (F6) mutant, as a parental strain. In the first round of evolution, two mutants were identified to have improved in comparison with the WT, mutants N312S and A306V. These two mutants had a single missense mutation and no or more than one silent mutation. Thus, these two positions (Asn and Ala306) were used for the saturation mutagenesis. In the second round, three mutants were identified that showed an improvement in comparison with the parental strain (N312S). These mutants were I56V, H374R, and N227D, and each mutant had a single missense mutation; these positions (Ala56, His374, and Asn227) were therefore used for site saturation mutagenesis. In the third round, mutant R109K was identified as the most improved mutant compared with the parental strain (I56V), and sequencing results revealed that there was only one missense mutation. This amino acid position Lys109 was also used for site saturation mutagenesis. Libraries were created and at least 93 mutants were screened for each position. This was to increase the chance of representing all 20 amino acids at the given position. Saturation mutagenesis for position Asn227 seemed to yield the most improved mutant of all the positions. Three mutants can be seen in Figure 6-21E, where the improvement achieved reached about two-fold in comparison with the parental strain.



Detailed characterisation performed after purification of mutants along the WT showed no significant improvement in the specific activity of mutants with ABTS assay. This might be because, the concentration of hydrogen peroxide used in the evolution was high, and consequently F6 and D4 mutants exhibited significantly improved resistant to oxidation by hydrogen peroxide. F6 and D4 mutants showed 10.6-fold and 14.1-fold improvement against inactivation by hydrogen peroxide compared to the WT. Taken together, these results suggest the potential use of Osmy-based secretion mechanism to facilitate the screening of mutants in directed evolution in *E. coli* not only to improve total activity, but also can be adjusted to increase the tolerance against hydrogen peroxide as shown in this study. It should also be investigated to improve other characteristics (e.g. pH and thermal stability, catalytic efficiency for lignin model substrates, natural mediators, and other industrial dyes).

## 6.5 Conclusion

DyPs, a newly discovered class of peroxidases, are of great biotechnological significance due to the oxidative reactions they catalyse and, more importantly, their role in the decomposition of lignin. This is an essential step towards the complete utilization of lignocellulose derived products: biofuel production and the synthesis of fine chemicals. The present work describes a colorimetric HTS assay to screen variants of DyP4 created by *ep*PCR and saturation mutagenesis, produced and secreted in the *E. coli* strain BL21 (DE3) using a novel secretory mechanism.

Four cycles of directed evolution using *ep*PCR and saturation mutagenesis led to the identification of two mutants (F6 and D4) with a significant improvement in resistant to oxidation by hydrogen peroxide compared the DyP4-WT. As seen from the results obtained in this study, this method facilitated the engineering of DyP4 towards better characterisation for industrial utilization. Importantly, this method might also facilitate the screening process in directed evolution studies for enzymes produced in bacteria; the most studied and characterized system for protein production and directed evolution studies.

## CHAPTER 7 Summary and future work

### 7.1 Summary

In Chapter 3, the successful packaging of DyP4 into an encapsulin nanocompartment from *Rhodococcus erythropolis* was described. Encapsulation was achieved in this study by fusing the C-terminal sequence (37 aa) of ReDypB to the C-terminus of DyP4. Our results showed that the packaged enzyme remained active and, due to low permeability, this activity was reduced in comparison to the activity of the purified enzyme. This is similar to previously reported studies of packaged peroxidases into ReEncapsulin when a decrease in activity is seen (Tamura et al., 2015). The presence of small pores in the nanocompartment prevents the permeability of large molecules, however small sized molecules still go through. The present study provides another evidence of packaging non-native proteins into ReEncapsulin nanocompartment.

It was suggested that encapsulin packages a single copy of DyP (Snijder et al., 2016). The specific targeting sequence can be used to encapsulate a protein of interest in the encapsulin nanocompartment with tremendous potential applications since the packaged proteins are protected from outside environment. In addition, the possibility of loading or release of that protein under certain conditions could have certain applications in nanobiotechnology (Rahmanpour and Bugg, 2013).

In Chapter 4, the use of site-directed mutagenesis to replace and substitute the three Met residues of DyP4 was reported. It was envisaged that this strategy would lead to an increase in the enzyme's oxidative stability, a major limitation of haem-containing peroxidases for industrial applications. Substitution of Met residues in haem containing peroxidases with residues that are more stable against oxidation by hydrogen peroxide using site directed mutagenesis approach have been reported in the literature and found to enhance oxidative stability of the peroxidase. In the present work, the results showed,

however, that there was no improvement in DyP4 stability for hydrogen peroxide. The activity of the WT and variants were similar at 8 mM of hydrogen peroxide as WT and M253F retained 53% and 47% of their activities, respectively. However, at the highest concentration of hydrogen peroxide used in this study (20 mM), WT and mutants retained approximately 26% of their activity.

It remains unknown whether any of these substitutes Met residues are located near the active site of the enzyme as the 3D model structure shares less than 50% of similarity with the relative template used. A possible explanation of why the substitution of Met residues, in this study, did not improve the hydrogen peroxide stability of Dy4 is that the Met residues are located on the surface, as suggested by the 3D model structure. While the stability against hydrogen peroxide has been reported in several studies in literature, the position of substituted residues was a critical factor and affected the results of site-directed mutagenesis, and led to an increase in the stability of the peroxidase (Ogola et al., 2010, Bao et al., 2014).

In Chapter 5, a number of assays were described for DyP4, including VA, the most widely used assay for lignin degradation. This result is in accordance with the ability of DyPs to directly catalyse the oxidation of non-phenolic substrates, unlike laccases and MnP, which require mediators. In addition, a number of colorimetric assays based on the oxidation of naturally occurring mediators were reported. Our results may facilitate the design of HTS assays to engineer DyP4 towards enhanced characteristics for more efficient conversion of plant biomass.

In Chapter 6, a detailed protocol for mutant construction, secretion, and screening in *E. coli* BL21 (DE3) was provided. The most commonly used method in directed evolution studies, *ep*PCR, was used here to construct the libraries. For the first time, the screening process was carried out in *E. coli* with help of an OsmY-based protein secretion

mechanism. To date, four generations of directed evolution led to a mutant with around a 10-fold improvement in total activity and a significant improvement of tolerance against hydrogen peroxidase inactivation, the later is a major limiting step of using peroxidases on industrial scale. More structural investigations needed to predict the effect of these mutations on the improvement of tolerance against hydrogen peroxide. If OsmY-based secretion system validated, it would largely simplify and facilitate the screening process in *E. coli*, however, more studies needed to confirm that this mechanism can be used widely in directed evolution studies in *E. coli*. The present study highlights the powerfulness of directed evolution in improving the tolerance of Pleos-DyP4 against inactivation by hydrogen peroxide and maybe other properties if the screening condition of mutants adjusted specifically to target certain priority. For instance, method might be adjusted to specifically improve a targeted property for example, catalytic efficiencies for lignin degradation similar to previously reported studies (Brissos et al., 2017). Upon the validation of this method in directed evolution studies in *E. coli*, the prospective are opened to use this method to improve properties of other enzymes of the DyP-type peroxidase family.

## 7.2 Prospective

It would be interesting to investigate whether packaged DyP4 will be released upon treatment of the specie at low pH. In addition, assembly and disassembly should be investigated by treatment at low and high pH where the enzyme is encapsulated and released. Also, dynamic light scattering can be performed to visualize the nanocompartment in the His-tagged form and the nanocompartment with DyP4 packaged inside.

A packaged DyP was found to show enhanced catalytic conversion compared to the purified protein and this might lead to an application of encapsulation in the deconstruction of lignocellulose (Rahmanpour and Bugg, 2013). This calls for investigation to see if packaged DyP4 would have increased catalytic conversion for lignin model substrates.

In order to prove the findings obtained in (Chapter 4) where no improvement against hydrogen peroxide was acquired, two different tests can be carried out. In the first test, the variants of DyP4 are incubated at different concentrations of H<sub>2</sub>O<sub>2</sub> (0-20 mM) H<sub>2</sub>O<sub>2</sub> at 37°C for 20 min to inactivate all variants. After 20 min, the inactivation is stopped by immediate dilution to lower concentration of hydrogen peroxide with incubation on ice. Residuals activity using the treated solutions is conducted with ABTS assay using the average of at least three replicates. In the second test, WT-Pleos-DyP4 and mutant variants can be incubated with 4 or 8 mM of hydrogen peroxide at room temperature and spectral measurements are conducted to determine the haem stability of the WT and mutants.

To further extend substrate spectrum of DyP4 and more importantly lignin related substrates, two lignin model compounds namely 1,3,5-trimethoxybenzene and guaiacylglycerol-beta-guaiacyl ether can be used similar to VA oxidation test. The

detection of the oxidation product however requires different mode of action for these two substrates, unlike VA. For 1,3,5-trimethoxybenzene and guaiacylglycerol-beta-guaiacyl ether, high-performance liquid chromatography (HPLC) will be used to analyse the oxidation reaction and detect the products.

Several further rounds of directed evolution might lead to the identification of a significantly improved mutant than that identified so far which has ~10-fold improvement in total activity. After three successful rounds of directed evolution, using *ep*PCR and a fourth round using saturation mutagenesis, a fifth round using *ep*PCR might lead to further improvement of total activity for DyP4.

The use of alternative assays to the ABTS assay to screen for improved mutants of DyP4 will facilitate the identification of mutants more suited for industrial utilization and specifically with improved catalytic conversion of plant biomass. For example, using the D4 mutant or the improved mutant from the fifth round of evolution using *ep*PCR, the acetosyringon or sinapic acid assays could be leveraged. This is informed by the fact that that these assays have already optimized for the 1-ml cuvette oxidation measurements using purified WT (Chapter 5). Upon the identification of a mutant where oxidation is detected in 96-well microtitre plates, optimization of HTS assays using these substrates can be performed. Other lignin model substrates such as DMP can also be optimized for the screening of mutants in HTS form. Fernandez-Fueyo *et al.* have assayed DMP using DyP4 and monitoring oxidation by the preformation of dimeric coerulignone at 469 nm (Fernandez-Fueyo *et al.*, 2015). On this note, this substrate can be used for oxidation test in 96-well microtitre plate using the D4 mutant. Detection of activity with this mutant will inform further evolution of DyP4.

OsmY based secretion mechanism should benefit the secretion of proteins in *E. coli* and facilitate the screening process in directed evolution studies by reducing the number of steps required. Using the WT-OsmY, the secretion in the 96-well microtiter plates was only possible with ABTS assay. No activity was detected when other assays such as acetyringon and sinapic acid were used. This could be due to insufficient amount of secreted enzyme required to catalyse the oxidation of these substrates in the 96-well microtiter plates. One of the possible solutions to overcome this issue is to evolve the OsmY protein to enhance its secretion potential and hence achieve several fold improvement of secretion of DyP4 in *E. coli*.



## References

- Achinas, S. & Euverink, G. J. W. 2016. Consolidated briefing of biochemical ethanol production from lignocellulosic biomass. *Electronic Journal of Biotechnology*, 23, 44-53.
- Adenipekun, C., Lawal, R. J. B. & Reviews, M. B. 2012. Uses of mushrooms in bioremediation: A review. 7, 62-68.
- Agbor, V. B., Cicek, N., Sparling, R., Berlin, A. & Levin, D. B. 2011. Biomass pretreatment: Fundamentals toward application. *Biotechnology Advances*, 29, 675-685.
- Aken, B. V., Correa, P. A. & Schnoor, J. L. 2010. Phytoremediation of polychlorinated biphenyls: new trends and promises. *Environ Sci Technol*, 44, 2767-2776.
- Alcalde, M., Ferrer, M., Plou, F. J. & Ballesteros, A. 2006. Environmental biocatalysis: from remediation with enzymes to novel green processes. *Trends in Biotechnology*, 24, 281-287.
- Alvira, P., Tomas-Pejo, E., Ballesteros, M. & Negro, M. J. 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresour Technol*, 101, 4851-4861.
- Aranda, E., Ullrich, R. & Hofrichter, M. 2010. Conversion of polycyclic aromatic hydrocarbons, methyl naphthalenes and dibenzofuran by two fungal peroxygenases. *Biodegradation*, 21, 267-281.
- Arnold, F. H. & Georgiou, G. 2003. *Directed enzyme evolution: screening and selection methods*, Springer Science & Business Media.
- Arora, D. S. & Gill, P. K. 2001. Comparison of two assay procedures for lignin peroxidase. *Enzyme and Microbial Technology*, 28, 602-605.
- Asgher, M., Bashir, F. & Iqbal, H. M. N. 2014. A comprehensive ligninolytic pre-treatment approach from lignocellulose green biotechnology to produce bio-ethanol. *Chemical Engineering Research & Design*, 92, 1571-1578.
- Ayala, M., Batista, C. V. & Vazquez-Duhalt, R. 2011. Heme destruction, the main molecular event during the peroxide-mediated inactivation of chloroperoxidase from *Caldariomyces fumago*. *Journal of Biological Inorganic Chemistry*, 16, 63-68.
- Baiocco, P., Barreca, A. M., Fabbrini, M., Galli, C. & Gentili, P. 2003. Promoting laccase activity towards non-phenolic substrates: a mechanistic investigation with some laccase-mediator systems. *Organic & Biomolecular Chemistry*, 1, 191-197.
- Balat, M. 2011. Production of bioethanol from lignocellulosic materials via the biochemical pathway: A review. *Energy Conversion and Management*, 52, 858-875.
- Bansal, N. & Kanwar, S. S. 2013. Peroxidase(s) in Environment Protection. *Scientific World Journal*.
- Bao, X., Huang, X. N., Lu, X. F. & Li, J. J. 2014. Improvement of hydrogen peroxide stability of *Pleurotus eryngii* versatile ligninolytic peroxidase by rational protein engineering. *Enzyme and Microbial Technology*, 54, 51-58.
- Battistuzzi, G., Bellei, M., Bortolotti, C. A. & Sola, M. 2010. Redox properties of heme peroxidases. *Archives of Biochemistry and Biophysics*, 500, 21-36.
- Biswas, R., Uellendahl, H. & Ahring, B. K. 2015. Wet explosion: a universal and efficient pretreatment process for lignocellulosic biorefineries. *BioEnergy Research*, 8, 1101-1116.
- Block, H., Maertens, B., Priestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J. & Schafer, F. 2009. Immobilized-Metal Affinity Chromatography (Imac): A Review. *Guide to Protein Purification, Second Edition*, 463, 439-473.

- Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. & Schwede, T. 2009. Protein structure homology modeling using SWISS-MODEL workspace. *Nature Protocols*, 4, 1-13.
- Bridgeman, T. G., Jones, J. M., Shield, I. & Williams, P. T. 2008. Torrefaction of reed canary grass, wheat straw and willow to enhance solid fuel qualities and combustion properties. *Fuel*, 87, 844-856.
- 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 Catalysis*, 7, 3454-3465.
- Brosse, N., Dufour, A., Meng, X., Sun, Q. & Ragauskas, A. 2012. Miscanthus: a fast-growing crop for biofuels and chemicals production. *Biofuels, Bioproducts and Biorefining*, 6, 580-598.
- Bugg, T. D. H., Ahmad, M., Hardiman, E. M. & Rahmanpour, R. 2011. Pathways for degradation of lignin in bacteria and fungi. *Natural Product Reports*, 28, 1883-1896.
- Bugg, T. D. H. & Rahmanpour, R. 2015. Enzymatic conversion of lignin into renewable chemicals. *Current Opinion in Chemical Biology*, 29, 10-17.
- Camarero, S., Canas, A. I., Nousiainen, P., Record, E., Lomascolo, A., Martinez, M. J. & Martinez, A. T. 2008. p-hydroxycinnamic acids as natural mediators for laccase oxidation of recalcitrant compounds. *Environmental Science & Technology*, 42, 6703-6709.
- Camarero, S., Ibarra, D., Martinez, A. T., Romero, J., Gutierrez, A. & del Rio, J. C. 2007. Paper pulp delignification using laccase and natural mediators. *Enzyme and Microbial Technology*, 40, 1264-1271.
- Camarero, S., Pardo, I., Canas, A. I., Molina, P., Record, E., Martinez, A. T., Martinez, M. J. & Alcalde, M. 2012. Engineering Platforms for Directed Evolution of Laccase from *Pycnoporus cinnabarinus*. *Applied and Environmental Microbiology*, 78, 1370-1384.
- Canas, A. I. & Camarero, S. 2010. Laccases and their natural mediators: Biotechnological tools for sustainable eco-friendly processes. *Biotechnology Advances*, 28, 694-705.
- Cara, C., Ruiz, E., Oliva, J. M., Saez, F. & Castro, E. 2008. Conversion of olive tree biomass into fermentable sugars by dilute acid pretreatment and enzymatic saccharification. *Bioresource Technology*, 99, 1869-1876.
- Chen, R. D. 2001. Enzyme engineering: rational redesign versus directed evolution. *Trends in Biotechnology*, 19, 13-14.
- Chen, Z. & Wan, C. X. 2017. Biological valorization strategies for converting lignin into fuels and chemicals. *Renewable & Sustainable Energy Reviews*, 73, 610-621.
- Cherry, J. R. & Lamsa, M. H. 2004. Screening for oxidative resistance. *Protein Engineering*, 388, 167-175.
- Cherry, J. R., Lamsa, M. H., Schneider, P., Vind, J., Svendsen, A., Jones, A. & Pedersen, A. H. 1999. Directed evolution of a fungal peroxidase. *Nature Biotechnology*, 17, 379-384.
- Christopher, L. P., Yao, B. & Ji, Y. 2014. Lignin Biodegradation with Laccase-Mediator Systems. *Frontiers in Energy Research*, 2, 1-13.
- Colpa, D. I., Fraaije, M. W. & van Bloois, E. 2014. DyP-type peroxidases: a promising and versatile class of enzymes. *Journal of Industrial Microbiology & Biotechnology*, 41, 1-7.

- Conesa, A., Punt, P. J. & van den Hondel, C. A. M. J. J. 2002. Fungal peroxidases: molecular aspects and applications. *Journal of Biotechnology*, 93, 143-158.
- 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. J. & Goulding, C. W. 2014. Characterization of a Mycobacterium tuberculosis Nanocompartment and Its Potential Cargo Proteins. *Journal of Biological Chemistry*, 289, 18279-18289.
- Couto, S. R. & Herrera, J. L. T. 2006. Industrial and biotechnological applications of laccases: A review. *Biotechnology Advances*, 24, 500-513.
- Dashtban, M., Schraft, H., Syed, T. A. & Qin, W. 2010. Fungal biodegradation and enzymatic modification of lignin. *Int J Biochem Mol Biol*, 1, 36-50.
- de Souza, S. M. A. G. U., Forgiarini, E. & de Souza, A. A. U. 2007. Toxicity of textile dyes and their degradation by the enzyme horseradish peroxidase (HRP). *Journal of Hazardous Materials*, 147, 1073-1078.
- Dijkerman, R., Bhansing, D. C., den Camp, H. J. O., van der Drift, C. & Vogels, G. D. 1997. Degradation of structural polysaccharides by the plant cell-wall degrading enzyme system from anaerobic fungi: an application study. *Enzyme and microbial technology*, 21, 130-136.
- Duong-Ly, K. C. & Gabelli, S. B. 2014. Using Ion Exchange Chromatography to Purify a Recombinantly Expressed Protein. *Laboratory Methods in Enzymology: Protein, Pt C*, 541, 95-103.
- Fawal, N., Li, Q., Savelli, B., Brette, M., Passaia, G., Fabre, M., Mathe, C. & Dunand, C. 2013. PeroxiBase: a database for large-scale evolutionary analysis of peroxidases. *Nucleic Acids Research*, 41, D441-D444.
- Fernandez-Fueyo, E., Davo-Siguero, I., Almendral, D., Linde, D., Baratto, M. C., Pogni, R., Romero, A., Guallar, V. & Martinez, A. T. 2018. Description of a Non-Canonical Mn(II)-Oxidation Site in Peroxidases. *Acs Catalysis*, 8, 8386-8395.
- Fernandez-Fueyo, E., Linde, D., Almendral, D., Lopez-Lucendo, M. F., Ruiz-Duenas, F. J. & Martinez, A. T. 2015. Description of the first fungal dye-decolorizing peroxidase oxidizing manganese(II). *Applied Microbiology and Biotechnology*, 99, 8927-8942.
- FitzPatrick, M., Champagne, P., Cunningham, M. F. & Whitney, R. A. 2010. A biorefinery processing perspective: Treatment of lignocellulosic materials for the production of value-added products. *Bioresource Technology*, 101, 8915-8922.
- Galzie, Z. 1991. What Is Protein Engineering. *Biochemical Education*, 19, 74-75.
- Garcia-Ruiz, E., Gonzalez-Perez, D., Ruiz-Duenas, F. J., Martinez, A. T. & Alcalde, M. 2012. Directed evolution of a temperature-, peroxide- and alkaline pH-tolerant versatile peroxidase. *Biochemical Journal*, 441, 487-498.
- Giessen, T. W. 2016. Encapsulins: microbial nanocompartments with applications in biomedicine, nanobiotechnology and materials science. *Current opinion in chemical biology*, 34, 1-10.
- Giessen, T. W. & Silver, P. A. 2017. Widespread distribution of encapsulin nanocompartments reveals functional diversity. *Nature Microbiology*, 2, 17029.
- Gonzalez-Perez, D. & Alcalde, M. 2018. The making of versatile peroxidase by directed evolution. *Biocatalysis and Biotransformation*, 36, 1-11.
- Gonzalez-Perez, D., Garcia-Ruiz, E., Ruiz-Duenas, F. J., Martinez, A. T. & Alcalde, M. 2014. Structural Determinants of Oxidative Stabilization in an Evolved Versatile Peroxidase. *Acs Catalysis*, 4, 3891-3901.

- Gonzalez-Perez, D., Mateljak, I., Garcia-Ruiz, E., Ruiz-Duenas, F. J., Martinez, A. T. & Alcalde, M. 2016. Alkaline versatile peroxidase by directed evolution (vol 6, pg 6625, 2016). *Catalysis Science & Technology*, 6, 8375-8375.
- Gutierrez, A., Babot, E. D., Ullrich, R., Hofrichter, M., Martinez, A. T. & del Rio, J. C. 2011. Regioselective oxygenation of fatty acids, fatty alcohols and other aliphatic compounds by a basidiomycete heme-thiolate peroxidase. *Arch Biochem Biophys*, 514, 33-43.
- Haffner, F. B., Mitchell, V. D., Arundale, R. A. & Bauer, S. 2013. Compositional analysis of *Miscanthus giganteus* by near infrared spectroscopy. *Cellulose*, 20, 1629-1637.
- Hamid, M. & Khalil-ur-Rehman 2009. Potential applications of peroxidases. *Food Chemistry*, 115, 1177-1186.
- Hatakka, A. & Hammel, K. E. 2010. Fungal Biodegradation of Lignocelluloses. *Mycota: Industrial Applications, Vol 10, Second Edition*, 10, 319-340.
- He, D. D., Hughes, S., Vanden-Hehir, S., Georgiev, A., Altenbach, K., Tarrant, E., Mackay, C. L., Waldron, K. J., Clarke, D. J. & Marles-Wright, J. 2016. Structural characterization of encapsulated ferritin provides insight into iron storage in bacterial nanocompartments. *Elife*, 5, 18972.
- Hofrichter, M., Ullrich, R., Pecyna, M. J., Liers, C. & Lundell, T. 2010. New and classic families of secreted fungal heme peroxidases. *Applied Microbiology and Biotechnology*, 87, 871-897.
- Hollmann, F. & Arends, I. W. C. E. 2012. Enzyme Initiated Radical Polymerizations. *Polymers*, 4, 759-793.
- Inui, H., Itoh, T., Yamamoto, K., Ikushiro, S. I. & Sakaki, T. 2014. Mammalian Cytochrome P450-Dependent Metabolism of Polychlorinated Dibenzo-p-dioxins and Coplanar Polychlorinated Biphenyls. *International Journal of Molecular Sciences*, 15, 14044-14057.
- Isikgor, F. H. & Becer, C. R. 2015. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry*, 6, 4497-4559.
- Janusz, G., Pawlik, A., Sulej, J., Swiderska-Burek, U., Jarosz-Wilkolazka, A. & Paszczynski, A. 2017. Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. *Fems Microbiology Reviews*, 41, 941-962.
- Jarvinen, J., Taskila, S., Isomaki, R. & Ojamo, H. 2012. Screening of white-rot fungi manganese peroxidases: a comparison between the specific activities of the enzyme from different native producers. *Amb Express*, 2, 62.
- Karigar, C. S. & Rao, S. S. 2011. Role of microbial enzymes in the bioremediation of pollutants: a review. *Enzyme Res*, 2011, 805187.
- Kasai, N., Ikushiro, S., Shinkyo, R., Yasuda, K., Hirose, S., Arisawa, A., Ichinose, H., Wariishi, H. & Sakaki, T. 2010. Metabolism of mono- and dichloro-dibenzo-p-dioxins by *Phanerochaete chrysosporium* cytochromes P450. *Applied Microbiology and Biotechnology*, 86, 773-780.
- Khan, H. U. 2012. The Role of Ion Exchange Chromatography in Purification and Characterization of Molecules. *Ion Exchange Technologies*, 331-342.
- Kim, D. J. M. 2018. Physico-chemical conversion of lignocellulose: Inhibitor effects and detoxification strategies: A mini review. 23, 309.
- Kim, H. C., Mickel, M. & Hampf, N. 2003. Molecular origin of the stability of violuric acid radicals at high pH-values. *Chemical Physics Letters*, 371, 410-416.

- Kim, T. H., Taylor, F. & Hicks, K. B. 2008. Bioethanol production from barley hull using SAA (soaking in aqueous ammonia) pretreatment. *Bioresource Technology*, 99, 5694-5702.
- Kornilowicz-Kowalska, T. & Rybczynska, K. 2014. Anthraquinone dyes decolorization capacity of anamorphic *Bjerkandera adusta* CCBAS 930 strain and its HRP-like negative mutants. *World Journal of Microbiology & Biotechnology*, 30, 1725-1736.
- Kumar, A. K. & Sharma, S. 2017. Recent updates on different methods of pretreatment of lignocellulosic feedstocks: a review. *Bioresour Bioprocess*, 4, p. 7.
- Kumar, P., Barrett, D. M., Delwiche, M. J. & Stroeve, P. 2009. Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Industrial & Engineering Chemistry Research*, 48, 3713-3729.
- Kumar, R., Singh, S. & Singh, O. V. 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of Industrial Microbiology & Biotechnology*, 35, 377-391.
- Kumar, S. 2010. Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert Opinion on Drug Metabolism & Toxicology*, 6, 115-131.
- Kunamneni, A., Camarero, S., Garcia-Burgos, C., Plou, F. J., Ballesteros, A. & Alcalde, M. 2008. Engineering and Applications of fungal laccases for organic synthesis. *Microbial Cell Factories*, 7, 32-49.
- Lambertz, C., Ece, S., Fischer, R. & Commandeur, U. 2016. Progress and obstacles in the production and application of recombinant lignin-degrading peroxidases. *Bioengineered*, 7, 145-154.
- Leemhuis, H., Kelly, R. M. & Dijkhuizen, L. 2009. Directed Evolution of Enzymes: Library Screening Strategies. *Iubmb Life*, 61, 222-228.
- Li, A. T., Acevedo-Rocha, C. G. & Reetz, M. T. 2018. Boosting the efficiency of site-saturation mutagenesis for a difficult-to-randomize gene by a two-step PCR strategy. *Applied Microbiology and Biotechnology*, 102, 6095-6103.
- Li, D. M., Li, N., Ma, B., Mayfield, M. B. & Gold, M. H. 1999. Characterization of genes encoding two manganese peroxidases from the lignin-degrading fungus *Dichomitus squalens*. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology*, 1434, 356-364.
- Li, H., Kim, N. J., Jiang, M., Kang, J. W. & Chang, H. N. 2009. Simultaneous saccharification and fermentation of lignocellulosic residues pretreated with phosphoric acid-acetone for bioethanol production. *Bioresour Technol*, 100, 3245-3251.
- Liers, C., Bobeth, C., Pecyna, M., Ullrich, R. & Hofrichter, M. 2010. DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes. *Applied Microbiology and Biotechnology*, 85, 1869-1879.
- Longoria, A., Tinoco, R. & Vazquez-Duhalt, R. 2008. Chloroperoxidase-mediated transformation of highly halogenated monoaromatic compounds. *Chemosphere*, 72, 485-490.
- Lorenz, P. & Eck, J. 2005. Metagenomics and industrial applications. *Nature Reviews Microbiology*, 3, 510-516.
- Maciel, M. J. M., Silva, A. C. E. & Ribeiro, H. C. T. 2010. Industrial and biotechnological applications of ligninolytic enzymes of the basidiomycota: A review. *Electronic Journal of Biotechnology*, 13, 14-15.

- 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. *Current Opinion in Biotechnology*, 20, 348-357.
- Menon, V. & Rao, M. 2012a. Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. *Progress in Energy and Combustion Science*, 38, 522-550.
- Menon, V. & Rao, M. 2012b. Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. *Progress in Energy and Combustion Science*, 38, 522-550.
- 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. *Scientific Reports*, 5, 8245.
- Min, K., Kim, S., Yum, T., Kim, Y., Sang, B. I. & Um, Y. 2013. Conversion of levulinic acid to 2-butanone by acetoacetate decarboxylase from *Clostridium acetobutylicum*. *Applied Microbiology and Biotechnology*, 97, 5627-5634.
- Miyazaki-Imamura, C., Oohira, K., Kitagawa, R., Nakano, H., Yamane, T. & Takahashi, H. 2003. Improvement of H<sub>2</sub>O<sub>2</sub> stability of manganese peroxidase by combinatorial mutagenesis and high-throughput screening using in vitro expression with protein disulfide isomerase. *Protein Engineering*, 16, 423-428.
- Miyazaki, C. & Takahashi, H. 2001. Engineering of the H<sub>2</sub>O<sub>2</sub>-binding pocket region of a recombinant manganese peroxidase to be resistant to H<sub>2</sub>O<sub>2</sub>. *Febs Letters*, 509, 111-114.
- Mohan, U., Kaushik, S. & Banerjee, U. C. 2011. PCR based random mutagenesis approach for a defined DNA sequence using the mutagenic potential of oxidized nucleotide products. *Open Biotechnol J*, 5, 21-27.
- Molina-Espeja, P., Garcia-Ruiz, E., Gonzalez-Perez, D., Ullrich, R., Hofrichter, M. & Alcalde, M. 2014. Directed evolution of unspecific peroxygenase from *Agrocybe aegerita*. *Applied and environmental microbiology*, 80, 3496-3507.
- Molina-Espeja, P., Ma, S., Mate, D. M., Ludwig, R. & Alcalde, M. 2015. Tandem-yeast expression system for engineering and producing unspecific peroxygenase. *Enzyme and Microbial Technology*, 73-74, 29-33.
- Morawski, B., Quan, S. & Arnold, F. H. 2001. Functional expression and stabilization of horseradish peroxidase by directed evolution in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 76, 99-107.
- Moreno, A. D., Ibarra, D., Alvira, P., Tomas-Pejo, E. & Ballesteros, M. 2015. A review of biological delignification and detoxification methods for lignocellulosic bioethanol production. *Crit Rev Biotechnol*, 35, 342-354.
- Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M. & Ladisch, M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*, 96, 673-686.
- Mrozik, A., Piotrowska-Seget, Z. & Labuzek, S. 2003. Bacterial degradation and bioremediation of polycyclic aromatic hydrocarbons. *Polish Journal of Environmental Studies*, 12, 15-25.
- Nannemann, D. P., Birmingham, W. R., Scism, R. A. & Bachmann, B. O. 2011. Assessing directed evolution methods for the generation of biosynthetic enzymes with potential in drug biosynthesis. *Future Medicinal Chemistry*, 3, 803-819.

- Nichols, R. J., Cassidy-Amstutz, C., Chaijarasphong, T. & Savage, D. F. 2017. Encapsulins: molecular biology of the shell. *Critical Reviews in Biochemistry and Molecular Biology*, 52, 583-594.
- Nousiainen, P., Kontro, J., Maijala, P., Uzan, E., Hatakka, A., Lomascolo, A. & Sipila, J. 2012. Lignin Model Compound Studies To Elucidate the Effect of "Natural" Mediators on Oxidoreductase-Catalyzed Degradation of Lignocellulosic Materials. *Functional Materials from Renewable Sources*, 1107, 229-242.
- Ogola, H. J. O., Ashida, H., Ishikawa, T. & Sawa, Y. 2015. Explorations and Applications of Enzyme-linked Bioremediation of Synthetic Dyes. *Advances in Bioremediation of Wastewater and Polluted Soil*, 111-144.
- Ogola, H. J. O., 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. *Applied Microbiology and Biotechnology*, 87, 1727-1736.
- Packer, M. S. & Liu, D. R. 2015. Methods for the directed evolution of proteins. *Nature Reviews Genetics*, 16, 379-394.
- Pardo, I., Chanaga, X., Vicente, A. I., Alcalde, M. & Camarero, S. 2013. New colorimetric screening assays for the directed evolution of fungal laccases to improve the conversion of plant biomass. *Bmc Biotechnology*, 13, 90.
- Pecyna, M. J., Ullrich, R., Bittner, B., Clemens, A., Scheibner, K., Schubert, R. & Hofrichter, M. 2009. Molecular characterization of aromatic peroxygenase from *Agrocybe aegerita*. *Appl Microbiol Biotechnol*, 84, 885-897.
- Pérez, J., Muñoz-Dorado, J., de la Rubia, T. & Martínez, J. 2002. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *International microbiology*, 5, 53-63.
- Piontek, K., Strittmatter, E., Ullrich, R., Gröbe, G., Pecyna, M. J., Kluge, M., Scheibner, K., Hofrichter, M. & Plattner, D. A. 2013. Structural basis of substrate conversion in a new aromatic peroxygenase cytochrome P450 functionality with benefits. *Journal of Biological Chemistry*, 288, 34767-34776.
- Piontek, K., Ullrich, R., Liers, C., Diederichs, K., Plattner, D. A. & Hofrichter, M. 2010. Crystallization of a 45 kDa peroxygenase/peroxidase from the mushroom *Agrocybe aegerita* and structure determination by SAD utilizing only the haem iron. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 66, 693-698.
- Plácido, J. & Capareda, S. 2015. Ligninolytic enzymes: a biotechnological alternative for bioethanol production. *Bioresources and Bioprocessing*, 2, 1-12.
- Pogni, R., Brogioni, B., Baratto, M. C., Sinicropi, A., Giardina, P., Pezzella, C., Sanna, G. & Basosi, R. 2007. Evidence for a radical mechanism in biocatalytic degradation of synthetic dyes by fungal laccases mediated by violuric acid. *Biocatalysis and Biotransformation*, 25, 269-275.
- Pollegioni, L., Tonin, F. & Rosini, E. 2015. Lignin-degrading enzymes. *Febs Journal*, 282, 1190-1213.
- Poraj-Kobielska, M., Kinne, M., Ullrich, R., Scheibner, K. & Hofrichter, M. 2012. A spectrophotometric assay for the detection of fungal peroxygenases. *Anal Biochem*, 421, 327-329.
- Porter, J. L., Rusli, R. A. & Ollis, D. L. 2016. Directed Evolution of Enzymes for Industrial Biocatalysis. *ChemBiochem*, 17, 197-203.

- Pourmir, A. & Johannes, T. W. 2012. Directed evolution: selection of the host organism. *Computational and structural biotechnology journal*, 2, e201209012.
- Qian, Z. G., Xia, X. X., Choi, J. H. & Lee, S. Y. 2008. Proteome-based identification of fusion partner for high-level extracellular production of recombinant proteins in *Escherichia coli*. *Biotechnology and Bioengineering*, 101, 587-601.
- Qin, X., Sun, X., Huang, H., Bai, Y., Wang, Y., Luo, H., Yao, B., Zhang, X. & Su, X. 2017. Oxidation of a non-phenolic lignin model compound by two *Irpex lacteus* manganese peroxidases: evidence for implication of carboxylate and radicals. *Biotechnology for biofuels*, 10, p. 103.
- Rahmanpour, R. & Bugg, T. D. H. 2013. Assembly in vitro of *Rhodococcus jostii* RHA1 encapsulin and peroxidase DypB to form a nanocompartment. *Febs Journal*, 280, 2097-2104.
- Reymond, J.-L. 2006. *Enzyme assays : high-throughput screening, genetic selection, and fingerprinting*, Weinheim, Germany, Wiley-VCH.
- Ribeiro, A. L., Mencia, M. & Hidalgo, A. 2018. A Brief Guide to the High-Throughput Expression of Directed Evolution Libraries. *Methods Mol Biol*, 1685, 131-143.
- 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. *Journal of Experimental Botany*, 60, 441-452.
- Rytioja, J., Hilden, K., Yuzon, J., Hatakka, A., de Vries, R. P. & Makela, M. R. 2014. Plant-Polysaccharide-Degrading Enzymes from Basidiomycetes. *Microbiology and Molecular Biology Reviews*, 78, 614-649.
- Sakaki, T. & Munetsuna, E. 2010. Enzyme systems for biodegradation of polychlorinated dibenzo-p-dioxins. *Applied Microbiology and Biotechnology*, 88, 23-30.
- Sakaki, T., Yamamoto, K. & Ikushiro, S. 2013. Possibility of application of cytochrome P450 to bioremediation of dioxins. *Biotechnology and Applied Biochemistry*, 60, 65-70.
- Sanchez, C. 2009. Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnology Advances*, 27, 185-194.
- Sarkar, N., Ghosh, S. K., Bannerjee, S. & Aikat, K. 2012. Bioethanol production from agricultural wastes: An overview. *Renewable Energy*, 37, 19-27.
- Scheller, H. V. & Ulvskov, P. 2010. Hemicelluloses. *Annual Review of Plant Biology*, Vol 61, 61, 263-289.
- Schnepf, R. & Yacobucci, B. D. Renewable Fuel Standard (RFS): overview and issues. CRS Report for Congress, 2010.
- Shahzadi, T., Mehmood, S., Irshad, M., Anwar, Z., Afroz, A., Zeeshan, N., Rashid, U., Sughra, K. J. A. i. B. & Biotechnology 2014. Advances in lignocellulosic biotechnology: A brief review on lignocellulosic biomass and cellulases. 5, 246-251.
- Shrestha, R., Huang, G. C., Meekins, D. A., Geisbrecht, B. V. & Li, P. 2017. Mechanistic Insights into Dye-Decolorizing Peroxidase Revealed by Solvent Isotope and Viscosity Effects. *Acs Catalysis*, 7, 6352-6364.
- Singh, R. & Eltis, L. D. 2015. The multihued palette of dye-decolorizing peroxidases. *Archives of Biochemistry and Biophysics*, 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. *Journal of Biological Chemistry*, 287, 10623-10630.

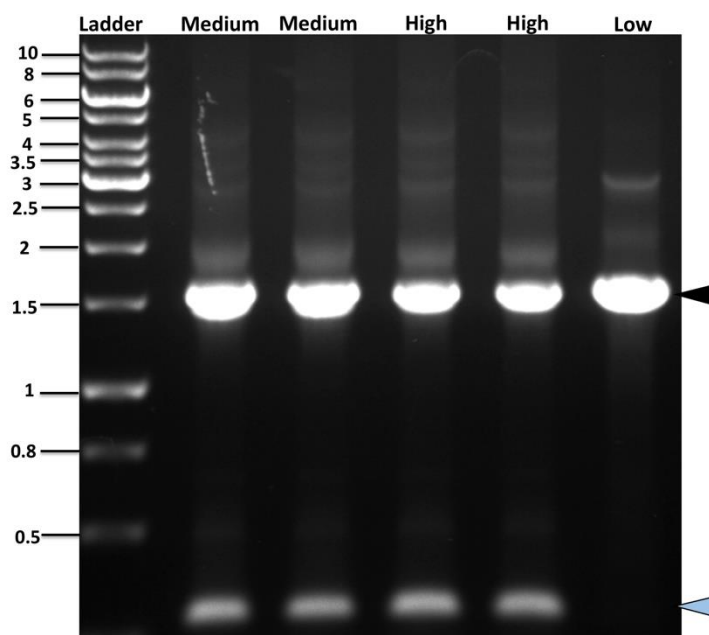


- Singh, R., Singh, P., Sharma, R., Selvalakshmi, S., Jayakumar, S., Ramachandran, V., Sayadi, M., Rezaei, M., Ghoneim, A. & Al-Modaihsh, A. 2014. Microorganism as a tool of bioremediation technology for cleaning environment: A review. *Proceedings of the International Academy of Ecology and Environmental Sciences*, 4, 18-29.
- Singh, R. L., Singh, P. K. & Singh, R. P. 2015. Enzymatic decolorization and degradation of azo dyes - A review. *International Biodeterioration & Biodegradation*, 104, 21-31.
- Snijder, J., Kononova, O., Barbu, I. M., Uetrecht, C., Rurup, W. F., Burnley, R. J., Koay, M. S. T., Cornelissen, J. J. L. M., Roos, W. H., Barsegov, V., Wuite, G. J. L. & Heck, A. J. R. 2016. Assembly and Mechanical Properties of the Cargo-Free and Cargo Loaded Bacterial Nanocompartment Encapsulin. *Biomacromolecules*, 17, 2522-2529.
- Sorek, N., Yeats, T. H., Szemenyei, H., Youngs, H. & Somerville, C. R. 2014. The Implications of Lignocellulosic Biomass Chemical Composition for the Production of Advanced Biofuels. *Bioscience*, 64, 192-201.
- Spiestersbach, A., Kubicek, J., Schafer, F., Block, H. & Maertens, B. 2015. Purification of His-Tagged Proteins. *Laboratory Methods in Enzymology: Protein, Pt D*, 559, 1-15.
- Steiner, K. & Schwab, H. 2012. RECENT ADVANCES IN RATIONAL APPROACHES FOR ENZYME ENGINEERING. *Computational and Structural Biotechnology Journal*, 2, e201209010.
- Sutter, M., Boehringer, D., Gutmann, S., Guenther, S., Prangishvili, D., Loessner, M. J., Stetter, K. O., Weber-Ban, E. & Ban, N. 2008. Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nature Structural & Molecular Biology*, 15, 939-947.
- Tamura, A., Fukutani, Y., Takami, T., Fujii, M., Nakaguchi, Y., Murakami, Y., Noguchi, K., Yohda, M., Odaka, M. J. B. & bioengineering 2015. Packaging guest proteins into the encapsulin nanocompartment from *Rhodococcus erythropolis* N771. 112, 13-20.
- Tee, K. L. & Wong, T. S. 2014. Directed evolution: A powerful algorithm for advancing synthetic biology. In: SINGH, V. (ed.) *Applied Synthetic Biology*. Gyan Books Pvt. Ltd.
- Turner, N. J. 2009. Directed evolution drives the next generation of biocatalysts. *Nature Chemical Biology*, 5, 568-574.
- Uihlein, A. & Schebek, L. 2009. Environmental impacts of a lignocellulose feedstock biorefinery system: An assessment. *Biomass & Bioenergy*, 33, 793-802.
- Valderrama, B., Ayala, M. & Vazquez-Duhalt, R. 2002. Suicide inactivation of peroxidases and the challenge of engineering more robust enzymes. *Chemistry & Biology*, 9, 555-565.
- van Bloois, E., Pazmino, D. E. T., Winter, R. T. & Fraaije, M. W. 2010. A robust and extracellular heme-containing peroxidase from *Thermobifida fusca* as prototype of a bacterial peroxidase superfamily. *Applied Microbiology and Biotechnology*, 86, 1419-1430.
- van Rantwijk, F. & Sheldon, R. A. 2000. Selective oxygen transfer catalysed by heme peroxidases: synthetic and mechanistic aspects. *Current Opinion in Biotechnology*, 11, 554-564.
- Walker, G. M. 2011. 125th Anniversary Review: Fuel Alcohol: Current Production and Future Challenges. *Journal of the Institute of Brewing*, 117, 3-22.

- Wan, C. & Li, Y. 2012. Fungal pretreatment of lignocellulosic biomass. *Biotechnology Advances*, 30, 1447-1457.
- Warburton, M., Omar Ali, H., Liong, W. C., Othusitse, A. M., Abdullah Zubir, A., Maddock, S. & Wong, T. S. J. A. B. 2015. OneClick: a program for designing focused mutagenesis experiments. 2, 126-143.
- Woodley, J. M. 2013. Protein engineering of enzymes for process applications. *Current Opinion in Chemical Biology*, 17, 310-316.
- Xu, Z. Y. & Huang, F. 2014. Pretreatment Methods for Bioethanol Production. *Applied Biochemistry and Biotechnology*, 174, 43-62.
- Yadav, M. & Yadav, H. S. 2015. Applications of ligninolytic enzymes to pollutants, wastewater, dyes, soil, coal, paper and polymers. *Environmental Chemistry Letters*, 13, 309-318.
- Yan, Z., Wei, X., Yuan, Y., Li, Z., Li, D., Liu, X. & Gao, L. 2016. Deodorization of pig manure using lignin peroxidase with different electron acceptors. *Journal of the Air & Waste Management Association*, 66, 420-428.
- Yang, Y. S., Zhou, J. T., Lu, H., Yuan, Y. L. & Zhao, L. H. 2011. Isolation and characterization of a fungus *Aspergillus* sp strain F-3 capable of degrading alkali lignin. *Biodegradation*, 22, 1017-1027.
- Ye, F. X., Zhu, R. F. & Li, Y. 2009. Deodorization of swine manure slurry using horseradish peroxidase and peroxides. *Journal of Hazardous Materials*, 167, 148-153.
- Zamocky, M., Gasselhuber, B., Furtmuller, P. G. & Obinger, C. 2014. Turning points in the evolution of peroxidase-catalase superfamily: molecular phylogeny of hybrid heme peroxidases. *Cellular and Molecular Life Sciences*, 71, 4681-4696.
- Zamocky, M., Hofbauer, S., Schaffner, I., Gasselhuber, B., Nicolussi, A., Soudi, M., Pirker, K. F., Furamuller, P. G. & Obinger, C. 2015. Independent evolution of four heme peroxidase superfamilies. *Archives of Biochemistry and Biophysics*, 574, 108-119.
- Zamocky, M. & Obinger, C. 2010. Molecular Phylogeny of Heme Peroxidases. *Biocatalysis Based on Heme Peroxidases*, 7-35.
- Zeng, Y. N., Zhao, S., Yang, S. H. & Ding, S. Y. 2014. Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Current Opinion in Biotechnology*, 27, 38-45.
- Zhang, J., Feng, M. Y., Jiang, Y. C., Hu, M. C., Li, S. N. & Zhai, Q. G. 2012. Efficient decolorization/degradation of aqueous azo dyes using buffered H<sub>2</sub>O<sub>2</sub> oxidation catalyzed by a dosage below ppm level of chloroperoxidase. *Chemical Engineering Journal*, 191, 236-242.
- Zhao, X., Zhang, L. & Liu, D. 2012. Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels, Bioproducts and Biorefining*, 6, 465-482.
- Zhou, C. H., Xia, X., Lin, C. X., Tong, D. S. & Beltramini, J. 2012. Catalytic conversion of lignocellulosic biomass to fine chemicals and fuels (vol 40, pg 5588, 2011). *Chemical Society Reviews*, 41, 8210-8210.

## Appendix

**Appendix (1):** Analysis of PCR products on a 1% (w/v) DNA agarose gel showing products of different mutagenic reactions of *epPCR* in the second round of evolution using N312S mutant as a template for *epPCR*. ◀ Main bands for Low, Medium and High mutagenic conditions (1.5 kb); Side products ◀. Low mutagenic condition contained 1.5 mM of MgCl<sub>2</sub>, 0.01 mM of MnCl<sub>2</sub>, 0.3 mM of dNTP mix, 4.5 pmol of forward and reverse primers, 3.5 ng/μL of DNA template, and 1.25 U of Taq DNA polymerase. Medium mutagenic condition contained 7 mM of MgCl<sub>2</sub>, 0.2 mM of dATP, 0.2 mM of dGTP, 1 mM of dTTP, 1 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/μL of DNA template, and 1.25 U of Taq DNA polymerase. High mutagenic condition contained 7 mM of MgCl<sub>2</sub>, 0.05 mM of MnCl<sub>2</sub>, 0.2 mM of dATP, 0.2 mM of dGTP, 0.2 mM of dTTP, 0.2 mM of dCTP, 20 pmol of forward and reverse primers, 50 ng/μL of DNA template, and 1.25 U of Taq DNA polymerase.



**Appendix (2)** List of the sequencing sequences used to sequence mutants of Pleos-DyP4

Sequence	Oligonucleotide sequences
Seq (1)	5'ACAAACTGCACGTTTCGTG 3'
Seq (2)	5'AAAACGCTGATGCGTTTA 3'
Seq (3)	5'AATGGTGCCGGAGTTTGA 3'
Seq (4)	5'GGTCTGCTGTTTCGTTTGCTA 3'

**Appendix (3)** DNA and protein sequences for Pleos-DyP4-tag

```
      10      20      30      40      50
      -      -      -      -      -
1
ATGACCACCCCGGCGCCGCCGCTGGACCTGAACAACATCCAGGGTGATAT
  M T T P A P P L D L N N I Q G D I

      60      70      80      90     100
      -      -      -      -      -
51
TCTGGGTGGCCTGCCGAAGCGTACCGAGACCTACTTCTTTTTTCGACGTGA
  L G G L P K R T E T Y F F F D V T

      110     120     130     140     150
      -      -      -      -      -
101
CCAACGTTGATCAATTTAAAGCGAACATGGCGCACTTCATCCCGCACATT
  N V D Q F K A N M A H F I P H I

      160     170     180     190     200
      -      -      -      -      -
151
AAGACCAGCGCGGGCATCATTAAAGACCGTGAGGCGATCAAGGAACACAA
  K T S A G I I K D R E A I K E H K

      210     220     230     240     250
      -      -      -      -      -
201
ACGTCAGAAGAAACCGGGTCTGGTGCCGATGGCGGCGGTGAACGTTAGCT
  R Q K K P G L V P M A A V N V S F

      260     270     280     290     300
      -      -      -      -      -
251
TTAGCCACCTGGGCCTGCAGAAGCTGGGTATCACCGACGATCTGAGCGAT
  S H L G L Q K L G I T D D L S D

      310     320     330     340     350
      -      -      -      -      -
301
AACGCGTTCACCACCGGCCAGCGTAAGGACGCGGAGATTCTGGGCGATCC
  N A F T T G Q R K D A E I L G D P

      360     370     380     390     400
      -      -      -      -      -
351
GGGTAGCAAAAACGGTGATGCGTTTACCCCGGCGTGGGAAGCGCCGTTCC
  G S K N G D A F T P A W E A P F L
```

410      420      430      440      450  
 -      -      -      -      -  
 401  
 TGAAGGACATCCACGGTGTGATTTTTGTTGCGGGCGATTGCCACGGTAGC  
 K D I H G V I F V A G D C H G S

460      470      480      490      500  
 -      -      -      -      -  
 451  
 GTGAACAAGAAACTGGACGAGATCAAACACATTTTCGGCGTTGGTACCAG  
 V N K K L D E I K H I F G V G T S

510      520      530      540      550  
 -      -      -      -      -  
 501  
 CCACGCGAGCATCAGCGAAGTGACCCATGTTTCGTGGCGACGTGCGTCCGG  
 H A S I S E V T H V R G D V R P G

560      570      580      590      600  
 -      -      -      -      -  
 551  
 GTGATGTTACGCGCACGAGCACTTTGGCTTCCTGGATGGTATTAGCAAC  
 D V H A H E H F G F L D G I S N

610      620      630      640      650  
 -      -      -      -      -  
 601  
 CCGGCGGTTGAACAGTTTGATCAGAACCCGCTGCCGGGTCAGGACCCGAT  
 P A V E Q F D Q N P L P G Q D P I

660      670      680      690      700  
 -      -      -      -      -  
 651  
 CCGTCCGGGTTTCATTCTGGCGAAGGAGAACGGTGACAGCCGTGCGGGCGG  
 R P G F I L A K E N G D S R A A A

710      720      730      740      750  
 -      -      -      -      -  
 701  
 CGCGTCCGGACTGGGCGAAAGATGGTAGCTTTCTGACCTTCCGTTACCTG  
 R P D W A K D G S F L T F R Y L

760      770      780      790      800  
 -      -      -      -      -  
 751  
 TTCAAATGGTGCCGGAGTTTGACGATTTCTGGAAAGCAACCCGATCGT  
 F Q M V P E F D D F L E S N P I V

810      820      830      840      850  
 -      -      -      -      -

801  
TCTGCCGGGCCTGAGCCGTAAGAGGGTAGCGAACTGCTGGGTGCGCGTA  
L P G L S R K E G S E L L G A R I

860 870 880 890 900  
- - - - -

851  
TTGTGGGCCGTTGGAAAAGCGGTGCGCCGATCGAGATTACCCCGCTGAAG  
V G R W K S G A P I E I T P L K

910 920 930 940 950  
- - - - -

901  
GACGATCCGAAACTGGCGGCGGATGCGCAGCGTAACAACAAGTTTGACTT  
D D P K L A A D A Q R N N K F D F

960 970 980 990 1000  
- - - - -

951  
CGGCGATAGCCTGGTTCGTGGTGACCAAACCAAGTGCCCGTTCGCGGGCGC  
G D S L V R G D Q T K C P F A A H

1010 1020 1030 1040 1050  
- - - - -

1001  
ACATCCGTAAAACCTACCCGCGTAACGATCTGGAAGGTCCGCCGCTGAAA  
I R K T Y P R N D L E G P P L K

1060 1070 1080 1090 1100  
- - - - -

1051  
GCGGACATCGATAACCGTCGTATCATTGTCGTGGCATTTCAGTTTGGTCC  
A D I D N R R I I R R G I Q F G P

1110 1120 1130 1140 1150  
- - - - -

1101  
GGAAGTGACCAGCCAAGAACACCACGACAAGAAAACCCACCACGGCCGTG  
E V T S Q E H H D K K T H H G R G

1160 1170 1180 1190 1200  
- - - - -

1151  
GTCTGCTGTTGTTTTGCTATAGCAGCAGCATCGACGATGGCTTTCACTTC  
L L F V C Y S S S I D D G F H F

1210 1220 1230 1240 1250  
- - - - -

1201  
ATTCAGGAAAGCTGGGCGAACGCGCCGAACCTTCCGGTGAACGCGGTTAC

I Q E S W A N A P N F P V N A V T

1260 1270 1280 1290 1300

- - - - -

1251

CAGCGCGGGTCCGATCCCGCCGCTGGATGGTGTGGTTCCGGGTTTCGATG

S A G P I P P L D G V V P G F D A

1310 1320 1330 1340 1350

- - - - -

1301

CGATCATTGGCCAGAAAGTGGGTGGCGGTATCCGTCAAATTAGCGGTACC

I I G Q K V G G G I R Q I S G T

1360 1370 1380 1390 1400

- - - - -

1351

AACCCGAACGACCCGACCACCAACATTACCCTGCCGGACCAGGATTTTGT

N P N D P T T N I T L P D Q D F V

1410 1420 1430 1440 1450

- - - - -

1401

GGTTCCGCGTGGCGGTGAGTACTTTTTTCAGCCCAGCATCACCGCGCTGA

V P R G G E Y F F S P S I T A L K

1460 1470 1480 1490 1500

- - - - -

1451

AGACCAAATTCGCGATTGGCGTTGCGAGCCC GGCGCCGCACAGCCAAGCG

T K F A I G V A S P A P H S Q A

1510 1520 1530 1540 1550

- - - - -

1501

CCGATCAGCGCGTTTCTGGATGATCCGCCGGATGCGCCGACCCGTCTGGT

P I S A F L D D P P D A P T R L V

1560 1570 1580 1590 1600

- - - - -

1551

TCCGGAAGCGACCTTTACCGCGCCGATCAGCGATGGTAGCCTGGGCATTG

P E A T F T A P I S D G S L G I G

1610 1620

- - - - -

1601 GTAGCCTGAAACGTAGCGCGCAGCAA

S L K R S A Q Q

## Signal peptide sequence

### Appendix (4) DNA and protein sequences for ReEncapsulin

```

      10      20      30      40      50
      -      -      -      -      -
1
ATGACCAACCTGCACCGTGATCTGGCGCCGATCAGCGCGGGCGGCGTGGGC
  M T N L H R D L A P I S A A A W A

      60      70      80      90     100
      -      -      -      -      -
51
GGAAATTGAGGAAGAGGGCGAGCCGTACCTTCAAGCGTCACGTGGCGGGTC
  E I E E E A S R T F K R H V A G R

      110     120     130     140     150
      -      -      -      -      -
101
GTCGTGTGGTTGACGTTGAAGGTCCGAGCGGTGATGATCTGGCGGGCGATC
  R V V D V E G P S G D D L A A I

      160     170     180     190     200
      -      -      -      -      -
151
CCGCTGGGTCACCAGGTGCCGATTAACCCGCTGGCGGATGGTGTATTGC
  P L G H Q V P I N P L A D G V I A

      210     220     230     240     250
      -      -      -      -      -
201
GCACGCGCGTCAGAGCCAACCGATCATTGAACTGCGTGTGCCGTTTACCG
  H A R Q S Q P I I E L R V P F T V

      260     270     280     290     300
      -      -      -      -      -
251
TTAGCCGTCAAGCGATTGACGATGTGGAGCGTGGTGC GAAAGACAGCGAT
  S R Q A I D D V E R G A K D S D

      310     320     330     340     350
      -      -      -      -      -
301
TGGCAGCCGTTAAGGACGCGGCGAAACAAATTGCGTTCGCGGAAGATCG
  W Q P V K D A A K Q I A F A E D R

      360     370     380     390     400
      -      -      -      -      -

```



351  
TGCGATCTTTGAGGGCTATCCGGCGGGCAGCATCACCGGTGTGCGTGCGA  
A I F E G Y P A A S I T G V R A S

410 420 430 440 450  
- - - - -

401  
GCGGCAGCAACCCGGAAGCTGCCGGTTGACGCGAAAGATTACCCG  
G S N P E L K L P V D A K D Y P

460 470 480 490 500  
- - - - -

451  
GAGGCGATCAGCCAGGCGATTACCAGCCTGCGTCTGGCGGGTGTGAACGG  
E A I S Q A I T S L R L A G V N G

510 520 530 540 550  
- - - - -

501  
TCCGTATAGCCTGCTGCTGAACGCGGACGCGTTCACCGCGATTAACGAAA  
P Y S L L L N A D A F T A I N E T

560 570 580 590 600  
- - - - -

551  
CCAGCGATCACGGCTACCCGATCCGTGAACACCTGCGTCGTGTTCTGGAC  
S D H G Y P I R E H L R R V L D

610 620 630 640 650  
- - - - -

601  
GGCGAGATCATTGGGGCGCCGGCGATTGATGGCGCGTTTCTGCTGAGCAC  
G E I I W A P A I D G A F L L S T

660 670 680 690 700  
- - - - -

651  
CCGTGGTGGCGACTACGAGCTGCACCTGGGTCAGGATCTGAGCATCGGCT  
R G G D Y E L H L G Q D L S I G Y

710 720 730 740 750  
- - - - -

701  
ATCTGAGCCACGATGCGAACAGCGTGGAAGTACTTCCAAGAGAGCATG  
L S H D A N S V E L Y F Q E S M

760 770 780 790 800  
- - - - -

751  
ACCTTTCTGATGTATAACCAGCGAAGCGGTGGTTAGCCTGGCGGAG **CACCA**

T F L M Y T S E A V V S L A E H H

810

801 CCACCACCACCAC  
H H H H

His-tag sequence

Appendix (5) : DNA and protein sequences for DyP4-WT sequence

```
      10      20      30      40      50
      -      -      -      -      -
1
ATGACCACCCCGGCGCCGCGCTGGACCTGAACAACATCCAGGGTGATAT
  M T T P A P P L D L N N I Q G D I

      60      70      80      90     100
      -      -      -      -      -
51
TCTGGGTGGCCTGCCGAAGCGTACCGAGACCTACTTCTTTTTTCGACGTGA
  L G G L P K R T E T Y F F F D V T

     110     120     130     140     150
      -      -      -      -      -
101
CCAACGTTGATCAATTTAAAGCGAACATGGCGCACTTCATCCCGCACATT
  N V D Q F K A N M A H F I P H I

     160     170     180     190     200
      -      -      -      -      -
151
AAGACCAGCGCGGGCATCATTAAAGACCGTGAGGCGATCAAGGAACACAA
  K T S A G I I K D R E A I K E H K

     210     220     230     240     250
      -      -      -      -      -
201
ACGTCAGAAGAAACCGGGTCTGGTGCCGATGGCGGCGGTGAACGTTAGCT
  R Q K K P G L V P M A A V N V S F

     260     270     280     290     300
      -      -      -      -      -
251
TTAGCCACCTGGGCCTGCAGAAGCTGGGTATCACCGACGATCTGAGCGAT
  S H L G L Q K L G I T D D L S D
```

310    320    330    340    350  
 -    -    -    -    -  
 301  
 AACGCGTTCACCACCGGCCAGCGT **AAG** GACGCGGAGATTCTGGGCGATCC  
 N A F T T G Q R **K** D A E I L G D P

360    370    380    390    400  
 -    -    -    -    -  
 351  
 GGGTAGCAAAAACGGTGTGATGCGTTTACCCCGGCGTGGAAGCGCCGTTCC  
 G S K N G D A F T P A W E A P F L

410    420    430    440    450  
 -    -    -    -    -  
 401  
 TGAAGGACATCCACGGTGTGATTTTTGTTGCGGGCGATTGCCACGGTAGC  
 K D I H G V I F V A G D C H G S

460    470    480    490    500  
 -    -    -    -    -  
 451  
 GTGAACAAGAACTGGACGAGATCAAACACATTTTCGGCGTTGGTACCAG  
 V N K K L D E I K H I F G V G T S

510    520    530    540    550  
 -    -    -    -    -  
 501  
 CCACGCGAGCATCAGCGAAGTGACCCATGTTTCGTGGCGACGTGCGTCCGG  
 H A S I S E V T H V R G D V R P G

560    570    580    590    600  
 -    -    -    -    -  
 551  
 GTGATGTTACGCGCACGAGCACTTTGGCTTCCTGGATGGTATTAGCAAC  
 D V H A H E H F G F L D G I S N

610    620    630    640    650  
 -    -    -    -    -  
 601  
 CCGGCGGTTGAACAGTTTGATCAGAACCCGCTGCCGGGTCAGGACCCGAT  
 P A V E Q F D Q N P L P G Q D P I

660    670    680    690    700  
 -    -    -    -    -  
 651  
 CCGTCCGGGTTTCATTCTGGCGAAGGAGAACGGTGACAGCCGTGCGGCGG  
 R P G F I L A K E N G D S R A A A

710    720    730    740    750  
 -    -    -    -    -

701  
CGCGTCCGGACTGGGCGAAAGATGGTAGCTTTCTGACCTTCCGTTACCTG  
R P D W A K D G S F L T F R Y L

760 770 780 790 800  
- - - - -

751  
TTCCAAATGGTGCCGGAGTTTGACGATTTCTGAAAGCAACCCGATCGT  
F Q M V P E F D D F L E S N P I V

810 820 830 840 850  
- - - - -

801  
TCTGCCGGGCCTGAGCCGTAAAGAGGGTAGCGAACTGCTGGGTGCGCGTA  
L P G L S R K E G S E L L G A R I

860 870 880 890 900  
- - - - -

851  
TTGTGGGCCGTTGGAAAAGCGGTGCGCCGATCGAGATTACCCCGCTGAAG  
V G R W K S G A P I E I T P L K

910 920 930 940 950  
- - - - -

901  
GACGATCCGAAACTGGCGGCGGATGCGCAGCGTAACAACAAGTTTGACTT  
D D P K L A A D A Q R N N K F D F

960 970 980 990 1000  
- - - - -

951  
CGGCGATAGCCTGGTTCGTGGTGACCAAACCAAGTGCCCGTTCGCGGGCGC  
G D S L V R G D Q T K C P F A A H

1010 1020 1030 1040 1050  
- - - - -

1001  
ACATCCGTAAAACCTACCCGCGTAACGATCTGGAAGGTCCGCCGCTGAAA  
I R K T Y P R N D L E G P P L K

1060 1070 1080 1090 1100  
- - - - -

1051  
GCGGACATCGATAACCGTCGTATCATTGTCGTGGCATTTCAGTTTGGTCC  
A D I D N R R I I R R G I Q F G P

1110 1120 1130 1140 1150  
- - - - -

1101  
GGAAGTGACCAGCCAAGAACACCACGACAAGAAAACCCACCACGGCCGTG

E V T S Q E H H D K K T H H G R G

1160 1170 1180 1190 1200

- - - - -

1151

GTCTGCTGTTTCGTTTGCTATAGCAGCAGCATCGACGATGGCTTTCACTTC

L L F V C Y S S S I D D G F H F

1210 1220 1230 1240 1250

- - - - -

1201

ATTCAGGAAAGCTGGGCGAACGCGCCGAACTTTCCGGTGAACGCGGTTAC

I Q E S W A N A P N F P V N A V T

1260 1270 1280 1290 1300

- - - - -

1251

CAGCGCGGGTCCGATCCCGCCGCTGGATGGTGTGGTTCCGGGTTTCGATG

S A G P I P P L D G V V P G F D A

1310 1320 1330 1340 1350

- - - - -

1301

CGATCATTGGCCAGAAAGTGGGTGGCGGTATCCGTCAAATTAGCGGTACC

I I G Q K V G G G I R Q I S G T

1360 1370 1380 1390 1400

- - - - -

1351

AACCCGAACGACCCGACCACCAACATTACCCTGCCGGACCAGGATTTTGT

N P N D P T T N I T L P D Q D F V

1410 1420 1430 1440 1450

- - - - -

1401

GGTTCGCGTGCGGGTGGTACTTTTTTCAGCCCGAGCATCACCGCGCTGA

V P R G G E Y F F S P S I T A L K

1460 1470 1480 1490 1500

- - - - -

1451

AGACCAAATTCGCGATTGGCGTTGCGAGCCCGGCGCCGCACAGCCAAGCG

T K F A I G V A S P A P H S Q A

1510

-

1501 CCGATCAGCGCG

P I S A

**Appendix (6):** DNA and protein sequence for DyP4 mutant (F6) N312S

```
      10      20      30      40      50
      -      -      -      -      -
1
ATGACCACCCCGGCGCCGCGCTGGACCTGAACAACATCCAGGGTGATAT
  M T T P A P P L D L N N I Q G D I

      60      70      80      90     100
      -      -      -      -      -
51
TCTGGGTGGCCTGCCGAAGCGTACCGAGACCTACTTCTTTTTTCGACGTGA
  L G G L P K R T E T Y F F F D V T

     110     120     130     140     150
      -      -      -      -      -
101
CCAACGTTGATCAATTTAAAGCGAACATGGCGCACTTCATCCCGCACATT
  N V D Q F K A N M A H F I P H I

     160     170     180     190     200
      -      -      -      -      -
151
AAGACCAGCGCGGGCGTCATTAAGACCGTGAGGCGATCAAGGAACACAA
  K T S A G V I K D R E A I K E H K

     210     220     230     240     250
      -      -      -      -      -
201
ACGTCAGAAGAAACCGGGACTGGTGCCGATGGCGGCGGTGAACGTTAGCT
  R Q K K P G L V P M A A V N V S F

     260     270     280     290     300
      -      -      -      -      -
251
TTAGCCACCTGGGCCTGCAGAAGCTGGGTATCACCGACGATCTGAGCGAT
  S H L G L Q K L G I T D D L S D

     310     320     330     340     350
      -      -      -      -      -
301
AACCGTTACACACCGGCCAGCGTAGGGACGCGGAGATTCTGGGCGATCC
  N A F T T G Q R R D A E I L G D P

     360     370     380     390     400
      -      -      -      -      -
```

351  
GGGTAGCAAAAACGGTGATGCGTTTACCCCGGCGTGGGAAGCGCCGTTCC  
G S K N G D A F T P A W E A P F L

410 420 430 440 450  
- - - - -

401  
TGAAGGACATCCACGGTGTGATTTTTGTTGCGGGCGATTGCCACGGTAGC  
K D I H G V I F V A G D C H G S

460 470 480 490 500  
- - - - -

451  
GTGAACAAGAAACTGGACGAGATCAAACACATTTTCGGCGTTGGTACCAG  
V N K K L D E I K H I F G V G T S

510 520 530 540 550  
- - - - -

501  
CCACGCGAGCATCAGCGAAGTGACCCATGTTTCGTGGCGACGTGCGTCCGG  
H A S I S E V T H V R G D V R P G

560 570 580 590 600  
- - - - -

551  
GTGATGTTACGCGCACGAGCACTTTGGCTTCCTGGATGGTATTAGCAAC  
D V H A H E H F G F L D G I S N

610 620 630 640 650  
- - - - -

601  
CCGGCGGTTGAACAGTTTGATCAGAACCCGCTGCCGGGTCAGGACCCGAT  
P A V E Q F D Q N P L P G Q D P I

660 670 680 690 700  
- - - - -

651 CCGTCCGGGTTTCATTCTGGCGAAGGAGAACGGTGACAGCCGTGCGGCGG  
R P G F I L A K E N G D S R A A A

710 720 730 740 750  
- - - - -

701 CGCGTCCGGACTGGGCGAAA**GAC**GGTAGCTTT**CTT**ACCTTCCGTTACCTG  
R P D W A K **D** G S F **L** T F R Y L

760 770 780 790 800  
- - - - -

751  
TTCCAAATGGTGCCGGAGTTTGACGATTTCTGGAAAGCAACCCGATCGT  
F Q M V P E F D D F L E S N P I V

810 820 830 840 850  
 - - - - -  
 801  
 TCTGCCGGGCCTGAGCCGTAAAGAGGGTAGCGAACTGCTGGGTGCGCGTA  
 L P G L S R K E G S E L L G A R I

860 870 880 890 900  
 - - - - -  
 851  
 TTGTGGGCCGTTGGAAAAGCGGTGCGCCGATCGAGATTACCCCGCTGAAG  
 V G R W K S G A P I E I T P L K

910 920 930 940 950  
 - - - - -  
 901  
 GACGATCCGAAACTGGCGGCGGATGCGCAGCGT **AGC** AACAAGTTTGACTT  
 D D P K L A A D A Q R **S** N K F D F

960 970 980 990 1000  
 - - - - -  
 951  
 CGGCGATAGCCTGGTTCGTGGTGACCAAACCAAGTGCCCGTTCGCGGGCGC  
 G D S L V R G D Q T K C P F A A H

1010 1020 1030 1040 1050  
 - - - - -  
 1001  
 ACATCCGTAAAACCTACCCGCGTAACGATCTGGAAGGTCCGCCGCTGAAA  
 I R K T Y P R N D L E G P P L K

1060 1070 1080 1090 1100  
 - - - - -  
 1051  
 GCGGACATCGATAACCGTCGTATCATTGTCGTGGCATTTCAGTTTGGTCC  
 A D I D N R R I I R R G I Q F G P

1110 1120 1130 1140 1150  
 - - - - -  
 1101  
 GGAAGTGACCAGCCAAGAACACCACGACAAGAAAACCCACCACGGCCGTG  
 E V T S Q E H H D K K T H H G R G

1160 1170 1180 1190 1200  
 - - - - -  
 1151  
 GTCTGCTGTTTCGTTTGCTATAGCAGCAGCATCGACGATGGCTTTCCTTC  
 L L F V C Y S S S I D D G F H F

1210 1220 1230 1240 1250  
 - - - - -



1201  
 ATTCAGGAAAGCTGGGCGAACGCGCCGAACCTTTCCGGTGAACGCGGTTAC  
 I Q E S W A N A P N F P V N A V T

1260 1270 1280 1290 1300  
 - - - - -

1251  
 CAGCGCGGGTCCGATCCCGCCGCTGGATGGTGTGGTTCCGGGTTTCGATG  
 S A G P I P P L D G V V P G F D A

1310 1320 1330 1340 1350  
 - - - - -

1301  
 CGATCATTGGCCAGAAAGTGGGTGGCGGTATT CGTCAAATTAGCGGTACC  
 I I G Q K V G G G I R Q I S G T

1360 1370 1380 1390 1400  
 - - - - -

1351  
 AACCCGAACGACCCGACCACCAACATTACCCTGCCGGACCAGGATTTTGT  
 N P N D P T T N I T L P D Q D F V

1410 1420 1430 1440 1450  
 - - - - -

1401  
 GGTTCGCGTGCGGGTGAGTACTTTTTTCAGCCCGAGCATCACCGCGCTGA  
 V P R G G E Y F F S P S I T A L K

1460 1470 1480 1490 1500  
 - - - - -

1451  
 AGACCAAATTCGCGATTGGCGTTGCGAGCCCGGCGCCGCACAGCCAAGCG  
 T K F A I G V A S P A P H S Q A

1510  
 -  
 1501 CCGATCAGCGCGTAA  
 P I S A \*

Silent mutation, Missense mutation