

Heterologous Expression and Characterisation of Unspecific Peroxygenases

Tamara Felicitas Mielke

PhD

University of York
Chemistry

December 2017

Abstract

In 2004, unspecific peroxygenases (UPOs) from fungi were first identified and showed to catalyse selective oxygenation reactions, with high turnovers, good stability and a broad reaction scope, providing a valuable alternative to established biocatalytic hydroxylation systems such as cytochrome P450s. At the start of this project, access to UPOs was limited to expression in their native fungi, hence studies were performed looking into heterologous expression of the enzymes from *Agrocybe aegerita* and *Agaricus bisporus*, followed by characterisation, and application of the enzymes.

UPO expression was attempted using *Escherichia coli* and the yeast *Pichia pastoris* as expression hosts. Expression in the latter yielded functional, His-tagged protein. rAaeUPO-H, an evolved nine-point mutant prepared for the expression in yeast by Molina *et al.*, was obtained and catalytic constants (k_{cat}/K_m) of $7,700 \pm 700$, and $340 \pm 70 \text{ s}^{-1} \text{ M}^{-1}$ for 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), and 1,2-(methylenedioxy)-4-nitrobenzene (NBD), respectively were determined. rAaeUPO-H was also expressed in industrial large-scale fermentations and lyophilisation of the protein was successfully performed, giving powdered, easy-to-handle catalyst. Crystal structures of rAaeUPO-H showed the limited effect the mutations had on the overall folding of the enzyme when compared to the wild-type. The structures also gave information about novel glycan structures, and revealed different molecules coordinating to the heme.

The UPO from *Agaricus bisporus* (H-AbiUPO) was also expressed in *P. pastoris*, and associated research into construct design, should enable the expression of further UPO homologs in the future. H-AbiUPO displayed a narrower substrate scope than rAaeUPO-H and reduced overall activity.

Lyophilised rAaeUPO-H was successfully applied to a 1 g scale hydroxylation of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine to the benzylic (*R*)-alcohol in 63% yield and with >85% *ee*. For other substrates, including a range of ethylbenzenes, conversions of greater than 90%, with excellent enantiomeric excesses, were observed.

I. Table of Contents

Abstract.....	III
I. Table of Contents.....	V
i. List of Figures.....	IX
ii. List of Tables.....	XXIV
II. Acknowledgments.....	XXIX
III. Author's Declaration.....	XXX
1. Introduction.....	1
1.1. Enzymes in Chemistry and Industry.....	1
1.1.1. Enzymatic Oxygenation and Hydroxylation Reactions.....	4
1.1.2. Peroxidases.....	9
1.1.3. Cytochrome P450 Monooxygenases.....	13
1.2. Unspecific Peroxygenases.....	17
1.2.1. Proposed Catalytic Cycle.....	19
1.2.2. Crystal Structure.....	21
1.2.3. Biotransformations and Applications.....	23
1.3. Aims.....	35
2. Bacterial Expression of <i>Aae</i> UPO and <i>Abi</i> UPO.....	37
2.1. Materials and Methods.....	38
2.1.1. Chemicals and Materials.....	38
2.1.2. Molecular Biology.....	38
2.1.3. Gene Design and Synthesis.....	40
2.1.4. Vector Design and Production.....	43
2.1.5. Cloning and Analysis.....	43
2.1.6. Bacterial Growth Curves.....	46

2.1.7.	Protein Production and Analysis.....	46
2.1.8.	Protein Purification and Characterisation	52
2.2.	Construct Design and Bacterial Growth.....	55
2.2.1.	N-terminal His-tag (pETYSBLIC-3C)	56
2.2.2.	C-terminal His-tag (pET22b).....	57
2.2.3.	pelB Secretion Sequence and C-terminal His-tag (pETFPP31).....	58
2.2.4.	Bacterial Transformations.....	59
2.2.5.	Growth Curves	59
2.3.	<i>E. coli</i> Test Expressions.....	61
2.3.1.	Expression over time.....	64
2.3.2.	Solubility Test.....	67
2.4.	Scale-up and Purification	69
2.5.	Summary, Discussion and Conclusion.....	76
3.	Yeast Expression and Characterisation of the Unspecific Peroxygenase from <i>Agrocybe aegerita</i>	79
3.1.	Materials and Methods.....	80
3.1.1.	Chemicals and Materials.....	80
3.1.2.	Molecular Biology	80
3.1.3.	Gene Design and Synthesis	83
3.1.4.	Vector Design and Production	84
3.1.5.	Cloning and Analysis	85
3.1.6.	Protein Production and Analysis.....	86
3.1.7.	Protein Purification and Characterisation	91
3.1.8.	Crystallisation of <i>AaeUPO</i>	97
3.2.	Selection and Expression of <i>AaeUPO</i>	100
3.2.1.	Construct Design	100

3.2.2.	Selection and Test Expression.....	102
3.2.3.	Shake-Flask Expression	104
3.2.4.	Fed-Batch Fermentation	106
3.2.5.	Fermentation Optimisation.....	115
3.3.	Purification and Biophysical Characterisation of <i>Aae</i> UPO.....	121
3.3.1.	Multistep Purification	121
3.3.2.	Protein Identification and Glycosylation Analysis.....	125
3.3.3.	Kinetic Characterisation	128
3.4.	Crystallisation and Structure Analysis of <i>Aae</i> UPO	130
3.5.	Summary, Discussion and Conclusion	137
4.	Yeast Expression and Characterisation of the Unspecific Peroxygenase from <i>Agaricus bisporus</i>	143
4.1.	Material and Methods.....	144
4.1.1.	Gene Design and Synthesis for <i>Agaricus bisporus</i> UPO (K5XIK0)	144
4.1.2.	Purification	145
4.2.	Selection and Expression of <i>Abi</i> UPO	146
4.2.1.	Construct Design	146
4.2.2.	Test Expression.....	147
4.2.3.	Shake Flask Expression and Purification Trials.....	148
4.2.4.	Fed-Batch Fermentation	153
4.3.	Purification and Biophysical Characterisation of <i>Abi</i> UPO.....	155
4.3.1.	Multistep Purification	157
4.4.	Summary, Discussion and Conclusion	159
5.	Biotransformations and Industrial Applications of UPOs	161
5.1.	Materials and Methods	162
5.1.1.	Chemicals and Materials	162

5.1.2.	Analytical Instruments	162
5.1.3.	Industrial Fermentation and Processing of <i>rAaeUPO-H</i>	162
5.1.4.	High Throughput Substrate and Condition Screening	164
5.1.5.	Stirred Small-Scale Reactions.....	165
5.1.6.	Reaction Optimisation	166
5.1.7.	Design of Experiments (DoE)	169
5.2.	Characterisation of Lyophilised <i>rAaeUPO-H</i>	173
5.3.	Substrate Screens.....	175
5.4.	Condition and Reaction Optimisation	178
5.4.1.	Hydrogen Peroxide Addition.....	181
5.4.2.	<i>In situ</i> Hydrogen Peroxide Generation	184
5.5.	Stirred Small-Scale Reactions.....	190
5.6.	Design of Experiments	194
5.6.1.	Scale-up Reaction.....	198
5.7.	Summary, Discussion and Conclusion.....	201
6.	Conclusion and Future Work	207
7.	Abbreviations.....	211
8.	References	217

i. List of Figures

Figure 1-1: Selected enzyme catalysed reactions used in fine and bulk chemistry. A – Formation of acrylamide from acrylonitrile using nitrile hydratase ⁹ ; B – Synthesis of (<i>R</i>)-epichlorohydrin from 1,3-dichloropropan-2-ol using halohydrin dehalogenase and an epoxide hydrolase. ^{11, 12}	2
Figure 1-2: Selected enzyme catalysed reactions as utilised by pharmaceutical industry. A – Synthesis of a boceprevir precursor using the monoamine oxidase (MAO) from <i>Apergillus niger</i> ¹⁴ ; B – Application of a keto-reductase (KRED) from <i>Lactobacillus kefir</i> in the synthesis of a montelukast intermediate ¹⁵ ; C – C-C-bond formation of a statin-drug intermediate catalysed by an engineered 2-deoxy-ribose-5-phosphate aldolase (DERA). ¹⁶	3
Figure 1-3: General mechanism for the oxygen activation in flavin-dependent monooxygenases. Upon the addition of oxygen, the reduced flavin converts into the flavin C4a-(hydro)peroxide which can react with the substrate to give the flavin C4a-hydroxide.	5
Figure 1-4: Selected examples of transformations catalysed by various flavin dependent monooxygenases.....	6
Figure 1-5: Examples of biotransformations catalysed by α -KG-dependent dioxygenases, with GloC - <i>trans</i> -4-proline hydroxylase, ³⁸ LipL - putative TauD enzyme from the A-90289 gene cluster, ³⁹ and DdaC – Dapdiamide C. ⁴⁰	7
Figure 1-6: Structure of the heme cofactor and general schemes for the formation of compound I in different heme enzymes.....	8
Figure 1-7: Proposed catalytic mechanism for the reactions catalysed by <i>CfuCPO</i> : top – chlorination reaction, bottom – peroxidase-like reactions.	11
Figure 1-8: Ribbon-structure of <i>CfuCPO</i> (1CPO), with the heme and Cys29 (insert) in cylinder representation (carbon – green, oxygen – red, nitrogen – blue, sulfur – yellow). Created using CCP4mg.	12
Figure 1-9: Classes of CYP electron transfer systems; with FeS – iron-sulfur cluster, FAD – Flavin adenine dinucleotide, FMN – Flavin mononucleotide.....	13
Figure 1-10: The catalytic cycle of oxygen activation by CYPs and their shunt pathways. Starting from the resting heme (A), the proximal water ligand is replaced in the presence of substrate (B). The supply of an electron allows for iron reduction	

yielding C, which is able to react with molecular oxygen. *Via* an ferri-superoxide intermediates (D) and upon addition of a second electron, a ferri-peroxo intermediate is formed (E). Protonation leads to the formation of compound 0 (F), and the loss of water generates compound I, the porphyrin π radical Fe^{IV} oxo intermediate (G). Through oxygen insertion into the substrate (H), the return to the ferric resting state is observed (A)..... 15

Figure 1-11: Superimposed ribbon-structures of P450_{cam} (2CPP) in gold, P450_{terp} (1CPT) in coral and P450_{BM3} (2HPD) in grey to illustrate the rigid and flexible elements of the structures, with a close-up of the active site heme and coordinating cysteine, with carbon – green, oxygen – red, nitrogen – blue, sulfur – yellow. Created using CCP4mg. 16

Figure 1-12: Oxidation of (+)-camphor catalysed by P450_{cam}..... 17

Figure 1-13: Amorphadiene oxidation to artemisinic acid. The formation of artemisinic alcohol catalysed by CYP71AV1 (a CYP), CPR1 (cognate reductase), and CYB5 (cytochrome b₅), followed by the oxidation to artemisinic aldehyde catalysed by ADH1 (alcohol dehydrogenase), and the final oxidation giving artemisinic acid using ALDH1 (artemisinic aldehyde dehydrogenase). 17

Figure 1-14: Selected reactions catalysed by *Aae*UPO in the presence of hydrogen peroxide. 18

Figure 1-15: Proposed catalytic cycle for reactions catalysed by UPOs. The water molecule binding to the heme in the resting state (A) is replaced by *m*CPBA (B). The active site glutamate abstracts a proton from *m*CPBA yielding in compound 0 (C). Compound I, the oxo-ferryl cation radical complex, is formed by the deprotonation of the glutamate and subsequent loss of *m*CBA and (D). Through a one-electron transfer facilitated by 3CP (3-carboxy-PROXYL) compound II (E) is formed with a pK_a of 10.0. The insertion of oxygen into the substrate leads back to the resting heme (A). 20

Figure 1-16: Structure of *Aae*UPO (2YOR) with A – ribbon structure of the full length protein in crimson, the heme in cylinders (green – carbons, blue – nitrogen, red – oxygen), and carbohydrates as glycoblocks (blue – *N*-acetylglucosamine, green – mannose); B – carbohydrate chains linked to Asn141, with the electron density corresponding to the $2F_o - F_c$ map downloaded from the PDB and contoured at 1σ X

(colour scheme as described for glycoblocks); C – close up of the active site highlighting the acid-base pair Glu196 and Arg189; D – cone-shaped appearance of the substrate channel giving access to the active site highlighted in grey. Created using CCP4mg.....22

Figure 1-17: Halogenation reactions catalysed by *Aae*UPO; with MCD for monochlorodimedone, and X for chloride or bromide.24

Figure 1-18: Conversion of veratryl alcohol (VA) to veratraldehyde and veratric acid using *Aae*UPO.....24

Figure 1-19: Reactions of naphthalene (left) and toluene (right) with *Aae*UPO showing the major products formed.....25

Figure 1-20: Structures of 2-phenoxypropionic acid (POPA), propranolol, and diclofenac with their hydroxylated products in red, as produced by *Aae*UPO.26

Figure 1-21: Structures of flavonoids which have been applied in reactions with *Aae*UPO, in red the position where hydroxylation occurred, if substrate was converted.27

Figure 1-22: Examples for aliphatic hydroxylations using *Aae*UPO.....27

Figure 1-23: Conversion of cyclohexane to cyclohexanol and cyclohexanol using UPOs.....28

Figure 1-24: Structures of vitamin D₂, D₃, and testosterone, with the hydroxylation/epoxidation sites as attacked by UPOs highlighted in red.....29

Figure 1-25: Reaction of dibenzothiophene (DBT) to the sulfoxide, sulfone and ring-hydroxylated products in the presence of UPOs.30

Figure 1-26: Examples for epoxidation reactions catalysed by *Aae*UPO.....30

Figure 1-27: Examples for ether cleavages catalysed by *Aae*UPO.....31

Figure 1-28: Structures of drug molecules dealkylated by *Aae*UPO and/or *Cra*UPO; with the preferred dealkylated chain in red.32

Figure 1-29: Structure of the human metabolite SAR548304 which is *N*-demethylated by *Mro*UPO (highlighted in red).....33

Figure 1-30: Structure of ABTS and DMP, two peroxidase substrates used to quantify UPOs; red arrows highlighting where the reaction occurs.33

Figure 1-31: Conversion of 5-nitro-1,3-benzodioxole (NBD) using UPOs.....34

Figure 2-1: Schematic set up for the protein transfer from SDS gels onto a nitrocellulose membrane.....52

Figure 2-2: Alignment of the amino acid sequences of *AaeUPO* (AAP) and *AbiUPO* (ABP) with the signal peptide in dark red and the pro-peptide in orange, the mature protein sequence is portrait in black.55

Figure 2-3: Overview of the different *AaeUPO* and *AbiUPO* gene constructs created for bacterial expression alongside the native genes.56

Figure 2-4: 1% agarose gels of double digested *AaeUPO* (left) and *AbiUPO* (right) in pETYSBLIC-3C with *NdeI* and *NcoI*: 1 – Marker; 2 – undigested control sample; 3 – 5:1 insert:vector ligation ratio; 4 – 1:1 insert:vector; 5 – 1:5 insert:vector.57

Figure 2-5: 1% agarose gel showing the results of the colony PCR of *AaeUPO* (A1-A4) and *AbiUPO* (B1-B4) after cloning into pET22b using In-Fusion cloning.58

Figure 2-6: 1% agarose gel of double digested *AaeUPO* (A1, A2) and *AbiUPO* (B1, B2) in pETFPP31 using *NdeI* and *XhoI*.59

Figure 2-7: Growth curves for BL21 (square), Rosetta 2 (triangle) and Origami 2 (diamond) transformed with *AaeUPO* carrying an N-terminal His-tag in LB (red), M9-glucose (orange) and M9 glycerol (grey) at 37 °C. Analysis was performed in duplicates, the error bars describe the standard deviation.60

Figure 2-8: 12 % SDS gel for the test expression of *AaeUPO* (approx. 38 kDa) with N-terminal His-tag (left) and immunoblot (right) using BL21 in M9-glc. M - Low molecular weight marker (BioRad); 1 - 16 °C insoluble; 2 - 16 °C soluble; 3 - 16 °C control insoluble; 4 - 16 °C control soluble; 5 - 30 °C insoluble; 6 - 30 °C soluble; 7 - 30 °C control insoluble; 8 - 30 °C control soluble; 9 - 37 °C insoluble; 10 – 37 °C soluble; 11 – 37 °C control insoluble; 12 – 37 °C control soluble.....62

Figure 2-9: Immunoblots for the production N-terminal His-tag *AaeUPO* (approx. 38 kDa) over time: left – using BL21 in M9-glc at 30 °C; right – using Rosetta2 in LB at 30 °C. Highlighted the different insoluble to soluble ratios with: “0” no (soluble) protein; “1” soluble < insoluble; “2” soluble = insoluble; “3” soluble > insoluble. ...65

Figure 2-10: 3D-blots showing the amount of soluble protein expressed over time at different temperatures for the N-terminal His-tag *AaeUPO* (top) and the *AaeUPO* with pelB secretion sequence and C-terminal His-tag (bottom) for the different *E. coli* strains and growth media tested.66

Figure 2-11: Immunoblots of the buffer screening using *AaeUPO* (top; approx. 38 kDa) and *AbiUPO* (bottom; approx. 38 kDa). M - Low molecular weight marker (BioRad); uneven numbers – insoluble fractions; even numbers – soluble fractions; respective buffer number printed in bold on top.....68

Figure 2-12: Purification of the N-terminal His-tag *AaeUPO* (approx. 38 kDa) expressed by BL21 in M9-glc at 16 °C: left: 12 % SDS gel and immunoblot; M - Low molecular weight marker; 1 – crude extract; 2 – flow-through I; 3 – flow-through II; 4 – wash I, 5 – wash II; 6 – fraction 7; 7 – fraction 8; 8 – fraction 9; 9 – fraction 10; 10 – fraction 11; 11 – fraction 12; 12 – fraction 13; 13 – fraction 15; 14 – fraction 20; right: UV/Vis spectra for the flow-through and fractions 8, 12, and 20 from the Ni-purification, as well as concentrated protein.71

Figure 2-13: Purification of the N-terminal His-tag *AaeUPO* (1-3) and the secreted C-terminally tagged *AaeUPO* (5-14) expressed by Rosetta2 in LB at 37 °C left: 12 % SDS gel and immunoblot: 1 – insoluble protein; 2 – soluble/load; 3 – Ni flow-through; M - low molecular weight marker; 5 – cells before lysis; 6 - insoluble; 7 – soluble/load; 8 – Ni flow-through; 9 – fraction 8; 10 – fraction 9; 11 – fraction 10; 12 – fraction 11; 13 – fraction 12; 14 – fraction 13; right: UV/Vis spectra of the Ni flow-through samples of both *AaeUPO* variants.....72

Figure 2-14: Purification of the secreted C-terminally tagged *AaeUPO* expressed by Rosetta2 in LB at 37 °C with added ALA/FeCl₃ (left to the marker) and without (right to the marker) ; 12 % SDS gel and immunoblot: 1 – sucrose fraction; 2 – Ni load (after dialysis); 3 – Ni flow-through; 4 – fraction 4; 5 – fraction 5; 6 – fraction 6; 7 – fraction 7; M - low molecular weight marker; 9 - sucrose fraction; 10 – Ni load (following dialysis); 11 – Ni flow-through; 12 – fraction 4; 13 – fraction 5; 14 – fraction 6; 15 – fraction 7.74

Figure 2-15: left: 12 % SDS gel and immunoblot of the ammonium sulfate precipitation and anion exchange purification of the N-terminally tagged *AaeUPO* (BL21 in M9-glc at 16 °C): M - low molecular weight marker; 1 – 40% AS fraction; 2 – 60% AS fraction; 4 – Q-column flow-through; 5 – Q-fraction 8; 6 - Q-fraction 9; 7 - Q-fraction 10; 8 - Q-fraction 11; 9 - Q-fraction 12; 10 - Q-fraction 13; 11 - Q-fraction 14; right: Elution profile obtained for the purification of *AaeUPO* using anion exchange chromatography; highlighted the fractions run on the gel (8 - 14).....75

Figure 3-1: Set-up of the dot-blot apparatus.....	91
Figure 3-2: Purification tree applied in the purification of <i>Aae</i> UPO in the absence of a His-affinity tag.....	92
Figure 3-3: <i>rAae</i> UPO construct with the signalling and pro-peptide region in orange and the mature protein in grey, and the mutations highlighted in red.	100
Figure 3-4: Comparison of the yeast constructs <i>rAae</i> UPO-H (top) and H- <i>rAae</i> UPO (bottom), with the mature protein in grey, the <i>rAae</i> UPO signalling- and pro-peptide in orange, the His-tag in red, the 3C protease site in yellow, and the linker regions in apricot, with description of the three amino acids used as linkers.....	101
Figure 3-5: 1% Agarose gels showing: left - the successful amplification of <i>rAae</i> UPO-H and creation of the N-terminal 3C site (1140 bp); right – amplification of pPICZ α , deleting α -MF and <i>c-myc</i> regions, retaining the His-tag (approx. 3000 bp).	101
Figure 3-6: 1% Agarose gels for the colony PCR analysis of <i>rAae</i> UPO-H (left) and H- <i>rAae</i> UPO (right).....	102
Figure 3-7: Luminol-based dot-blot analysis of <i>P. pastoris</i> transformants during the process of the microfermentation; <i>rAae</i> UPO-H on the left and H- <i>rAae</i> UPO on the right, samples were taken after 0, 24, 48, and 72 h or 24, 48, 72, and 96 h, respectively.	104
Figure 3-8: 12% SDS gel and immunoblot for the 100 mL scale fermentation and Ni-purification of <i>rAae</i> UPO-H using pre-stained protein standard (Biorad) with: 1 – 24 h, 2 – 48 h, 3 – 72 h, 4 – dialysed sample, 5 – flow-through, 6 – wash unbound I, 7 – wash unbound II, 8 – wash I, 9 – wash II, 10 – fraction 11, 11 – fraction 12, 12 – fraction 13. The arrows are indicating the molecular weight at which the glycosylated protein is observed (55-60 kDa).	105
Figure 3-9: 12% SDS gel and immunoblot for the 100 mL scale fermentation of H- <i>rAae</i> UPO (left) and 12% SDS gel and ABTS-based activity assay for the Ni-purification (right) with: 1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h, 5 – 96 h, 6 – Ni load, 7 – Ni flow-through, 8 – wash, 9 – wash 30 mM imidazole, 10 – fraction 9, 11 – fraction 11, 12 – fraction 13. All gels were run using pre-stained protein standards and the ABTS reaction was run with a no-enzyme control (ctrl) and various samples from the purification.....	106

Figure 3-10: Schematic representation of the steps of a *Pichia* fermentation, as performed in this work.107

Figure 3-11: Fermentation-log for the rAaeUPO-H fermentation run at 0.2 L scale. The pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated are the end of the glycerol batch and beginning of the glycerol-fed batch phase, the beginning of the methanol fed-batch phase, and two examples for DO spike responses to feed-pauses.109

Figure 3-12: 12% SDS gels using the low molecular weight marker, for left: the fermentation of rAaeUPO-H at 0.2 L scale with supernatant samples taken during the methanol fed-batch phase at : 1 – 0 h, 2 – 13 h, 3 – 21 h, 4 – 37 h; right: the Ni-affinity purification of rAaeUPO-H harvested after 45 h, with: 1 – sample load, 2 – flow-through, 3 – wash, 4 – fraction 8, 5 – fraction 9, 6 – fraction 10, 7 – fraction 11, 8 – fraction 12, 9 – fraction 13, 10 – fraction 14, 11 – fraction 15, 12 – fraction 16, 13 – fraction 17, 14 – fraction 18. These fractions are highlighted in the chromatogram insert (top right) with the absorbance at 280 nm in green and at 420 nm in orange. The expected molecular weight for rAaeUPO-H is indicated using the red arrows.111

Figure 3-13: left: 12% SDS gel and immunoblot for samples taken during the methanol fed-batch phase of the 2 L fermentation and Ni-affinity purification after 38 h and 62 h fermentation for rAaeUPO-H (approx. 60 kDa, highlighted with arrows) with : 1 – 0 h, 2 – 14 h, 3 – 22 h, 4 – 38 h, 5 – 45 h, 6 – 68 h, 7 – flow-through (38 h), 8 – fraction 5 (38 h), 9 – fraction 6 (38h), 10 – fraction 7 (38 h), 11 – flow-through (62 h), 12 – fraction 5 (62 h), 13 – fraction 6 (62 h), 14 – fraction 7 (62 h); right: NBD-based activity tests run on the flow-through and peak fractions of the Ni purifications run after 38 h (top) and 62 h (bottom) fermentation showing a control (ctrl) on the left, followed by the flow-through (FT) and the fractions 5, 6, and 7.112

Figure 3-14: Fermentation-log for the H-rAaeUPO fermentation run at 0.2 L scale. The pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not

correspond to the actual volume added. Indicated is the beginning of the methanol fed-batch phase. 114

Figure 3-15: 12% SDS gel and immunoblot for the time course of expression for the 0.2 L scale fermentation of H-rAaeUPO with: 1 – 0 h, 2 – 24 h, 3 – 42 h, 4 – 48 h, 5 – 66 h, 6 – 72 h; gels were run using pre-stained protein standards. 115

Figure 3-16: top: harvested fermentations of rAaeUPO-H performed at varying pH, showing different colours of the fermentation supernatants against the cell pellet in the centrifugation tubes; bottom: activity assays using NBD (left) and ABTS (right) to compare the activity of rAaeUPO after 60 h fermentation at different pH. 117

Figure 3-17: 12% SDS-gel with pre-stained ladder for the fermentation (1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h) and Ni-affinity purification (5 – column load, 6 – flow-through, 7 – wash, 8 – wash with 30 mM imidazole, 9 – fraction 9, 10 – fraction 10, 11 – fraction 11, 12 – fraction 12, 13 – fraction 13) of rAaeUPO-H grown in the presence of casamino acids; arrows are indicating the expected molecular weight around 65 kDa. Insert: ABTS activity assays performed with the Ni-affinity purification samples in the presence of a control (ctrl). 118

Figure 3-18: Fermentation-log for the rAaeUPO-H fermentation run at 0.2 L scale with a 96 h methanol fed-batch phase; the cells were added to the fermenter at t = 18 h (start of the glycerol batch phase). The pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated are the beginning of the methanol fed-batch phase, as well as DO spikes. 119

Figure 3-19: top left: Chromatogram for the HIC purification of rAaeUPO-H, with the 280 nm absorbance in green, the 420 nm absorbance in orange, the concentration of buffer B (no salt buffer) in black, and the conductivity in brown; top right: reaction rate for the conversion of NBD plotted against the HIC fraction number; bottom: 12% SDS gel with low molecular weight marker showing the purification of rAaeUPO using HIC with: 1 – column load, 2 – flow-through, 3 – fraction 2, 4 – fraction 4, 5 – fraction 6, 6 – fraction 8, 7 – fraction 10, 8 – fraction 12, 9 – fraction 14, 10 – fraction 16, 12

– fraction 18, 12 – fraction 20, 13 – fraction 22, 14 – fraction 24; highlighted with arrows the target protein (approx. 60 kDa).....122

Figure 3-20: top: Chromatogram for the Q purification of *rAaeUPO-H*, with the 280 nm absorbance in green, the 420 nm absorbance in orange, the concentration of buffer B (high salt buffer) in black, and the conductivity in brown; bottom left: reaction rate for the conversion of NBD plotted against the Q fraction number; bottom right: 12% SDS gel with low molecular weight marker showing the purification of *rAaeUPO* using Q chromatography with: 1 – column load, 2 – wash, 3 – fraction 3, 4 – fraction 5, 5 – fraction 7, 6 – fraction 8, 7 – fraction 9, 8 – fraction 10, 9 – fraction 11, 10 – fraction 13, 12 – fraction 15, 12 – fraction 17, 13 – fraction 19, 14 – fraction 21; highlighted with arrows the target protein (approx. 55-60 kDa). 123

Figure 3-21: left: Chromatogram for the analytical SEC purification of *rAaeUPO-H*, with the 280 nm absorbance in green, the 420 nm absorbance in orange; right: 12% SDS gel with low molecular weight marker showing the purification of *rAaeUPO* using SEC with: 1 – fraction 8, 2 – fraction 9, 3 – fraction 10, 4 – fraction 14; highlighted with arrows the target protein (approx. 55-60 kDa).124

Figure 3-22: UV/Vis spectrum for the purified and concentrated *rAaeUPO-H*.125

Figure 3-23: 12% SDS gel run with a pre-stained ladder showing the denatured *rAaeUPO* (lane 1), the deglycosylation mixture (lane 2) and PNGaseF (lane 3); the molecular weights for the proteins were calculated using the migration of the ladder to create a calibration curve.126

Figure 3-24: Glycan fragmentation for two glycan chains determined in LC-MS experiments, displayed is the C₄ elution under standard source energy; with the m/z plotted against the signal intensity, the heme at 618 m/z , H for a hexose moiety (approx. 160 Da).....127

Figure 3-25: Linear MALDI-MS spectrum for *rAaeUPO-H*, showing peaks for the glycosylated and deglycosylated protein, as well as their higher-charge species...128

Figure 3-26: Michaelis-Menten plots for the reactions of purified *rAaeUPO-H* with ABTS, NBD, and VA; with the reaction rate plotted against the substrate concentration. Reactions were run in duplicates and reported with standard deviation.....128

Figure 3-27: pH profiles for the relative activity of the fermentation supernatant (crude, red) and the purified *rAaeUPO-H* (orange), reacting with ABTS (0.2 mM, left) and NBD (1 mM, right). Reactions were run in duplicates and are reported with standard deviations. 130

Figure 3-28: *rAaeUPO-H* protein crystals showing different morphologies, obtained in initial 96-well based screens, with the name and composition of the condition given..... 131

Figure 3-29: *rAaeUPO-H* protein crystals with different morphologies grown in 0.1 M Bis-Tris pH 5.5-7.5, with or without 0.2 M NaCl, PEG-3350 concentrations between 18-32%, using 6.4 mg mL⁻¹ purified protein. 132

Figure 3-30: Ribbon-structure of the solved *rAaeUPO-H_ACT* (5898); with the protein backbone in dark red, the sugars represented in glycoblocks (*N*-acetylglucosamine (GlcNAc) in blue, mannose (Man) in green), and the heme and acetate ligand coordinating to the iron in cylinder representation (carbons in green, oxygen in red, nitrogen in blue). Created using CCP4mg. 133

Figure 3-31: Ribbon-structure of *rAaeUPO-H_sugars* (4181) in dark red, with the heme and mutated side-chains in cylinder representation (C - green, O - red, N - blue), overlaid with the wild-type *AaeUPO* (2YOR) with crimson cylinders representing the mutated side-chains. Created using CCP4mg. 134

Figure 3-32: left: Ribbon-structure of *rAaeUPO-H_sugar* (4181) highlighting the sugar-chain off Asn286 (cylinder, C - green, O - red, N - blue); the protein backbone in dark red ribbons, the sugars represented in glycoblocks (insert) or cylinders, displaying GlcNAc in blue, and Man in green, and with the electron density corresponding to the $2F_o-F_c$ omit map contoured at 1σ ; right: schematic representation of the different sugar-chains observed in the structure. Created using CCP4mg. 134

Figure 3-33: Ribbon-structure of *rAaeUPO-H* in dark red, with the heme, acetate, Glu196, Arg169, and Arg189 in cylinder representation (C - green, O - red, N - blue) and the electron density corresponding to the $2F_o-F_c$ omit map contoured at 1σ ; left: *rAaeUPO-H_ACT* (5898) structure showing the acetate ligand coordinating to the Glu196; right: *rAaeUPO-H_sugar* (4181) with an acetate ligand in the presence of

Glu196 in two conformations, interacting with Arg189 and Arg169. Created using CCP4mg.	135
Figure 3-34: Ribbon-structure of <i>rAaeUPO-H</i> in dark red, with the heme, ligand, Glu196, and Cys36 in cylinder representation (C - green, O - red, N - blue, S - yellow) and the electron density corresponding to the $2F_o-F_c$ omit map contoured at 1σ ; left: <i>rAaeUPO-H_NHI</i> (7995) obtained coordinating to an unidentified ligand; right: <i>N</i> -hydroxyimidazole modelled as possible ligand into <i>rAaeUPO-H_NHI</i> . Created using CCP4mg.	136
Figure 4-1: Comparison of the yeast constructs <i>AbiUPO-H</i> (top) and <i>H-AbiUPO</i> (bottom), with the mature protein in grey, the <i>rAaeUPO</i> secretion sequence in orange, the His-tag in red, the 3C protease site in yellow, and the linker region in apricot, with description of the three amino acids used as linker.	147
Figure 4-2: Dot-blot analysis of two microfermentations with <i>AbiUPO-H</i> over time using: left – an ECL-based stain, showing response only on the outside of the functionalised dots; middle - SigmaFast tablet-based stain, moments after addition of the stain to the membrane, highlighting gas evolution; right - SigmaFast tablet-based stain after 12 h incubation, showing the typical stain response.....	148
Figure 4-3: Dot-blot analysis of the test expression with <i>H-AbiUPO P. pastoris</i> transformants; samples taken after 24, 48, 72, and 96 h protein expression, and visualised using ECL stain.	148
Figure 4-4: 12% SDS gel with low molecular weight marker showing the supernatant samples taken over time from the 100 mL shake flask fermentation of <i>AbiUPO-H</i> using <i>Pichia</i> (lane 1 – 24 h, 2 – 48 h, 3 – 72 h) and the dialysis (lane 4) and Ni-flow through sample (lane 5).....	149
Figure 4-5: left: 12% SDS gel with the low molecular weight marker showing the 100 mL shake flask fermentation of <i>AbiUPO-H</i> over time (lane 1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h), the dialysed sample (lane 5) and the analysis of the HIC purification, with: 6 – flow-through, 7 – fraction 8, 8 – fraction 10, 9 – fraction 12, 10 – fraction 14, 11 – fraction 16, 12 – fraction 18, 13 – fraction 20, 14 – fraction 22; right: chromatogram obtained for the purification of <i>AbiUPO-H</i> using HIC, with the 280 nm absorbance in green (scale 0-800 mAU) and the 420 nm absorbance in orange (scale 0-160 mAU), with indications of different fractions.....	150

Figure 4-6: left: 12% SDS gel with the low molecular weight marker showing the Q-purification of *AbiUPO*-H, with: 1 – column load, 2 – flow-through, 3 – fraction 5, 4 – fraction 6; right: chromatogram obtained for the Q-purification of *AbiUPO*-H, with the 280 nm absorbance in green (scale 0-250 mAU) and the 420 nm absorbance in orange (scale 0-250 mAU)..... 151

Figure 4-7: 12% SDS-gels with pre-stained markers for left: the 100 mL shake flask fermentation of H-*AbiUPO* (transformant 6) over time, with: 1 – 24 h, 2 – 48 h, 3 – 72 h, 4 – 96 h; right: the Ni-affinity purification of H-*AbiUPO* with the load (lane 1), flow-through (2), wash without (3) and with imidazole (30 mM, lane 4) and the peak fractions 9 (lane 5), 11 (lane 6) and 13 (lane 7)..... 152

Figure 4-8: 12% SDS gel run with a pre-stained marked and corresponding immunoblot visualised using ECL for the 100 mL shake flask fermentation of H-*AbiUPO* (transformant 1) over time, with: 1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h, 5 – 96 h. 153

Figure 4-9: Fermentation-log for the H-*AbiUPO* fermentation run at 0.2 L scale with a 67 h methanol fed-batch phase; with the pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated are the beginning of the methanol fed-batch phase, as well as a DO spike. 154

Figure 4-10: 12% SDS gel run with a pre-stained marked and corresponding immunoblot visualised using ECL for the fed-batch fermentation of H-*AbiUPO* (transformant 1) during the methanol fed-batch phase, with: 1 – 0 h, 2 – 13 h, 3 – 22 h, 4 – 37 h, 5 – 45 h, 6 – 61 h, 7 – 67 h. 155

Figure 4-11: left: 12% SDS gel with pre-stained marker showing the purification of H-*AbiUPO* using Ni-affinity chromatography, with: 1 – column load, 2 – flow-through, 3 – fraction 21, 4 – fraction 22, 5 – fraction 23, 6 – fraction 24, 7 – fraction 25, 8 – fraction 26, 9 – fraction 27, 10 – fraction 28, 11 – fraction 29, 12 – fraction 30; right: chromatogram for the Ni-affinity purification of H-*AbiUPO*, with the 280 nm absorbance in green, the 420 nm absorbance in orange, and the % imidazole buffer in black. 156

Figure 4-12: UV/Vis spectrum recorded for H- <i>Abi</i> UPO fraction 26 from the Ni-affinity purification.....	156
Figure 4-13: Chromatograms recorded for the purification of fed-batch fermented H- <i>Abi</i> UPO using left: Ni-affinity chromatography, with the 280 and 420 nm in green and orange, respectively, and the imidazole buffer percentage in black; right: size exclusion chromatography, with the 280 nm and 420 nm absorbance in blue and red, respectively.	157
Figure 4-14: UV/Vis spectrum recorded for two-step purified and concentrated H- <i>Abi</i> UPO.	158
Figure 5-1: 4-12% precast SDS gel with Novex Sharp ladder showing the concentrated fermentation supernatant (10X) of <i>rAae</i> UPO-H (lane 1) and the lyophilised <i>rAae</i> UPO-H sample (lane 2), running at around 60 kDa (arrow).	174
Figure 5-2: GC-chromatogram recorded after 2.5 h reaction of <i>rAae</i> UPO-H with ethylbenzene 1 (0.71 min) yielding in 1-phenylethanol 2 (1.24 min) and acetophenone 3 (1.27 min), with the traces for the lyophilised <i>rAae</i> UPO-H in green and pink and for the protein in solution in blue and red.....	175
Figure 5-3: Overview over the different screening reactions run with 10 mM substrate in the presence of <i>rAae</i> UPO-H (conversions in black) or H- <i>Abi</i> UPO (conversions in dark red); the substituents of the toluene derivatives cannot be further identified. Reactions were incubated overnight (20-24 h), except for the reaction of <i>rAae</i> UPO with 4 (1 h) and 6 (5 h).	176
Figure 5-4: GC chromatograms obtained for reactions of 1-benzylpiperidin-4-ol 22 (marker in blue, 2.71 min) in the presence and absence (ctrl) of <i>rAae</i> UPO-H yielding in piperidin-4-ol 23 (marker in red, 4.03 min).	178
Figure 5-5: GC-chromatograms for the conversion of ethylbenzene 1 to 1-phenylethanol 2 and acetophenone 3 using <i>rAae</i> UPO-H in the presence of different solvents, with acetonitrile in pink, acetone in blue, DMSO in red, ethanol in purple, ethyl acetate in green, and MTBE in light green.....	181
Figure 5-6: Conversions of ethylbenzene 1 to 1-phenylethanol 2 (dark red) and acetophenone 3 (orange) in the presence of <i>rAae</i> UPO-H over time at different hydrogen peroxide addition rates: left – every 15 min; right – every 30 min. All	

reactions were run in duplicates and the errors are reported as standard deviation.

..... 182

Figure 5-7: Conversions for the formation of methyl *p*-tolyl sulfoxide 37 from the corresponding sulfide 36 in the presence (dark red) and absence (orange) of *rAaeUPO-H* over time at different amounts and addition rates of hydrogen peroxide: top left – 1 mM hydrogen peroxide added every 10 min; top right – 1 mM hydrogen peroxide added every 30 min; bottom left – 2 mM hydrogen peroxide added every 10 min; bottom right – 2 mM hydrogen peroxide added every 30 min. All reactions were run in duplicates and the errors are reported as standard deviation..... 184

Figure 5-8: Conversion of ethylbenzene 1 to 1-phenylethanol 2 by *rAaeUPO-H*, using an *in situ* hydrogen peroxide generation system based on the conversion of methanol by *PpAOx*. 185

Figure 5-9: Conversions for the formation of 1-phenylethanol 2 and acetophenone 3 plotted over time at: top left: varying *PpAOx* concentration at 2 U mL⁻¹ *rAaeUPO-H* and 200 mM MeOH; top right: varying *rAaeUPO-H* concentration at 2 U mL⁻¹ *PpAOx* and 200 mM MeOH; bottom left: varying enzyme loading at 1:2 *rAaeUPO-H* to *PpAOx* ratio using 200 mM MeOH; bottom right: varying *rAaeUPO-H* to *PpAOx* ratios at 200 mM MeOH. Reactions were run in duplicates and the errors are reported as standard deviation. 186

Figure 5-10: Conversions for the formation of 1-phenylethanol 2 and acetophenone 3 plotted over time at: left: varying MeOH concentration at 1 U mL⁻¹ *rAaeUPO-H* and 2 U mL⁻¹ *PpAOx*; right: varying solvent and solvent concentration after 2 h incubation with 1 U mL⁻¹ *rAaeUPO-H*, 2 U mL⁻¹ *PpAOx* and 200 mM MeOH. All reactions were run in duplicates and the errors are reported as standard deviation. 188

Figure 5-11: Chiral GC-chromatogram for the separation of 1-(pyridine-2-yl)ethan-1-ol 5 into the (*R*)- (5a) and (*S*)- (5b) enantiomers, with the standard for 5a in blue, and the reactions catalysed by *rAaeUPO-H* in green and red. 189

Figure 5-12: Overview over the different ethylbenzene and thioanisole derivatives used in reactions with *rAaeUPO-H* and the determined conversions and enantiomeric purity. 190

Figure 5-13: GC-conversions of ethylbenzene 1 (left) and methyl *p*-tolyl sulfide 36 (right) in the presence of *rAaeUPO-H* over time yielding in 1-phenylethanol 2 and

methyl *p*-tolyl sulfoxide 37, respectively. The data represents the mean calculated from independent three reactions. All reactions were run in triplicates and the errors are reported as standard deviation.192

Figure 5-14: Chiral GC-chromatograms for samples containing 1-(4-fluorophenyl)ethanol 33 (left) and methyl phenyl sulfoxide 35 (right), obtained in the presence (top) or absence (bottom) of *rAae*UPO-H after 180 min reaction time. .193

Figure 5-15: *rAae*UPO-H catalysed conversion of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine 8 in the presence of the *Pp*AOx and methanol.194

Figure 5-16: Half-Normal plots for the responses 1, 2, 4, and 5 determined using the Design-Expert software, representing the influences of the different factors on the conversion to for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-5-ol 9 after 20 h (top left), and after extraction (bottom left), and the sum of conversion rates for the by-products 10-12 after 20 h (top right), and after extraction (bottom right).197

Figure 5-17: Overlay plot showing the *rAae*UPO-H to *Pp*AOx ratio plotted against the substrate concentration aiming to maximise the GC-based conversion towards the desired product 9 (>40% after 20 h, >50% after extraction), and minimising the by-product formation (<20% after 20 h, <15% after extraction) which is represented by the area highlighted in yellow.198

Figure 5-18: 1 g reaction of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine 8 with *rAae*UPO-H before extraction: left - GC chromatogram; right – ¹H-NMR zoomed to the area of interest between 5.0 and 5.5 ppm, showing the peaks for the protons bound to the hydroxylated carbon.199

Figure 5-19: ¹H-NMR spectrum obtained for the ‘pure’-sample after column chromatography, with the peak assignments for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol 9 and 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-7-ol 11. .200

Figure 5-20: Chiral GC-chromatogram for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol 9: left – racemic mixture (red) and (*S*)-9 (blue) as standards; middle – samples taken before (blue) and after (red) extraction; right – final samples for the fractions combined for ‘pure’ (blue) and ‘mixed’ (red).200

ii. List of Tables

Table 2-1: Composition of the different bacterial growth media used.....	38
Table 2-2: Overview of bacterial strains used and their respective antibiotic resistance.....	39
Table 2-3: Composition of Tris base, acetic acid, EDTA (TAE) buffer used for agarose gel electrophoresis experiments.	40
Table 2-4: Gene sequence optimised for <i>E. coli</i> and list of primers used for amplification of <i>AaeUPO</i> for the different cloning vectors.	41
Table 2-5: Gene sequence optimised for <i>E. coli</i> and list of primers used for amplification of <i>AbiUPO</i> for the different cloning vectors.	42
Table 2-6: Primers used for inverse PCR amplification of pETFPP31 from pETFPP30.	43
Table 2-7: left: Components and amounts used in PCR amplification reactions. right: Temperature, time, and cycle details describing the PCR amplification program....	44
Table 2-8: left: Pipetting scheme for the T4 DNA polymerase digest; right: Temperature, time and cycle details for the T4 digest.....	44
Table 2-9: left: Pipetting scheme for the In-Fusion cloning; right: Temperature, time and cycle details for the reaction.	45
Table 2-10: Overview of the restrictions enzymes and buffers required for enzymatic digests for the different vectors used.....	45
Table 2-11: left: Components and amounts used in the colony-PCR amplification; right: Temperature, time, and cycle details describing the PCR program.	46
Table 2-12: Composition of buffer used for the resuspension of bacterial cell pellets. The buffer was made up with ultra-pure water and filter over 0.2 µm filter paper before use.	47
Table 2-13: List and composition of buffers used for the solubility test.....	49
Table 2-14: Solutions prepared for SDS-PAGE.....	50
Table 2-15: Recipes for the resolving and stacking gel 12% SDS-PA gels.....	51
Table 2-16: Buffers used in immunoblot analysis.....	51
Table 2-17: Composition of buffers used in bacterial protein preparation and purification.....	53

Table 2-18: Antibiotic resistance and working concentrations for the different bacterial strains and plasmids used.....	59
Table 2-19: Results of the expression tests with the different <i>Aae</i> UPO constructs in different expression strains using various growth media at 16, 30, and 37 °C; ✓ - soluble protein; ✕ - insoluble or no protein; “-” – condition not tested.	62
Table 2-20: Results of the expression tests with the different <i>Abi</i> UPO constructs in different expression strains using various growth media at 16, 30, and 37 °C; ✓ - soluble protein; ✕ - insoluble or no protein; “-” – condition not tested.	63
Table 2-21: Summary of all expression test (incl. measurements over time) for the different <i>Aae</i> UPO constructs in different expression strains using various growth media at 16, 30, and 37 °C; ✓ - soluble protein; ✕ - insoluble or no protein; “-” – condition not tested.....	67
Table 2-22: Summarized results for the solubility screen performed with <i>Aae</i> UPO and <i>Abi</i> UPO; “✓” -soluble protein, “✓✓” – higher amounts of soluble protein, “✕” – no soluble protein.	69
Table 3-1: Composition of low-salt LB used in bacterial experiments with the yeast vector pPICZ.	80
Table 3-2: Pipetting scheme and temperature settings for the <i>Pme</i> I digest of DNA.	81
Table 3-3: Composition of the different growth media used in yeast experiments.	82
Table 3-4: Gene sequence and list of primers used for amplification of <i>Aae</i> UPO creating a C-terminal 3C site and an N-terminal His-tag and 3C cleavage site.	83
Table 3-5: Primers used for inverse PCR amplification of pPICZ deleting the C-terminal signal sequence and creating an N- or C-terminal 3C protease site.....	85
Table 3-6: left: Components and amounts used in PCR amplification reactions. right: Temperature, time, and cycle details describing the PCR amplification program. ...	85
Table 3-7: left: Components and amounts used in the colony-PCR amplification; right: Temperature, time, and cycle details describing the PCR program.	86
Table 3-8: Primers designed for sequencing experiments in pPICZ vector variants.	86
Table 3-9: List of components and amounts needed for the preparation of buffered glycerol and buffered methanol complex medium for test-expressions using <i>P. pastoris</i>	87

Table 3-10: Composition of basal salt and PTM ₁ trace salts used in the fermentation of <i>P. pastoris</i>	88
Table 3-11: Fermentation settings, limits and pump setting for the small scale (0.5 L) fermentations.	89
Table 3-12: Composition of buffers used in the protein preparation and purification.	92
Table 3-13: Compositions for the buffers used in activity assays.	95
Table 3-14: Parameters needed for the calculation of the specific activity towards ABTS and NBD.	95
Table 3-15: Substrate concentrations for ABTS, NBD, and VA used for the determination of K_m and k_{cat}	96
Table 3-16: Data collection and refinement statistics for rAaeUPO-H, where the numbers in brackets refer to data of the highest resolution shells.	99
Table 3-17: Specific activities towards NBD and ABTS during the 96 h methanol fed-batch phase of the fermentation of rAaeUPO-H.	120
Table 3-18: Kinetic parameters for the purified rAaeUPO-H (this work) and literature values for the <i>P. pastoris</i> expressed rAaeUPO. ¹³³	129
Table 3-19: Summary of the specific activities towards ABTS and NBD determined in the supernatants of various fermentations of rAaeUPO variants performed in this work.	138
Table 3-20: Biochemical and spectroscopic data for the <i>P. pastoris</i> expressed rAaeUPO-H in comparison to literature known data. ¹³³	140
Table 4-1: Gene sequence for the not optimised <i>AbiUPO</i> and list of primers used for amplification of for pPICZ α with N- and C-terminal His-tag and 3C site.	144
Table 4-2: Composition of buffers used in the protein preparation and purification.	145
Table 5-1: Overview of the different lyophilisation methods, regular, short, and weekend, used in this work, with the temperature, pressure, and duration times of the different steps.	163
Table 5-2: GC method and condition used for the achiral analysis (HP5) of reactions run at GSK. Numbers refer to structures are displayed in Figure 5-3.	164

Table 5-3: GC methods and conditions used for the analysis of the ethylbenzene and thioanisole derivatives, with achiral analysis using an HP5 column (numbers only), C-prefix referring to chiral analysis using a cyclosil B column, and B-prefix refers to a chiral BGB-175 column. The numbers refer to the structures displayed in Figure 5-12.	166
Table 5-4: <i>rAaeUPO-H</i> and <i>PpAOx</i> amounts used in experiments run to determine the best protein ratio for the <i>in situ</i> generation of hydrogen peroxide.	168
Table 5-5: GC method and condition used for the chiral analysis (HYDRODEX β -TBDAC) of 1-(pyridin-2-yl)ethan-1-ol 5 run at GSK.	169
Table 5-6: Set-up of the ten DoE runs, with the specifics of the factors 1 to 4, and the space type, with f for factorial and c for centre.	171
Table 5-7: GC method and condition used for the chiral analysis (cyclosil B) of 6,7-dihydro-5 <i>H</i> -cyclopenta[<i>b</i>]pyridin-5-ol 9 run at GSK.	171
Table 5-8: Responses determined for the ten DoE runs, with the conversions for 6,7-dihydro-5 <i>H</i> -cyclopenta[<i>b</i>]pyridin-5-ol 9, and the by-products 6,7-dihydro-5 <i>H</i> -cyclopenta[<i>b</i>]pyridin-5-one 10, 6,7-dihydro-5 <i>H</i> -cyclopenta[<i>b</i>]pyridin-6-ol 12, and 6,7-dihydro-5 <i>H</i> -cyclopenta[<i>b</i>]pyridin-7-ol 11, before and after extraction, as well as the crude isolated yield.....	195

II. Acknowledgments

I want to start by thanking Gideon Grogan and Jared Cartwright for supervising me throughout this project. Without their input and practical experience, I would have not been able to accomplish the work presented here. Their continued support and positivity kept me on track during periods where all I could see was empty Westerns. I would like to further acknowledge Anne Duhme-Klair, who acting as IPM during my PhD, always had an open ear and valuable advice for me.

I would also like to thank the industrial affiliates from the CoEBio³ funding scheme for their interest and input into the project. I especially want to thank Pete Sutton and Joe Adams for acting as my industrial mentors and for inviting me to spend time at GSK. Further, I would like to thank Alba Diaz-Rodriguez for looking after me during my time there and for bringing my project to a new level.

A big thank you to all the admin and technical staff in YSBL and M2, you made my life a lot easier and I appreciate all the work you put into running the labs.

I want to thank the whole of YSBL and the SynBio group from GSK for many thought-provoking conversations and countless attempts to teach me new practical skills. Without these networks of knowledge, I would have not been able to push the project to where it is now. A special thanks goes to the Grogan Group, past and present members, I want to thank you for the fun distractions, group outings, and quiz nights!

All of my friends, old and new, from near and far, I want to let you know how much I appreciate you being there for me during this time! Each and every one of you played a role on this journey, you kept me sane and made me smile and I will never forget that.

Finally, I want to thank my family for their love and support throughout, and my best friend, who has always been there, provided me with chocolate, and a shoulder to lean on. You make me a better person, and this is for you, too!

III. Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author. Any contributions are acknowledged below or within the text. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

All mass spectrometry experiments were performed by the Metabolomics & Proteomics branch of the Technology Facilities within the Department of Biology at the University of York. Johan P. Turkenburg and Sam Hart collected all X-ray data in house and at the Diamond Light Source. Alison Dann performed the fermentation of *rAaeUPO-H* on industrial scale at GSK Worthing and Cyril Boudet assayed and concentrated the protein on site. Alba Diaz-Rodriguez processed ¹H-NMR spectra recorded at GSK Stevenage and helped with the data analysis for the Design of Experiments.

1. Introduction

The selective functionalisation of unactivated carbon-hydrogen bonds has been a major challenge for organic chemists, as the reactions often involve harsh conditions or toxic reagents. Biocatalysts have shown to provide a greener alternative for a range of synthetic reactions so far. For oxygenation and hydroxylation reactions the literature suggests that the use of cytochrome P450 monooxygenases (CYPs), peroxidases or the rather recently discovered unspecific peroxygenases (UPOs) can provide a solution with respect to selectivity and sustainability. While CYPs are well studied and display a wide substrate scope; they are also characterised by their poor stability and are limited to small-scale laboratory use. UPOs on the other hand display good stability and tolerance of organic solvents, but access to the protein has proven difficult since the enzyme was first identified.

The following chapter is an introduction to the application of enzymes in chemical reactions and their potential for industrial processes. It will outline the different enzymes capable of hydroxylation and oxygenation reactions, with a special focus on the UPO from *Agrocybe aegerita*, its known homologs from the literature, and closest relatives amongst other enzyme families.

1.1. Enzymes in Chemistry and Industry

Organic synthesis is an enabling science giving access to compounds with new, and valuable properties. It provides molecules of interest to the pharmaceutical industry, and the material and agricultural sciences for new biological applications. In recent years the environmental impact of industrial processes has become more important, so companies are aiming to replace toxic and expensive chemicals and reduce overall cost and waste.¹

The application of enzymes in the transformation of organic compounds is termed as biocatalysis and was initially driven by the interest in understanding biochemical pathways and enzyme mechanisms.² As for any synthetic method it is essential that it is robust, provides a reliable yield, has a broad applicability, and also high selectivity. Limited performance has long been a downside to biocatalysis, where reactions were tailored to enzymes and not *vice versa*. In recent years however, a

series of innovations allowed this principle to be reversed, and work on enzymes to tailor them to reactions of interest has undergone rapid development.³ The quick and cheap sequencing of DNA and whole genomes has allowed the exploration and understanding of sequence-function relationships.⁴ Advances in gene synthesis have led to optimised reaction conditions and parallel synthesis, which helped to lower the cost of synthetic genes.⁵ A range of bioinformatics tools available today, such as multi-sequence alignment and homology searches, have helped with the deeper understanding of evolutionary relationships and conservation of structurally or catalytically important domains. As a result the preparation of small libraries now contains a high proportion of catalytically active variants.⁶ Coupled with high-throughput approaches, providing simple and affordable testing of whole libraries, the analysis time could be reduced drastically. Together, these approaches make the search and use of proteins for specific reactions more attractive. The increase in solved protein crystal structures further paved the road to structure guided evolution projects, providing insight into mechanism, selectivity and stability. These rational design approaches can be easily complemented by directed evolution experiments which cycle through rounds of random *in vitro* mutagenesis screening for and selecting by improvements in stability, specificity, and selectivity.⁷

An overview of enzyme-catalysed reactions in industry is given in the following paragraphs, highlighting examples taken from recent articles on the industrial application of enzyme catalysis.^{3, 8} Fine and bulk chemistry utilise enzymes in various reactions, including the production of acrylamide (nitrile hydratase)⁹, 1,3-propanediol (glycerol dehydratase and oxidoreductase isozyme)¹⁰, and (*R*)-epichlorohydrin (halohydrin dehalogenase and epoxide hydrolase; see Figure 1-1).^{11, 12}

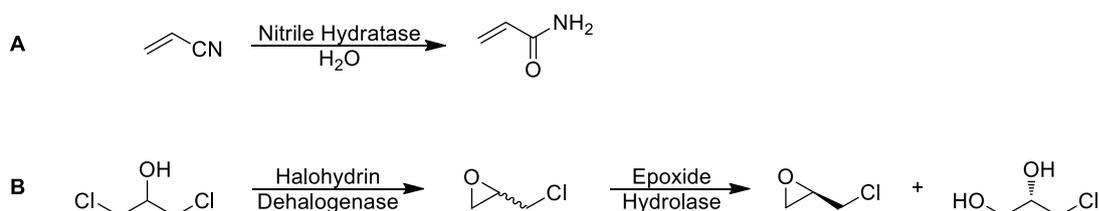


Figure 1-1: Selected enzyme catalysed reactions used in fine and bulk chemistry. **A** – Formation of acrylamide from acrylonitrile using nitrile hydratase⁹; **B** – Synthesis of (*R*)-epichlorohydrin from 1,3-dichloropropan-2-ol using halohydrin dehalogenase and an epoxide hydrolase.^{11, 12}

In the pharmaceutical industry transaminases are applied in the production of sitagliptin, a type II diabetes drug (Januvia by Merck).¹³ The monoamine oxidase from *Aspergillus niger* is involved in the synthesis of boceprevir, a precursor in the synthesis of a chronic Hepatitis C drug (Victrelis by Merck; see Figure 1-2 A).¹⁴ Keto-reductases (KREDs) are involved in the synthesis of various drug intermediates, such as precursors for montelukast, an anti-asthma drug (Singulair by Merck; see Figure 1-2 B) or the cholesterol-lowering atorvastatin (Lipitor by Pfizer) which is part of a multi-enzyme cascade.¹⁵

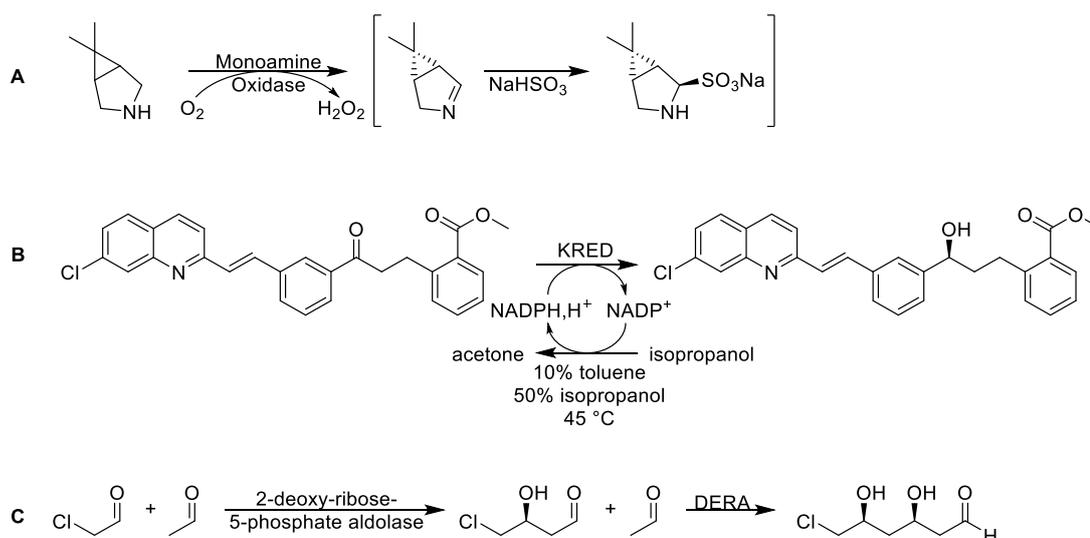


Figure 1-2: Selected enzyme catalysed reactions as utilised by pharmaceutical industry. **A** – Synthesis of a boceprevir precursor using the monoamine oxidase (MAO) from *Aspergillus niger*¹⁴; **B** – Application of a keto-reductase (KRED) from *Lactobacillus kefir* in the synthesis of a montelukast intermediate¹⁵; **C** – C-C-bond formation of a statin-drug intermediate catalysed by an engineered 2-deoxy-ribose-5-phosphate aldolase (DERA).¹⁶

One-pot cascade reactions are of great interest as they allow the minimisation of steps involved in a reaction, as well as starting from achiral, often cheaper materials.¹⁷ Cascade examples often include redox enzymes which can be employed in cofactor recycling.¹⁸ An example for a cascade reaction giving access to norephedrine and pseudo-norephedrine utilises acetoxyacid synthase (AHAS-I) to couple pyruvate to benzaldehyde and selective transaminases (ATAs) to convert the ketone into the (*R*)- or (*S*)-amine, resulting in pseudo-norephedrine and norephedrine, respectively.¹⁹ Increased interest in carbon-carbon bond formation intensified the exploration of aldolases, with 2-deoxy-ribose-5-phosphate aldolase (DERA), being one example. An engineered version of DERA is able to produce key

intermediates for statin-type drugs like atorvastatin (as above for example Lipitor by Pfizer; see Figure 1-2 C).¹⁶

Other industries are also taking advantage of greener and eco-friendlier enzymatic approaches. The food industry employs enzymes in the production of prebiotics and artificial and low-calorie sweeteners.²⁰ The cosmetic industry uses enzymes in the production of arbutin, a skin lightener²¹ and emollient esters, which have moisturizing properties.²² The textile and pulp and paper industry is able to reduce the amount of harsh chemicals used in their processes by applying enzymes. A cellulase from *Trichoderma viride* for example is used in the staining processes of jeans²³ and laccases are able to remove lignin impurities in paper, reducing the need for chlorine.²⁴

In addition to the limited number of enzymes that are used in industrial applications, there is an even wider spectrum available for laboratory use, which is constantly expanding.

1.1.1. Enzymatic Oxygenation and Hydroxylation Reactions

Focussing on the ability to oxygenate and hydroxylate organic compounds, the following section will give an overview of different enzyme classes that are able to address these reactions. Oxidoreductases, sub-grouped into oxidases, peroxidases, oxygenases/hydroxylases, and dehydrogenases/reductases, are able to facilitate the electron transfer needed for these reactions often by utilising cofactors. Depending on the mode of binding, cofactors are termed as prosthetic groups in the case of covalent binding, or coenzymes if lower modes of affinity are observed.²⁵

1.1.1.1. Monooxygenases

Monooxygenases utilise molecular oxygen to transfer one atom of oxygen to the substrate while the second oxygen is reduced to water.²⁶ To achieve this reaction, monooxygenases have to activate the oxygen through electron donation, and allow the transition from a unreactive triplet state to a singlet state. The type of reaction depends on the cofactor present.²⁷

Flavin-dependent monooxygenases catalyse reactions ranging from hydroxylations and sulfoxidations to Bayer-Villiger oxidations, epoxidations, halogenations and oxidative decarboxylations.²⁸ Their versatility is also reflected in the wide range of

biological processes in which these enzymes are involved, including for example the biosynthesis of hormones, and vitamins. The formation of a flavin C4a-oxygen adduct is important for the oxygen activation and allows for the incorporation of oxygen into the substrate (see Figure 1-3).

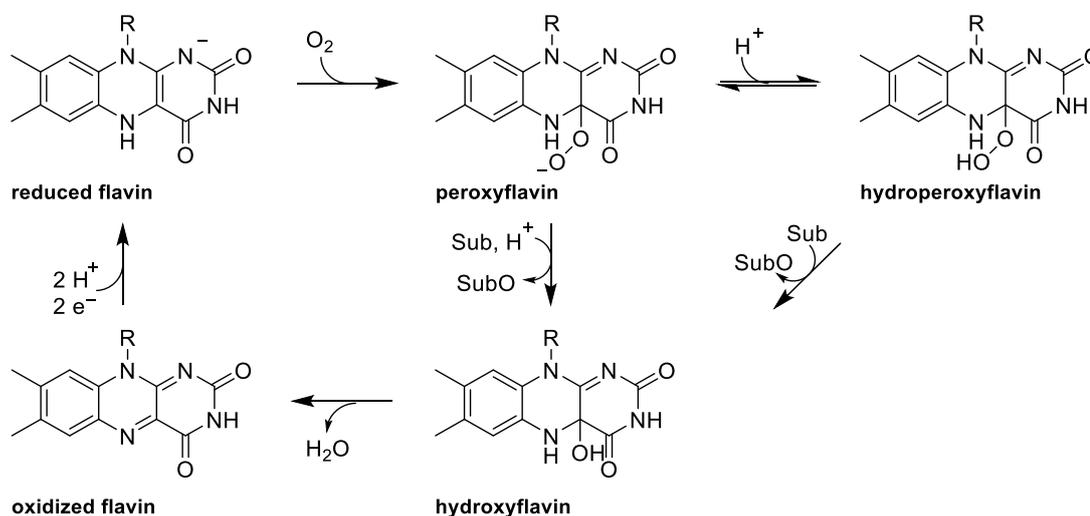


Figure 1-3: General mechanism for the oxygen activation in flavin-dependent monooxygenases. Upon the addition of oxygen, the reduced flavin converts into the flavin C4a-(hydro)peroxide which can react with the substrate to give the flavin C4a-hydroxide.

Depending on their electron donor, protein fold and specific reaction, flavin-dependent monooxygenases are classed into groups. Group A and B rely on an NAD(P)H as an external electron donor and are known for regioselective hydroxylations of phenolic compounds (group A), ester and lactone formations from ketones (group B, Bayer-Villiger monooxygenases, BVMOs), and heteroatom oxygenations (flavo-protein monooxygenases, *N*-hydroxylating monooxygenases). One example for a group A member is the kynurenine 3-monooxygenase (KMO). KMO catalyses conversion of L-kynurenine to 3-hydroxykynurenine, an intermediate of the tryptophan catabolism, and its activity has been connected to neurodegenerative diseases (see Figure 1-4).²⁹ Members of the groups C-F rely on external reductases for the generation of reduced flavins. Examples include the bacterial luciferase, with its unique light-emitting ability, alkanesulfonate monooxygenase, which catalyses the desulfonation of alkanes, styrene monooxygenases, which are able to produce epoxides, and tryptophan 7-halogenases, which regioselectively halogenate tryptophan through an hypochlorous acid intermediate (see Figure 1-4).²⁸ The last groups G and H are internal flavoprotein monooxygenases which achieve flavin reduction through substrate

oxidation/decarboxylation. The decarboxylation of tryptophan yielding indole-3-acetamide by tryptophan 2-monooxygenase, and the oxidation of L-lactate to acetate, water and carbon dioxide by lactate 2-monooxygenase are two examples of internal flavoprotein monooxygenases (see Figure 1-4).

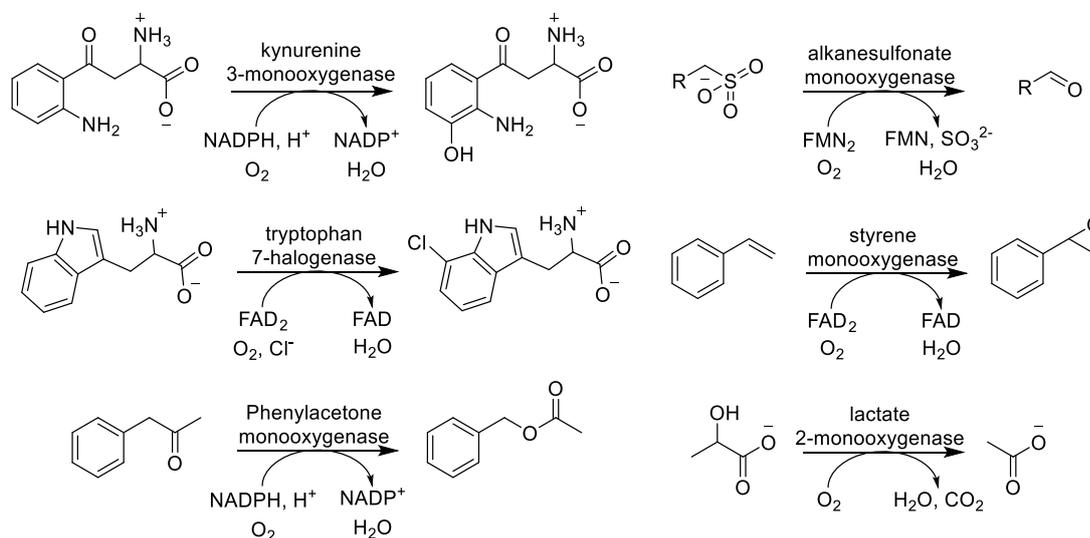


Figure 1-4: Selected examples of transformations catalysed by various flavin dependent monooxygenases.

While flavin-dependent monooxygenases cover a wide spectrum of oxygenations, each individual enzyme often prefers a specific type of reaction, mainly influenced by its fold.²⁷ A similar trend can be observed for the substrates that are converted. Enzymes with high sequence similarities might be able to convert the same substrate into different regio- and stereoisomers. In recent years, genome mining, especially for BVMOs, revealed many possible flavin-dependent monooxygenases. The access to some of those enzymes, as well as the maintenance of the cofactor/coenzyme system are known limitations for these enzymes.³⁰

Besides flavin-dependent monooxygenases, subgroups utilising other cofactors are also known. Non-heme iron-dependent monooxygenases are mostly bacterial enzymes which carry two iron atoms as cofactor and consist of three components, the monooxygenase, the reductase and a regulatory protein.³¹ The mechanism proceeds *via* a diiron(IV)-bis- μ -oxo intermediate, which is similar to the reactive intermediate found in copper-dependent monooxygenases.^{32, 33} The most prominent monooxygenases are cytochrome P450 monooxygenases (CYPs), which coordinate a heme cofactor in the active site and have been reported active in a wide range of reactions, with broad substrate scope. CYPs will be discussed in more detail (1.1.3).

1.1.1.2. Dioxygenases

Another class of enzymes that utilise molecular oxygen as oxygen donor for redox reactions are dioxygenases. As the name suggests, these enzymes are able to insert both oxygen atoms into the substrate or a primary and a co-substrate.³⁴ Ferrous iron and α -ketoglutarate-dependent dioxygenases (α KGDs) are the largest subgroups of mononuclear non-heme iron-dependent dioxygenases, able to catalyse reactions ranging from hydroxylations, dealkylations, desaturations, epoxidations to halogenations and cyclisations, to name a few (see Figure 1-5).³⁵ In the active site of α KGDs Fe(II) is coordinated by three amino acids (2x His, Asp/Glu), leaving the three remaining sites occupied by water.³⁶ Oxygen binding occurs only after the coordination of the primary substrate.³⁷ The full reaction requires four electrons, two electrons are provided by the primary substrate and further two electrons, needed for the activation of the oxygen, are obtained from the conversion of α KG into succinate and carbon dioxide.

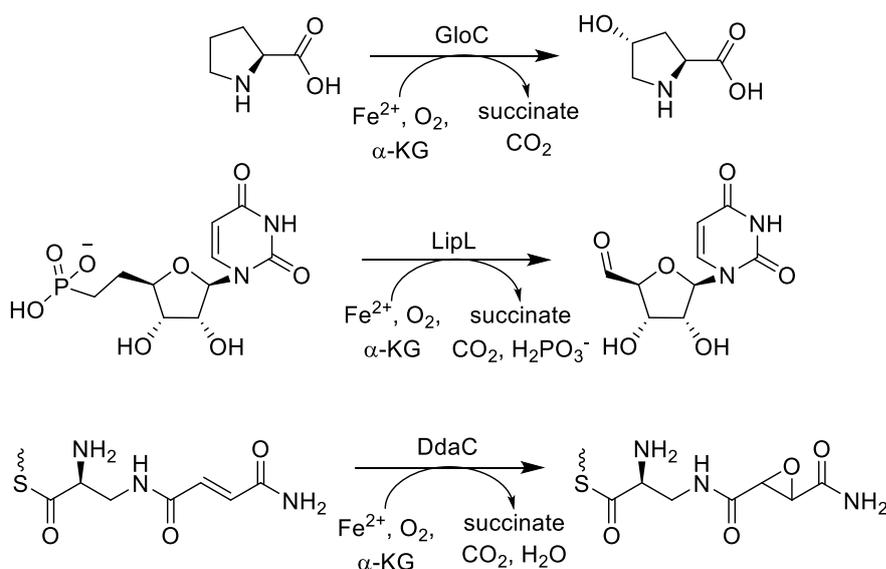


Figure 1-5: Examples of biotransformations catalysed by α -KG-dependent dioxygenases, with GloC - *trans*-4-proline hydroxylase,³⁸ LipL - putative TauD enzyme from the A-90289 gene cluster,³⁹ and DdaC - Dapdiamide C.⁴⁰

α KGDs are routinely active towards amino acids, but they are also involved in the synthesis of precursors of non-ribosomal peptide products and in the modification of natural products. It has been shown that amongst others they can accept DNA, RNA, oligosaccharides, polyketides, and terpenoids as substrates.³⁵ Whilst α KGDs are a very versatile enzyme class many limitations, like poor catalytic efficiency, enzyme stability, and substrate selectivity are problems which still need to be overcome.

1.1.1.3. Heme containing enzymes

Heme containing enzymes are one of the best known classes of enzymes catalysing the functionalisation of inert carbon-hydrogen bonds. A protoporphyrin IX coordinates an iron which further coordinates/binds to an active site residue (proximal) leaving a distal coordination site for ligand/substrate binding. The nature of the proximal ligand dictates whether peroxide or dioxygen is utilised in the reaction mechanism.⁴¹ All reactions proceed *via* an Fe(IV)-oxo- π -radical known as compound I (see Figure 1-6). Depending on the enzyme subfamily this intermediate is sequentially utilised. Peroxidases catalyse oxidative coupling using one-electron abstractions,⁴² haloperoxidases perform halide oxidations to facilitate the halogenation of sp^2 hybridised C-H bonds,⁴³ and CYPs can catalyse hydroxylation reactions *via* hydrogen abstraction, electrophilic aromatic substitution, or direct oxo transfer.⁴⁴

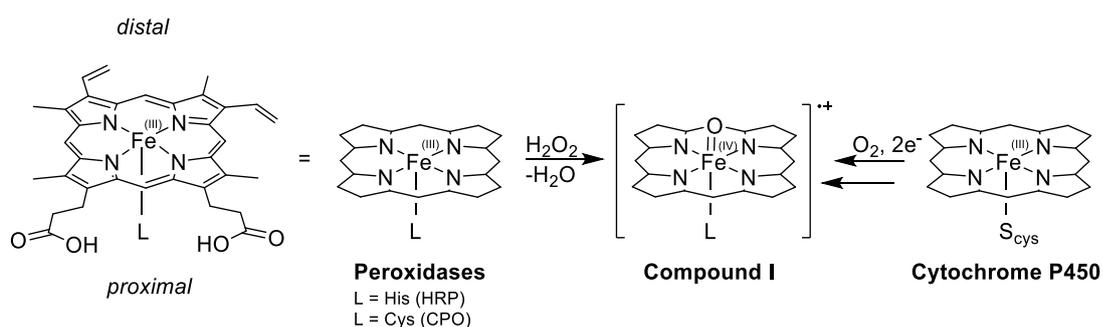


Figure 1-6: Structure of the heme cofactor and general schemes for the formation of compound I in different heme enzymes.

Within the peroxidase family structural variations can lead to differences in the reactivity. In horseradish peroxidases (HRP) for example the proximal ligand is a histidine coordinating to the iron *via* a nitrogen atom. This imidazole character is believed to induce a preference for one-electron transfers over two-electron transfers. As a result, it enables HRP to couple electron-rich (hetero)arenes or (hetero)arenes and heteroatoms.⁴⁵ Haloperoxidases, like the chloroperoxidase from *Caldariomyces fumago* (*CfuCPO*) are unique amongst peroxidases due to a heme-thiolate coordination originating from a cysteine. The strongly donating axial ligand encourages heterolytic O-O cleavage and allows lowering the reduction potential of the iron.⁴⁶ In addition to halogenation reactions, CPO is also able to catalyse oxidative transformations, which will be discussed in further detail below.

A proximal cysteine ligand is also characteristic for cytochrome P450s, which get their name from the absorption maximum when coordinating carbon monoxide as distal ligand. In comparison to haloperoxidases the active site of CYPs is much more hydrophobic and welcomes a wide range of substrates including those from biosynthetic and degradative pathways.³⁶ CYPs can activate dioxygen to give compound I, but they are also able to utilise hydrogen peroxide *via* the shunt-pathway (see Section 1.1.3). Transformations catalysed by CYPs include epoxidations, deformylations, rearrangements and many more.⁴⁷

Related to both *CfuCPO* and CYPs are the unspecific peroxygenases (UPOs) which were discovered about a decade ago. They also show coordination to the heme-iron *via* a cysteine residue and are able to catalyse a broad range of reactions ranging from aromatic, aliphatic and allylic hydroxylations, to heteroatom oxidations and dealkylations, as well as epoxidations, and selected halogenations.⁴⁸ They are often referred to as the link between peroxidases and monooxygenases as they are able to combine catalytic features from both classes, allowing for one- and two-electron transfer. While structural similarities are present, no sequence similarities have been observed. The self-sufficient nature and high stability of UPOs makes them interesting alternatives to the well-studied *CfuCPO* and CYPs and will be discussed in detail (1.1.2 and 1.1.3, respectively).

1.1.2. Peroxidases

Peroxidases catalyse hydrogen peroxide-mediated one- and two-electron transfer reactions. There are four distinct reaction types that peroxidases are able to catalyse, the first being the reduction of peroxide coupled to the one-electron oxidation of donors (AH_2) to their respective radicals (AH^\cdot). Halides are two electron donors and are oxidised to the corresponding hypohalous acid (HOX) during the reduction of peroxide to water. A few peroxidases also exhibit catalase activity, which allows them to derive dioxygen from two molecules of peroxide. The selective functionalisation of organic molecules using oxygen derived from peroxides, the peroxygenation reaction, is the last distinct activity peroxidases are able to display.⁴⁹ The following will focus on fungal heme-thiolate peroxidases, like the chloroperoxidase from

Caldariomyces fumago (*CfuCPO*), which shows unique features within the peroxidase superfamily.

CfuCPO has been studied extensively since it was discovered in the 1960s as the first halogenating enzyme.⁵⁰ It has often been described as a hybrid between peroxidases and CYPs due to its similarities to both enzyme families. Like CYPs *CfuCPO* coordinates the heme *via* a sulfur atom provided by a cysteine, rather than a nitrogen donated by histidine as observed in other peroxidases. The distal pocket however is lined with more polar residues giving it a stronger peroxidase character. Structural studies also found that, unlike other peroxidases, in *CfuCPO* the heme is accessible for substrates to directly interact with the compound I (cpd I).⁵¹ While the primary function of *CfuCPO* is the participation in chlorination reactions, for example the production of the antibiotic caldariomycin, in the absence of halides it exhibits other oxidative reaction potentials. Reported reactions include the oxidation of phenols and anilines, styrene epoxidation, and sulfoxidations.⁵² The mechanism by which *CfuCPO* halogenates its substrates is still not fully understood. The formation of compound I is similar to other heme peroxidases, though glutamic acid and not histidine is involved in an acid/base mechanism. The next step is the formation of an iron(III) hypohalite intermediate, called compound X (see Figure 1-7, top).⁵³ It is proposed that compound X rapidly decays into the resting enzyme and hypohalous acid. The hypohalous acid in turn oxidises the substrate *via* halogenation. It is unknown whether this step occurs within the enzymes' active site, and therefore questions about the enzymes contribution towards regio- and stereoselectivity are often discussed. In the absence of halides *CfuCPO* acts like other peroxidases and from compound I the enzyme returns to its resting state *via* two one-electron transfer steps and a compound II intermediate (see Figure 1-7, bottom).

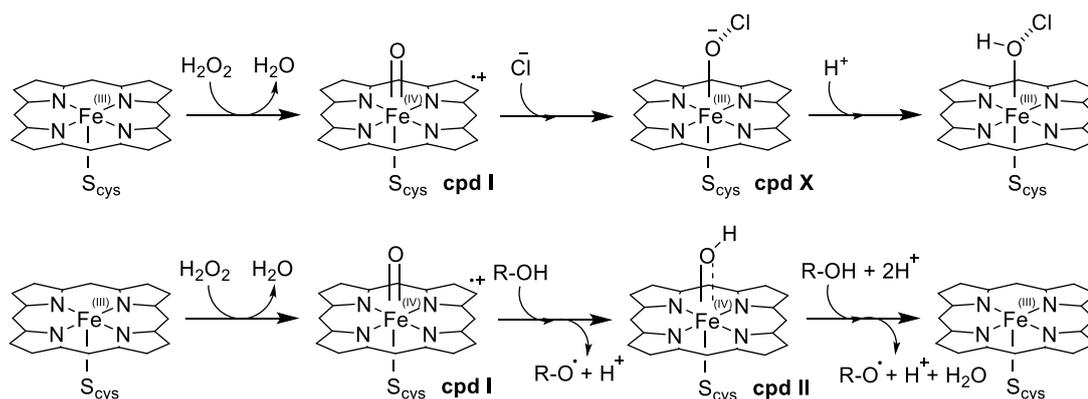


Figure 1-7: Proposed catalytic mechanism for the reactions catalysed by *CfuCPO*: **top** – chlorination reaction, **bottom** – peroxidase-like reactions.

The structure of *CfuCPO* was solved in the mid-1990s and helped not only the understanding of the enzymes' mechanism but also provided information about binding sites and key residues.⁵¹ The *CfuCPO* crystal structure revealed a novel tertiary fold with approx. 50% helical character. The structure of the active site shows resemblance to other heme proteins. The protein shows one disulfide bond, multiple *cis*-Pro, and the post-translational modifications include three *N*- and eleven *O*-glycosylation sites, as well as the cyclisation of an *N*-terminal glutamic acid. The Cys-29 coordinates to the heme-iron at a distance of 2.3 Å. The sulfur atom also forms hydrogen bonds to the neighbouring amides at 3.6 Å, which is described as an unusual distance and might be important for how the iron uses the cysteine ligand. The heme itself is bowl shaped and the iron is displaced by 0.14 Å towards the proximal ligand. Interactions with the protein backbone fix the heme in its position. In addition, one of the heme propionate residues coordinates a manganese cation, which shows an octahedral coordination sphere occupied by the heme propionate, amino acid residues and well as water ligands. The distal ligand of the iron is at 3.4 Å distance and confirms the expected high-spin ferric state. The distal binding site resembles the polar binding sites observed in peroxidases. From the active site a substrate channel is observed with a mostly hydrophobic character towards the active site, for organic substrate coordination and a more polar character towards the top. A glutamic acid residue, Glu183, has been identified in the proximity of the peroxide binding site as the only polar and charged amino acid and is therefore believed to participate in an acid/base mechanism. In theory ligands with weaker acidity than glutamic acid ($pK_a \approx 4.1$) are required in order to achieve protonation of

the ligand at lower pHs and allow for the coordination to the deprotonated Glu183 in the enzymes active site. More recent studies gave more insight into halide binding in *CfuCPO*.⁵⁴ Docking was observed in a dent on the surface, in close proximity to a second channel accessing the active site. Within this narrow channel two more halide binding sites were observed, with one being specific for iodide. Substrate complexes were also obtained in the study; however, the substrate was not found in a favourable orientation. Cyanide soaks were performed successfully to block the axial binding site but attempts to reorient the substrate into a more favourable position were unsuccessful.

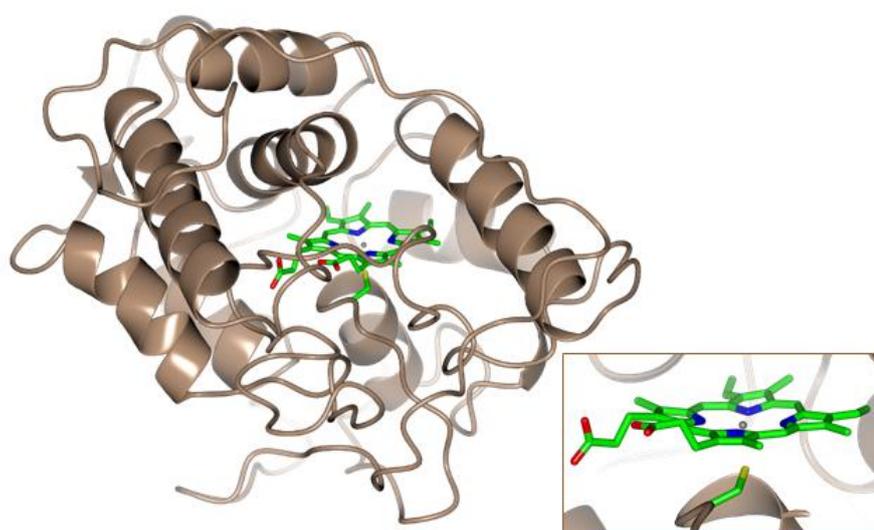


Figure 1-8: Ribbon-structure of *CfuCPO* (1CPO), with the heme and Cys29 (insert) in cylinder representation (carbon – green, oxygen – red, nitrogen – blue, sulfur – yellow). Created using CCP4mg.

The expression of *CfuCPO* in its native strain has been found to give yields as high as 500 mg L⁻¹, when working with fructose induced cultures.⁵⁵ Attempts to express the enzyme recombinantly have not been very successful. *E. coli* produces the protein in its apo-form in inclusion bodies, insect cells give extracellular inactive enzyme, and *S. cerevisiae* and *A. nidulans* did not produce any soluble protein. Only *A. niger* has been reported to produce fully active *CfuCPO* in a recombinant system.⁵⁶ Whilst the expression of *CfuCPO* is not a problem, there are other factors limiting the usage of the enzyme. Like many other heme proteins, *CfuCPO* is easily inactivated by an excess of hydrogen peroxide. This is due to the formation of superoxide, O₂⁻, as part of the reaction with the peroxide, leading to oxidative damage caused to the porphyrin moiety.⁵⁷ This drawback can be overcome by regulating the peroxide concentration,

either through feed-on-demand, *in situ* generation, or through the addition of radical scavengers.

1.1.3. Cytochrome P450 Monooxygenases

Cytochrome P450 monooxygenases (P450s or CYPs) are the most prominent class of heme-dependent monooxygenases with more than 21,000 members identified by 2014.⁵⁸ This sheer number allows for the catalysis of a broad range of synthetically challenging oxidation reactions, while specific enzymes often display a very narrow substrate scope. CYPs are physiologically involved in the biosynthesis of steroids, vitamins, lipids, and in eukaryotic systems they have also been found to be involved in biodegradation, as for example of drug metabolites.⁵⁹ Besides hydroxylation reactions CYPs also show activity in alcohol, and *N*-, *O*-, and *S*-dealkylations, C-C-bond cleavage and epoxidation reactions. The name pigment 450 (P450) originates from the enzymes absorption spectrum when complexing carbon monoxide as the iron's axial ligand, with a distinct maximum at 450 nm. The oxygen activation in CYPs is dependent on the supply of electrons derived from a NAD(P)H cofactor. The electron transfer to the active site is facilitated by a range of electron transfer systems, which can be classed in ten distinct classes, depending on the number of independent components participating. Four of these classes are depicted below (see Figure 1-9).⁶⁰

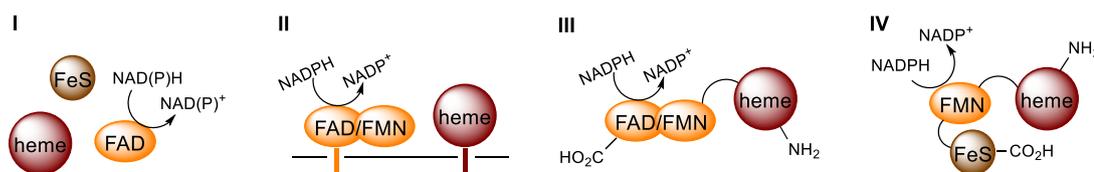


Figure 1-9: Classes of CYP electron transfer systems; with FeS – iron-sulfur cluster, FAD – Flavin adenine dinucleotide, FMN – Flavin mononucleotide.

Most prokaryotic CYPs are three component systems comprised of a NAD(P)H-dependent FAD-containing reductase, an iron-sulfur ferredoxin, and the monooxygenase protein (class I, example: P450_{cam}). Eukaryotic CYPs are often two component systems containing a membrane bound NAD(P)H-dependent diflavin [FAD/FMN] reductase (CPR) in addition to the heme protein (class II, example: human liver P450s). Class III describes a fusion protein containing a CYP and a CPR domain (example: P450_{BM3}), while class IV proteins possess an additional iron-sulfur cluster

domain (example: P450_{RhF}). The two types of self-sufficient single component CYPs are the most attractive platforms for development.⁵⁹

The electrons derived from the reduced pyridine nucleotide (NAD(P)H) are channelled to the iron in the active site to facilitate the oxygen activation and allow for substrate modification. The catalytic cycle is initiated by the substrate replacing the water coordinating to the iron in the active site (**A**), inducing a shift in the ferric spin state, from low-spin to high-spin (see Figure 1-10, **B**). This change in redox potential allows for the first electron to reduce the iron to the ferrous (Fe^{II}) state (**C**). The dioxygen can now bind to the iron and cycles through a conversion of triplet oxygen (³O₂) to singlet oxygen (¹O₂) and finally settles in a ferri-superoxo anion (Fe^{III}-O₂⁻) intermediate (**D**).⁶¹ The addition of the second electron at this stage generates a nucleophilic di-negatively charged ferri-peroxo intermediate, with one negative charge on the oxygen and the second delocalised over the cysteine ligand (**E**). Protonation leads to the formation of compound 0 (**F**), a ferric hydroperoxo form, which immediately dissociates by O-O bond scission, producing water and compound I (**G**), a porphyrin π-radical Fe^{IV}-oxo intermediate. From here the activated oxygen is inserted into the substrate and the heme regenerates to its ferric state (**A**).

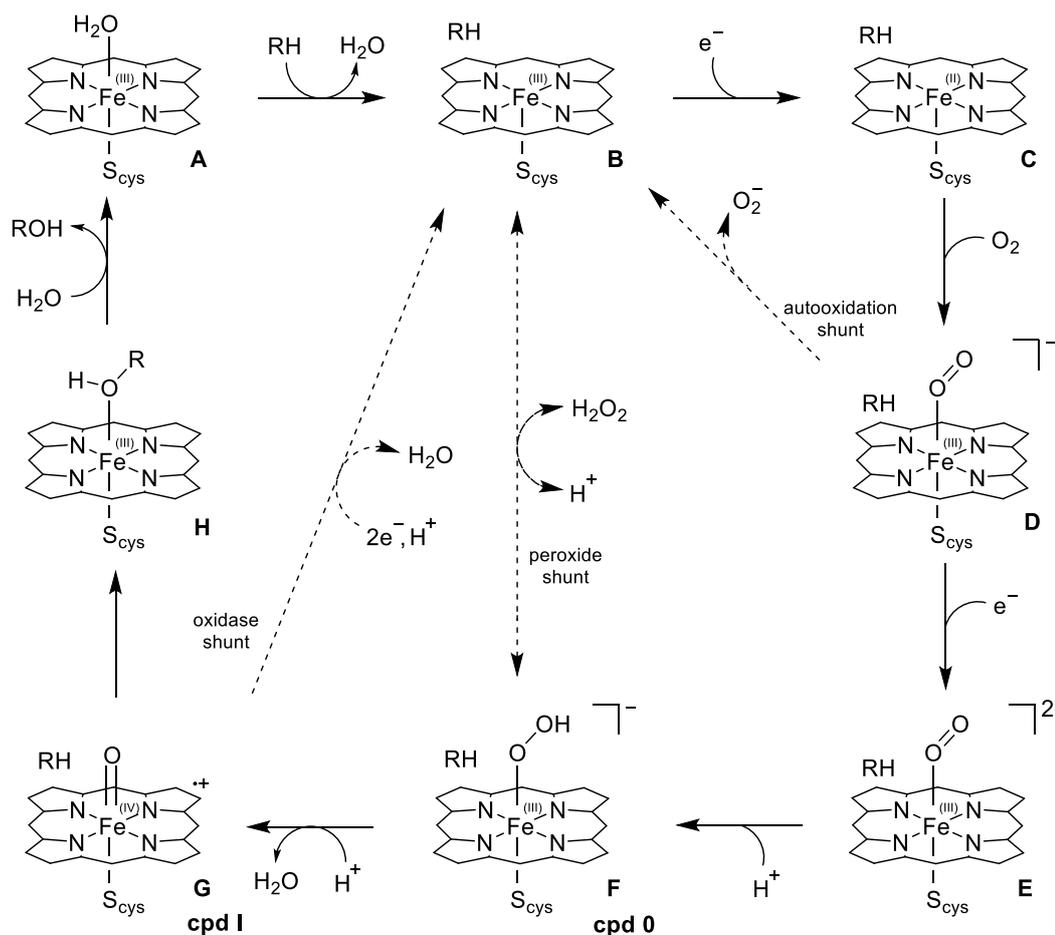


Figure 1-10: The catalytic cycle of oxygen activation by CYPs and their shunt pathways. Starting from the resting heme (A), the proximal water ligand is replaced in the presence of substrate (B). The supply of an electron allows for iron reduction yielding C, which is able to react with molecular oxygen. Via an ferri-superoxide intermediates (D) and upon addition of a second electron, a ferri-peroxo intermediate is formed (E). Protonation leads to the formation of compound 0 (F), and the loss of water generates compound I, the porphyrin π radical Fe^{IV} oxo intermediate (G). Through oxygen insertion into the substrate (H), the return to the ferric resting state is observed (A).

The mechanism by which CYPs activate dioxygen can also lead to the formation of various reactive oxygen species, through uncoupling reactions. Various intermediates have an unstable character which permit the release of singlet oxygen, hydrogen peroxide, and hydroxyl radicals, at the expense of electrons provided by NAD(P)H. These uncoupling reactions describe one limitation of CYPs. However, the ability of CYPs to activate oxygen derived from hydrogen peroxide through the peroxide shunt pathway also has advantages, as it eliminates the need for the extensive electron transfer systems.^{61, 62} Compared to the electron-assisted oxygen activation, the peroxygenase activity derived from the peroxide shunt pathway is lower and reactions are less efficient.⁶³

Many crystal structures of CYPs have been solved over the years. While the overall fold is maintained, the exact positioning of the substrate-binding loops can vary.⁶⁴ The closer the residues to the heme-containing active site, the more conserved they are. The two helices in direct contact with the heme are the main contributors to substrate specificity. The segment carrying the cysteine residue coordination, the heme-iron displays a high degree of rigidity in order to fix the residue in place (see Figure 1-11). Similar to the *CfuCPO*, the sulfur forms hydrogen bonds to the neighbouring amides, which regulates the iron's redox potential.

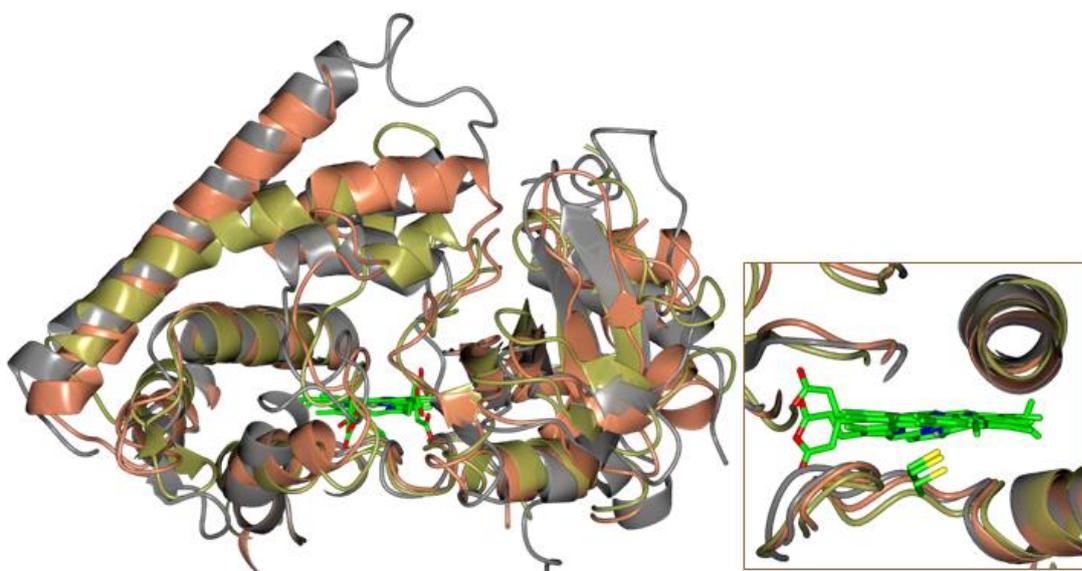


Figure 1-11: Superimposed ribbon-structures of P450_{cam} (2CPP) in gold, P450_{terp} (1CPT) in coral and P450_{BM3} (2HPD) in grey to illustrate the rigid and flexible elements of the structures, with a close-up of the active site heme and coordinating cysteine, with carbon – green, oxygen – red, nitrogen – blue, sulfur – yellow. Created using CCP4mg.

It is widely accepted that the solvent in the active site acts as proton donor. In some cases, threonine residues have been recorded at an appropriate distance, which can replace the role of the solvent.⁶⁵ Structures with substrates bound within the active site have confirmed the switch from a low-spin (substrate free) to a high-spin (substrate bound) state in the iron (P450_{cam}) due to a change in the distance between the metal and its ligand.⁶⁶ Furthermore, the closing of the heme access channel and the loss of solvent in the channel have been observed upon substrate coordination for a P450_{BM3} variant.⁶⁷

Both P450_{cam} and P450_{BM3} have been widely studied and are subject to extensive engineering attempts, due to their high catalytic activity, as well as their good solubility and expression in *E. coli*, and the availability of structural information.⁶⁰

P450_{cam} is a class I CYP and catalyses the oxidation of (+)-camphor to 5-*exo*-hydroxycamphor (see Figure 1-12). P450_{BM3} is the most prominent self-sufficient CYP (class III) which is known for the terminal hydroxylation of fatty acids.

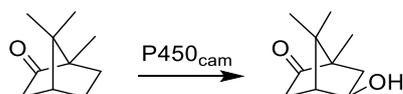


Figure 1-12: Oxidation of (+)-camphor catalysed by P450_{cam}.

Both enzymes have been engineered using a mixture of site-directed, and random mutagenesis, as well as directed evolution approaches, to accept a wider range of substrates and to tune their regio- and stereoselectivities. Furthermore, protein engineering and the development of appropriate assays allowed addressing other limitations, which included improving stability and solvent tolerance, changing electron transfer-systems and swapping cofactor selectivities.⁵⁹ As a result the potential industrial and biotechnological applications of CYPs have increased.⁵⁸ The *de novo* synthesis of artemisinic acid (see Figure 1-13) in *S. cerevisiae* using CYP71AV1 from *Artemisia annua* is a good example of how CYPs have been adapted to make it into industrial applications.⁶⁸

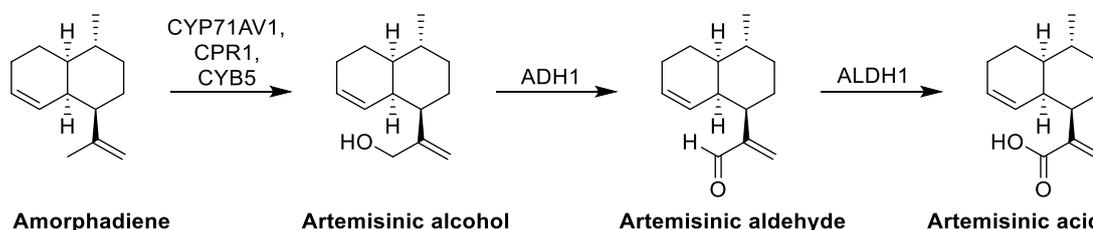


Figure 1-13: Amorphaadiene oxidation to artemisinic acid. The formation of artemisinic alcohol catalysed by CYP71AV1 (a CYP), CPR1 (cognate reductase), and CYB5 (cytochrome b₅), followed by the oxidation to artemisinic aldehyde catalysed by ADH1 (alcohol dehydrogenase), and the final oxidation giving artemisinic acid using ALDH1 (artemisinic aldehyde dehydrogenase).

1.2. Unspecific Peroxygenases

The unspecific peroxygenase from *Agrocybe aegerita* (*AaeUPO*) was first reported in 2004 and was originally classed as an aromatic peroxygenase (APO in older publications).⁶⁹ Since then it was studied in more detail and the reaction scope extended from the hydroxylation of toluene and naphthalene to a wide range of substrates.⁴⁸ In addition to aromatic hydroxylation reactions, alkane and alkyl hydroxylations have been observed, as well as epoxidations, heteroatom oxidations and dealkylations, and selected halogenation reactions (see Figure 1-14).^{52, 70} As a co-

substrate UPOs utilise hydrogen peroxide which acts both as electron acceptor and oxygen-donor.⁷¹

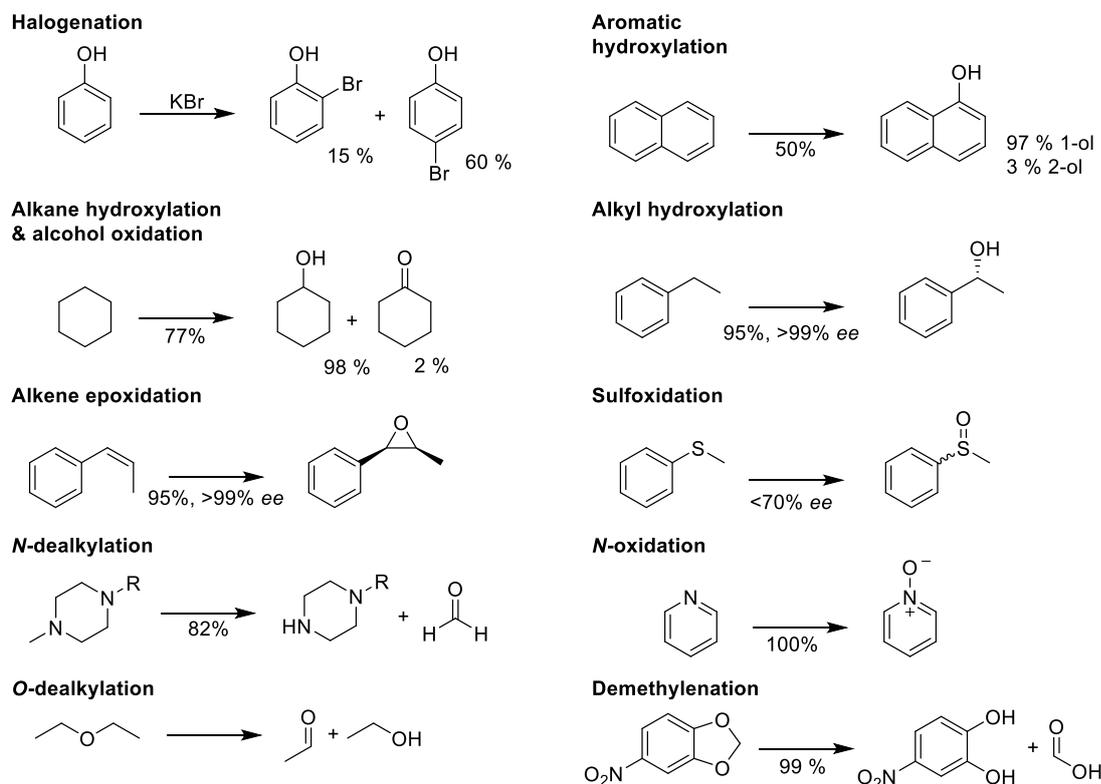


Figure 1-14: Selected reactions catalysed by *AaeUPO* in the presence of hydrogen peroxide.

The physiological role of UPOs has not yet been determined, though they are thought to be involved in the detoxification of soil and degradation of lignin. As secreted enzymes with the ability to perform oxyfunctionalisations, UPOs are able to change the chemical properties of environmental pollutants, like polycyclic aromatic hydrocarbons and dibenzothophene, which allows for the incorporation into humic substances through solubilisation.^{72, 73} The high level of *N*-glycosylation (up to 40%) of the high mannose type stabilises the protein in the extracellular environment.^{70, 74} UPOs contain a prosthetic heme group where the iron coordinates to a cysteine which is part of a conserved -ProCysPro- motif.⁴⁸ In complex with carbon monoxide the UV/Vis spectrum of UPOs shows an absorption maximum between 445 and 450 nm, similar to Soret bands observed in CYPs and *CfuCPO*.^{75, 76}

Together with *CfuCPO*, UPOs are classified in a novel heme peroxidase superfamily, characterised by the catalysed reactions, the cysteine as axial heme ligand, their structure, and protein sequence. The sequence similarities between UPOs and *CfuCPO* are around 40% and mainly stretch over the N-terminus containing the

conserved heme-thiolate region and distal heme pocket.⁷⁴ Within the UPO family sequence similarities of up to 74% can be found and the screening of sequence databases revealed more than a thousand homologous sequences encoding for putative UPO proteins. The potential proteins are widespread across the fungal kingdom with the exception of yeast, and indications were made that they are part of an evolutionary and phylogenetically old superfamily.^{48, 52} In addition to the UPO from *Agrocybe aegerita*, three other homologs have been described in the literature until 2014 and more in patents, including the enzyme from *Coprinellus radians*,⁷⁷ *Marasmius rotula*,⁷⁸ and *Coprinopsis cinerea*.⁷⁹ The biggest challenge up to date is the development of a heterologous expression system for UPOs, as currently access to the enzyme is limited to expression by the native fungus. The production yield in shake flasks and bioreactors varies between a few milligrams (*AaeUPO* and *CraUPO*) to a few hundred milligrams (*MroUPO*). The only successful heterologous expression reported at the start of this work has been described for *CciUPO* which used *Aspergillus oryzae* as production host.⁸⁰

1.2.1. Proposed Catalytic Cycle

The ability of UPOs to catalyse reactions with a peroxygenase (two-electron transfer) or a peroxidase (two one-electron transfers) character first led to the proposal of two different catalytic cycles.⁵² While oxygen insertions were believed to follow the peroxide-shunt pathway described for CYPs (see Figure 1-10), the oxidation of phenols for example was believed to follow the catalytic cycle proposed for peroxidases (see Figure 1-7) describing two one-electron transfer reactions *via* a compound II intermediate. In both cases the reactions proceed *via* a compound 0 intermediate in which hydrogen peroxide replaces the water in the resting state and a spin shift from low-spin to high-spin occurs. Compound 0 rapidly converts into compound I, the oxo-ferryl cation radical complex, which then proceeds *via* a two-electron or two one-electron transfer, yielding the product and the resting protein. The existence of a compound I in *AaeUPO* (*AaeUPO*-I) was first published in 2012, in which a study of the reaction of ferric *AaeUPO* (resting enzyme) with *m*-chloroperoxybenzoic acid (*mCPBA*) was performed using stopped-flow techniques.⁸¹ The monitoring of the UV/Vis spectrum of *AaeUPO* showed a decrease in the Soret-

band at 417 nm and the Q-bands at 538 and 571 nm, which are indicative for the resting, low-spin ferric enzyme, and the appearance of new bands at 361 nm and 694 nm upon the mixing of enzyme with *m*CPBA. The weaker, blue-shifted Soret band and the absorbance at 694 nm strongly suggest the presence of the high-spin *Aae*UPO-I, which is in accordance with findings reported for CYPs and *Cfu*CPO.⁸¹ The same study further included insight into the participation of the active-site glutamate in the proton-transfer during catalysis. Glu196 was first identified as potential acid-base catalytic residue when the crystal structure of *Aae*UPO was solved in 2010.⁸² A representation of the generation of *Aae*UPO-I using *m*CPBA in the presence of glutamate is presented in Figure 1-15, steps **A** to **D**. The same study also postulated that the C-H hydroxylation was the rate-limiting step in this reaction by studying the intermolecular isotope effect (k_H/k_D) using THF and THF- d_8 .

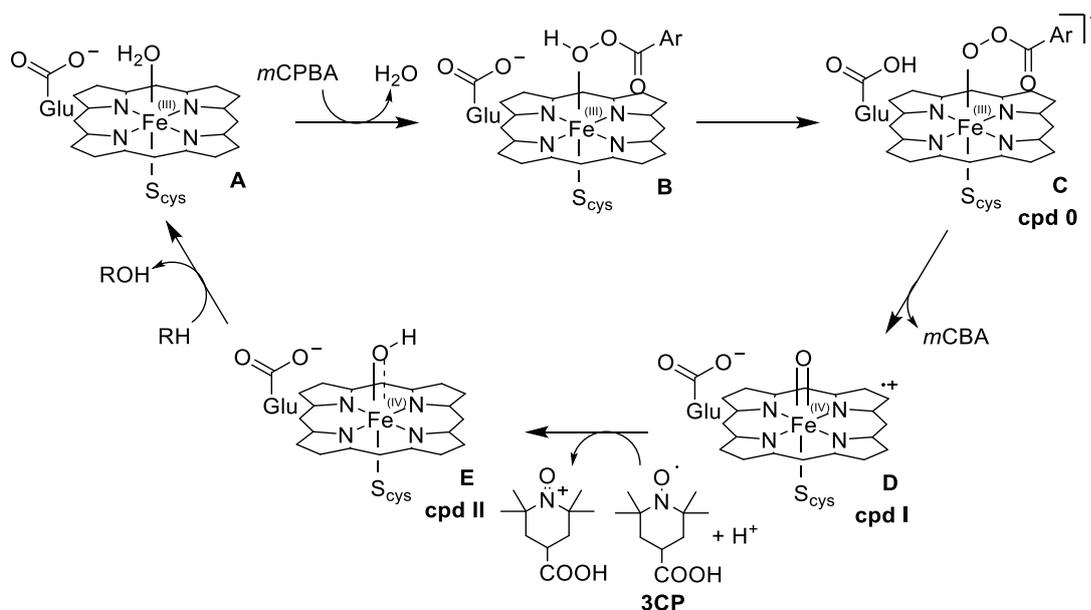


Figure 1-15: Proposed catalytic cycle for reactions catalysed by UPOs. The water molecule binding to the heme in the resting state (**A**) is replaced by *m*CPBA (**B**). The active site glutamate abstracts a proton from *m*CPBA yielding in compound 0 (**C**). Compound I, the oxo-ferryl cation radical complex, is formed by the deprotonation of the glutamate and subsequent loss of *m*CBA and (**D**). Through a one-electron transfer facilitated by 3CP (3-carboxy-PROXYL) compound II (**E**) is formed with a pK_a of 10.0. The insertion of oxygen into the substrate leads back to the resting heme (**A**).

In a more recent study, the same group obtained experimental evidence for an *Aae*UPO compound II intermediate (*Aae*UPO-II, labelled **E** in Figure 1-15) using a similar stopped-flow set up and 3-carboxy-PROXYL (3CP) for the one electron transfer step from *Aae*UPO-I.⁸³ The UV/Vis spectrum of *Aae*UPO-II displays a split Soret band (370 and 428 nm) and two Q bands (535 and 567 nm), which are slightly blue shifted

to those observed for the ferric enzyme. The analysis of experiments performed at different pHs indicates a protonation event at higher pH, which was attributed to the deprotonation of *Aae*UPO-II. This finding strengthens earlier claims, which suggested a basic, protonated *Aae*UPO-II intermediate.^{81, 84} The hydroxyl group in *Aae*UPO-II was found to have a pK_a of 10.0 which lies between the pK_a for *Cfu*CPO-II, estimated ≥ 8.2 , and the pK_a of CYP-II measured to be 11.9. The basic character of the compound II in these heme-thiolate proteins was attributed to the electron-push effect arising from the cysteine ligand.⁸⁵ Horseradish peroxidase or myoglobin, which carry a neutral histidine as axial iron ligand, are estimated to have a $pK_a \leq 4$ suggesting they maintain the ferryl oxo form. The basic character of *Aae*UPO-II is further regarded as the driving force in proton abstractions and might be essential for difficult C-H scissions, as needed in the case of unactivated C-H bonds.⁸³

1.2.2. Crystal Structure

The crystal structure of *Aae*UPO was first reported in 2010.⁸² Crystals of the purified proteins were obtained at different pH under high ammonium sulfate conditions. An in-depth analysis of the two structures was published a few years later.⁷⁰ *Aae*UPO contains 328 amino acids in its mature form. The crystal structures cover the polypeptide starting at the leucine at residue four to the C-terminus at position 328 (see Figure 1-16, **A**). Six different glycosylation sites were identified in the structure, showing up to eight carbohydrates in the high mannose type conformation (see Figure 1-16, **B**).⁸⁶ A disulfide bridge between Cys278 and Cys319 appears to stabilise the C-terminus. Cys36 coordinates to the proximal site of the heme iron at a distance of 2.3 Å. One of the heme propionates is involved in the coordination of a magnesium-ion, which completes its octahedral coordination sphere by complexing the protein backbone and two water molecules. In the distal heme pocket Glu196 and Arg189 are within hydrogen-bonding distance (2.9 Å) of each other and are proposed to form the acid-base pair involved in the formation of *Aae*UPO-I (see Figure 1-16, **C**).⁸¹ The substrate channel is lined with hydrophobic amino acid residues and displays a cone shape with a length of 17 Å and a diameter of 10 Å at the entrance and 8.5 Å at the active site (see Figure 1-16, **D**). As axial ligands, acetate and 5-(hydroxymethyl)imidazole were observed coordinating to the iron.

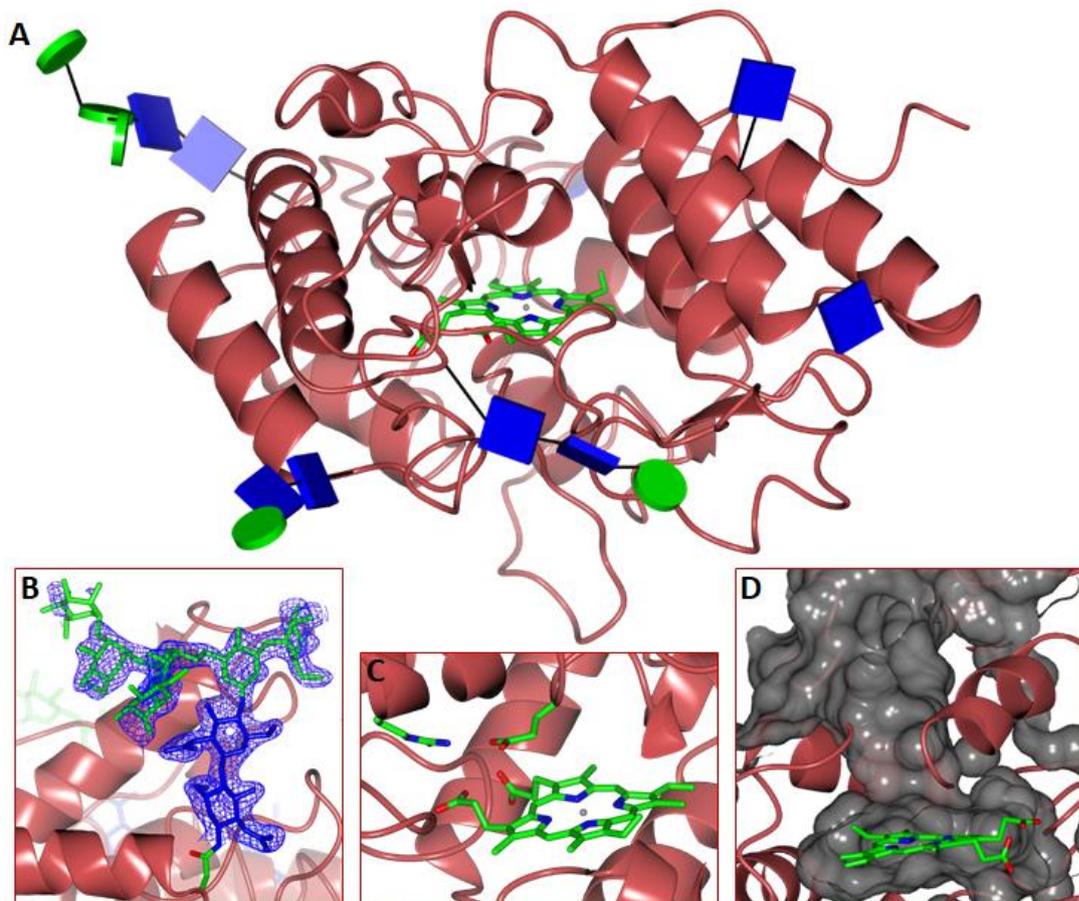


Figure 1-16: Structure of *AaeUPO* (2YOR) with **A** – ribbon structure of the full length protein in crimson, the heme in cylinders (green – carbons, blue – nitrogen, red – oxygen), and carbohydrates as glycoblocks (blue – *N*-acetylglucosamine, green – mannose); **B** – carbohydrate chains linked to Asn141, with the electron density corresponding to the $2F_o - F_c$ map downloaded from the PDB and contoured at 1σ (colour scheme as described for glycoblocks); **C** – close up of the active site highlighting the acid-base pair Glu196 and Arg189; **D** – cone-shaped appearance of the substrate channel giving access to the active site highlighted in grey. Created using CCP4mg.

The structure further allowed a more precise comparison to *CfuCPO*. The structural similarities are mainly restricted to the heme-binding site of the proteins. Both enzymes are postulated to use a distal acid-base pair, which is reported to consist of a glutamic acid and histidine for *CfuCPO*, while in *AaeUPO* the basic part is an arginine residue. Both enzymes also show the coordination of a cation in proximity to the heme, coordinated by one of the heme propionates. The magnesium ion found in *AaeUPO* has been postulated to have a structure stabilising role, similarly to calcium ions in lignin peroxidases.⁸⁷ In *CfuCPO* a manganese ion was proposed be present, though the identity of the ion is still discussed.⁷⁰ The moderate resemblance with distinct changes in some key residues might be able to give more insight into the catalytic workings of these two related enzymes.

Solving the crystal structure of *Aae*UPO also paved the road to structure guided evolution experiments to adjust the regio- and stereoselectivity of the protein. However, attempts to crystallise the protein with a substrate ligand were unsuccessful, leaving docking experiments to help identify residues important to the reaction.⁷⁰ Polycyclic aromatic hydrocarbons (PAH) were modelled into the binding pocket, highlighting interactions with five phenylalanines lining the substrate channel and showing size restrictions. These results were in line with the reported activity of *Aae*UPO towards a range of PAHs.⁸⁸ To date, a substrate complex of UPOs has not been reported.

1.2.3. Biotransformations and Applications

Due to the versatile reaction scope of UPOs their application in biotransformations has been widely studied. The following will give an overview about the different products which were accessed using UPOs over time, and how UPOs compare to other enzymes. Further, the limitations and approaches to overcoming these will be discussed.

1.2.3.1. Halogenations

When UPOs were first discovered in 2004 they were considered haloperoxidases, due to their similarity to the well-studied *Cfu*CPO.⁶⁹ The chlorination and bromination of monochlorodimedone (MCD, see Figure 1-17), a typical *Cfu*CPO substrate, were successful with *Aae*UPO. However, the activities for these reactions were 21- and 14-fold lower than those recorded for *Cfu*CPO using chloride and bromide, respectively. While *Aae*UPO and *Cfu*CPO are both able to halogenate MCD, there are noticeable differences in substrate scope, specificity and pH behaviour. One example for this was published in 2005 studying the halogenation of phenol.⁷⁶ In the presence of potassium bromide *Aae*UPO converts phenol into 2-bromo- and 4-bromophenol in a 1:4 ratio, with 15% and 60% yield, respectively. *Cfu*CPO only formed traces of bromophenols (<3%) under the same conditions. In the presence of potassium chloride *Aae*UPO generated *p*-benzoquinone (39%) as the major product, with traces of 2-chlorophenol (1%) present. *Cfu*CPO, on the other hand, converted phenol into both 2-chloro- and 4-chlorophenol, as well as dichlorophenols. These inverse

activities towards bromide and chloride allow for a clear distinction between *AaeUPO* and *CfuCPO*.



Figure 1-17: Halogenation reactions catalysed by *AaeUPO*; with MCD for monochlorodimedone, and X for chloride or bromide.

CraUPO was the first described in 2007.⁷⁷ Initial studies showed that, similarly to *AaeUPO*, *CraUPO* was also able to halogenate phenol. The bromination pattern showed a similar 1:4 ratio between the 2-bromo- and 4-bromophenol, with yields of 10 and 40%, slightly lower than those reported for *AaeUPO*. Chlorination only produced trace amounts of 2-chlorophenol.

MroUPO was later described in 2011 and was found not to chlorinate or brominate phenol.⁷⁸ The slow oxidation of iodide to triiodide was observed and indicates a weak halogenating activity for *MroUPO*.

1.2.3.2. Oxidation Reactions

Upon its discovery *AaeUPO* was found to catalyse the oxidation of aryl alcohols and aldehydes.⁶⁹ It was found to convert veratryl alcohol (VA) to veratraldehyde and on to veratric acid, and a similar oxidation pattern was observed for benzyl alcohol (see Figure 1-18). The specific activities of *AaeUPO* for benzyl alcohol (234 U mL⁻¹) and benzaldehyde (237 U mL⁻¹) are almost identical and highlight one difference between the UPO and *CfuCPO*, which is able to catalyse the oxidation of benzyl alcohol (18 U mL⁻¹), but does not accept benzaldehyde as substrate. As the conversion of VA to the aldehyde is UV/Vis traceable (change of absorption at 310 nm), VA is used in spectrophotometric assays to determine the specific activity of UPOs and allow quantification.⁸⁹

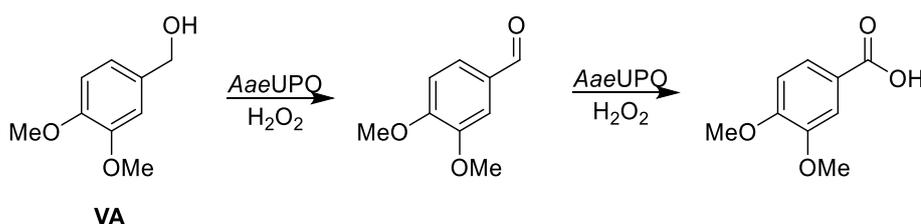


Figure 1-18: Conversion of veratryl alcohol (VA) to veratraldehyde and veratric acid using *AaeUPO*.

1.2.3.3. Aromatic Hydroxylations

In 2005 aromatic hydroxylations catalysed by *Aae*UPO were first described using toluene and naphthalene substrates.⁷⁶ While the reaction with naphthalene produced 1-naphthol as major product (64%) with only traces of 2-naphthol (2%) and 1,4-naphthoquinone (1%) present, the reaction of *Aae*UPO with toluene gave a mixture of products including benzyl alcohol (37%) and its oxidation products (benzaldehyde 12%, benzoic acid 4%), cresol (*para* 2%, *ortho* 4%) and methyl-*p*-benzoquinone (23%, see Figure 1-19). *Cra*UPO and *Mro*UPO were also tested for their aromatic hydroxylation activities using naphthalene.^{77, 78} For *Cra*UPO it was reported that naphthalene was converted at approximately 50%, with a regioselectivity of 60:1 for the 1- and 2-naphthol, an improved selectivity when compared to the results obtained for *Aae*UPO.⁷⁷ *Mro*UPO showed a full conversion of naphthalene with a 12:1 regioselectivity favouring the 1-naphthol.⁷⁸ As a comparison, *Cfu*CPO is not able to catalyse the hydroxylation of aromatic rings. However, paired with spectral data obtained for *Aae*UPO indicating the presence of a heme-thiolate, a similarity to CYPs was suggested. In this context the possibility of UPOs being monooxygenases was examined, but no activity was observed when incubating *Aae*UPO with an electron-donor in the presence of molecular oxygen. Hence, UPOs were further described as a missing link between heme-thiolate haloperoxidases and CYPs.

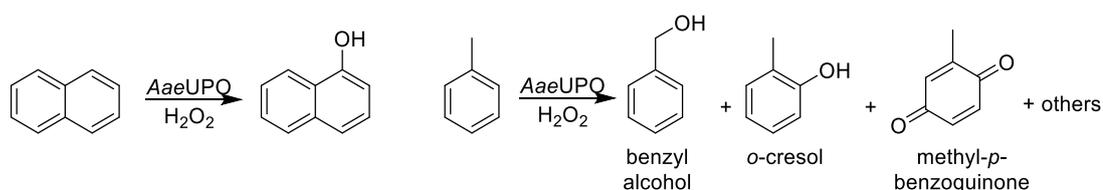


Figure 1-19: Reactions of naphthalene (left) and toluene (right) with *Aae*UPO showing the major products formed.

The conversion of 2-phenoxypropionic acid (POPA) to 2-(4-hydroxyphenoxy) propionic acid (HPOPA), an herbicide precursor, is an early example highlighting the applicability of *Aae*UPO in industrially relevant reactions (see Figure 1-20).⁸⁹ Starting with the racemic POPA, conversions of 43% could be obtained at 60% *ee*, favouring the (*R*)-enantiomer. The publication also discussed the addition of ascorbic acid as radical scavenger, to avoid polymerisation of the phenolic compounds. Other examples for aromatic hydroxylations include the conversions of propranolol (1-naphthalen-1-yloxy-3-(propan-2-ylamino)-propan-2-ol, a beta blocker) and

diclofenac (2-[2-[(2,6-dichlorophenyl)amino]phenyl]acetic acid, an anti-inflammatory drug) to give their respective metabolites 5-hydroxypropranolol (20% yield) and 4-hydroxydiclofenac (65% yield).^{90, 91} The hydroxylation of propranolol has previously been studied using an evolved CYP.⁹² Under comparable conditions *Aae*UPO (0.6 μ M) gave 13.6% conversion to a single product within 2 min, while the CYP (5 μ M) gave 0.5% conversion to four different products in 180 min reaction time.

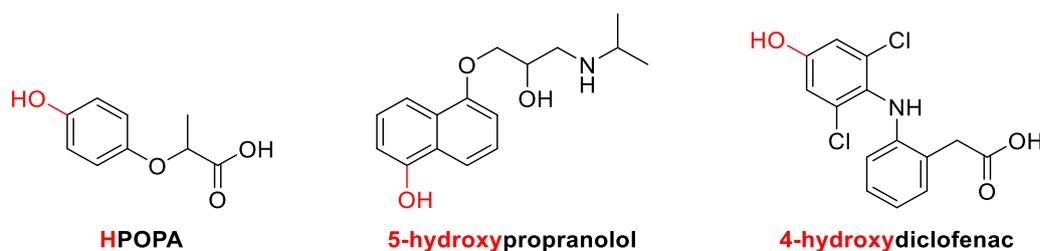


Figure 1-20: Structures of 2-phenoxypropionic acid (POPA), propranolol, and diclofenac with their hydroxylated products in red, as produced by *Aae*UPO.

The ability of UPOs to perform aromatic hydroxylations on polycyclic aromatic compounds was studied using *Aae*UPO and *Cra*UPO.⁸⁸ Both UPOs accepted ring sizes of up to four aromatic units as substrates. Generally, *Cra*UPO was not as active towards the polycyclic aromatic hydrocarbons as *Aae*UPO, and it was shown to prefer the hydroxylation of non-aromatic carbon centres where possible (examples 1- and 2-methylnaphthalene, and fluorine). The regioselectivity of *Cra*UPO was also improved over *Aae*UPO which showed higher ring hydroxylation activity leading to multiple hydroxylations on the ring scaffold, with up to eleven different products formed in the case of dibenzofuran. In a comparison to CYPs, UPOs show higher specific activity towards the tested substrates, and the same is true for CPOs with the exception of pyrene.

Steric hindrance of larger substrates was also observed in a study focusing on the hydroxylation of flavonoids using *Aae*UPO.⁹³ The diglycoside rutin was not accepted as a substrate, potentially due its bulky sugar residue. A range of flavonoids, including quercetin, and daidzein, were selectively hydroxylated in the C6-position with conversions around 94 and 73%, respectively (see Figure 1-21). The analysis of apigenin allowed the comparison to CYPs, which showed a seven-fold reduced catalytic efficiency when compared to *Aae*UPO (with $k_{cat}/K_m = 1.25 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$ and $8.97 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$, respectively).

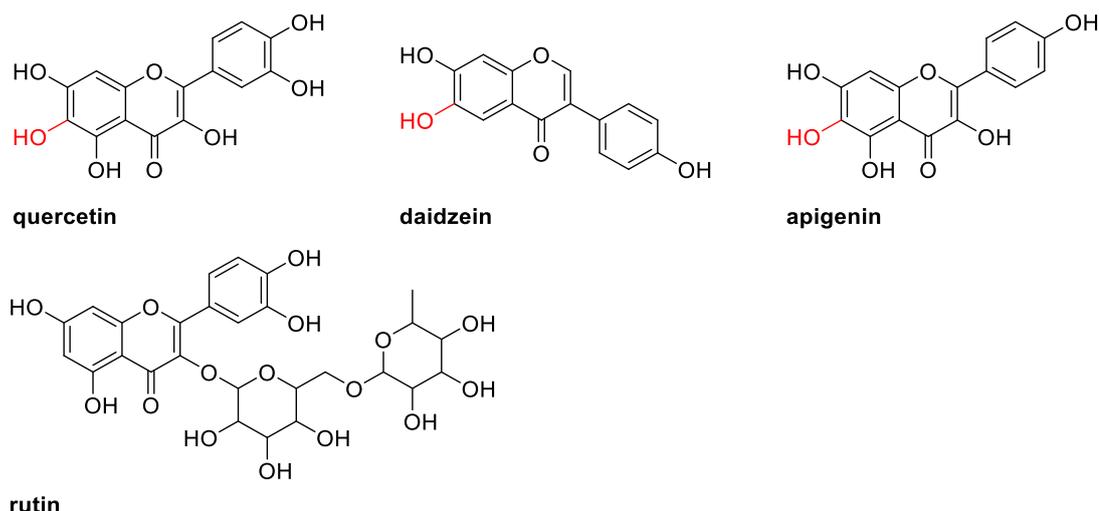


Figure 1-21: Structures of flavonoids which have been applied in reactions with *AaeUPO*, in red the position where hydroxylation occurred, if substrate was converted.

1.2.3.4. Benzylic and Aliphatic Hydroxylations

In addition to aromatic hydroxylation reactions catalysed by UPOs, studies also focused on aliphatic substrates and aromatic systems with sidechains, where non-aromatic hydroxylations could occur. The conversion of ethylbenzene and its derivatives with *AaeUPO*, showed hydroxylations in the alkyl position with good to excellent *ees* (*p*-chloroethylbenzene >99% (*R*)).⁹⁴ Ethylbenzene was found to be converted at 95% with an enantiomeric excess of the (*R*)-1-phenylethanol of >99%.⁹⁵ The loss of enantiomeric selectivity in alkylbenzenes carrying longer chains was often due to the overoxidation of the alcohol to the ketone.

UPOs were also shown to hydroxylate drug compounds like ibuprofen into 2-hydroxyibuprofen (see Figure 1-22).⁹¹ *CraUPO* converted the substrate at 72%, with 75% regioselectivity. *AaeUPO* on the other hand displayed a conversion around 20% with only 24% regioselectivity.

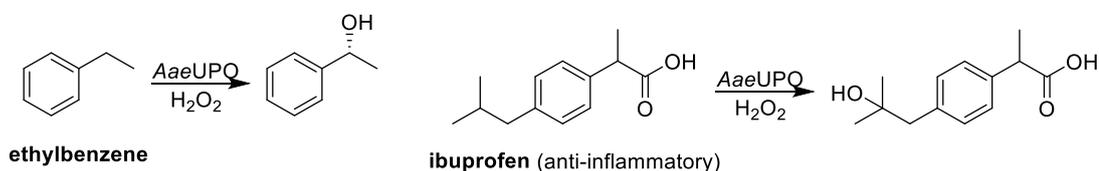


Figure 1-22: Examples for aliphatic hydroxylations using *AaeUPO*.

Comprehensive studies were performed looking at the conversion of aliphatic compounds like fatty acids, and fatty alcohols by UPOs.^{79,96} *AaeUPO* and the homolog from *Coprinopsis cinerea* (*CciUPO*) were applied in reactions. *CciUPO* generally showed higher efficiencies, however different regioselectivities were recorded for

the two homologs. Both enzymes showed conversions of fatty acids to the monohydroxylated products, at the ω -1 or ω -2 position. The overoxidation to the ketones was also observed in non-limiting hydrogen peroxide conditions. *Cci*UPO showed a dependence of preference on the length of the alkyl chain, similar to CYPs, with ω -1 preference for chains $\leq C_{14}$ and ω -2 preference for chains $\geq C_{16}$. *Aae*UPO showed a preference for the ω -1 position independent of the chain length. Myristoleic acid ($C_{14:1}$), an unsaturated fatty acid, was converted by both *Aae*UPO and *Cci*UPO solely to the ω -2 hydroxy product. Oleic acid ($C_{18:1}$) was converted with a preference of ω -1 in case of *Aae*UPO, and with a preference for ω -2 in case of *Cci*UPO.

The hydroxylation of cyclohexane to cyclohexanol and the oxidation of cyclohexanol to cyclohexanone was studied with three different UPO homologs in parallel, *Aae*UPO, *Cci*UPO, and *Mro*UPO.⁹⁷ The hydroxylation reaction of cyclohexane was performed with similar efficiencies (k_{cat}/K_m) around $10^4 \text{ M}^{-1} \text{ s}^{-1}$. In case of the alcohol oxidation step, the efficiency of *Aae*UPO and *Cci*UPO were found in the range of $10^2 \text{ M}^{-1} \text{ s}^{-1}$, while *Mro*UPO showed an efficiency two magnitudes larger. The interest in accessing cyclohexanone comes from a range of industrial applications it is involved in, as for example the synthesis of nylon *via* ϵ -caprolactam.

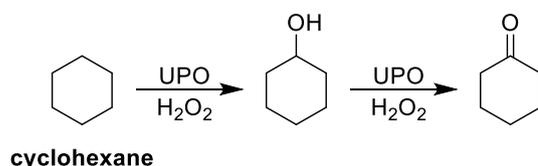


Figure 1-23: Conversion of cyclohexane to cyclohexanol and cyclohexanone using UPOs.

Further aliphatic hydroxylation reactions were performed with pharmaceutically relevant compounds, such as vitamin D₂ and D₃, and testosterone (see Figure 1-24).⁹⁸⁻¹⁰⁰ It was found that different UPO homologs showed changes in regioselectivity and efficiency of converting these complex compounds. *Cci*UPO showed 90-100% conversions for vitamin D₂ and D₃ with 100% regioselectivity for the C₂₅ hydroxylation, for *Aae*UPO and *Mro*UPO the conversion rates were between 45-77% under the same conditions.^{98, 100} The conversion of testosterone was reported for a new UPO homolog from *Chaetomium globosum* (*Cgl*UPO) and no reaction took place using *Aae*UPO or *Mro*UPO.⁹⁹ The reaction yielded in 90% epoxide and 10% hydroxylation product, with isolated purities >96%.

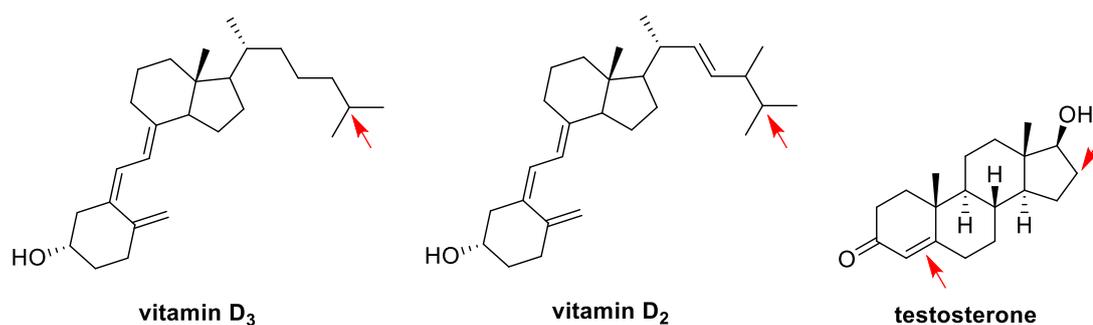


Figure 1-24: Structures of vitamin D₂, D₃, and testosterone, with the hydroxylation/epoxidation sites as attacked by UPOs highlighted in red.

1.2.3.5. N-Oxidations

N-Oxidation reactions using *Aae*UPO were first reported in 2008, using pyridine (PY) and its derivatives as substrates.¹⁰¹ Full conversion was observed for PY and a K_m of 69 μ M and a k_{cat} of 0.21 s^{-1} were determined. The presence of an electronegative nitrogen gives to an electron deficient aromatic system making it less activated and leading to the low k_{cat} , when comparing the values to other *Aae*UPO substrates like 2,6-dimethoxyphenol (DMP) with a k_{cat} of 108 s^{-1} . The relatively low K_m is indicative of a high affinity of *Aae*UPO to PY, potentially originating from the ability of the nitrogen to coordinate to the heme. Not many literature examples are available looking at the *N*-oxidation of unsubstituted PY. The only known examples include methane monooxygenases (MMOs, specific activity of 0.029 U mg^{-1}) and human/animal CYPs which have been used *in vivo* and involved urine analysis to determine conversion rates (specific activities 0.002-0.004 U mg^{-1}).^{102, 103} *Aae*UPOs specific activity for PY was found to be about 100-fold higher than those previously reported for monooxygenases.¹⁰¹

1.2.3.6. Sulfoxidations

The first sulfoxidations were reported in 2009 looking at the ability of *Aae*UPO and *Cra*UPO to convert dibenzothiophene (DBT).¹⁰⁴ Experiments were prepared *in vivo* and *in vitro* showing very different outcomes. *In vivo* *Aae*UPO was able to convert DBT completely within ten days, small amounts of ring-hydroxylations, the sulfone as major product (42%) and the sulfoxide as minor product (15%) were detected (see Figure 1-25). *In vitro* *Aae*UPO produced a range of mono-, di-, tri-, and tetra-hydroxylated products but only traces of the sulfoxidation products were observed. *Cra*UPO was able to produce the sulfoxidation products, with only traces of

hydroxylated ring-products, *in vivo* as well as *in vitro*, however, conversion rates were at 60% and 80%, respectively.

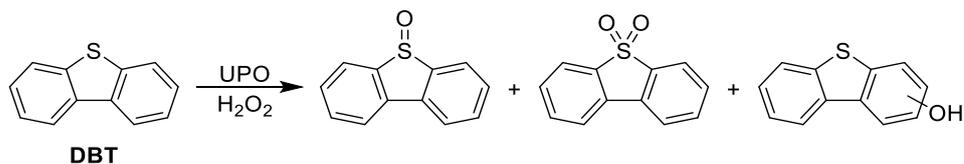


Figure 1-25: Reaction of dibenzothiophene (DBT) to the sulfoxide, sulfone and ring-hydroxylated products in the presence of UPOs.

Besides the conversion of DBT, *AaeUPO* was described to catalyse the sulfoxidation of thioanisol with low turnover frequencies of 35 min⁻¹ and an enantiomeric excess of <70%.⁹⁴

1.2.3.7. Epoxidation

The epoxidation of styrene and its derivatives by *AaeUPO* was first reported in 2011.⁹⁴ Depending on the substitution pattern, the activity and selectivity varied. Terminal alkenes showed acceptable conversion rates, however, with low enantiomeric specificity. The best results were obtained for *cis*- β -methyl styrene with a turnover of 228 min⁻¹ and >99% *ee* (2*S*, 3*R*). Aliphatic alkenes showed epoxidation and the conversion to the allylic alcohol in a 2:1 ratio. A more in depth analysis of epoxidation reactions catalysed by *AaeUPO* in 2012 revealed similar results for styrene derivatives, with conversions ranging from 20-95% and excellent enantiomeric excess for non-terminal alkenes.⁹⁵ The work also included data on the epoxidation of cycloalkenyl benzenes at conversions over 90%, however with poor enantioselectivity (<30%). The reaction with 1,2-dihydronaphthalene in particular was subjected to more studies.¹⁰⁵ It was found that in addition to the formation of the epoxide, in a side reaction *AaeUPO* catalyses the allylic hydroxylation of the compound leading to the aromatisation forming naphthalene and subsequent hydroxylation giving 1- and 2-naphthol.

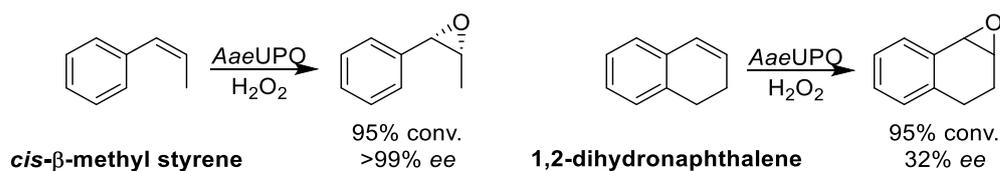


Figure 1-26: Examples for epoxidation reactions catalysed by *AaeUPO*.

The epoxidation of linear, branched and cyclic alkenes was studied in another publication.¹⁰⁶ The efficiency of *AaeUPO* to convert branched and cyclic alkenes was 30

generally higher than those reported for linear molecules, the lowest efficiency was observed with cycloienes. Propene, butadiene and branched alkenes show excellent regioselectivity, while allylic hydroxylation was observed in the presence of a free CH₂ moiety adjacent to the double-bond. In the case of cyclohexene, the ratio between epoxidation and allylic hydroxylation was 1.2:1, for 1-methyl-1-cyclohexene this ratio was with 2.3:1 favouring the epoxidation reaction in a more pronounced way.

1.2.3.8. N- and O-Dealkylation Reactions

The ability of *Aae*UPO to cleave ethers was first reported in 2009.¹⁰⁷ The oxidative cleavage of a range of ethers, including environmentally significant ethers, was described to yield in a carbonyl compound and an alcohol (see Figure 1-27). Solvents like methyl *t*-butyl ether (MTBE) and tetrahydrofuran (THF) reacted to give *t*-butanol and formaldehyde or the ring-opened 4-hydroxybutanal, respectively. Methyl 3,4-dimethoxybenzyl ether was another substrate which was studied and due to its spectrophotometric properties, it allowed for the determination of kinetic parameters. For *Aae*UPO a K_m of 1.4 mM was determined and a k_{cat} of 720 s⁻¹. When comparing this to literature reported values for CYPs, it becomes clear that the K_m values are within the same range, the k_{cat} values for CYPs however are <0.1 s⁻¹, making *Aae*UPO a much more efficient catalyst for this reaction. Studies focussing on larger, macromolecular ethers only reported poor results, suggesting the active site of *Aae*UPO is not able to accommodate these types of substrates.

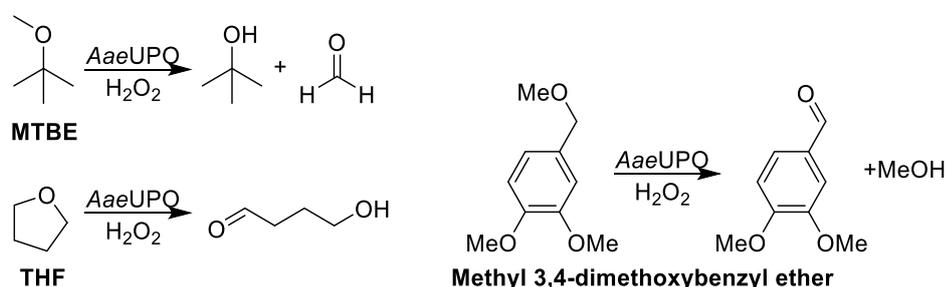


Figure 1-27: Examples for ether cleavages catalysed by *Aae*UPO.

A screening of drug molecules also revealed a range of dealkylation reactions catalysed by UPOs (some examples see Figure 1-28).⁹¹ *O*-Demethylations were observed for metoprolol and naproxen, and a *O*-deethylation was observed in phenacetin. The conversions using *Aae*UPO were around 17, 57, and 23%,

respectively, with regioselectivities as high as 95%. *Cra*UPO catalysed the same reactions at 4, 9, and 13% conversion and regioselectivities up to 85%. Selective *N*-dealkylations were catalysed for sildenafil and lidocaine, amongst others. For *Aae*UPO the best result was an 82% conversion and 99% regioselectivity for sildenafil and 25% conversion (41% regioselectivity) for the mono-deethylation and 18% (30% regioselectivity) for the di-deethylation of lidocaine. *Cra*UPO showed only low conversion for the first compound (4% conversion, 5% regioselectivity), but displayed a clear preference for the mono-deethylated lidocaine with 32% conversion over 5% conversion to the di-deethylation. *Cra*UPO was also able to catalyse an ester cleavage, a special form of *O*-dealkylation. A selective cleavage of oseltamivir gave oseltamivir carboxylate and acetaldehyde with 71% conversion and 88% selectivity.

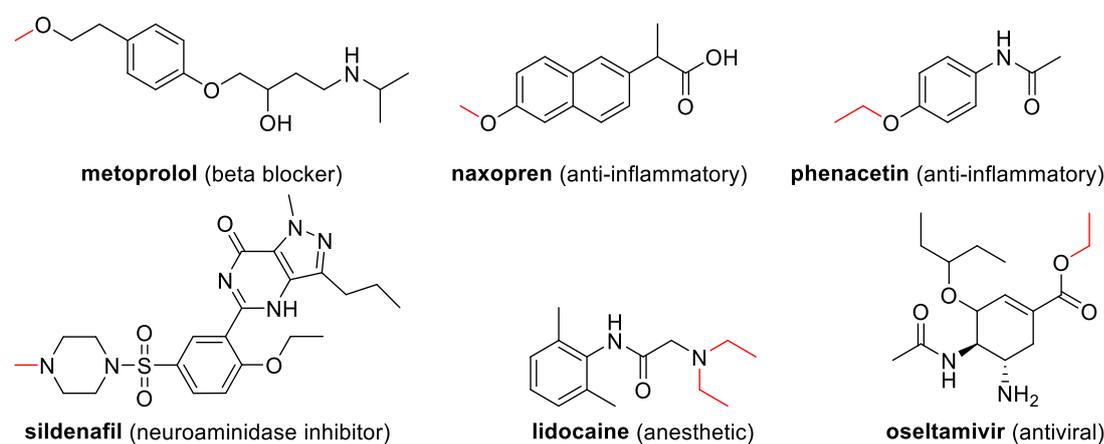
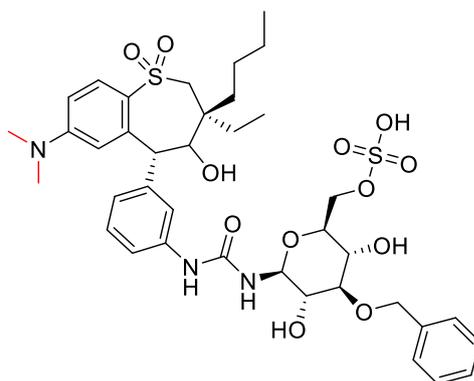


Figure 1-28: Structures of drug molecules dealkylated by *Aae*UPO and/or *Cra*UPO; with the preferred dealkylated chain in red.

The *N*-dealkylation of the human metabolite SAR548304 was studied using *Aae*UPO and *Mro*UPO.¹⁰⁸ It showed that *Aae*UPO was able to dealkylate smaller metabolite precursors, however, only *Mro*UPO showed sufficient activity in the di-demethylation of the compound.



SAR548304 (human metabolite)

Figure 1-29: Structure of the human metabolite SAR548304 which is *N*-demethylated by *Mro*UPO (highlighted in red).

1.2.3.9. One-Electron Transfer Reaction

The ability of UPOs to perform one-electron transfer reactions allowed the application of well-known peroxidase substrates as quantification tools. The activity towards substrates, like 2,2'-Azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) and 2,6-dimethoxyphenol (DMP), was used over the years to compare peroxidase activity of different UPO homologs.^{69, 77, 78, 99} Both compounds can be studied by recording the change of absorption at either 420 nm or 469 nm, respectively. In ABTS, the colour change is a result of the formation of a radical cation on the ring-nitrogen (see Figure 1-30, left, indicated with an arrow), while in DMP a one-electron mediated hydroxylation occurs (see Figure 1-30, right).

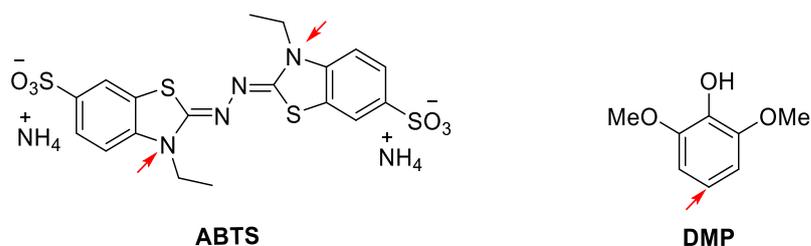


Figure 1-30: Structure of ABTS and DMP, two peroxidase substrates used to quantify UPOs; red arrows highlighting where the reaction occurs.

1.2.3.10. Other Applications

One spectrophotometric assay developed for UPOs takes advantage of the demethylenation of 5-nitro-1,3-benzodioxole (NBD) yielding 4-nitrocatechol (see Figure 1-31).¹⁰⁹ At neutral pH the product displays a yellow colour, while at pH 12 a red colour can be detected. The advantage of this assay over others, like ABTS, is that it is specific for peroxygenase activity. An application of this assay during the

fermentation of UPOs eliminated the possibility of interference originating from native peroxidase, which might be secreted by the organism.

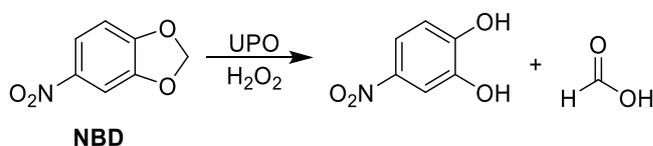


Figure 1-31: Conversion of 5-nitro-1,3-benzodioxole (NBD) using UPOs.

Due to the sensitivity of UPOs to hydrogen peroxide, which made addition in small portions over time necessary, different approaches were taken to generate the co-substrate *in situ*. One example would be the use of glucose oxidase and glucose, which has been described to work with *Aae*UPO.⁶⁹ However, the poor atom efficiency and ethical questions arising from using glucose, meant that this generation system was not applied routinely. A system utilising the photochemical induced reduction of oxygen in the presence of flavin adenine mononucleotide (FMN) and a co-substrate like ethylenediaminetetraacetic acid (EDTA) has also been reported.⁹⁴ It was found that the *in situ* generation of photoexcited FMN was rate-limiting and that higher concentrations of FMN showed no significant increase in hydroxylation rate. Two more recent studies from 2016 focused on the electrochemical generation of hydrogen peroxide, and on an enzymatic system which oxidizes methanol fully to carbon dioxide using a four enzyme system, starting with the commercially available alcohol oxidase from *Pichia pastoris* (*Pp*AOx).^{110, 111}

In addition to the applications of UPO in biotransformations, there have been reports where UPOs were used as biosensors.^{112, 113} The immobilisation of *Mro*UPO and *Aae*UPO at chitosan-capped gold-nanoparticle modified glassy carbon electrodes generated recognition elements capable of responding to a range of metabolic reactors, including aromatic and drug molecules.

1.3. Aims

Unspecific peroxygenases have a great potential for incorporation into industrial processes due to their broad versatility, excellent selectivity, and good stability in the presence of organic solvents. However, in the first instance the access to the catalyst has to be improved, and a quick and simple expression system with high yields needs to be established. Following on from this, putative UPO homologs can be screened for activity and evolution experiments, guided by crystal structures, can help to tailor the reaction selectivity.

Starting from the well-studied *AaeUPO*, this work looks at gene expression in bacteria (*E. coli*) and yeast (*P. pastoris*). In parallel, the homolog from *Agaricus bisporus* (*AbiUPO*) was also subjected to expression trials. Crystallisation trials were performed with the purified and characterised *AaeUPO* to aid structure guided evolution. The application of UPOs in biotransformations was studied. Substrate screens and scale-up reactions were performed to define the UPOs' scope and applicability in industrially-relevant processes.

The aims of this work were to:

- Establish a heterologous expression system for *AaeUPO* using bacterial and yeast hosts
- Test the applicability of the expression route for sequence-related UPOs
- Purify and characterise the obtained protein and compare to the enzyme from the native expression host
- Perform crystallisation trials to advise protein engineering
- Screen and identify novel substrates of industrial interest
- Scope reaction conditions and scalability of biotransformations

2. Bacterial Expression of *AaeUPO* and *AbiUPO*

The increased interest in UPOs in recent years regarding their ability to functionalise unactivated carbon-bonds, highlights the need for a simple heterologous expression system. To date, expression of *AaeUPO* has only been described in the native fungus, a process which takes multiple weeks and results in only low yields, even after process development (9 mg L⁻¹ from a 4 L bioreactor).⁶⁹ A similar result was reported for *CraUPO* (7 mg L⁻¹ obtained in a 0.2 L shake-flask experiment).⁷⁷ These time-consuming, low-yielding processes are not industrially viable and often hard to incorporate into research laboratories. With regard to tailoring enzymes for specific purposes, the use of native strains often does not allow for gene manipulation and targeted expression.

Bacterial expression systems allow for simple genetic manipulation, fast expression, and good yields. Bacterial systems are simple and do not involve post-translational modifications of protein products, and the choice of expression strain can influence for example the folding of the protein, through aiding the formation of disulfide-bonds. There are various commercially available *Escherichia coli* (*E. coli*) expression strains, that could be used for expression optimisation. The heterologous expression of enzymes from a wide range of organisms using bacteria has been described in the literature. Even enzymes carrying prosthetic groups like heme have been shown to express in good yields.¹¹⁴

This chapter is an investigation into the possibility of expressing *AaeUPO* and *AbiUPO* using *E. coli*. For this purpose, different strains of *E. coli* were studied, one of which mimics the proteins' native secretion process.

2.1. Materials and Methods

2.1.1. Chemicals and Materials

2.1.1.1. Materials

Chemicals used in this study were purchased from Alfa Aesar (Heysham, UK), Fisher Scientific UK Ltd. (Loughborough, UK), Fluorochem Ltd. (Glossop, UK), Geneflow Ltd. (Lichfield, UK), Generon (Slough, UK), Insight Biotechnology Ltd. (Wembley, UK), Merck Chemicals Ltd. (Nottingham, UK), Scientific Laboratory Supplies Ltd. (Nottingham, UK), Sigma-Aldrich Company Ltd. (Dorset, UK), Takara Bio Europe Clontech (St Germain-en-Laye, France), Tokyo Chemical Industry UK Ltd. (Oxford, UK), and VWR International Ltd. (Lutterworth, UK). Restriction enzymes were bought from New England Biolabs (Ipswich, UK), Promega UK Ltd. (Southampton, UK), and Thermo Fisher Scientific Biosciences GmbH (St. Leon-Rot, Germany). PCR primers were synthesised by Eurofins Scientific (Wolverhampton, UK) and genes by GeneArt (now Life Technologies Ltd., Paisley, UK).

2.1.2. Molecular Biology

For bacterial growth different media, Lysogeny broth (LB), super optimal broth with catabolite suppression (SOC) and minimal medium M9, were prepared according to the information presented below (Table 2-1).

Table 2-1: Composition of the different bacterial growth media used.

LB	SOC
5 g L ⁻¹ Yeast Extract	5 g L ⁻¹ Yeast Extract
10 g L ⁻¹ Tryptone	20 g L ⁻¹ Tryptone
10 g L ⁻¹ NaCl	0.5 g L ⁻¹ NaCl
	0.19 g L ⁻¹ KCl
	4.8 g L ⁻¹ MgSO ₄
	3.6 g L ⁻¹ Glucose
10X M9 Salts	M9 Medium
60 g L ⁻¹ Na ₂ HPO ₄	1X M9 salts
30 g L ⁻¹ KH ₂ PO ₄	0.4% C-source
5 g L ⁻¹ NaCl	2 mM MgSO ₄
10 g L ⁻¹ NH ₄ Cl	100 μM CaCl ₂

All solutions were autoclaved before use; glucose and glycerol in case of the SOC and M9 medium were added after autoclaving *via* sterile filtration.

2.1.2.1. Bacterial Transformation

All work was carried out using aseptic techniques. 1 μL plasmid was added to 25 μL competent cells and was kept on ice for 30 min. The mixture was then heated to 42 $^{\circ}\text{C}$ for 45 s and placed on ice for further 2 min. 1 mL of LB medium was added to all bacterial expression strains (BL21-D3, Rosetta2, Origami2) and 125 μL SOC medium was added to NovaBlue or Stellar cells before incubation for 1 h at 37 $^{\circ}\text{C}$, 180 rpm. 100 μL of the cells were plated out on LB-agar plates (LB recipe plus 16 g L^{-1} agar) containing the appropriate antibiotics and incubated overnight at 37 $^{\circ}\text{C}$ (see Table 2-2). After this the plates were stored at 4 $^{\circ}\text{C}$ and sealed with parafilm.

Table 2-2: Overview of bacterial strains used and their respective antibiotic resistance.

Bacterial Strain	Antibiotic resistance
NovaBlue Singles	According to plasmid
Stellar	According to plasmid
BL21(D3)	According to plasmid
Rosetta2 pLysS	Chloramphenicol (34 $\mu\text{g mL}^{-1}$), plus according to plasmid
Origami2 pLysS	According to plasmid

2.1.2.2. Starter Cultures

Under aseptic conditions 5 mL LB cultures were prepared in 50 mL falcon tubes. The appropriate antibiotics were added to the medium and a single colony was transferred from an LB-agar plate using a pipette tip. The tube was sealed and incubated overnight at 37 $^{\circ}\text{C}$ in an orbital shaker at 180 rpm.

2.1.2.3. Preparation of Glycerol Stocks

In a 1.5 mL screw-top vial 700 μL starter culture was mixed with 300 μL 80% glycerol (final glycerol concentration 25%) using aseptic techniques. Glycerol stocks were prepared in duplicates and stored at -80 $^{\circ}\text{C}$.

2.1.2.4. Agarose Gel Analysis and Extractions

For the analysis of DNA samples using agarose gel electrophoresis TAE buffer is prepared using deionised water (see Table 2-3).

Table 2-3: Composition of Tris base, acetic acid, EDTA (TAE) buffer used for agarose gel electrophoresis experiments.

TAE buffer	
40 mM	Tris base
20 mM	Acetic acid
1 mM	Ethylenediaminetetraacetic acid (EDTA)

Agarose gels (1%) were prepared by heating agarose in TAE buffer to dissolve the solid. 1 μ L per 100 mL SybrSafe was added to the cooled solution, before casting the gel. Once polymerised, the gel was transferred into gel running chamber and covered with TAE buffer. The DNA samples were prepared using 6X DNA dye, and 10 kb DNA ladder was used for referencing. The gels were run at 110 V for 60 – 90 min to isolate the bands, which were then excised and the DNA was extracted using a GenElute™ Gel Extraction Kit (Sigma) with pre-heated water (60 °C) as elution medium. The concentration of the DNA was determined using a UV-spectrophotometer; water was used as reference.

2.1.2.5. DNA Extraction from Culture

To extract DNA samples from cultures, 5 mL starter cultures were grown according to the protocol described. The DNA was extracted using the GenElute™ Plasmid Miniprep Kit (Sigma). The elution was performed by adding pre-heated water (60 °C) to the immobilised sample and incubation at room temperature for 15 min before the final centrifugation step. The concentration was determined using a UV-spectrophotometer with water as reference.

2.1.3. Gene Design and Synthesis

2.1.3.1. APO1 from *Agrocybe aegerita* (B9W4V6)

For bacterial expression, the *AaeUPO* gene sequence was codon optimized for *E. coli* and ordered in full length. Amplification of the gene was performed using PCR

methods with the following primers. The primers were designed to amplify only the mature protein, i.e. not including the signalling and pro-peptide.

Table 2-4: Gene sequence optimised for *E. coli* and list of primers used for amplification of *AaeUPO* for the different cloning vectors.

AaeUPO codon optimised for *E. coli*

ATGAAATACTTTCCGCTGTTTCCGACCCTGGTTTTTGCAGCACGTGTTGTTGCATTTCCGG
CATATGCAAGCCTGGCAGGTCTGAGCCAGCAAGAAGCTGGATGCAATTATCCGACACTG
GAAGCACGTGAACCGGGTCTGCCTCCGGGTCCGCTGGAAAATTCAAGCGCAAAACTGG
TTAATGATGAAGCACATCCGTGGAACCGCTGCGTCCGGGTGATATTCGTGGTCCGTGT
CCGGGTCTGAATACCCTGGCAAGCCATGGTTATCTGCCTCGTAATGGTGTGCCACACC
GGTTCAGATTATTAACGCAGTTCAAGAAGGCCTGAACTTTGATAATCAGGCAGCAGTTT
TTGCAACCTATGCAGCACATCTGGTTGATGGCAATCTGATTACCGATCTGCTGAGCATTG
GTCGTAACCCGTCTGACCGGTCCGGATCCGCCTCCGCCTGCAAGCGTTGGTGGTCTG
AATGAACATGGCACCTTTGAAGGTGATGCAAGCATGACCCGTGGTGTGATGCATTTTTGG
TAACAACCACGATTTTAACGAAACCCTGTTTGAACAGCTGGTTCGATTATTCAAATCGTTT
TGGTGGTGGCAAATACAATCTGACCGTTGCAGGCGAACTGCGCTTTAACGTATTCAGG
ATAGCATTGCAACCAACCCGAATTTAGCTTTGTGGATTTTCGTTTCTTTACCGCCTATGG
TGAAACCACCTTTCCGGCAAACCTGTTTGTGATGGTCGTGATGATGGTCAGCTGG
ATATGGATGCAGCCCGTAGCTTTTTTTCAGTTTAGCCGTATGCCGGATGATTTTTTTCGTG
CACCGAGTCCGCGTAGCGGCACCGGTGTTGAAGTTGTTATTTCAGGCACATCCGATGCAG
CCTGGTTCGTAATGTTGGTAAAATCAATAGCTATACCGTTGATCCGACCAGCAGCAGTTTT
AGCACCCCGTGTCTGATGTATGAAAAATTTGTGAACATCACCGTGAAAAGCCTGTATCC
GAATCCGACCGTTCAGCTGCGTAAAGCACTGAATACCAATCTGGATTTTTTCTTTCAGGG
TGTTGCAGCAGGTTGTACCCAGGTTTTTCCGTATGGTCGTGAT

Primers for cloning into pETYSBLIC-3C

Forward	CCAGGGACCAGCAATGGAACCGGGTCTGCCTCCGGT
Reverse	GAGGAGAAGGCGCGTTAATCACGACCATACGGAAAAACCTGGGTACAAC

Primers for cloning into pET22b

Forward	TAAGAAGGAGATATACATATGGAACCGGGTCTGCCTCCG
Reverse	GTGGTGGTGGTGGTGTCTGAGATCACGACCATACGGAAAAACCTG

Primers for cloning into pET3FPP31

Forward	CAGCCGGCGATGGCCGAACCGGGTCTGCCTCCG
Reverse	GAACAGAACCTCGAGATCACGACCATACGGAAAAACCTG

2.1.3.2. *Unspecific Peroxygenase from Agaricus bisporus (K5XIK0)*

The gene sequence of the *AbiUPO* homolog was also codon optimized for *E. coli* and ordered through GeneArt. Only the mature protein was amplified using the primers listed below (see Table 2-5).

Table 2-5: Gene sequence optimised for *E. coli* and list of primers used for amplification of *AbiUPO* for the different cloning vectors.

AbiUPO codon optimised for *E. coli*

```
ATGTTTAGCCTGCTGAATTTTGTACCCCTGGGTCTGGCATGTACCTGGTCAGTTCTGGCA
TTTCCGAGCACCTATACCAGCCTGGGTGGTCTGCCTCGTGAAGAACTGGATCGTATTCT
GCCGAGCCTGCAGTATCGTAGTCCGGGTGCACCGCCTGGTCCGCTGAAATTCAATGGCA
CCAACTGGTTAATGATGATCAGCATCCGTGGAAACCGCTGAAACATGGTGATATGCGT
GGTCCGTGTCCGGGTCTGAATACCCTGGCAAGCCATGGTTATCTGCCACGTAATGGTAT
TGCAACACCGGTTCAAGATTATAATGCAGTTCAAGAGGGCTTTAACATGGAAAATTGAG
TTGCACGTCTGGTTACCTATGCAGCACATCTGGTTGATGGCAATCTGATTACCGATAAAC
TGAGCATTGGTGGTAAAAGTCCGCTGACCGGTCCGAGCCCTCCGGCACCGGCAAATGC
AGCAGGTCTGAACACCCATGCACTGTTTGAAGGTGATGTTAGCATGACCCGTGCAGATG
CATTTTTTGGTGATAATCACAGCTTTAACGAAACCCTGTTTGTGATGAATTTACCGCCTTTAG
CAATCAGTTTGGCGCAGGCAAATATAACCTGACCGTTGCAGCAGAATATCGCTTTCATC
GTATTCAAGAAAGCATTGCCACCAATCCGAATTTTAGCTTTGTTAGCCCTCGTTTTTTTAC
CGCATATGCCGAAAGCGTTTTTCCGATCAACTTTTTTATCGATGGTCGTCAGGGTGATGG
TCAGCTGGATCTGGATGTTGCACGTGGTTTTTTTTCAGAATATGCGTATGCCGGATGGTTT
TCATCGTGCAAGCATGCCGACCGGTCTGGAAGGTCTGGCAGAAATTGCAAGCGTTCATC
CGATTTACCGGGTGCAAATGTTAATGGTGTAAATACCTATACCTTTGATCCGAGCAGCG
CAGATTTTACCACATTTTGTCTGCTGTATGTGAACTTTGTTAATCAGACCGTTCGTAGCCT
GTATCCGGAACCGACCGGTAACCTGAAAAAGCACTGAAAAAAACCTGGAATTCCTGT
ATGGTCCGTTTAGCGATCAGTGTAGCCAGGTGTTTCCGATGGTAAAGATAAC
```

Primers for cloning into pETYSBLIC-3C

Forward	CCAGGGACCAGCAATGAGTCCGGGTGCACCGCCTGGT
Reverse	GAGGAGAAGGCGCGTTAGTTATCTTTACCATACGGAAACACCTGGCTACA

Primers for cloning into pET22b

Forward	TAAGAAGGAGATATACATATGAGTCCGGGTGCACCGCC
Reverse	GTGGTGGTGGTGGTCTCGAGGTTATCTTTACCATACGGAAACAC

Primers for cloning into pETFPP31

Forward	CAGCCGGCGATGGCCAGTCCGGGTGCACCGCC
Reverse	GAACAGAACCTCGAGGTTATCTTTACCATACGGAAACAC

2.1.4. Vector Design and Production

2.1.4.1. *pETYSBLIC-3C*

Starter cultures were prepared using glycerol stocks of NovaBlue cells transformed with *pETYSBLIC-3C*. The DNA was extracted from the culture using as described before (see Section 2.1.2.5). For linearization 50 µg vector DNA was incubated with 50 µL *Bse*RI, 100 µL NEB2 buffer (10X) and water (final volume 1 mL) and incubated at 37 °C for 110 min. The linearized vector was isolated on a 1% agarose gel and extracted from the gel (GenElute™ Gel Extraction Kit, Sigma).

2.1.4.2. *pET22b*

NovaBlue cells were transformed with the target vector and overnight cultures were grown as described before (see 2.1.2.2). The DNA was extracted and 2 µg of isolated vector DNA was incubated with 2 µL *Nde*I, 2 µL *Xho*I, 5 µL CutSmart Buffer (10X) and water (final volume 50 µL) at 37 °C for 3 h. The now linearized vector was isolated on a 1% agarose gel and extracted from the gel using the GenElute™ Plasmid Miniprep Kit.

2.1.4.3. *pETFPP31*

The *pETFPP31* vector was amplified and linearized using inverse PCR starting from the *pETFPP30* vector. The amplicon was *Dpn*I treated to digest methylated DNA, before the DNA was isolated and extracted from a gel.

Table 2-6: Primers used for inverse PCR amplification of *pETFPP31* from *pETFPP30*.

<i>pETFPP31</i>	
Forward	CTCGAGGTTCTGTTCCAGGGAC
Reverse	GGCCATCGCCGGCTGGGCAG

2.1.5. Cloning and Analysis

2.1.5.1. *PCR Amplification*

Genes were amplified using the following procedure. All components were pipetted on ice in the given order. Before KOD polymerase was added, the sample was centrifuged (1 min, 16,300 × g). After amplification all constructs were isolated on 1% agarose gels, excised and the DNA was isolated (GenElute™ Gel Extraction Kit, Sigma).

Table 2-7: left: Components and amounts used in PCR amplification reactions. **right:** Temperature, time, and cycle details describing the PCR amplification program.

ddH ₂ O	33 μ L					
KOD buffer (10X)	5 μ L					
dNTPs (2 mM)	5 μ L	94 $^{\circ}$ C	2 min	Initial denature		
MgSO ₄	3 μ L	94 $^{\circ}$ C	30 s	Denature	Repeat 35 x	
Template	1 μ L	50 $^{\circ}$ C	30 s	Anneal		
Forward primer (20 μ M)	1 μ L	72 $^{\circ}$ C	30 s	Extend		
Reverse primer (20 μ M)	1 μ L	72 $^{\circ}$ C	3 min	Final extension		
KOD Polymerase	1 μ L	10 $^{\circ}$ C	∞	cool		

2.1.5.2. Ligation-Independent Cloning (LIC)

The PCR amplified insert and the linearized vector were digested using T4 DNA polymerase to prepare them for ligation-independent cloning. The following components were pipetted together on ice and brought to a final volume of 20 μ L. The polymerase was added last.

Table 2-8: left: Pipetting scheme for the T4 DNA polymerase digest; **right:** Temperature, time and cycle details for the T4 digest.

Insert/linearized vector	0.2 pmol				
T4 buffer (10X)	2 μ L				
dATP (25 mM)	2 μ L				
DTT	0.5 μ L	20 $^{\circ}$ C	30 min	Digest	
T4 DNA polymerase	0.5 μ L	75 $^{\circ}$ C	20 min	Inactivation	
ddH ₂ O	X μ L (to 20 μ L)	10 $^{\circ}$ C	∞	cool	

For the annealing reaction, 3 μ L digested insert were mixed with 3 μ L digested vector and incubated at room temperature for 10 min. 0.5 μ L EDTA (25 mM) were added and the mixture was incubated for further 10 min. 25 μ L of NovaBlue cells were transformed with 1 μ L of the annealed plasmid and grown at 37 $^{\circ}$ C overnight on LB-agar plates containing the appropriate antibiotics.

2.1.5.3. In-Fusion Cloning

The gene of interest was PCR amplified using the In-Fusion primers. The expression vector was linearized according to the procedures mentioned, carrying sticky ends complementary to the ones of the insert. The annealing of the insert and vector was performed as outlined below (see Table 2-9). Stellar cells were transformed with the

annealed plasmid. 25 μL of cells were mixed with 2.5 μL In-Fusion plasmid and colonies were grown on LB-agar plates at 37 °C overnight.

Table 2-9: left: Pipetting scheme for the In-Fusion cloning; right: Temperature, time and cycle details for the reaction.

5X In-Fusion Cloning Mix	1 μL			
Vector	75 ng			
Insert	60 ng			
ddH ₂ O	To 5 μL	50 °C	15 min	Annealing

2.1.5.4. Restriction Enzyme Digest

Restriction enzyme digests were performed to check the efficiency of the annealing reaction. Multiple colonies from the transformed plasmid were picked and starter cultures were prepared. The DNA from these cultures was extracted. Depending on the vector used, different restriction enzymes were used in combination with the appropriate buffer (Table 2-10). The total volume was 10 μL ; consisting of 7 μL annealed sample, 1 μL buffer and 1 μL of each restriction enzyme. The sample was routinely incubated at 37 °C for up to 3 h. The enzymes were not inactivated prior to analysis on 1% agarose gels.

Table 2-10: Overview of the restrictions enzymes and buffers required for enzymatic digests for the different vectors used.

pETYSBL-LIC3C		pET22b		pETFPP31	
Restriction enzymes	buffer	Restriction enzymes	buffer	Restriction enzymes	buffer
<i>NdeI</i>		<i>NdeI</i>		<i>NdeI</i>	
<i>NcoI</i>	3.1	<i>XhoI</i>	3.1	<i>XhoI</i>	3.1

2.1.5.5. Colony-PCR

Colony-PCR, just like restriction enzyme digests, is a method to analyse the efficiency of the annealing of vector and target gene. From the LB-agar plates of the annealed and transformed plasmid, multiple colonies were screened. A colony was picked and transferred onto a new LB-agar plate/quadrant of a plate before washing the tip in the colony PCR mix (see Table 2-11). The PCR program was run according to the information provided to amplify the target gene and the samples were analysed on 1% agarose gels. The newly infected plates were grown at 37 °C. Colonies which gave

a positive amplification result were then used to prepare a started culture. The DNA from the culture was extracted using a plasmid miniprep kit.

Table 2-11: **left:** Components and amounts used in the colony-PCR amplification; **right:** Temperature, time, and cycle details describing the PCR program.

ddH ₂ O	13.8 μL	94 °C	2 min	Initial denature	Repeat 25 x
dNTPs (2 mM)	2 μL	94 °C	30 s	Denature	
Taq buffer (10X)	2 μL	50 °C	30 s	Anneal	
Forward primer (20 μM)	1 μL	72 °C	40 s	Extend	
Reverse primer (20 μM)	1 μL	72 °C	5 min	Final extension	
Taq Polymerase	0.2 μL	10 °C	∞	cool	

2.1.5.6. DNA Sequencing

After checking the efficiency of the annealing, the gene was sequenced to ensure the target sequence has not been altered. 20 μL of the extracted DNA were sent to GATC for sequencing; using the T7 forward and pET-RP reverse primers provided.

2.1.6. Bacterial Growth Curves

Starter cultures from glycerol stocks of all expression cell lines (BL21, Rosetta2 and Origami2) were grown in LB. These cultures were used to a) inoculate fresh 5 mL LB cultures to monitor the growth at 37 °C or b) to inoculate 5 mL M9 starter cultures, which were then used to inoculate the M9-growth cultures the following day. The growth cultures were inoculated to a starting OD₆₀₀ of 0.1 and the cell density (OD₆₀₀) was monitored over 24 h when grown at 37 °C. The growth cultures were set up in duplicates. The growth rate, μ , and the doubling time, t_D , were calculated using Equation 2-1 and Equation 2-2. The time interval chosen for these calculations represented the bacterial exponential growth phase.

Equation 2-1

$$\mu = \frac{\ln OD_{600}(2) - \ln OD_{600}(1)}{t(2) - t(1)}$$

Equation 2-2

$$t_D = \frac{\mu}{\ln 2}$$

2.1.7. Protein Production and Analysis

In order to find the optimum conditions for protein expression a range of experiments were performed. Test-expressions were performed in different media and the protein production was monitored over time. In addition, a solubility screen using thirty buffers was performed.

2.1.7.1. Test-Expression in LB Medium

Starter cultures were prepared from glycerol stocks as described elsewhere (see 2.1.2.2). Six 5 mL LB cultures with the appropriate antibiotic were inoculated with starter culture (1:100) and grown at 37 °C to an OD₆₀₀ = 0.6-0.8. Three cultures were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated at 16, 30 and 37 °C; the other three cultures were used as controls at the given temperatures. After shaking over night the cell pellet was collected by centrifugation (15 min, 3,900 × g) and resuspended in 0.5 mL resuspension buffer (see Table 2-12). The suspensions were sonicated individually whilst being kept on ice. The pulse cycle was run three times choosing a 30 s on - 30 s off rhythm. Multiple samples were sonicated in bulk using a 24-well sonicator. A 1 s on - 1 s off pulse was run three times over a period of 30 s. The soluble and insoluble fractions were obtained upon centrifugation (5 min, 16,300 × g). The insoluble pellet was resuspended in 0.5 mL resuspension buffer. All samples were run on 12% gels and analysed using immunoblots.

Table 2-12: Composition of buffer used for the resuspension of bacterial cell pellets. The buffer was made up with ultra-pure water and filter over 0.2 μm filter paper before use.

Resuspension Buffer	
100 mM	Tris-Cl
200 mM	NaCl
pH	8.2

2.1.7.2. Test-Expression in M9-Minimal Medium

Starter cultures were grown from glycerol stocks in LB and used to inoculate 5 mL M9 starter cultures. From these M9 starter cultures expression tests were performed as described for the LB samples. Induction was performed at OD₆₀₀ = 0.6-0.8 using 1 mM IPTG along with 0.5 μM FeCl₃ and 0.1 μM 5-aminolevulinic acid (ALA).

2.1.7.3. Bacterial Expression over Time

In addition to the expression tests, the expression levels were monitored over time. A 50 mL culture was inoculated with starter culture to an OD₆₀₀ = 0.1 and grown at 37 °C to an OD₆₀₀ = 0.6-0.8 before induction with IPTG or IPTG, FeCl₃ and ALA, depending on the growth medium. After induction and before incubation at the

chosen temperature a sample was taken ($t = 0$ h). Further time points were taken over the course of 24 h. To allow a quantitative analysis of the samples, the weight of the pellets was kept constant, by measuring the OD_{600} and calculating the volume needed to obtain a density of 1.5. The cell pellet was collected *via* centrifugation (2 min, $16,300 \times g$) and resuspended in 0.5 mL resuspension buffer. The suspension was sonicated in a 24-well sonicator (3X 30 s; 1 s on, 1 s off) and the soluble and insoluble fractions were obtained upon centrifugation (5 min, $16,300 \times g$). Insoluble pellets were resuspended in 0.5 mL buffer and analysed alongside the soluble fraction using immunoblots.

2.1.7.4. Scale-up Gene Expression

Scale-up expression was routinely performed in 500 mL growth medium, inoculated with starter culture of the same medium (1:100). The cells were grown to a density between 0.6 and 0.8 and induced with 1 mM IPTG, or 1 mM IPTG, 0.5 mM $FeCl_3$ and 0.5 mM ALA, depending on the medium. Expression at a chosen temperature was performed overnight (20-22 h) unless stated otherwise. The cell pellet was collected upon centrifugation (20 min, $5,000 \times g$) and resuspended in resuspension buffer (5 mL g^{-1} pellet). The cells were disrupted using a cell disrupter at 25 kpsi and cell lysate was obtained upon centrifugation ($50,000 \times g$, 30 min).

2.1.7.5. Periplasmic Extraction

For scale-up experiments of the secreted UPO, the protein was extracted from the periplasm using an osmotic shock procedure. The cell pellet was resuspended over 20 min at $4^\circ C$ in a sucrose buffer (50 mM Tris-Cl pH 7.4, 1 mM EDTA, 20% sucrose; 30 mL g^{-1} pellet). After resuspension the suspension was centrifuged (15 min, $8,000 \times g$), the supernatant was collected and the pellet was resuspended slowly at $4^\circ C$ in a low salt buffer (10 mM Tris-Cl pH 8.2, 5 mM $MgCl_2$, 10 mM NaCl, 1% glycerol). The sample was centrifuged (20 min, $30,000 \times g$) and the supernatant was collected. Samples were dialysed into a Tris-based buffer before conducting further experiments (100 mM Tris-Cl pH 8.2, 200 mM NaCl).

2.1.7.6. Solubility Tests

The solubility of the protein in different buffers was studied following the procedure described by G. Lindwall *et al.*¹¹⁵ Protein was expressed in 500 mL cultures and harvested as described above. The cell pellet was resuspended in buffer giving a final volume of 30 mL (10 mM Tris pH 8.5, 100 mM NaCl, 1 mM EDTA) and the solution was aliquoted into 30 x 1 mL fractions. The pellets were collected again (2 min, 16,300 × g) and each pellet was resuspended in 1 mL buffer from Table 2-13. The suspensions were sonicated on ice (3X 30 s on, 30 s off) before separation of the soluble and insoluble fraction (2 min, 16,300 × g). The insoluble pellets were resuspended in 0.5 mL of the respective buffer. All samples were analysed using immunoblots.

Table 2-13: List and composition of buffers used for the solubility test.

#	Buffer Conditions
1	100 mM HEPES, 1 M MgSO ₄ , pH 7.0
2	100 mM HEPES, 100 mM sodium glutamate, 5 mM DTT, pH 7.0
3	100 mM HEPES, 100 mM KCl, pH 7.0
4	100 mM HEPES, 50 mM LiCl, 0.1% CHAPS, pH 7.0
5	100 mM HEPES, 50 mM LiCl, 0.1% deoxycholate, pH 7.0
6	100 mM HEPES, 50 mM (NH ₄) ₂ SO ₄ , 20% glycerol, pH 7.0
7	100 mM K ₂ HPO ₄ /KH ₂ PO ₄ , 2.5 mM ZnCl ₂ , pH 4.3
8	100 mM K ₂ HPO ₄ /KH ₂ PO ₄ , 50 mM (NH ₄) ₂ SO ₄ , 0.05% dextran sulfate, pH 6.0
9	100 mM K ₂ HPO ₄ /KH ₂ PO ₄ , 50 mM (NH ₄) ₂ SO ₄ , 1% Triton X-10, pH 6.0
10	250 mM K ₂ HPO ₄ /KH ₂ PO ₄ , 0.1% CHAPS, pH 6.0
11	100 mM potassium acetate, 50 mM NaCl, 0.05 % dextran sulfate, 0.1% CHAPS, pH 5.5
12	100 mM sodium acetate, 1 M MgSO ₄ , pH 5.5
13	100 mM sodium acetate, 100 mM glutamine, 10 mM DTT, pH 5.5
14	100 mM sodium acetate, 100 mM KCl, 0.1% <i>n</i> -octyl-β-D-glucoside, pH 5.5
15	100 mM sodium acetate, 50 mM LiCl, 5 mM CaOAc, pH 5.5
16	100 mM TEA, 100 mM KCl, 0.05% dextran sulfate, pH 8.5
17	100 mM TEA, 100 mM KCl, 10 mM DTT, pH 8.5
18	100 mM TEA, 100 mM sodium glutamate, 0.02% <i>n</i> -octyl-β-D-glucoside, 10% glycerol, pH 8.5

- 19 100 mM TEA, 50 mM (NH₄)₂SO₄, 10 mM MgSO₄, pH 8.5
- 20 100 mM TEA, 50 mM LiCl, 5 mM EDTA, pH 8.5
- 21 100 mM Tris, 1 M (NH₄)₂SO₄, 10 mM DTT, pH 8.2
- 22 100 mM Tris, 20% glycerol, pH 7.6
- 23 100 mM Tris, 100 mM KCl, 0.1% deoxycholate, 25% glycerol, pH 7.6
- 24 100 mM Tris, 100 mM KCl, 2 mM EDTA, 1% Triton X-100, pH 8.2
- 25 100 mM Tris, 100 mM sodium glutamate, 10 mM DTT, pH 8.5
- 26 100 mM Tris, 2 M NaCl, 0.1% *n*-octyl-β-D-glucoside, pH 7.6
- 27 100 mM Tris, 50 mM NaCl, 10% *i*PrOH, pH 8.2
- 28 100 mM Tris, 50 mM NaCl, 100 mM urea, pH 8.2
- 29 100 mM Tris, 50 mM NaCl, 5 mM CaOAc, pH 7.6
- 30 100 mM Tris, 50 mM LiCl, pH 7.6

2.1.7.7. SDS-PAGE

For the analysis of the protein samples using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) the following buffers were used (Table 2-14).

Table 2-14: Solutions prepared for SDS-PAGE.

Resolving Buffer	Stacking Buffer
1.5 M Tris-Cl	0.5 M Tris-Cl
0.4% SDS	0.4% SDS
pH 8.8	pH 6.8
4X SDS running buffer	Magic Dye
767 mM Glycine	60 g L ⁻¹ G250 Blue Dye
100 mM Tris base	1% HCl
2X SDS PAGE Sample Buffer	5X SDS PAGE Sample Buffer
60 mM Tris-Cl pH 6.8	250 mM Tris-Cl pH 6.8
10% Glycerol	30% Glycerol
2% SDS	10% SDS
0.05% Bromophenol Blue	0.02% Bromophenol Blue
715 mM β-mercaptoethanol	710 mM β-mercaptoethanol

The resolving and stacking gel buffer make up components of the gel, the running buffer was utilised during the electrophoresis process and the magic dye was used to visualise the protein samples on the SDS gels. Protein analysis was performed using

12% acrylamide SDS gels. The gels were prepared using the recipes below (see Table 2-15). Routinely 15-well gels were poured. Mini-PROTEAN TGX precast gels were obtained containing 12% acrylamide from BioRad, with 15 wells.

Table 2-15: Recipes for the resolving and stacking gel 12% SDS-PA gels.

Resolving Gel		Stacking Gel	
3.2 mL	Water	3.2 mL	Water
2.5 mL	Resolving Gel Buffer	1.3 mL	Stacking Gel Buffer
4.2 mL	Acrylamide Stock	0.5 mL	Acrylamide Stock
		10 μ L	Bromophenol Blue (1%)
75 μ L	10% APS	50 μ L	10% APS
8 μ L	TEMED	8 μ L	TEMED

Samples for SDS-PAGE analysis were prepared with 2X or 5X SDS-PAGE sample buffer to a total volume of 15 or 20 μ L, for either bought-in or self-cast gels, respectively. The low molecular weight marker or precision plus protein pre-stained standard (dual colour) from Biorad was run as standard on every gel. The samples were heated to 95 °C for 5 min to ensure the denaturation of the sample, before loading onto the gel. The running buffer was made up from stock and by addition of 0.1% SDS before use. Gels were run at 200 V for 40 min or 55 min, for bought-in or self-cast gels, respectively. The visualisation was performed using Magic Dye.

2.1.7.8. Immunoblot Analysis (Western Blot)

For the immunoblot analysis of protein samples the transfer buffer and Tris-buffered saline solution with Tween-20 (TBST) were prepared following the details given below (Table 2-16).

Table 2-16: Buffers used in immunoblot analysis.

Transfer Buffer		10X TBST	
25 mM	Tris-Cl	100 mM	Tris-Cl
192 mM	Glycine	1 M	NaCl
20 %	MeOH	1 %	Tween-20
pH	8.3	pH	7.5

Samples were prepared and run on an SDS gel as described above, without the final staining step. The gel was rinsed in transfer buffer and transferred into the blotting machine following the schematics shown below (Figure 2-1). The membrane was

activated in methanol for 2 min and then incubated in transfer buffer for at least 15 min. The chromatography paper was incubated in transfer buffer for at least 15 min before assembly. The transfer was performed at 25 V for 30 or 50 min, for bought-in or self-cast gels, respectively.

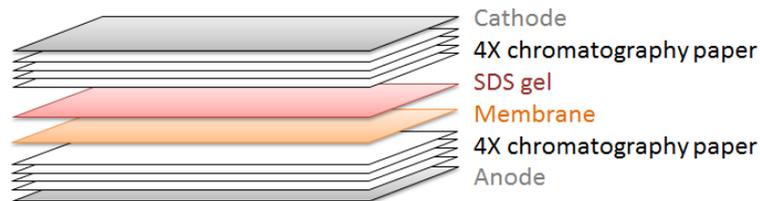


Figure 2-1: Schematic set up for the protein transfer from SDS gels onto a nitrocellulose membrane.

The membrane was stained with Ponceau S and the molecular weight standards were marked with pencil. The membrane was de-stained using a small amount of 5% milk (low-fat) in TBST over 5 min. It was then incubated for 1 h, shaking at rt in fresh 25 mL of 5% milk/TBST. The blocking of the membrane was followed by incubation with the antibody in 5% milk/TBST (1 h, rt; monoclonal anti-polyhistidine-peroxidase antibody produced in mouse, diluted according to delivery instructions). The membrane was washed multiple times with TBST before the visualization with 3,3'-diaminobenzidine tablets.

2.1.8. Protein Purification and Characterisation

The following buffers were prepared for protein isolation and purification (Table 2-17). All buffers are based on Tris-base at pH 8.2, with variables concentrations of sodium chloride and imidazole, depending on the purification method. Ultra-pure water was used in the preparation of the buffers and a filtration over 0.2 μm filters was performed prior to use.

Table 2-17: Composition of buffers used in bacterial protein preparation and purification.

Salt-free Buffer		Ion-Exchange Elution Buffer	
100 mM	Tris-Cl	100 mM	Tris Cl
pH	8.2	1 M	NaCl
		pH	8.2

Ni-Purification Buffer	
100 mM	Tris-Cl
200 mM	NaCl
500 mM	Imidazole
pH	8.2

2.1.8.1. Ni-Affinity Chromatography

5 mL HisTrap FF Crude columns were routinely used for purification of protein *via* Ni-affinity. The columns were equilibrated with resuspension buffer (5 CV) before loading with cell lysis using a peristaltic pump. The flow-through was collected for SDS gel analysis. The column was attached to a FPLC and the following method was run. Two washing steps with no and 30 mM imidazole were performed over 5 CV each, before a gradient was run from 30 to 500 mM imidazole over 20 CV. A final 5 CV wash at 500 mM imidazole concentration was performed. Fractions of interest were collected and analysed *via* SDS gel or immunoblot. The column was regenerated following the procedure provided by the supplier (GE).

2.1.8.2. Ammonium Sulfate Precipitation and Desalting

Precipitation of protein using ammonium sulfate (AS) was performed at 4 °C. Fractions were collected in 20% AS steps (20, 40, 60, and 80%). The corresponding amount of AS was slowly added to the stirring solution. Once added, the precipitated proteins were collected *via* centrifugation (10 min, 3,900 × g). The supernatant was treated with the next portion of AS and the pellet was resuspended in a small amount of buffer and desalted. For small sample volumes G25 desalting columns (5 mL CV, GE, used according to supplier instructions) were used to reduce the salt concentration. Larger samples were dialysed (3 kDa cut-off) overnight.

2.1.8.3. Anion Exchange

Samples for anion exchange chromatography were prepared in salt-free buffer, either by resuspension of the cell pellet in the respective buffer or by dialysis. The

protein was eluted from Q-Sepharose Fast Flow columns (5 mL, GE) after a 5 CV wash with no NaCl using the gradient to 1 M NaCl over 20 CV, with a final wash step at 1 M NaCl lasting 5 CV. The column was washed and stored according to the information provided by the supplier. A buffer exchange *via* concentration (30 kDa cut-off) or dialysis was performed before further purification.

2.1.8.4. Protein Characterisation using UV/Vis Spectrometry

UV/Vis wavelength scans were routinely performed to characterise the protein due to the heme's absorbance maximum at 420 nm. Baseline corrections were performed. The scan range was from 350 to 600 nm, with a 1 nm s⁻¹ read.

chromatography, all clones were carrying a His-tag. The position of, and linker to, the His-tag can influence the protein folding and the functional appearance. As a result, one construct was chosen with an N-terminal His-tag, separated from the mature protein by a 3C-protease cleavage site, and another construct carried a C-terminal tag, directly fused to the mature protein. A vector system with an N-terminal secretion sequence (*pelB*, pETFP31) and a C-terminal His-tag was chosen to mimic the secretion as it occurs in the native fungal strains. The different constructs created for bacterial expression are displayed in Figure 2-3 together with schematics for the native genes.

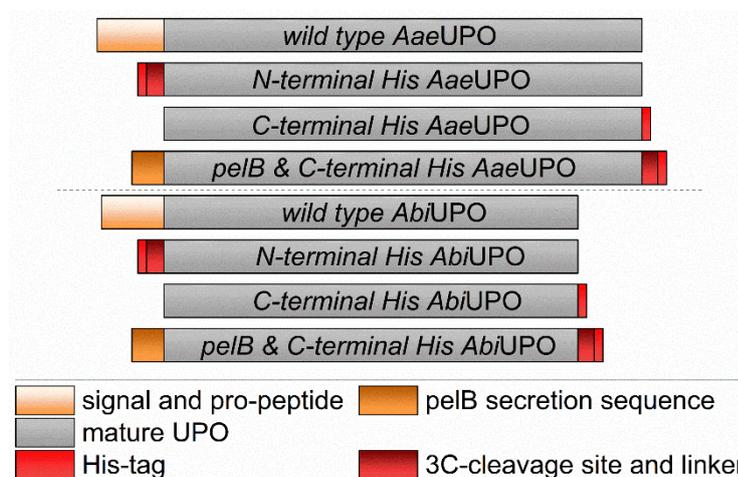


Figure 2-3: Overview of the different *AaeUPO* and *AbiUPO* gene constructs created for bacterial expression alongside the native genes.

2.2.1. N-terminal His-tag (pETYSBLIC-3C)

In order to create an N-terminal His-tagged protein, the amplified UPO genes were cloned into the pETYSBLIC-3C vector using ligation-independent cloning (LIC). The amplified insert was T4-treated to make it compatible with the linearized and T4 treated pETYSBLIC-3C vector. Different insert to vector ratios were tested during ligation. To check the efficiency of the cloning, NovaBlue cells were transformed with the plasmid and the DNA was extracted from a 5 mL starter culture. The extracted DNA was analysed through a double digest using *NdeI* and *NcoI* as restriction enzymes. Compared to an undigested sample, both homologs show multiple fragment bands on an agarose gel, indicating successful cloning for all insert to vector ratios tested (see Figure 2-4). The restriction enzymes isolate the vector at about 6 kbp, the *AaeUPO* fragment runs at about 1 kbp and the *AbiUPO* fragment can be observed at 0.5 kbp. The reason no full length UPO genes can be observed, with 1116

and 1128 bp for *AaeUPO* and *AbiUPO*, respectively, is that both genes carry an additional restriction site for the restriction enzymes used in this experiment. Fragments smaller than 0.5 kbp cannot be clearly identified on the gel. Before the transformation of *E. coli* expression strains with the constructs, the DNA samples were sequenced. Analysis of the gene sequence confirmed that both the *AaeUPO* and the *AbiUPO* were successfully cloned into the pETYSBLIC-3C vector system.

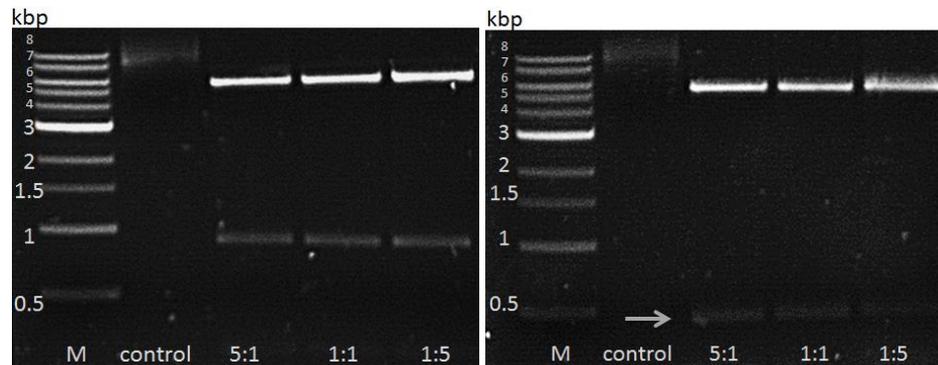


Figure 2-4: 1% agarose gels of double digested *AaeUPO* (left) and *AbiUPO* (right) in pETYSBLIC-3C with *NdeI* and *NcoI*: 1 – Marker; 2 – undigested control sample; 3 – 5:1 insert:vector ligation ratio; 4 – 1:1 insert:vector; 5 – 1:5 insert:vector.

2.2.2. C-terminal His-tag (pET22b)

An expression construct carrying a non-cleavable C-terminal His-tag was created using In-Fusion cloning inserting the *AaeUPO* and *AbiUPO* genes into pET22b. The amplified genes were mixed with the linearized vector before transforming NovaBlue cells with the cloning mixture. Four colonies from each plate were picked and analysed using colony PCR. The results are shown in Figure 2-5. All *AbiUPO* colonies (B1-B4) show a strong band around 1 kb, indicating a successful insertion of the gene into the vector system. For the *AaeUPO* variant (A1-A4) only three of the four picked colonies show amplification of the gene of interest. It is noticeable, that upon amplification of the *AaeUPO* gene a second smaller fragment is produced (approx. 0.7 kbp). This was due to the primer annealing with another part of the gene, allowing the formation of a second, shorter gene fragment. This has also been observed for *AaeUPO* after the PCR gene amplification, where two fragments around 1 kbp and 0.7 kbp were observed (data not shown). Solely DNA of the larger fragment was purified and used for further experiments. To check that the gene sequence cloned into the new vector system was still conserved, samples were sent for

sequencing. The results showed a successful cloning of *AaeUPO* and *AbiUPO* into pET22b.

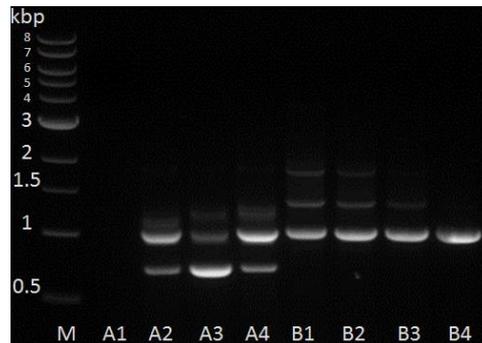


Figure 2-5: 1% agarose gel showing the results of the colony PCR of *AaeUPO* (A1-A4) and *AbiUPO* (B1-B4) after cloning into pET22b using In-Fusion cloning.

2.2.3. *pelB* Secretion Sequence and C-terminal His-tag (pETFPP31)

The aim of cloning the *AaeUPO* and *AbiUPO* genes into pETFPP31 was to obtain a gene construct with an N-terminal *pelB* secretion sequence in addition to a cleavable C-terminal His-tag for purification purposes. The idea behind including the secretion sequence was to mimic the secretion of the protein as it occurs under native conditions. The secretion into the periplasm and its altered oxidative environment is also thought to aid the protein folding and potentially the formation of the disulfide bond that has been reported for *AaeUPO*. The *pelB* secretion sequence is cleaved upon secretion of the protein into the periplasm and leaves the mature protein with its C-terminal tag. In-Fusion cloning was performed with the linearized vector and the amplified gene constructs. The reaction mixture was used to transform NovaBlue cells and the DNA of a starter culture was extracted. A double digest was performed to analyse the extracted DNA, using *NdeI* and *XhoI* as restriction enzymes. The results are shown in Figure 2-6 and confirm that for both, *AaeUPO* and *AbiUPO*, the cloning into the pETFPP31 vector was successful. Similar to the results obtained for the double digest of the genes cloned into pETYSBLIC-3C vector, multiple DNA fragment can be observed for *AbiUPO*, attributed to additional restriction enzyme sites on the gene sequence. The sequence validation was last step before conducting further studies. Both the forward and the reverse analysis of the construct confirmed the correct sequences for the *AaeUPO* and *AbiUPO* genes.

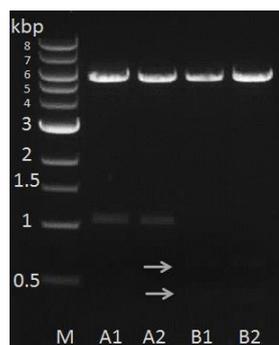


Figure 2-6: 1% agarose gel of double digested *Aae*UPO (A1, A2) and *Abi*UPO (B1, B2) in pETFPP31 using *Nde*I and *Xho*I.

2.2.4. Bacterial Transformations

The three different constructs for both *Aae*UPO and *Abi*UPO were used to transform three different *E. coli* strains, BL21, Rosetta2, and Origami2. The BL21 strain is a standard strain used in protein expression; Rosetta cells are generally used to enhance translation, especially in the presence of rare codons, and the Origami strain is known to improve protein folding. A standard heat-shock procedure was used, giving colonies after overnight incubation at 37 °C. The choice of antibiotics was determined by the plasmid and cell line used; an overview is given below (see Table 2-18).

Table 2-18: Antibiotic resistance and working concentrations for the different bacterial strains and plasmids used.

<i>E. coli</i> strain	Antibiotic	Plasmid	Antibiotic
BL21 (D3)	-	pETYSBLIC-3C	kan (35 mg mL ⁻¹)
Rosetta2, pLysS	cam (34 mg mL ⁻¹)	pET22b	amp (100 mg mL ⁻¹)
Origami2, pLysS	-	pETFPP31	amp (100 mg mL ⁻¹)

Overnight cultures were grown from the colonies on the plates and glycerol stocks (25%) were prepared. These glycerol stocks were used for subsequent experiments.

2.2.5. Growth Curves

The purpose of the following experiment was to study the growth behaviour of the different bacterial strains in different growth media. The N- and C-terminal His-tag constructs were studied in BL21 using LB as growth medium alongside each other. The growth was not influenced by the construct the cells were transformed with (data not shown). All further experiments were performed with the *Aae*UPO N-terminal His-tag construct. The growth of three different *E. coli* strains was

monitored (BL21, Rosetta 2, and Origami 2) in three different types of growth media (LB, M9-glucose, and M9-glycerol). The results are shown in Figure 2-7.

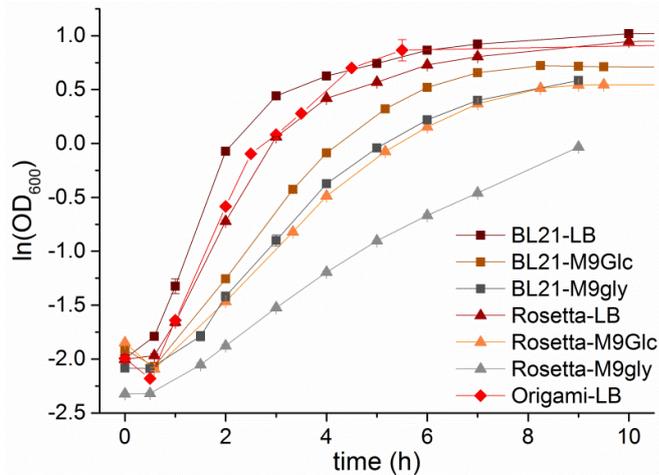


Figure 2-7: Growth curves for BL21 (square), Rosetta 2 (triangle) and Origami 2 (diamond) transformed with *AaeUPO* carrying an N-terminal His-tag in LB (red), M9-glucose (orange) and M9 glycerol (grey) at 37 °C. Analysis was performed in duplicates, the error bars describe the standard deviation.

All growth curves followed the expected trend; starting with the lag-phase, followed by the exponential growth and the stationary phase. A strong dependence on the growth medium is visible and can be confirmed by the doubling time. For the BL21 strain the doubling time in LB is around 30 min, in M9 with glucose as carbon source it takes approx. 65 min and using M9 with glycerol as carbon source the doubling time is approx. 100 min. A similar trend is observed for the Rosetta2 strain, where the doubling time in LB is 45 min, in M9-glucose it is 85 min and for M9-glycerol 120 min were determined. Using the Origami2 strain growth was only observed in LB medium with a doubling time of 40 min. Growing on a complex medium, such as LB, provides the cells with different nutrients which are easily utilized by the cells. When grown on minimal medium however, the nutrition sources are limited and growth also depends on the uptake and utilisation rates of the specific carbon source; an increase in doubling times was observed using minimal medium. Between the different *E. coli* strains a trend in the doubling time was also determined; in LB BL21 cells double in 30 min, Origami2 takes 40 min (approx. 30% increase) and Rosetta2 needs 45 min (approx. 50% increase). The presence of the pLysS plasmid in the Rosetta2 and Origami2 cells is one potential reason for this phenomenon. The use of a second antibiotic, which is added when using Rosetta2 cells, is another factor which can influence the cell growth and therefore effects the doubling time.

2.3. *E. coli* Test Expressions

In the process of finding a bacterial expression system for UPOs various different conditions were screened. The three afore mentioned constructs were used in the BL21, Rosetta2, and Origami2 strains. Gene expression was performed and studied in three different media; the first one being LB, the second and third are a M9-minimal medium. The advantage of using M9 medium arises from the option of adding different carbon-sources, as well as further additives, depending on the growth requirements. Here glucose and glycerol were tested as carbon sources in the process of gene expression; the uptake and processing of both molecules varies within the cell. The transcription was induced using IPTG and when using minimal medium, the heme-precursor ALA and FeCl₃ were also added upon induction. The growth to a density between 0.6 and 0.8 was performed at 37 °C; the gene expression was monitored at 16, 30, and 37 °C and was compared to a non-induced control sample for each condition. As SDS gel analysis did not allow the identification of the protein of interest at around 38 kDa (example see Figure 2-8, left), immunoblot analysis using an anti-His antibody allowed the visualisation of the His-tag carrying UPO (Figure 2-8, right). Figure 2-8 shows the results of the test expression carried out using the N-terminally tagged *Aae*UPO in BL21 cells when grown on M9-glc. Insoluble protein was visible at all three temperatures; soluble protein can only be observed when expressed at 16 °C (grey box, Figure 2-8). Furthermore, the Western blot showed the presence of a small amount of degradation products in the form of bands with decreased molecular weight (<38 kDa, insoluble fractions only). When studying *Abi*UPO the same trends regarding the need of immunoblot analysis were observed; no strong overexpression bands were visible in the SDS gels. The immunoblots, in contrast to the *Aae*UPO variant, show stronger degradation product and unspecific binding to a protein with a molecular weight around 66 kDa (data not shown). However, the His-tagged *Abi*UPO exhibits the strongest signal.

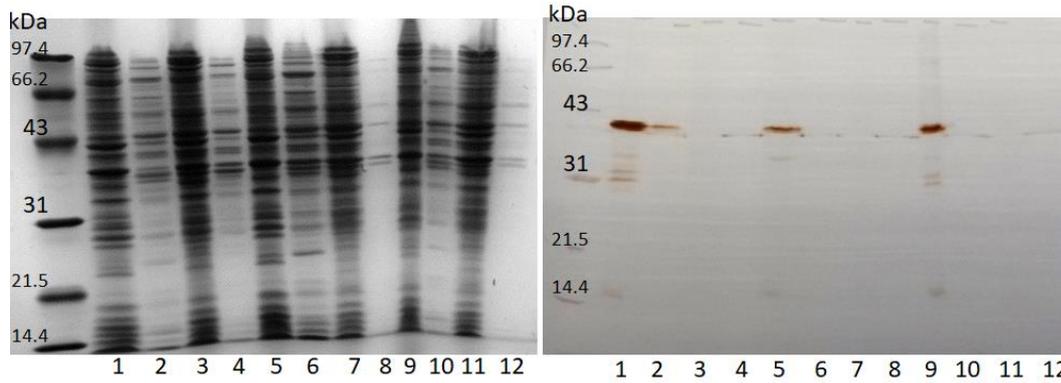


Figure 2-8: 12 % SDS gel for the test expression of *AaeUPO* (approx. 38 kDa) with N-terminal His-tag (left) and immunoblot (right) using BL21 in M9-glc. M - Low molecular weight marker (BioRad); 1 - 16 °C insoluble; 2 - 16 °C soluble; 3 - 16 °C control insoluble; 4 - 16 °C control soluble; 5 - 30 °C insoluble; 6 - 30 °C soluble; 7 - 30 °C control insoluble; 8 - 30 °C control soluble; 9 - 37 °C insoluble; 10 - 37 °C soluble; 11 - 37 °C control insoluble; 12 - 37 °C control soluble.

A summary of all the results obtained for *AaeUPO* and *AbiUPO* under the different conditions can be found in Table 2-19 and Table 2-20, respectively. The “✓” is only an indication for soluble protein and does not specify the amount of soluble protein found compared to the insoluble fraction. The “✗” is used to describe conditions where only insoluble or no protein was found.

Table 2-19: Results of the expression tests with the different *AaeUPO* constructs in different expression strains using various growth media at 16, 30, and 37 °C; ✓ - soluble protein; ✗ - insoluble or no protein; “-” - condition not tested.

<i>AaeUPO</i>		N-terminal His ₆								
strain		BL21			Rosetta2			Origami2		
T / °C		37	30	16	37	30	16	37	30	16
LB		✗	✗	✗	✓	✗	✗	✗	✗	✗
M9-glc		✗	✗	✓	✗	✗	✗	-	-	-
M9-gly		✗	✗	✗	✗	✗	✗	-	-	-

<i>AaeUPO</i>		C-terminal His ₆								
strain		BL21			Rosetta2			Origami2		
T / °C		37	30	16	37	30	16	37	30	16
LB		✗	✗	✗	✗	✗	✗	✗	✗	✗
M9-glc		✗	✗	✗	✗	✗	✗	-	-	-
M9-gly		✗	✗	✗	✗	✗	✗	-	-	-

<i>AaeUPO</i>		pelB and C-terminal His ₆								
strain		BL21			Rosetta2			Origami2		
T / °C		37	30	16	37	30	16	37	30	16
LB		✓	✓	✓	✓	✓	✗	✗	✗	✗
M9-glc		✗	✗	✗	✗	✗	✗	-	-	-
M9-gly		✗	✗	✗	✗	✓	✗	-	-	-

Table 2-20: Results of the expression tests with the different *AbiUPO* constructs in different expression strains using various growth media at 16, 30, and 37 °C; ✓ - soluble protein; ✗ - insoluble or no protein; “-” – condition not tested.

<i>AbiUPO</i>		N-terminal His ₆								
strain	BL21			Rosetta2			Origami2			
T / °C	37	30	16	37	30	16	37	30	16	
LB	✗	✗	✗	✓	✗	✗	✗	✗	✗	
M9-glc	✓	✓	✓	✗	✗	✗	-	-	-	
M9-gly	✓	✓	✓	✗	✗	✗	-	-	-	

<i>AbiUPO</i>		C-terminal His ₆								
strain	BL21			Rosetta2			Origami2			
T / °C	37	30	16	37	30	16	37	30	16	
LB	✗	✗	✗	✗	✗	✗	✗	✗	✗	
M9-glc	✗	✗	✗	✗	✗	✗	-	-	-	
M9-gly	✗	✗	✗	✗	✗	✗	-	-	-	

<i>AbiUPO</i>		pelB and C-terminal His ₆								
strain	BL21			Rosetta2			Origami2			
T / °C	37	30	16	37	30	16	37	30	16	
LB	✓	✓	✓	-	-	-	-	-	-	
M9-glc	✗	✗	✗	-	-	-	-	-	-	
M9-gly	✗	✗	✗	-	-	-	-	-	-	

The results obtained for both homologs during the test expression are mostly consistent. The N-terminal His-tag construct in BL21 shows soluble expression for the *AbiUPO* homolog in all minimal media conditions; the *AaeUPO* only showed soluble expression when grown on glucose as carbon source at 16 °C. The same construct in Rosetta2 cells showed soluble protein for both homologs when using LB at 37 °C. The non-cleavable C-terminal His-tag construct did not yield in any protein, soluble or insoluble. The pelB secretion construct carrying a cleavable C-terminal His-tag showed soluble expression using BL21 in conjunction with LB medium for both *AaeUPO* and *AbiUPO*. Furthermore, soluble expression is observed for *AaeUPO* when using Rosetta2 cells at higher temperatures in LB (30 and 37 °C) and in minimal medium with glycerol as carbon source (30 °C).

Both the BL21 and Rosetta2 strain gave soluble protein when paired with the right construct and growth medium. A clear trend was not visible. The only consistent condition when comparing different construct is the expression with Rosetta2 cells

in LB at 37 °C. The amount of soluble protein that was observed during these experiments never exceeded the amount of insoluble protein, when comparing the intensity of the bands visible from immunoblot analysis. No soluble expression was observed using Origami2 cells.

2.3.1. Expression over time

As the results of the expression test did not show any trends towards expression temperature, medium, cell line or construct further experiments were performed. These focused on monitoring the amount of protein expressed over time. The reasoning behind this experiment is that, especially at higher temperatures, the cells are likely to run out of nutrients and could potentially degrade the expressed protein over time. The conditions that were tested are the same as ones described earlier for the expression tests. Only the *AaeUPO* homolog was tested. The analysis was based on the Western blot analysis of the samples, comparing the intensity of the bands for the soluble and insoluble fractions. To allow a quantitative analysis of the different time points, the weight of the cell pellet was kept the same by determining the OD₆₀₀ and adjusting the sample volume analysed. Some examples are shown in Figure 2-9 for the N-terminally tagged *AaeUPO* when expressed by BL21 in LB at 30 °C (left) or when expressed by Rosetta2 in M9-glc at 30 °C (right). At t = 0 h, straight after induction, no protein band was visible on the blots. Both examples showed a similar development over time. The amount of soluble protein was first increasing, before degradation started to occur. The BL21-M9 example showed less soluble than insoluble protein at the beginning (t = 4 h, classed as "1") moving on to more soluble than insoluble protein (t = 17 h, classed as "3") and no soluble protein at the end of the experiment (t = 24 h, classed as "0"). The Rosetta2-LB sample started with equal amounts of soluble and insoluble protein after 2 h (classed as "2"), after 4 and 8 h more soluble than insoluble protein was detected ("3") and full degradation of all protein was observed in the final samples (t = 20/24 h, "0").

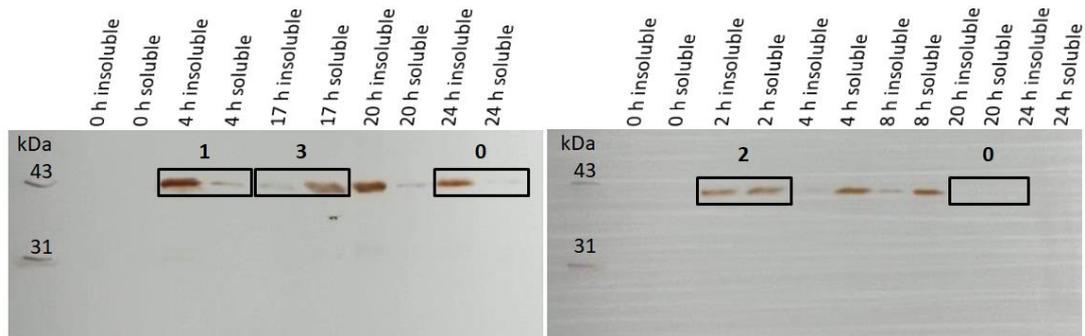


Figure 2-9: Immunoblots for the production N-terminal His-tag *AaeUPO* (approx. 38 kDa) over time: **left** – using BL21 in M9-glc at 30 °C; **right** – using Rosetta2 in LB at 30 °C. Highlighted the different insoluble to soluble ratios with: “0” no (soluble) protein; “1” soluble < insoluble; “2” soluble = insoluble; “3” soluble > insoluble.

The same analytical method was used to interpret the immunoblots for all different conditions (data not shown). The amount of soluble protein was determined in respect to the insoluble fraction and plotted against the time (see Figure 2-10). The third axis serves to compare the different expression temperatures under otherwise same conditions (same *E. coli* strain, same growth medium). The top half of the figure shows the results representing the N-terminally tagged protein, the bottom half the *AaeUPO* with pelB secretion sequence and C-terminal His-tag. Overall, the expression of soluble protein was more efficient using the N-terminal His-tag construct. The decreased amounts of soluble protein for the other construct could originate from the secretion process the protein undergoes. For the N-terminal His-tag construct, expression at higher temperatures peaks after a few hours of incubation (between 2 to 17 h) and showed degradation of all soluble protein towards the end of the measurements (24 h). The expression at lower temperatures did not yield high amounts of soluble protein; however, the observed degradation was reduced.

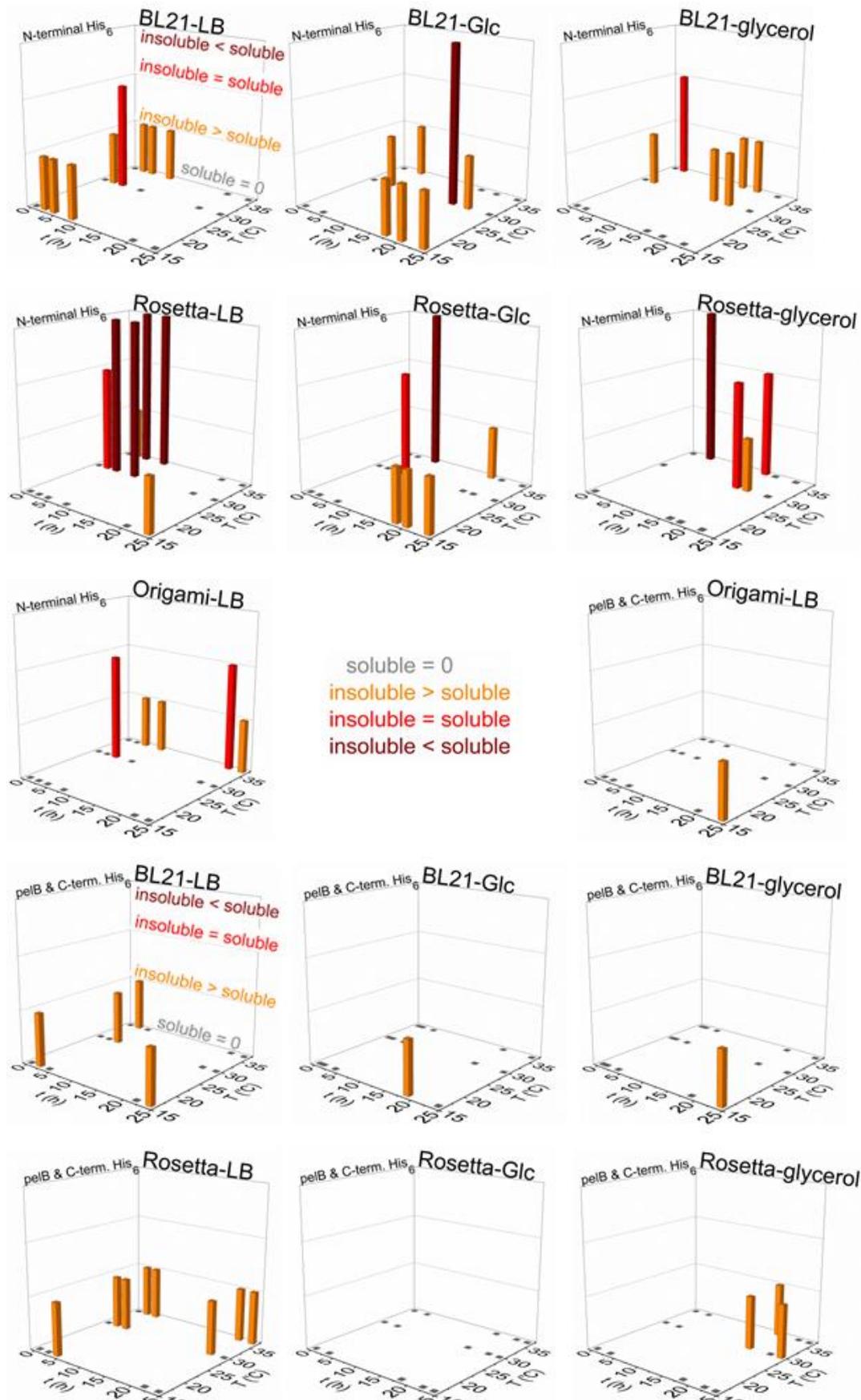


Figure 2-10: 3D-blots showing the amount of soluble protein expressed over time at different temperatures for the N-terminal His-tag *AaeUPO* (top) and the *AaeUPO* with pelB secretion sequence and C-terminal His-tag (bottom) for the different *E. coli* strains and growth media tested.

When analysing soluble protein obtained with the pelB/C-terminal His-tag construct the amounts were much smaller than those obtained with the other construct. Most of the lower temperature tests showed soluble protein towards the end of the measurements. For the pelB/C-terminal His-tag construct only three conditions showed expression of soluble protein at higher temperatures (BL21-LB, Rosetta2-LB, Rosetta2-M9-gly). The best overall condition for both constructs is the same. The Rosetta2 cells in conjunction with LB as growth medium showed for 30 and 37 °C expression-temperature high amounts of soluble protein, when compared to other conditions tested for the respective construct (enlarged graphs).

An updated version of Table 2-19 (*AaeUPO*), identifying those conditions where soluble protein is produced, is shown below (Table 2-21); highlighted in bold are conditions already known to produce *AaeUPO*. The number of conditions with soluble protein increased from 13% (8 conditions) to 46% (29 conditions) for a total of 63 conditions.

Table 2-21: Summary of all expression test (incl. measurements over time) for the different *AaeUPO* constructs in different expression strains using various growth media at 16, 30, and 37 °C; ✓ - soluble protein; ✗ - insoluble or no protein; “-” – condition not tested.

<i>AaeUPO</i>		N-terminal His ₆								
strain	BL21			Rosetta2			Origami2			
T / °C	37	30	16	37	30	16	37	30	16	
LB	✓	✓	✓	✓	✓	✓	✓	✓	✗	
Glc	✓	✓	✓	✓	✓	✓	-	-	-	
Gly	✓	✓	✗	✓	✓	✗	-	-	-	

<i>AaeUPO</i>		pelB and C-terminal His ₆								
strain	BL21			Rosetta2			Origami2			
T / °C	37	30	16	37	30	16	37	30	16	
LB	✓	✓	✓	✓	✓	✓	✗	✗	✓	
Glc	✗	✗	✓	✗	✗	✗	-	-	-	
Gly	✗	✗	✓	✓	✓	✗	-	-	-	

2.3.2. Solubility Test

To decrease the amount of insoluble protein and to obtain a better soluble to insoluble protein ratio solubility tests were performed. A buffer library was created following the procedure described by G. Lindwall and coworkers.¹¹⁵ The tests were performed using the N-terminal tagged *AaeUPO* and *AbiUPO* and the genes were

expressed by BL21 cells using M9-glc as medium at 16 °C. The obtained cell pellets were separated into thirty equal portions and each portion was paired with one buffer from the list (Table 2-13). After cell disruption the soluble and insoluble fractions were separated and analysed using immunoblots (see Figure 2-11).

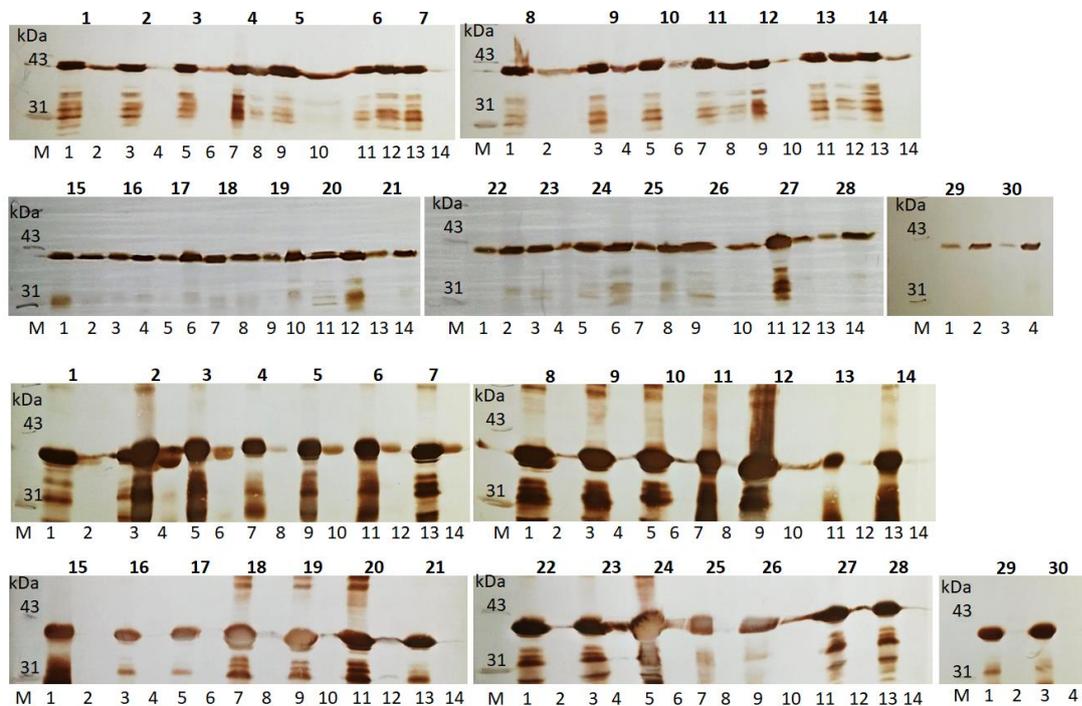


Figure 2-11: Immunoblots of the buffer screening using *AaeUPO* (top; approx. 38 kDa) and *AbiUPO* (bottom; approx. 38 kDa). M - Low molecular weight marker (BioRad); uneven numbers – insoluble fractions; even numbers – soluble fractions; respective buffer number printed in bold on top.

The solubility of *AaeUPO* in the tested buffers was with 66% soluble hits (twenty buffers) a lot higher than the solubility observed for *AbiUPO* with 40% (twelve buffers). The amount of soluble protein in the *AaeUPO* samples was equal or greater than the amount of insoluble protein in ten out of the twenty conditions showing any soluble protein. The most soluble protein was observed in TAE and Tris buffers (for TAE: buffers number 17, 19, and 20; for Tris: buffers number 21, 22, 29, and 30). Overall the *AaeUPO* was soluble in all different types of buffer tested over a wide pH range. Besides glycerol as a stabilisation agent no other preferences for salts and additives could be observed. However, in the presence of dithiothreitol (DTT, buffer numbers 2, 13, 17, 21, and 25), a reducing agent, some of the best results were obtained for *AaeUPO* (buffer numbers 13, 17, 21, and 25) and *AbiUPO* (buffer number 2). This is interesting as DTT could interfere with the disulfide bond formation expected in UPOs. For the *AbiUPO* homolog none of the conditions

showed equal amounts of soluble and insoluble protein or an excess of soluble protein. The best results for *Abi*UPO were obtained using buffer 2 (100 mM HEPES, 100 mM sodium glutamate, 5 mM DTT, pH 7.0). For the other conditions giving soluble protein, no clear trend towards buffer salt, pH, salt, or additives could be observed. A summary of the solubility of *Aae*UPO and *Abi*UPO in the different buffers is given below (Table 2-22); highlighted in bold are the conditions where an equal amount or excess of soluble over insoluble protein was detected.

Table 2-22: Summarized results for the solubility screen performed with *Aae*UPO and *Abi*UPO; “✓” - soluble protein, “✓” - higher amounts of soluble protein, “✗” - no soluble protein.

Buffer #	1	2	3	4	5	6	7	8	9	10
<i>Aae</i> UPO	✓	✗	✗	✓	✓	✓	✗	✗	✓	✗
<i>Abi</i> UPO	✓	✓	✓	✗	✓	✓	✓	✗	✗	✗
Buffer #	11	12	13	14	15	16	17	18	19	20
<i>Aae</i> UPO	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓
<i>Abi</i> UPO	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓
Buffer #	21	22	23	24	25	26	27	28	29	30
<i>Aae</i> UPO	✓	✓	✗	✓	✓	✗	✗	✓	✓	✓
<i>Abi</i> UPO	✗	✗	✓	✓	✗	✓	✓	✗	✗	✗

2.4. Scale-up and Purification

The scale-up expression and purification was performed solely with the *Aae*UPO homolog. The N-terminal His-tag construct and the pelB and C-terminal His-tag construct were expressed under different conditions and analysed using affinity chromatography. Additionally, different methods of extracting the secreted C-terminal His-tag construct from the media were tried and the protein was attempted to be purified using Ni-affinity chromatography. Further studies focused on other purifications methods, including ammonium sulfate precipitations and anion exchange chromatography (Q).

First attempts to purify *Aae*UPO were performed with the N-terminally tagged protein. The protein was expressed in BL21 using M9-glc as growth medium at 16 °C. The cell lysate was applied to a 5 mL His-Trap FF crude column. The flow-through was collected and analysed together with the wash and fractions of interest on a SDS gel (see Figure 2-12, left). Combined with the Western blot results, it was clear that

AaeUPO was not binding well to the Ni-column and the majority of the protein was located in the flow-through (gel lanes 2 and 3). The elution of UPO started at an imidazole concentration of approximately 80 mM (fraction 8, gel lane 7 onwards). From the Western blot it was visible that traces of protein were still eluted at higher volumes (fraction 20, 70 mL elution volume). Furthermore, the SDS-PAGE indicated multiple impurities; the most notable being a protein at approximately 70 kDa (lanes 6, 7). Since this protein did not appear in the western blot, it was concluded that it is not carrying a His-tag and a separation might be possible by increasing the wash volume. To study if the binding capacity of the Ni-column was the reason for protein eluting in the flow-through, an experiment where only a fraction of the crude extract was loaded onto the column was performed. The obtained results were comparable to the data shown below and indicated that poor binding of the protein to the resin and not overloading was the reason for the presence of *AaeUPO* in the flow-through (results not shown). Misfolded protein or poor affinity towards nickel as the complexing ion were potential reasons for the findings discussed so far.

In addition to the immunoblot analysis, UV/Vis wavelength scans for the different fractions were recorded (see Figure 2-12, right). The flow-through exhibits an absorption maximum at 421 nm, which undergoes a blue-shift during the course of the purification. This matches the data previously published ($\lambda_{\max} = 417$ nm) and describes the Soret-band specific for the ferric *AaeUPO*.⁸⁴ Other porphyrin specific absorption bands (Q-bands between 500 and 600 nm) or the oxo-Fe^{IV}-porphyrin radical cation band (compound I, $\lambda_{\max} = 361$ nm) were not decisive in these spectra. The wavelength shift over the course of the purification could originate from imidazole coordinating to the iron centre. These enzyme-imidazole complexes have been reported for CYPs, another group of heme-thiolate containing enzymes, before and complexes with imidazole derivatives have previously been reported for *AaeUPO*.⁷⁰

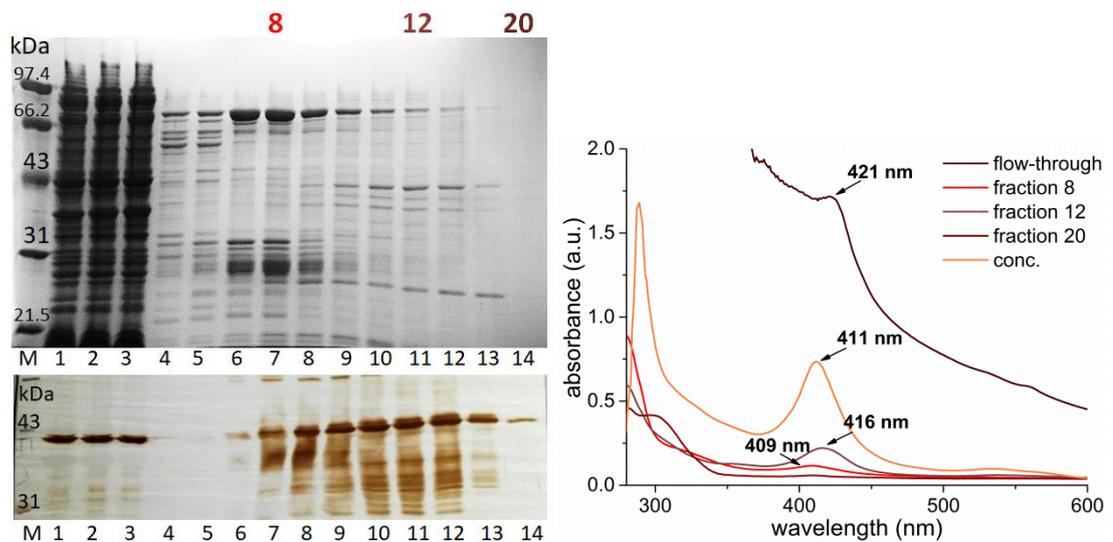


Figure 2-12: Purification of the N-terminal His-tag AaeUPO (approx. 38 kDa) expressed by BL21 in M9-glc at 16 °C: left: 12 % SDS gel and immunoblot; M - Low molecular weight marker; 1 – crude extract; 2 – flow-through I; 3 – flow-through II; 4 – wash I, 5 – wash II; 6 – fraction 7; 7 – fraction 8; 8 – fraction 9; 9 – fraction 10; 10 – fraction 11; 11 – fraction 12; 12 – fraction 13; 13 – fraction 15; 14 – fraction 20; right: UV/Vis spectra for the flow-through and fractions 8, 12, and 20 from the Ni-purification, as well as concentrated protein.

The next purification examined the N-terminal and secreted C-terminal His-tag versions of AaeUPO in comparison. In both cases the protein was produced by Rosetta2 cells, grown in LB at 37 °C, and harvested after 6 h. The SDS gel and immunoblot are shown below (Figure 2-13), with the results for the N-terminal His-tag UPO on the left of the marker (lanes 1 – 3) and the secreted C-terminal version on the right hand side (lanes 5 – 14). During the resuspension of the cell pellet, which was stored at -20 °C after harvesting, the secreted C-terminal AaeUPO showed partial cell lysis, a sample was taken for gel analysis (lane 5). It was visible from the Western blot that this fraction contained soluble protein. After cell disruption samples for the insoluble and soluble (Ni-load) parts were taken before charging the Ni-column. In both cases the chromatogram does not show any protein elution when studying the UV-trace (280 nm). The fractions analogous to previous purifications (Figure 2-12, fraction 8 onwards) were run on the gel and studied as well (data not shown for the N-terminal His-tag AaeUPO). No protein was observed in the elution fractions using either SDS-PAGE or immunoblot analysis. The flow-through samples for both AaeUPO variants show high levels of protein, which was comparable to the previous results (lanes 3 and 8). When studying the amounts of protein found in the soluble and insoluble fractions, most of the UPO was located in the latter (lanes 1 & 2, and 6 & 7). This indicates that either the scale up in growth medium (from 50 mL to

500 mL) or the method of cell lysis has an influence on the amount of soluble protein produced, when comparing these ratios to the results obtained from the monitoring the expression over-time. The immunoblot further showed that the amount of soluble protein decreased during the cell disruption, when comparing the sample before (lane 5) and the soluble fraction after cell disruption (lane 7). The UV/Vis spectra for the flow-through samples were collected (Figure 2-13, right). Both the N-terminally tagged and the secreted C-terminally tagged *AaeUPO* show an absorption band around 420 nm, which described the Soret-band specific for the ferric *AaeUPO*, as discussed earlier.

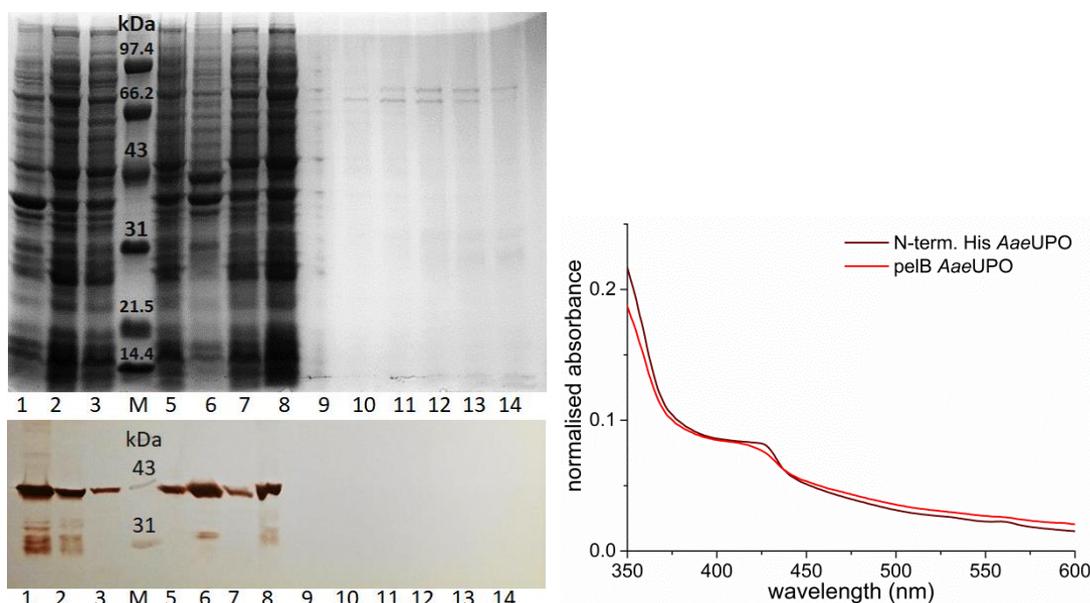


Figure 2-13: Purification of the N-terminal His-tag *AaeUPO* (1-3) and the secreted C-terminally tagged *AaeUPO* (5-14) expressed by Rosetta2 in LB at 37 °C left: 12 % SDS gel and immunoblot: 1 – insoluble protein; 2 – soluble/load; 3 – Ni flow-through; M - low molecular weight marker; 5 – cells before lysis; 6 - insoluble; 7 – soluble/load; 8 – Ni flow-through; 9 – fraction 8; 10 – fraction 9; 11 – fraction 10; 12 – fraction 11; 13 – fraction 12; 14 – fraction 13; right: UV/Vis spectra of the Ni flow-through samples of both *AaeUPO* variants.

The next trial focused on the purification of the secreted C-terminal His-tag *AaeUPO* isolated from the periplasm without lysing the cells. This was done to solely obtain secreted enzyme and study whether the secretion has an influence on the folding of the protein and the binding to the Ni-resin. The protein was isolated from the periplasm using an osmotic shock procedure. The sucrose/EDTA buffer in the first step permeates the bacterial membranes and the consequent addition of low salt buffer triggers an osmotic transition, which releases the protein. First experiments extracting the periplasmic UPO showed that most of the enzyme was found in the sucrose/EDTA fraction and only small amounts were released in the osmotic shock

fraction (data not shown). Having identified protein carrying fraction, it was then dialysed and applied to a Ni-chromatography column. The results shown below are for the *AaeUPO* expressed by Rosetta2 in LB at 37 °C (see Figure 2-14). The difference between the two samples was the absence (right) or presence (left) of ALA/FeCl₃ added upon induction. This aspect was studied to ensure sufficient supply of the heme precursor and making sure the heme synthesis was not a limiting step in the enzyme production. The SDS gel shows the sucrose/EDTA fraction before and after dialysis (lane 1, 2; lane 9, 10). The corresponding immunoblots showed that for the sample with ALA/FeCl₃ the amount of protein decreases after dialysis; no precipitation had been observed. The decreased intensity could be due to the increase in sample volume. No further signals could be obtained from this sample; neither in the flow-through nor in any of the fractions collect (lane 3, and lanes 4 to 7, respectively). The protein visible with a molecular weight around 45 kDa does not represent the target protein; this is confirmed by the missing response on the immunoblot. Both samples exhibited similar chromatograms; the fractions analysed on the gel were the same as the previously described ones containing protein (see Figure 2-12; the change in fraction number is due to a different collection method). The sample grown without the ALA/FeCl₃ additives showed slightly stronger bands on the immunoblot. Most of the sample was conserved during the dialysis step. The performed affinity chromatography showed only protein in the flow-through indicating that the secreted *AaeUPO* was still misfolded and the His-tag was not accessible or able to bind to the chromatography resin.

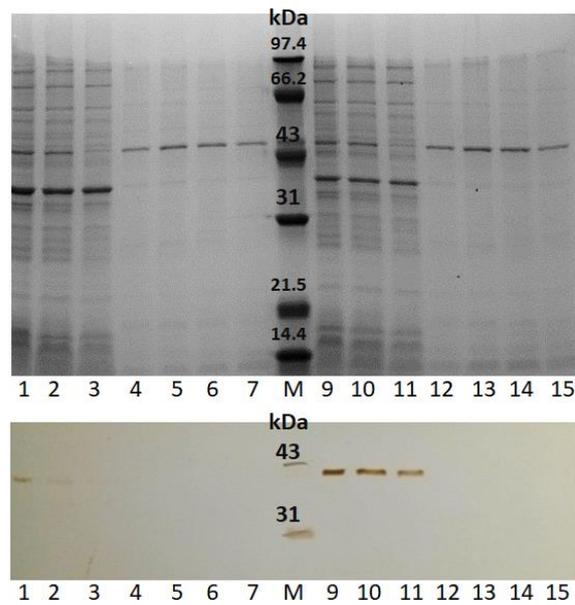


Figure 2-14: Purification of the secreted C-terminally tagged *AaeUPO* expressed by Rosetta2 in LB at 37 °C with added ALA/FeCl₃ (left to the marker) and without (right to the marker) ; 12 % SDS gel and immunoblot: 1 – sucrose fraction; 2 – Ni load (after dialysis); 3 – Ni flow-through; 4 – fraction 4; 5 – fraction 5; 6 – fraction 6; 7 – fraction 7; M - low molecular weight marker; 9 - sucrose fraction; 10 – Ni load (following dialysis); 11 – Ni flow-through; 12 – fraction 4; 13 – fraction 5; 14 – fraction 6; 15 – fraction 7.

In addition to the Ni-affinity chromatography, different purification methods were tested to isolate *AaeUPO*. The following studies were performed with the N-terminally tagged *AaeUPO* expressed by BL21 in M9-glc at 16 °C. Ammonium sulfate (AS) precipitations were performed in 20% steps to try and identify the correct concentration needed to obtain and isolate the target protein. The SDS gel and immunoblot of the 40% and 60% AS fractions is shown below (Figure 2-15; lane 1, 2). The majority of *AaeUPO* was precipitated using 40% AS, a small amount precipitated in the 60% AS fraction. The AS could be regarded as a first step in the process of isolating protein however a broad spectrum of other proteins was still detectable in the SDS gel. The analysed resuspended pellet was desalted using a G25 desalting column and purification *via* anion exchange chromatography (Q) was trialled (data not shown). No binding of the protein to the Q-column was observed, and protein was solely detected in the column flow-through. Further studies into anion exchange chromatography were also performed. For the following results, the cell pellet was resuspended in salt-free Tris buffer before cell lysis and the supernatant was loaded onto a Q-column. The chromatogram of the purification and the analysis of the fractions are shown in Figure 2-15 (lanes 4 to 11). The elution profile showed

overlapping peaks of various proteins without clear separation. This was confirmed by SDS gel analysis where high levels of impurity were found. The immunoblot analysis further showed, that the protein did not bind to the column and was located in the flow-through; a weak signal for fraction 12 (lane 9) was detected. These results are in agreement with the findings mentioned above attempting the purification of the AS fractions. The low affinity towards the anion exchange resins could be due to aggregation of the protein, which would inhibit interaction with the resin. In the chromatogram, the peak eluting between 80 and 100 mL corresponds to nucleic acid fragments, which are known to elute at higher salt concentrations.

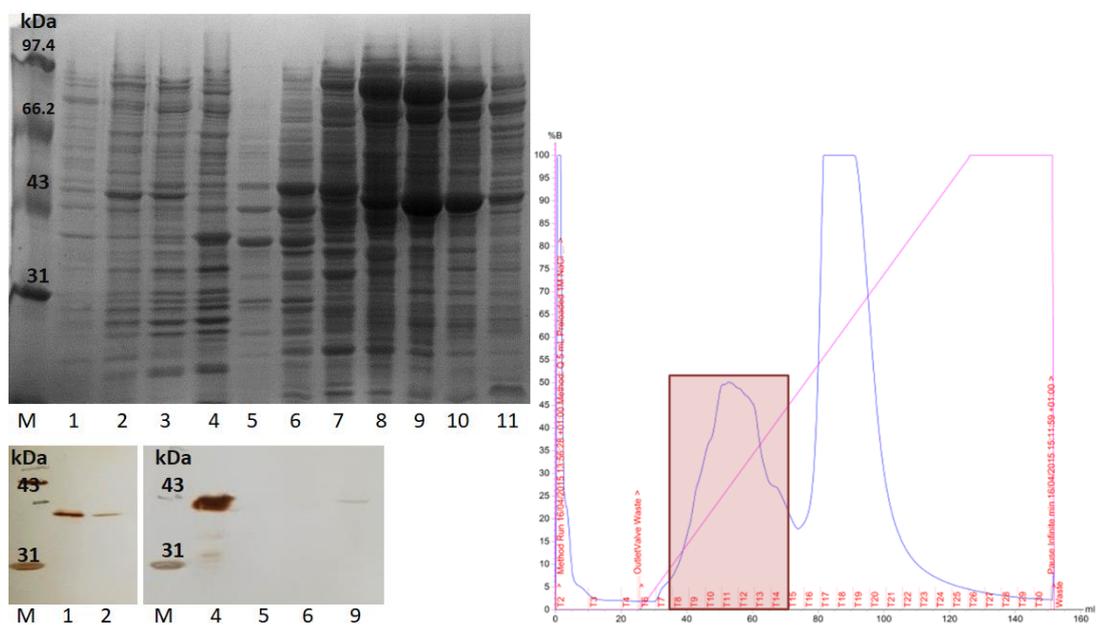


Figure 2-15: left: 12 % SDS gel and immunoblot of the ammonium sulfate precipitation and anion exchange purification of the N-terminally tagged *AaeUPO* (BL21 in M9-glc at 16 °C): M - low molecular weight marker; 1 – 40% AS fraction; 2 – 60% AS fraction; 4 – Q-column flow-through; 5 – Q-fraction 8; 6 - Q-fraction 9; 7 - Q-fraction 10; 8 - Q-fraction 11; 9 - Q-fraction 12; 10 - Q-fraction 13; 11 - Q-fraction 14; right: Elution profile obtained for the purification of *AaeUPO* using anion exchange chromatography; highlighted the fractions run on the gel (8 - 14).

2.5. Summary, Discussion and Conclusion

This chapter summarises the successful cloning of two unspecific peroxigenase homologs, from *Agrocybe aegerita* and from *Agaricus bisporus*, into three different bacterial cloning vectors. Constructs with an N-terminal or C-terminal His-tag were created, and the addition of a periplasmic secretion sequence for one of the constructs allowed studies with secreted protein (overview given in Figure 2-3). The His-tag was chosen to allow purification of the protein *via* affinity chromatography; the secretion sequence was to mimic the secretion of the protein as it occurs in the native fungal strain. All constructs were successfully transformed into various *E. coli* strains and the gene expression was studied in different growth media.

First expression tests performed here were run for 24 h and identified a small number of conditions able to produce soluble *AaeUPO* and *AbiUPO*. In contrast to this, a PhD thesis published in 2013 by Wang, where *AaeUPO* was cloned into a bacterial expression vector using the outer membrane protein A (OmpA) as signalling peptide, found the protein to be solely located in catalytically inactive inclusion bodies.¹¹⁶ Further knowledge about the amount of soluble protein was obtained through time course experiments using the *AaeUPO* homolog. It was found that an increased number of conditions yielded in soluble protein, from eight conditions found in the first 24 h test to 29 conditions after the time-resolved measurements. Especially at higher temperatures the amount of nutrients decreased over time and degradation of the protein of interest was preventing the cells from going into starvation. Studies focussed on the bacterial expression of *CfuCPO*, as structurally similar protein to UPOs, previously reported proteolytic degradation of the protein over time.¹¹⁷ This phenomenon was also reported to be temperature dependent, where degradation started earlier at higher expression temperatures. In this work, degradation was mainly observed for the N-terminally tagged protein, located within the cell. The secreted protein generally showed lower expression levels and amounts of soluble protein, when directly compared to the N-terminally tagged variant. The best condition to obtain soluble protein was consistent among the two constructs and found to be Rosetta2 in LB at 37 °C, when harvested after approx. 6 h (see Table 2-21 and Figure 2-10).

Solubility studies were performed by scanning thirty buffers towards their ability to stabilise the protein in solution. The *AaeUPO* and the *AbiUPO* homolog were studied together. While the *AaeUPO* showed good solubility in two thirds of the tested buffers with excellent results in about ten of these, *AbiUPO* shows lower levels of solubility throughout, with only one promising result (see Figure 2-11 and Table 2-22). Clear trends towards buffers used, pH preferences or salts and additives were not present.

Having identified good conditions to express soluble protein, scale-up experiments were performed, and the purification of the protein was attempted. The N-terminal His-tag *AaeUPO* was tested using BL21, M9-glc at 16 °C and also alongside the secreted C-terminal His-tag construct using Rosetta2, in LB at 37 °C. All conditions showed the expression of soluble protein. When compared to the expression tests however, the ratio between insoluble and soluble protein was leaning more towards the insoluble fraction (see Figure 2-13). It is possible that the scale of the expression system, which is coupled to various factors like the aeration, and the method of cell disruption influenced the production and isolation of soluble protein. It is further possible, that protein aggregated in inclusion bodies, which yielded in an inactive form of protein, as it was reported before for heme-containing peroxidases^{114, 117} and *AaeUPO*.¹¹⁶ The soluble protein which was obtained was subject to purification trials; the first ones focussed on Ni-affinity chromatography. Attempts were made with the different constructs, and as well as with the protein extracted from the periplasm. All experiments showed the same results; poor binding of the protein to the resin (see Figure 2-12, Figure 2-13, and Figure 2-14). These findings suggested the misfolding or aggregation of protein. The majority of the protein was found in the flow-through and only small amounts of the N-terminal His-tag *AaeUPO* expressed by BL21 showed any binding to the resin (see Figure 2-12). When studying the flow-through and fractions obtained using UV/Vis spectrophotometry, the wavelength scans showed a strong absorption band around 420 nm, similar to the typical ferric *AaeUPO* absorption band reported in previous works.⁸⁴ This points towards the successful construction of the active site, carrying the reported heme-thiolate. First tests analysing the activity of the obtained *AaeUPO* towards different substrates reported no activity (data not shown). Other studies have reported that it is possible to achieve

heme-binding in peroxidases expressed in *E. coli* without retaining the full catalytic activity.¹¹⁷ Further purification methods which were tested include ammonium sulfate precipitation and anion exchange chromatography (see Figure 2-15). AS concentrations around 40% were necessary to precipitate the majority of the protein. A successful purification of the *Aae*UPO was not accomplished using protein expressed by various *E. coli* strains. Anion exchange chromatography, similar to the results obtained for the Ni-affinity chromatography, showed no binding to the resin. As highly glycosylated proteins, UPOs might not be able to fold in a catalytically active form without the post translational modifications. *Cfu*CPO, which is also post translationally modified, was reported to express in its apo-form, when expressed in *E. coli*.¹¹⁷ A tedious high-pressure assisted refolding procedure was established, yielding in 1% refolded and active holoenzyme. While the catalytic activity was altered compared to the native enzyme, these experiments showed that glycosylation was not essential for *Cfu*CPO activity. However, this might not be the case for UPOs.

Bringing this to a conclusion, it was possible to express soluble UPOs using bacteria. However, the protein displays signs of misfolding and preliminary activity test were negative, suggesting that *E. coli* is not able to produce functional unspecific peroxygenases. The utilisation of eukaryotic expression system, capable of post translational modifications, would be a possibility to gain access to UPOs. *Aspergillus niger* was reported to produce functionally active fungal metalloproteins, including the afore mentioned *Cfu*CPO^{118, 119} and in 2013 Babot and co-workers reported the use of an UPO from *Coprinopsis cinerea* which was the first unspecific peroxygenase to be heterologously expressed, using *Aspergillus oryzae* (Novozymes patent WO/2008/119780).⁷⁹

3. Yeast Expression and Characterisation of the Unspecific Peroxygenase from *Agrocybe aegerita*

Following on from expression studies in bacteria further hosts were considered. There are limited cases described in the literature where UPOs have been expressed in heterologous hosts. *Cci*UPO was functionally expressed using a fungal host, *Aspergillus oryzae* (Novozymes patent WO/2008/119780), and more recently, in 2014, a study was published looking at directed evolution experiments to allow the functional expression of *Aae*UPO in *Saccharomyces cerevisiae*.^{79, 120} Within the latter study, a nine-point mutant was identified (*rAae*UPO), which showed good expression in the yeast host, with an increase in secreted protein from 0.007 mg L⁻¹ for the wild-type to 8 mg L⁻¹ for the mutant. Four of the point mutations were located within the signalling peptide of *Aae*UPO, while the other five mutations were found on the mature protein. The characterised *rAae*UPO showed very similar spectrophotometric and kinetic behaviour when compared to the wild-type *Aae*UPO. In 2015, the same group went on to publish a study in which they were able to show that *rAae*UPO is also functionally expressed in *Pichia pastoris*, with secretion levels as high as 217 mg L⁻¹.

As *Pichia* is a well-studied laboratory expression system, which compared to *S. cerevisiae* allows for higher cell density and increased protein production, it was chosen for establishing a heterologous expression route of *Aae*UPO in this work. The introduction of purification anchors was one aim to be addressed here. Furthermore, a full characterisation and the crystallisation of *rAae*UPO will be described, revealing novel insight into this unspecific peroxygenase.

3.1. Materials and Methods

3.1.1. Chemicals and Materials

Chemicals used in this study were purchased from Alfa Aesar (Heysham, UK), Fisher Scientific UK Ltd. (Loughborough, UK), Fluorochem Ltd. (Glossop, UK), Geneflow Ltd. (Lichfield, UK), Generon (Slough, UK), Insight Biotechnology Ltd. (Wembley, UK), Merck Chemicals Ltd. (Nottingham, UK), Scientific Laboratory Supplies Ltd. (Nottingham, UK), Sigma-Aldrich Company Ltd. (Dorset, UK), Takara Bio Europe Clontech (St Germain-en-Laye, France), Tokyo Chemical Industry UK Ltd. (Oxford, UK), and VWR International Ltd. (Lutterworth, UK). Restriction enzymes were bought from New England Biolabs (Ipswich, UK), Promega UK Ltd. (Southampton, UK), and Thermo Fisher Scientific Biosciences GmbH (St. Leon-Rot, Germany). PCR Primers were synthesised by Eurofins Scientific (Wolverhampton, UK) and genes by GeneArt (now Life Technologies Ltd., Paisley, UK).

3.1.2. Molecular Biology

Unless stated otherwise all molecular biology work concerning the preparation of recombinant DNA molecules in bacterial samples was performed as described before (Section 2.1.2). Similarly, the details for agarose gel electrophoresis and SDS-PAGE analysis have been reported elsewhere.

3.1.2.1. Bacterial Transformation

For bacterial growth low-salt lysogeny broth (LB) was prepared according to the information presented below (Table 3-1). Low-salt LB agar plates were prepared with 20 g L⁻¹ agar added to the broth recipe. The solutions were made up with ultra-pure water and were autoclaved prior to use.

Table 3-1: Composition of low-salt LB used in bacterial experiments with the yeast vector pPICZ.

Low salt LB	
5 g L ⁻¹	Yeast Extract
10 g L ⁻¹	Tryptone
5 g L ⁻¹	NaCl

Stellar cells were transformed with the genes of interest using the heat-shock procedure described elsewhere (Section 2.1.2.1). The cells were incubated with

150 μL of low salt LB prior to plating out on low salt LB plates containing 25 $\mu\text{g mL}^{-1}$ Zeocin.

3.1.2.2. *Bacterial Starter Cultures*

Under aseptic conditions 5 mL low salt LB cultures with Zeocin (25 $\mu\text{g mL}^{-1}$) were prepared in 50 mL falcon tubes. A single colony was transferred from the LB-agar plate using a pipette tip. The tube was sealed and incubated overnight at 37 °C in an orbital shaker at 180 rpm.

3.1.2.3. *DNA Extraction from Bacterial Culture*

To extract DNA samples from cultures, 5 mL starter cultures were grown according to the protocol described. The DNA was extracted using the GenElute™ Plasmid Miniprep Kit. For large scale DNA extractions, 50 mL low salt LB culture was inoculated with 500 μL starter culture and grown to an OD_{600} around 2-3 before harvest and extraction using a GeneJET Plasmid Midiprep Kit (Thermo Scientific). The concentration was determined using a UV-spectrophotometer with water as reference. The water was removed *in vacuo* and the dried sample was stored at -20 °C until further utilisation.

3.1.2.4. *Yeast Transformation*

The plasmid of interest was linearized for transformation using *PmeI* according to the specifications given below (Table 3-2). After inactivation the reaction was purified using the NucleoSpin® Gel and PCR Clean-up kit (MACHEREY-NAGEL GmbH & Co. KG), the concentration was determined and the solvent was removed as described above (3.1.2.3).

Table 3-2: Pipetting scheme and temperature settings for the *PmeI* digest of DNA.

DNA	15 μg			
<i>PmeI</i> (10,000 U mL^{-1})	10 μL			
Cutsmart buffer (10X)	5 μL	37 °C	1 h	Digest
ddH ₂ O	To 50 μL	65 °C	20 min	Inactivation

For the yeast transformation and further experiments, yeast extract-peptone based media (YP) was prepared according to information provided (see Table 3-3). Agar plates were prepared with 20 g L^{-1} agar added to the respective broth recipe. All

solutions were prepared using ultra-pure water and were autoclaved prior to use. Glucose was added to the medium after autoclaving *via* sterile filtration.

Table 3-3: Composition of the different growth media used in yeast experiments.

YPD		YPDS	
10 g L ⁻¹	Yeast extract	10 g L ⁻¹	Yeast extract
20 g L ⁻¹	Peptone	20 g L ⁻¹	Peptone
2% (w/v)	Glucose	2% (w/v)	Glucose
		182.2 g L ⁻¹	Sorbitol

The preparation of the *P. pastoris* strain X-33 followed the instructions provided in the EasySelect™ Pichia Expression Kit User Manual (Manual part no. 25-0172, Invitrogen). A glycerol stock was used to inoculate X-33 on an YPD agar plate and incubated for 72 h to obtain single colonies. A 5 mL YPD culture was inoculated with a single colony from the selection plate and incubated at 30 °C overnight. The culture was then used to inoculate 100 mL YPD in a baffled 500 mL flask in a 1:100 ratio. The culture was incubated at 30 °C until it reached an OD₆₀₀ of 1.3-1.5. The cells were collected (5 min, 1,500 × g, 4 °C) and resuspended in 100 mL ice-cold, sterile water. The cells were centrifuged as described and the resulting pellet was resuspended in 50 mL of ice-cold, sterile water. After centrifugation (as described before) the cell pellet was resuspended in 5 mL ice-cold sorbitol (1 M, sterile) and a final centrifugation was performed. The final pellet was resuspended in 200 µL ice-cold sorbitol (1 M, sterile). A transformation mixture containing 80 µL of cells and 5-10 µg of linearized DNA (made up to 1 µg µL⁻¹ with sterile water) was transferred to a cooled 0.2 cm electroporation cuvette and incubated on ice for 5 min. The cells were pulsed using a Gene Pulser II, fitted with a Capacitance Extender Plus and a Pulse Controller Plus (BioRad) with the capacitance set to 25 µF, the resistance set to 200 Ω and the volt setting at 2.0 kV. Ice-cold sorbitol (1 mL, 1 M) was added to the cuvette and the contents were transferred to a 15 mL sterile tube and incubated at 30 °C without shaking for 1-2 h. 50-150 µL of cells were spread on YPDS plates containing Zeocin (100 µg mL⁻¹) and incubated at 30 °C for 72 h or until colonies appeared. 5-10 colonies were selected, streaked for single colonies on YPD(S) plus Zeocin (100 µg mL⁻¹) plates and incubated at 30 °C for 2-3 days.

3.1.3. Gene Design and Synthesis

3.1.3.1. *apo1* from *Agrocybe aegerita* (B9W4V6)

For the expression of *AaeUPO* in yeast, a literature described 9-point mutant of the *apo1* gene was obtained. The following mutations are located on the signalling peptide: Phe12Tyr, Ala14Val, Arg15Gly and Ala21Asp. The other five mutations are located on the mature protein: Val57Ala, Leu67Phe, Val75Ile, Ile248Val and Phe311Leu. The full sequence is shown below (Table 3-4). The primers used to create a construct carrying a C-terminal 3C cleavage site and primers to prepare an N-terminally His-tagged protein with a 3C site are further presented below.

Table 3-4: Gene sequence and list of primers used for amplification of *AaeUPO* creating a C-terminal 3C site and an N-terminal His-tag and 3C cleavage site.

AaeUPO 9-point mutant for yeast expression

```
ATGAAATATTTTCCCCTGTTCCCAACCTTGGTCTACGCAGTGGGGGTCGTTGCTTTTCCT
GACTACGCCTCATTGGCCGGCCTCAGCCAGCAGGAATTGGACGCTATAATCCCAACACT
CGAGGCCCGAGAGCCAGGATTACCTCCTGGTCCTCTCGAGAATAGCTCTGCAAAGTTGG
TGAACGACGAGGCTCACCCATGGAAGCCGCTTCGACCTGGCGATATTCGTGGACCTTGC
CCTGGTCTCAATACTCTGGCATCTCACGGGTACCTCCCGAGAAATGGCGTTGCAACCCC
GGCGCAAATAATAAACGCGGTTTCAGGAAGGATTCAATTTGACAATCAAGCCGCAATCT
TCGCCACATATGCGGCCACCTTGTGGACGGCAATCTCATTACGGACTTGCTGAGCATC
GGACGCAAGACGCGGCTCACTGGGCCTGATCCACCACCCCCGCTTCCGTTGGTGGACT
CAATGAGCATGGCACCTTCGAAGGCGACGCCAGTATGACCCGAGGTGACGCATTCTTTG
GCAACAACCACGATTTCAATGAGACGCTCTTCGAACAGTTGGTTGACTACAGCAACCGA
TTTGGAGGAGGAAAATACAATCTTACCGTCGCGGGGGAGCTCCGTTTCAAGCGCATTCA
AGACTCCATTGCGACCAACCCCAATTTCTCCTTTGTTGACTTTAGGTTCTTTACTGCTTAC
GGCGAGACCACCTTCCCCGCGAATCTTTTTGTGGATGGGCGCAGGGACGACGGCCAGC
TAGATATGGATGCTGCACGGAGTTTTTTCCAATTCAGCCGTATGCCTGACGATTTCTTCC
GCGCACCCAGCCCCGAGAAGTGGCACAGGAGTCGAGGTAGTTGTACAGGCTCATCCTAT
GCAGCCCCGAAGAAATGTCGGCAAGATCAACAGCTACACCGTCGACCCAACATCCTCTG
ACTTTTCCACCCCCTGCTTGATGTACGAGAAATTCGTC AACATAACGGTCAAGTCACTCT
ACCCGAATCCGACGGTGCAGCTTCGCAAAGCCCTTAATACGAATCTCGATTTCTTATTCC
AGGGAGTCGCCGCTGGATGTACCCAGGTCTTCCCATACGGGCGAGATTGA
```

Creation of C-terminal 3C-site (I)

Forward	ATGAAATATTTTCCCCTGTTCCCAAC
Reverse	TGGTCCCTGGAACAGAACCTCGAGATCTCGCCCGTATGGGAAGAC

Creation of C-terminal 3C-site (II) and In-Fusion overhangs

Forward	TAATTATTCGAAACGATGAAATATTTTCCCCTGTTCCCAAC
Reverse	ATGATGGTTCGACGGCTGGTCCCTGGAACAGAACCTC

Creation of N-terminal 3C-site & His (I)

Forward	CATCACCATCTCGAGGTTCTGTTCCAGGGACAAGAGCCAGGATTACCTCC
Reverse	ATCTCGCCCGTATGGG

Creation of N-terminal His (II) & In-Fusion overhangs

Forward	ACACTCGAGGCCCGAGAGTCCTCCCATCACCACCATCACCATCTCGAGG
Reverse	ATGATGGTTCGACGGCTCAATCTCGCCCGTATGGG

Creating In-Fusion overhangs for pPICZ-His3C

Forward	CTGTTCCAGGGACAAGAGCCAGGATTACCTCC
Reverse	ATGATGGTTCGACGGCTCAATCTCGCCCGTATGGG

3.1.4. Vector Design and Production

3.1.4.1. *pPICZ α*

Inverse PCR was used to amplify the pPICZ α B vector. The primers were designed to delete the N-terminal α -factor signal sequence as well as the C-terminal *c-myc* epitope (primer sequences shown below, inverse PCR creating pPICZ). Further experiments were included the generation of a pPICZ variant carrying a C-terminal 3C-His site (pPICZ-3CHis) and another variant with an N-terminal His-3C (pPICZ-His3C). All amplicons were *DpnI* treated, isolated on a gel and extracted.

Table 3-5: Primers used for inverse PCR amplification of pPICZ deleting the C-terminal signal sequence and creating an N- or C-terminal 3C protease site.

Creation of pPICZ	
Forward	GCCGTCGACCATCATCATC
Reverse	CGTTTCGAATAATTAGTTGTTTTTG
Creation of pPICZ-3CHis (C-terminal)	
Forward	GCCGTCGACCATCATCATCATC
Reverse	TCGGCCTCGAGTGTGGG
Creation of pPICZ-His3C (N-terminal)	
Forward	GCCGTCGACCATCATCATCATC
Reverse	TTGTCCCTGGAACAGAACCTCG

3.1.5. Cloning and Analysis

3.1.5.1. PCR Amplification

Genes were amplified using the following procedure. All components were pipetted on ice in the given order. For the amplification of the *P. pastoris* compatible vectors as well as for the amplification of *AaeUPO* creating In-Fusion overhangs for pPICZ-His3C a gradient PCR was run; ranging from 50-65 °C. The extension time for vector amplifications was increased to 3 min. All amplified constructs were isolated on 1% agarose gels and excised as described before (Section 2.1.2.4).

Table 3-6: left: Components and amounts used in PCR amplification reactions. right: Temperature, time, and cycle details describing the PCR amplification program.

ddH ₂ O	21 µL					
Q5 buffer (5X)	10 µL					
Reaction enhancer (5X)	10 µL	98 °C	30 s	Initial denature		
dNTPs (2 mM)	5 µL	98 °C	10 s	Denature	Repeat 35 x	
Template	1 µL	60 °C	30 s	Anneal		
Forward primer (20 µM)	1.25 µL	72 °C	30 s	Extend		
Reverse primer (20 µM)	1.25 µL	72 °C	3 min	Final extension		
Q5 Polymerase	0.5 µL	10 °C	∞	cool		

In-Fusion cloning was chosen for the integration of the gene of interest into the target vector. The annealed product was used to transform cloning cells using the heat-shock procedure. All experimental details are described elsewhere (see Section 2.1.5.3).

3.1.5.2. Colony-PCR

From the low salt LB-agar plates of the annealed and transformed plasmid, multiple colonies were screened. A colony was picked and transferred to a new LB-agar plate/quadrant of a plate before washing the tip in the colony PCR mix (see Table 3-7). The forward and reverse sequencing primers (AOX1, described in 3.1.5.3) were used for this experiment. The PCR program was run according to the information provided to amplify the target gene and the samples were analysed on 1% agarose gels. The newly prepared plates were grown at 37 °C. Colonies which gave a positive amplification result were then used to prepare a starter culture. The DNA from the culture was extracted using a plasmid miniprep kit (Sigma or New England Biolabs).

Table 3-7: left: Components and amounts used in the colony-PCR amplification; right: Temperature, time, and cycle details describing the PCR program.

	ddH ₂ O	18.65 µL				
	Taq buffer (10X)	2.5 µL	94 °C	2 min	Initial denature	
	MgSO ₄ (50 mM)	0.75 µL	94 °C	30 s	Denature	Repeat 25 x
	dNTPs (2 mM)	2.5 µL	60 °C	30 s	Anneal	
	Forward primer (20 µM)	0.25 µL	72 °C	70 s	Extend	
	Reverse primer (20 µM)	0.25 µL	72 °C	2 min	Final extension	
	Platinum-Taq	0.1 µL	10 °C	∞	cool	

3.1.5.3. DNA Sequencing

After the annealing was confirmed the gene was sequenced to ensure the target sequence has not been altered. 20 µL of the extracted DNA (30-50 ng µL⁻¹) were send to GATC for sequencing along with 20 µL of the following primers (10 µM, Table 3-8). The primers were designed to match those described in the EasySelect™ *Pichia* Expression Kit User Manual (Manual part no. 25-0172, Invitrogen) for reads of pPICZ and vector variants.

Table 3-8: Primers designed for sequencing experiments in pPICZ vector variants.

AOX1

Forward	GACTGGTTCCAATTGACAAGC
Reverse	GCAAATGGCATTCTGACATCC

3.1.6. Protein Production and Analysis

For the fermentation of protein using *Pichia pastoris* on various scales the following solutions were prepared. For the *P. pastoris* based test expressions the cells were

grown in buffered glycerol complex medium (BMGY) and expression was induced using buffered methanol complex medium (BMMY). The preparation instructions for the solutions are given below. The 2X stock solution of yeast extract (YE, 10 g L⁻¹) and peptone (P, 20 g L⁻¹; final solution YEP), the potassium phosphate buffer (KPi, 1 M, pH 6.0), glycerol (10% v/v) and water were sterilised by autoclaving before use. The remaining components, yeast nitrogen base (YNB, 10X, 134 g L⁻¹), biotin (500X, 200 mg L⁻¹) and methanol (5% v/v) were filter-sterilised prior to use.

Table 3-9: List of components and amounts needed for the preparation of buffered glycerol and buffered methanol complex medium for test-expressions using *P. pastoris*.

Component	Stock	BMGY	BMMY	Final conc.
YEP	2X	25 mL	50 mL	1% YE, 2% P
KPi	1 M	5mL	10 mL	100mM KPi, pH 6.0
YNB	10X	5 mL	10 mL	1.34% YNB
Biotin	500X	0.1 mL	0.2 mL	4x10 ⁻⁵ %
CasA		0.5 g	1.0 g	1%
Glycerol	10%	5ml	-	1%
MeOH	5%	-	10 mL	0.5%
ddH ₂ O		9.9 mL	29.8 mL	
Total volume		50 mL	100 mL	

The following solutions were prepared for fed-batch fermentation experiments. The basal salts were added to the fermentation vessel and autoclaved prior to use. The PTM₁ salts were sterile filtered and stored covered in foil until use.

Table 3-10: Composition of basal salt and PTM₁ trace salts used in the fermentation of *P. pastoris*.

Basal Salts		PTM ₁ Trace Salts	
26.7 mL L ⁻¹	H ₃ PO ₄ , 85%	6.0 g L ⁻¹	CuSO ₄ × 5 H ₂ O
1.17 g L ⁻¹	CaSO ₄ × 2 H ₂ O	0.08 g L ⁻¹	NaI
18.2 g L ⁻¹	K ₂ SO ₄	3.0 g L ⁻¹	MnSO ₄ × H ₂ O
14.9 g L ⁻¹	MgSO ₄ × 7 H ₂ O	0.2 g L ⁻¹	Na ₂ MoO ₄ × 2 H ₂ O
4.13 g L ⁻¹	KOH	0.02 g L ⁻¹	H ₃ BO ₃
40.0 g L ⁻¹	Glycerol	0.5 g L ⁻¹	CoCl ₂
		20.0 g L ⁻¹	ZnCl ₂
		65 g L ⁻¹	FeSO ₄ × 7 H ₂ O
		0.5 g L ⁻¹	Biotin
		5.0 mL L ⁻¹	H ₂ SO ₄

3.1.6.1. 48-Well Test-Expression

P. pastoris based test-expressions were performed in 48-deepwell blocks. Up to 24 samples were analysed at a time. 2 mL buffered glycerol complex medium (BMGY) were inoculated with freshly grown colonies from selection plates. The block was sealed with Airpore tape to avoid contamination and incubated at 30 °C in a block shaker (660 rpm, Labnet 56 Vortemp) overnight. The OD₆₀₀ was monitored until it reached a value between 10-20 AU. At this point glycerol stocks were prepared (100 µL sample mixed with 100 µL 80% glycerol). 2 mL of fresh BMMY were aliquoted into the remaining 24 well and inoculated with 100 µL of BMGY overnight culture. The solutions were mixed and a t = 0 h sample was taken; 100 µL were centrifuged (5 min, 5,000 × g) and the supernatant and pellet were stored separately at -80 °C. The 48-well block was returned to the block shaker for 24 h. The next sample was taken (t = 24 h) and 100 µL of 10% methanol in BMMY were added. The same procedure was followed to obtain the t = 48 h and t = 72 h samples. For the final sampling the volume was increased to 1.5 mL. All samples can be analysed by SDS-PAGE, western blot and dot-blot analysis.

3.1.6.2. Scale-up Gene Expression in Shake Flasks

Scale-up expression was routinely performed in 100 mL buffered complex medium. A 5 mL YPD starter culture was inoculated with a freshly grown colony (from glycerol stock replated on YPD(S), without antibiotic) and incubated overnight at 30 °C, 250 rpm. 50 mL of BMGY in a 500 mL baffled flask were inoculated with 1 mL of

starter culture and incubated at 30 °C, 250 rpm until the OD₆₀₀ reached 2. The cells were collected *via* centrifugation (5 min, 3,900 × g) and resuspended in 100 mL BMMY. A 500 µL t = 0 h sample was taken (5 min, 5,000 × g); supernatant and pellet were stored at -80 °C. The culture was incubated at 30 °C, 250 rpm. Twice daily 0.5 mL methanol (100%, syringe-filtered) were added and samples were taken at t = 24, 48 and 72 h. After 72 the cells were spun down (25 min, 3,900 × g) and the supernatant was filtered (0.8 µm pore size). The samples were stored at -80 °C or dialysed into salt free buffer for purification.

3.1.6.3. Fed-Batch Fermentation

The procedure used here followed the instructions provided in *Pichia* Fermentation Process Guidelines (Invitrogen). Either a 7 L glass vessel fermenter or 500 mL glass vessel fermenters were used in this work, charged with 2 L or 200 mL basal salts, respectively. The basal salts were prepared to the specifications given above and autoclaved before use (Table 3-10). Prior to autoclaving the pH electrode was calibrated using a two-point method (pH 4 and pH 7) and the electrolyte in the dissolved oxygen (DO) electrode was renewed. The vessel was fitted with three feed-tubes for the base-, antifoam- and carbon-source addition and the sampling port was fitted with either a sampling tube (7 L vessel) or a sampling syringe (500 mL vessel). The condenser and the air/O₂ lines were fitted with filters. The following set-points and limits were selected in the software (Table 3-11).

Table 3-11: Fermentation settings, limits and pump setting for the small scale (0.5 L) fermentations.

	Set-point	Limits	Pump output* /%
Temperature /°C	25-30	15-60	
pH	5	3-8	25
Stirrer		200-1750	
Airflow /mL min ⁻¹	100		
DO /%	>20	Stirrer, then O ₂	30-75
Condenser			75
Antifoam			15

* for fermentation in 500 mL vessels.

The airflow, condenser, stirring and temperature loops were started first after autoclaving. The antifoam tubing was primed and the PTM₁ trace salts were added

(4.35 mL L⁻¹). The fermentation log was started prior to adjusting the pH of the basal salts and calibrating the DO electrode (1-point method). Cells from an overnight culture in YPD, grown from a single colony picked from a fresh plate, were collected (10 min, 3,500 × g), resuspended in water (1 mL per 5 mL culture) and used to inoculate the fermentation medium in a 1:10 ratio. The culture was grown until all the glycerol was consumed, as indicated by an increase of the DO to 100%, and the glycerol fed-batch phase was initiated. A feed of 50% w/v glycerol containing PTM₁ trace salts (12 mL L⁻¹) was started at a rate of 18.15 mL h⁻¹ L⁻¹ initial fermentation volume. The feed is continued for 4 h. Once all the glycerol was consumed, the methanol fed-batch phase was initiated. A feed of 100% methanol with PTM₁ trace salts (12 mL L⁻¹) was started at 3.6 mL h⁻¹ L⁻¹ initial fermentation volume for 4 h or overnight. When the culture was fully adapted with a steady DO read and a fast response after stopping the methanol feed, usually less than 2 min; the feed rate was increased to 7.3 mL h⁻¹ L⁻¹ initial fermentation volume. When the culture had adapted to the new rate the feed rate could be further increased to 10.9 mL h⁻¹ L⁻¹ initial fermentation volume. The entire methanol fed-batch process took approx. 72 h. The supernatant was collected upon centrifugation (20 min, 10,000 × g) and filtered (0.8 μm) before storage of the solution at -80 °C. Samples were taken twice daily for the duration of the entire fermentation process to determine the wet-cell weight and for SDS-PAGE analysis. The pellet and the supernatant were separated *via* centrifugation (2 min, 16,300 × g) and stored at -80 °C.

3.1.6.4. Immunoblot Analysis (Western Blot)

SDS-PAGE experiments were run as described before. For immunoblot analysis gels were run with Precision Plus Protein™ Dual Color Standards (Bio-Rad). The procedure for the blotting and blocking of the membrane has been described earlier (Section 2.1.7.8). The staining step with Ponceau S could be eliminated due to the use of a pre-stained standard. The visualization was achieved with either 3,3'-diaminobenzidine tablets or the luminol-based Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare) based on chemiluminescence reacting with the monoclonal anti-polyhistidine-peroxidase antibody.

3.1.6.5. Dot-Blot Analysis

Dot-blot analysis is based on the same principles as the western blot described earlier. While the samples for the western blot analysis are routinely denatured prior to the separation on the gel, native samples are used in this experiment and immobilised on the membrane rather than being transferred from a gel. A PVDF membrane was activated with 100% methanol (2 min) before incubation in transfer buffer for 15 min prior to use. The apparatus was assembled as shown below (Figure 3-1), connected to a vacuum pump and 100 μ L of resuspension buffer were run through the wells. Into each well 50 μ L of sample or buffer were added and the solvent was pulled through. The functionalisation was followed by a wash step using 50 μ L buffer and the membrane was dried under vacuum. The membrane was then blocked with 5% milk in TBST for 1 h, followed by incubation with a monoclonal anti-polyhistidine-peroxidase antibody in 5% milk in TBST for 1 h. The visualisation was achieved using the Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare).

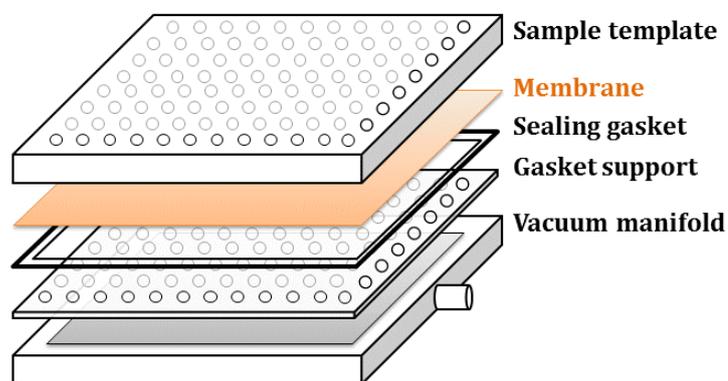


Figure 3-1: Set-up of the dot-blot apparatus.

3.1.7. Protein Purification and Characterisation

Different purification methods were tested to identify the most efficient way to purify the *P. pastoris* expressed and secreted rAaeUPO. In the presence of a His-tag the protein was purified *via* nickel affinity chromatography followed by a gel-filtration column. In the absence of the affinity tag, the protein was purified *via* the route shown below (Figure 3-2). The protein solution was brought to a 40% ammonium sulfate (AS) saturation and purified using hydrophobic interaction chromatography (HIC). After dialysis the protein was then purified using an anion exchange (Q) step, dialysed and applied to a gel-filtration column. All

chromatography steps were performed using FPLC equipment, monitoring the absorbance at 280, and when possible 420 nm.

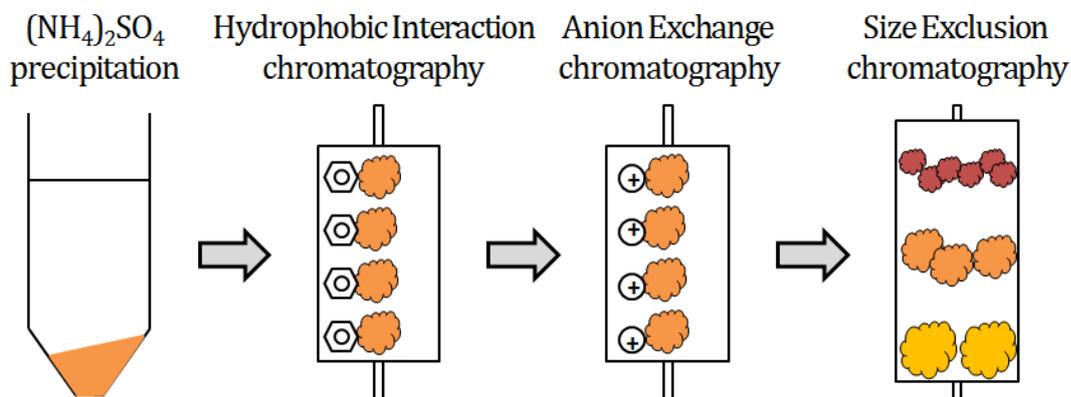


Figure 3-2: Purification tree applied in the purification of *AaeUPO* in the absence of a His-affinity tag.

The following buffers were prepared for protein isolation and purification (Table 3-12). All are based on Tris-base at pH 7.0, with variables salts and concentrations depending on the purification method. All were made up with ultra-pure water and filter over 0.2 µm filters prior to use.

Table 3-12: Composition of buffers used in the protein preparation and purification.

Resuspension Buffer	Hydrophobic Interaction Buffer
20 mM Tris-Cl	20 mM Tris-Cl
40 mM NaCl	40% (NH ₄) ₂ SO ₄
pH 7.0	pH 7.0
Ni-Purification Buffer	Anion-Exchange Buffer
20 mM Tris-Cl	20 mM Tris Cl
40 mM NaCl	2 M NaCl
500 mM Imidazole	pH 7.0
pH 7.0	
Salt-free Buffer	
20 mM Tris-Cl	
pH 7.0	

3.1.7.1. Ni-Affinity Chromatography

HisTrap FF Crude columns (1 mL or 5 mL column volume, GE Healthcare) were routinely used for purification of protein *via* nickel affinity chromatography. The column was equilibrated with resuspension buffer (5 CV) and loaded with the fermentation supernatant using the FPLC system (NGC™ Chromatography System, 92

Bio-Rad). The flow-through was collected for SDS gel analysis. Two washing steps with no and 30 mM imidazole were run over 5 CV each, before a gradient was run from 30 to 500 mM imidazole over 10 or 20 CV. Fractions of interest were collected and analysed. The column was regenerated following the procedure provided by the supplier.

3.1.7.2. Hydrophobic Interaction Chromatography

The fermentation supernatant was dialysed overnight into resuspension buffer and was subject to an ammonium sulfate (AS) precipitation step using 40% w/v AS. The precipitate was collected *via* centrifugation (10 min, 3,900 × g) and the supernatant was loaded onto an equilibrated HiTrap Phenyl HP (1 or 5 mL, GE Healthcare, equilibrated with 5 CV HIC buffer) using the NGC™ Chromatography System equipped with a sample pump (Bio-Rad). The column was washed with 10-15 CV HIC buffer before initiating a gradient from 40% AS to salt-free buffer over 15 CV. Flow-through, wash and fractions were collected and analysed *via* SDS-PAGE and activity assays. The column was regenerated following the procedure provided by the supplier. Fractions of interest were pooled and dialysed overnight into salt-free buffer.

3.1.7.3. Anion Exchange Chromatography

Samples for anion exchange chromatography were prepared in salt-free buffer. The protein was loaded onto the equilibrated Q-Sepharose Fast Flow column (5 mL, GE Healthcare, equilibrated with 5 CV salt-free buffer) using the NGC™ Chromatography System equipped with a sample pump (Bio-Rad). The column was washed with 10-15 CV salt-free buffer and the elution was achieved by running a sodium chloride gradient. Over 10 CV the salt concentration was increased to 300 mM, followed by an increase to 2 M salt over 4 CV. The column was washed and stored according to the information provided by the supplier. Flow-through, wash and fractions were collected for analysis. The fractions of interest were pooled and dialysed into resuspension buffer overnight.

3.1.7.4. Size Exclusion Chromatography

Prior to loading the sample onto the size exclusion chromatography column, the volume was reduced to approx. 500 μL for analysis using the Superdex 200 10/300 GL column (GE Healthcare) and 2 mL for purification on HiLoad 16/600 Superdex 200 prep grade columns (GE Healthcare) using Vivaspin 20 concentrators with a 30 kDa cut-off. The sample was loaded *via* an appropriate loop onto the column equilibrated with resuspension buffer. The elution was achieved by running 1.2 CV of buffer through the column. The column was cleaned and stored according to the information provided by the supplier. Fractions of interest were collected, analysed and pooled where applicable. The volume was reduced to achieve higher protein concentration using Vivaspin 20 concentrators and samples were flash-frozen prior to storage at $-80\text{ }^{\circ}\text{C}$.

3.1.7.5. Protein Characterisation using UV/Vis Spectrometry

UV/Vis wavelength scans were recorded to characterise the protein, taking advantage of the heme-thiolate absorbance maximum at 418 nm. Baseline corrections were performed. The scan range was from 250 to 600 nm, with a 1 nm s^{-1} read. The Reinheitszahl (Rz) of the protein was determined through the ratio of the absorbance at 418 nm and the absorbance at 280 nm (A_{418}/A_{280}). The protein concentration was determined either using the absorption reading at 280 nm with $\epsilon_{280} = 36,000\text{ M}^{-1}\text{ cm}^{-1}$ or using the absorption read at 420 nm with $\epsilon_{420} = 115,000\text{ M}^{-1}\text{ cm}^{-1}$.

3.1.7.6. UV/Vis based Specific Activity Assays

The following buffers were prepared for activity assays performed with UPOs (Table 3-13); the difference in pH allows for separation of peroxidase (pH 4.4, one-electron transfer) and peroxygenase (pH 7.0, two-electron transfer) activity. The qualitative analysis of fermentation or purification samples to identify UPO containing fractions was performed without using a spectrophotometer. All reactants were mixed together and any observed colour change was compared to the control samples.

Table 3-13: Compositions for the buffers used in activity assays.

Phosphate Buffer		Citrate-Phosphate Buffer	
100 mM	KP _i	200 mM	Na ₂ HPO ₄
pH	7.0	100 mM	Citric acid
		pH	4.4

For the quantitative analysis of UPO containing samples, all assays were performed in 1 mL quartz glass cuvettes and buffer was run as blank. The peroxidase activity was determined using 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) as substrate. The aqueous reaction contained 50 mM citrate-phosphate buffer pH 4.4, 0.2 mM ABTS (from 5 mM aqueous stock), 10 µL enzyme sample (for concentrated samples 1 µL), and were started by the addition of 2 mM hydrogen peroxide (from 100 mM aqueous stock). The change in absorption was monitored at 418 nm for 1 min. The peroxygenase activity was determined using 1,2-(methylenedioxy)-4-nitrobenzene (NBD) as substrate. The aqueous reaction contained 50 mM phosphate buffer pH 7.0, 1 mM NBD (from 10 mM stock in acetonitrile), 10 µL enzyme sample (for concentrated samples 1 µL), and were started by the addition of 2 mM hydrogen peroxide (from 100 mM aqueous stock). The change in absorption was monitored at 420 nm for 1 min.

For the calculation of the specific activity the change in absorbance was plotted against the time, and the reaction rate k_{obs} (au min⁻¹) was determined. The enzyme activity, U (µmol min⁻¹), was calculated using Equation 3-1 and from that the specific activity, U mL⁻¹, was obtained using Equation 3-2.

Equation 3-1

$$U (\mu\text{mol min}^{-1}) = \frac{k_{obs} (\text{au min}^{-1})}{\varepsilon (\text{M}^{-1} \text{cm}^{-1}) \times l (\text{cm})} \times V_{rxn} (\mu\text{L})$$

Equation 3-2

$$\frac{U}{\text{mL}} = \frac{U (\mu\text{mol min}^{-1})}{V_{UPO} (\text{mL})}$$

The following parameters were used in the calculations.

Table 3-14: Parameters needed for the calculation of the specific activity towards ABTS and NBD.

ε_{418} (M ⁻¹ cm ⁻¹)	36,000
ε_{420} (M ⁻¹ cm ⁻¹)	9,700
V_{rxn} (µL)	1000
V_{UPO} (mL)	0.01 (0.001)

3.1.7.7. UV/Vis based Enzyme Kinetics

For the determination of the UPOs kinetic parameters, K_m and k_{cat} , for ABTS, NBD and veratryl alcohol (VA) assays were run with varying substrate concentrations (see Table 3-15).

Table 3-15: Substrate concentrations for ABTS, NBD, and VA used for the determination of K_m and k_{cat} .

ABTS (μM)	NBD (mM)	VA (mM)
10	0.1	1
25	0.25	2
50	0.5	4
75	0.75	8
100	1	10
150	1.5	20
200		40
500		60

As before, all assays were performed in 1 mL quartz glass cuvettes, with buffer as a blank. The reactions contained 50 mM buffer (phosphate at pH 7.0 for NBD and VA, and citrate-phosphate at pH 4.4 for ABST), 0.5 μL purified enzyme (3.9 $\mu\text{g mL}^{-1}$), 10% acetonitrile in the case of NBD, and were started with hydrogen peroxide (2 mM for ABTS and VA, 1 mM for NBD). The reactions were monitored for 1 min at 418 nm for ABTS, 420 nm for NBD, and 310 nm for VA.

Using Michaelis-Menten kinetics (Equation 3-3), k_{obs} was determined as described in 3.1.7.6 and plotted against the substrate concentration, K_m (mM) and V_{max} (au s^{-1}) were obtained from the graph. Equation 3-4 is for the determination of k_{cat} (s^{-1}), where E_T (M) describes the enzyme concentration; the units of V_{max} were converted into M s^{-1} by dividing V_{max} (au s^{-1}) by the extinction coefficient (see Table 3-14, for VA $\epsilon_{310} = 9,300 \text{ M}^{-1} \text{ cm}^{-1}$).

Equation 3-3

$$k = \frac{V_{max} \times [S]}{K_m + [S]}$$

Equation 3-4

$$k_{cat} = \frac{V_{max} (\text{M s}^{-1})}{E_T (\text{M})}$$

3.1.7.8. Deglycosylation using PGaseF

To determine the *N*-glycosylation degree of a protein, the 5-10 μg of purified protein were incubated with 1 μL glycoprotein denaturing buffer (10X) and the volume was

adjusted to 10 μ L. The solution was heated to 100 °C for 10 min, cooled and centrifuged for 10 s. To this, 2 μ L glyco-buffer 2 (10X), 2 μ L NP-40 (10%), and water was added to a total volume of 20 μ L. 1 μ L of PNGaseF was added and the reaction mixture was gently mixed, before incubation at 37 °C for 1 h. The obtained sample was compared to the untreated protein using SDS-PAGE analysis.

3.1.7.9. Mass Spectrometry Experiments

For protein identification experiments the protein of interest was run on a denaturing gel, the band of interest was excised and sent for analysis at the Technology Facilities, Metabolomics & Proteomics, University of York. There the sample underwent a trypsin digest before analysis of the fragments using MALDI-TOF. The results were analysed for sequence similarity using MASCOT.

LC-MS and linear MALDI-MS experiments were performed to characterise the glycoprotein. Purified protein samples were diluted to a concentration of 1 mg mL⁻¹ and sent for analysis at the Technology Facilities, Metabolomics & Proteomics, University of York. For the LC-MS experiments the protein was run over a C4 trap column using a short HPLC gradient, a maXis qTOF mass spectrometer was used for MS detection. Myoglobin and BSA were run as standards. In case of the linear MALDI-MS experiments, 1 μ L protein sample was applied to the steel target plate, followed 1 μ L 2,5-dihydroxybenzoic acid (10 mg mL⁻¹) in 50% aqueous acetonitrile containing 0.1% trifluoroacetic acid. Positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III in linear mode, over a mass range of m/z 10,000-60,000. flexAnalysis 3.3 was used to generate average masses using a sum detection algorithm S/N threshold of 1.75, a relative intensity threshold of 20%, and a minimum intensity threshold of 25.

3.1.8. Crystallisation of AaeUPO

Purified protein in 20 mM Tris pH 7.0 and 40 mM sodium chloride was used to set up initial 96-well crystallisation screens at 3 and 9 mg mL⁻¹; the screens used included Hampton Research Index and Crystal Screen HT, Molecular Dimensions Pact, and Qiagen Ammonium Sulfate. 150 nL mother liquor and 150 nL enzyme solution were pipetted together into a sitting drop using a mosquito robot. 24-well plates were set up using the hanging-drop technique for the best conditions with varying pH, salt and

PEG content. The following conditions were chosen for optimisation: A - 0.2 M sodium acetate, 0.1 M sodium cacodylate pH 6.5, 30% PEG-8k; B - 0.2 M sodium acetate, 0.1 M Tris pH 8.5, 30% PEG-4k; C - 0.2 M magnesium chloride, 0.1 M Tris pH 8.5, 30% PEG-4k. Scale-up crystallisations were further performed in a 48-well sitting drop format. The two best conditions were chosen (A - 0.2 M sodium acetate, 0.1 M sodium cacodylate pH 6.5, 30% PEG-8k; B - 0.2 M sodium chloride, 0.1 M Bis-Tris pH 6.5, 25% PEG-3350) with pH, salt and PEG content varying. The protein concentration was at 2 and 8 mg mL⁻¹. The co-crystallisation with ligands was attempted by incubating the protein with the ligand 30 min prior to setting up the screens and by performing soaking experiments. For the latter crystals were incubated with (*R*)-1-phenylethanol, NBD and hydrogen peroxide overnight.

Crystals were flash-frozen in liquid nitrogen. Full datasets for were collected on beamlines I04, and I04-1 at the Diamond Light Source, Didcot, UK. Data was processed and integrated using XDS and scaled using AIMLESS and POINTLESS within the Xia2 processing system.¹²¹⁻¹²⁵ Data collection statistics can be found in Table 3-16. The structures were solved with PHASER using the structure of *Aae*UPO (PDB code 2YOR) as molecular replacement model.¹²⁶ The structures were refined using iterative cycles of the programmes COOT and REFMAC5 using the CCP4i2 interface.¹²⁷⁻¹²⁹ After building of the protein backbone and side chains, residual density of the omit map revealed different ligands and sugar residues in each structure. Coordinate and library files for *N*-hydroxy imidazole (NHI) were prepared using PRODRG.¹³⁰ Carbohydrate structures were validated using PRIVATEER.^{131, 132} Refinement statistics for all structures can be found in Table 3-16.

Table 3-16: Data collection and refinement statistics for rAaeUPO-H, where the numbers in brackets refer to data of the highest resolution shells.

Name	rAaeUPO-H_NHI	rAaeUPO-H_sugars	rAaeUPO-H_ACT
Condition	0.1 M Bis-Tris pH 6.5 26% PEG-3350 2 mg mL ⁻¹ protein	0.1 M Bis-Tris pH 6.5 0.2 M NaCl 26% PEG-3350 8 mg mL ⁻¹ protein	0.1 M Bis-Tris pH 6.5 0.2 M NaCl 28% PEG-3350 8 mg mL ⁻¹ protein
Data collection date	120217	120217	121216
Crystal number	7995	4181	5898
Beamline	Diamond i04	Diamond i04	Diamond i04-1
Wavelength (Å)	0.97950	0.97950	0.92819
Resolution (Å)	67.07-1.61 (1.65-1.61)	67.29-1.56 (1.60-1.56)	46.21-2.20 (2.27-2.20)
Space Group	<i>P</i> ₃ ₁ ₂	<i>P</i> ₃ ₁ ₂	<i>P</i> ₃ ₁ ₂
Unit cell (Å)	a=b=77.45; c=124.01 $\alpha=\beta=90^\circ$; $\gamma=120^\circ$	a=b=77.70; c=123.91 $\alpha=\beta=90^\circ$; $\gamma=120^\circ$	a=b=78.75; c=125.33 $\alpha=\beta=90^\circ$; $\gamma=120^\circ$
Molecules in the asymmetric unit	1	1	1
Unique reflections	56305 (4134)	62183 (4536)	23579 (2024)
Completeness (%)	99.7 (99.4)	99.8 (99.6)	100.0 (99.9)
R _{merge}	0.06 (2.43)	0.07 (2.52)	0.08 (0.707)
R _{pim}	0.03 (1.18)	0.04 (1.25)	0.04 (0.34)
Multiplicity	9.9 (10.0)	9.8 (9.8)	9.9 (10.3)
<I/σ(I)>	15.9 (1.0)	15.1 (1.0)	19.0 (3.6)
CC _{1/2}	1.00 (0.81)	1.00 (0.73)	1.00 (0.95)
R _{cryst} /R _{free} (%)	0.19/0.23	0.16/0.21	0.18/0.24
rmsd 1-2 bonds (Å)	0.018	0.021	0.017
rmsd 1-3 angles (°)	1.90	2.02	1.989
Main chain (Å ²)	39.3	23.3	43.0
Side chain (Å ²)	41.3	25.8	46.3
Water (Å ²)	48.3	41.0	48.2
Ligand (Å ²)	65.9	28.0	69.3
% Residues in favoured, allowed, and outlier regions of Ramachandran plot	97.5, 2.5, 0	98, 2, 0	96.3, 3.7, 0

3.2. Selection and Expression of *AaeUPO*

3.2.1. Construct Design

In 2014/15, the expression of *AaeUPO* in *Saccharomyces cerevisiae* and *Pichia pastoris* was described in the literature.^{120, 133} Directed evolution experiments identified the final 9-point mutant (further referred to as *rAaeUPO*, for recombinant *AaeUPO*) which shows increased secretion and increased activity. Four of the mutations are located on the signalling peptide of *AaeUPO* and are therefore the most likely the reason for the improved secretion of the protein into the fermentation medium (Figure 3-3).

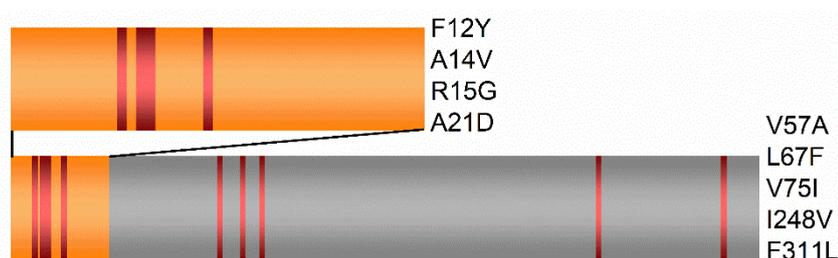


Figure 3-3: *rAaeUPO* construct with the signalling and pro-peptide region in orange and the mature protein in grey, and the mutations highlighted in red.

Using this information, the 9-point *AaeUPO* mutant was obtained and prepared for cloning into pPICZ α . Like most expression vectors used in *P. pastoris* pPICZ α is a shuttle vector allowing plasmid maintenance in *E. coli*. Expression is linked to the alcohol oxidase I (*AOX1*) promoter which is induced by methanol. The plasmid features the *S. cerevisiae* derived α -MF, which was removed in this study, it further allows Zeocin selection and also has the potential for fusing proteins to a C-terminal His-tag. In order to allow purification and also the analysis of the protein production at low levels over time, the His-tag was maintained at the C-terminus of the protein. The option of eliminating the His-tag in case of interference with the protein activity and for crystallisation purposes was given through the addition of a 3C-protease site linking the protein to the His-tag. A second construct was created which placed the His-tag and a 3C cleavage site on the N-terminus of the mature protein (Figure 3-4). The first amino acids from the mature *rAaeUPO* were introduced between the signalling sequence and the His-tag to ensure secretion and cleavage recognition. In summary, two *rAaeUPO* constructs were created either carrying an N-terminal (H-UPO) or a C-terminal (UPO-H) cleavable His-tag (Figure 3-4).

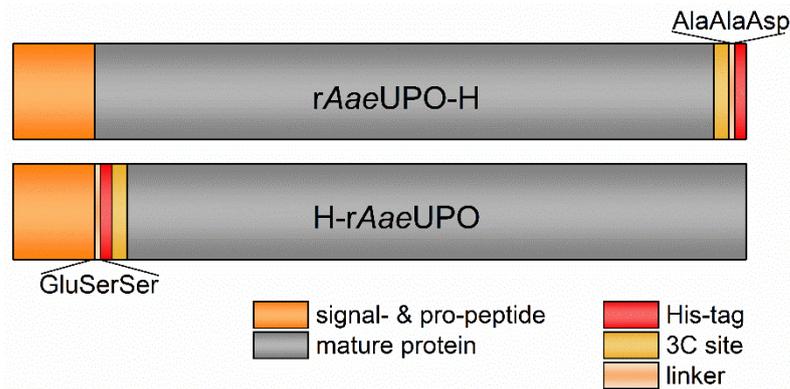


Figure 3-4: Comparison of the yeast constructs *rAaeUPO-H* (top) and *H-rAaeUPO* (bottom), with the mature protein in grey, the *rAaeUPO* signalling- and pro-peptide in orange, the His-tag in red, the 3C protease site in yellow, and the linker regions in apricot, with description of the three amino acids used as linkers.

During the amplification of *rAaeUPO-H* the 3C site was created in the first step and a second amplification allowed the creation of overhangs to facilitate the In-Fusion reaction. In between the two amplification reactions the DNA was purified and isolated through agarose gel electrophoresis. The agarose gels showed the production of inserts of the correct size (1140 bp; 1116 bp *rAaeUPO* and 24 bp 3C-site) and one example is shown below (see Figure 3-5). The vector was prepared for insertion of the UPO-H construct by inverse PCR, which deleted the α -MF secretion sequence and the C-terminal *c-myc* epitope, but retained the His-tag (see Figure 3-5).

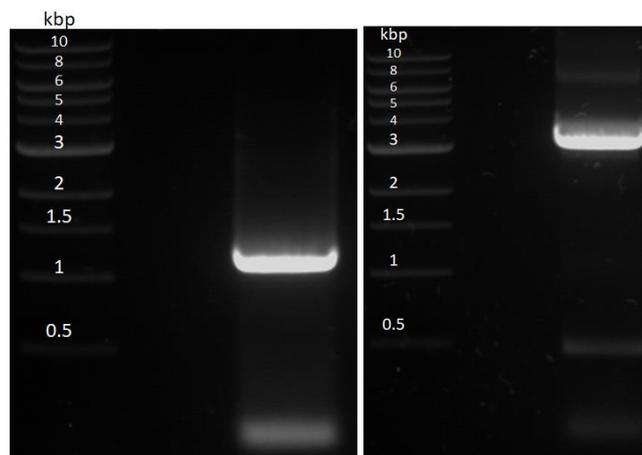


Figure 3-5: 1% Agarose gels showing: **left** - the successful amplification of *rAaeUPO-H* and creation of the N-terminal 3C site (1140 bp); **right** – amplification of pPICZ α , deleting α -MF and *c-myc* regions, retaining the His-tag (approx. 3000 bp).

For the preparation of *H-rAaeUPO* three PCR amplification reactions were run to create the 3C cleavage site, the His-tag and the In-Fusion overhangs on the N-terminus of the mature *rAaeUPO*. In addition, a stop codon before the C-terminal His-tag was introduced. The vector was again prepared through inverse PCR, using

the *rAaeUPO-H* construct as starting material. The primers were designed to retain the N-terminal signalling sequence and the C-terminal His-tag.

The prepared vector and insert constructs were annealed using In-Fusion cloning and the products were used to transform competent Stellar cells. The selection of positive transformants was facilitated by the Zeocin resistance introduced by pPICZ. The colony PCR reactions for H-*rAaeUPO* were run using the *AOX1* primers, designed for sequencing according to the EasySelect™ Pichia Expression Kit User Manual (Invitrogen), which allowed the identification of successful incorporation of the plasmid into the cells. The use of these primers led to the addition of extra bases at both the N- and C-terminus of the construct, increasing the length to approx. 1400 bp. The results for the colony PCRs run for *rAaeUPO-H* and H-*rAaeUPO* are shown below, identifying four out of four positive transformants and eight out of twelve positive transformant carrying the genes of interest, respectively (see Figure 3-6). The plasmid of two positive transformants was amplified in an overnight culture, isolated, and sequencing confirmed the production of the desired gene construct.

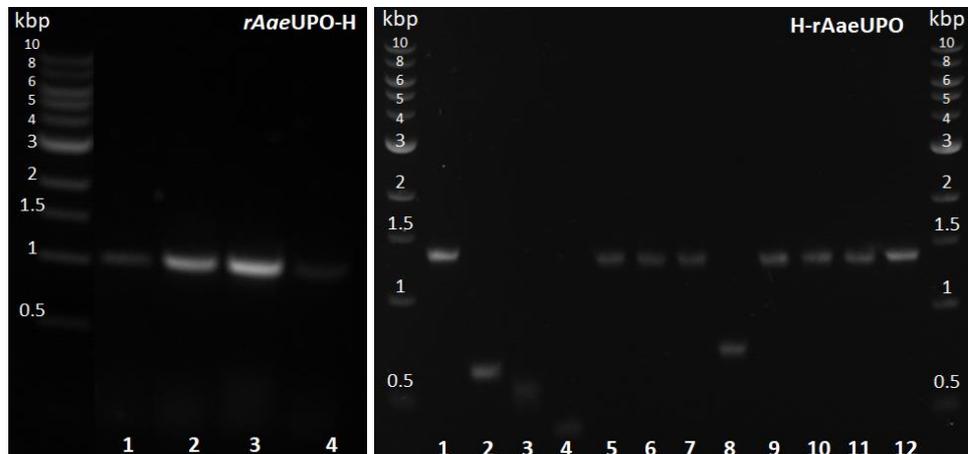


Figure 3-6: 1% Agarose gels for the colony PCR analysis of *rAaeUPO-H* (left) and H-*rAaeUPO* (right).

3.2.2. Selection and Test Expression

The *P. pastoris* strain used in the following experiments is X-33 which does not require supplements to grow on minimal medium. Electroporation was the method of choice to transform the yeast. It generally has high transformation frequencies (10^5 per μg), is easily performed, and allows multicopy integration of the gene. The use of Zeocin resistant strains translates to an easy selection tool for multicopy transformants.¹³⁴ Increased concentrations of the drug lead to an enrichment in

colonies of interest. However, multicopy vector transformants represent a minority of the transformed colonies, with around 10%. Screening of up to 100 colonies might be necessary in order to identify one carrying greater than five copies of the vector of interest. Due to the flocculent nature of *P. pastoris*, describing the tendency to form multi cell clumps, it is important to restreak colonies after transformation to obtain single colonies containing only one transformed yeast strain.

Per UPO construct up to eight separate colonies, depending on the transformation efficiency, were selected and analysed in 48-deepwell microfermentations to allow the identification of expression and potential multi-copy transformants. The colonies were picked from restreaked, freshly grown plates and incubated in buffered minimal medium containing glycerol to promote growth overnight. Once the OD₆₀₀ reached a value around 10-20, expression was induced by transferring the cultures into methanol containing medium. Samples were taken throughout the whole 72-96 h expression and fresh methanol was fed to the system on a regular basis. The final analysis of the transformants and their ability to express UPOs was based on immunoblot analysis. The supernatant samples collected over time were immobilised on a PVDF membrane and after blocking, the membrane was incubated with an anti-His antibody, to bind to the His-tag introduced to the UPOs. The visualisation was based on an electrochemiluminescence (ECL) assay, which showed a higher sensitivity towards the antibody than the previously used SigmaFast tablet based dye. The results obtained for rAaeUPO-H and H-rAaeUPO are shown below (Figure 3-7). Studying the signal intensity for one transformant, a change over time was noted. Whilst it would be expected to observe an increase of intensity, from no signal at t = 0 h to the strongest signal in the final sample, the intensity decreased after the 24 h or 48 h sample, as visible for transformant 1 and 4 (rAaeUPO-H). In other cases, the signal intensity increased again after an initial drop and came to levels similar to the one observed after 24 h (transformant 2 and 5 for rAaeUPO-H). Similar trends were also observed for the N-terminally tagged H-rAaeUPO, where the functionalisation of the membrane was not as effective as for the other construct, visible through the loss of definition. However, it was still observed that the intensity of the signal decreases after 24 or 48 h and disappeared almost completely in the 72 and 96 h samples. These changes in amount of His-tagged protein could be due to

various reasons. Different studies have previously described the presence of proteases in the *Pichia* fermentation medium, which could affect the protein stability.^{135, 136} However, using this analytical method it was not possible to further comment on the protein integrity; further test were needed. The results from different transformants clearly display varying levels of expression. While it is not possible to determine the exact number of copies taken up using this method, the intensity of the signal helped differentiating between different quality transformants. For the subsequent experiments with rAaeUPO-H the transformant 2 and for H-rAaeUPO the transformant 3 were chosen.

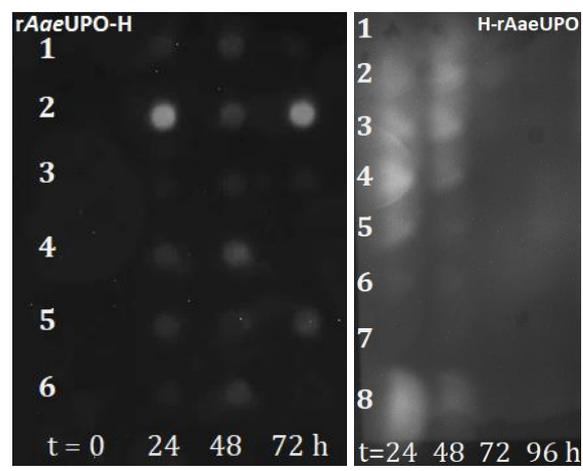


Figure 3-7: Luminol-based dot-blot analysis of *P. pastoris* transformants during the process of the microfermentation; rAaeUPO-H on the left and H-rAaeUPO on the right, samples were taken after 0, 24, 48, and 72 h or 24, 48, 72, and 96 h, respectively.

3.2.3. Shake-Flask Expression

Initial scale-up experiments were performed with the selected transformants on a 100 mL scale. The restreaked cells were grown on glycerol to an OD₆₀₀ of 2, harvested and used to inoculate a fresh medium containing methanol as induction source. The culture was maintained for 72 to 96 h and methanol was added twice daily. Samples were taken once daily for the course of the expression to allow monitoring the secretion of the protein *via* SDS-PAGE. In addition, simple spectroscopic assays were run with the fermentation supernatant. The 100 mL fermentation was usually followed by a Ni-purification trial. The analysis of the rAaeUPO-H is shown below (Figure 3-8). The SDS gel for rAaeUPO-H showed an intensification of the band between 50 and 75 kDa during the fermentation. A band of the same size was observed in the Ni-flow through and with decreasing intensity the wash fractions of the Ni-purification. More defined bands were found in the Ni-fractions 11 to 13. The

immunoblot analysis of the same samples, stained with SigmaFast tablets, showed a response to the His-tag in the 30 mM imidazole wash sample and in the peak fractions (10 to 12). No signal was observed for the samples of the expression over time and for the flow-through sample. It is possible that the samples during the expression of the protein were too dilute to show response to this particular stain. The missing response for the flow-through sample, which showed a band on the gel, could also be due to the concentration of the sample, but other factors are more likely. The chromatogram of the purification (data not shown) showed that the column was not overloaded with protein sample. The activity of the protein was analysed using an NBD-based assay, which displayed the distinct colour change, indicating the presence of active *rAaeUPO* (data not shown). These findings strengthened the statement made at the initial expression test stage about the presence of proteases in the media. The introduced 3C-protease site is a potential target for the native proteases secreted by *P. pastoris* and cleavage at this position would explain the observed activity of the correctly folded protein, whilst losing the His-tag and therefore the ability to bind to the Ni-purification column.

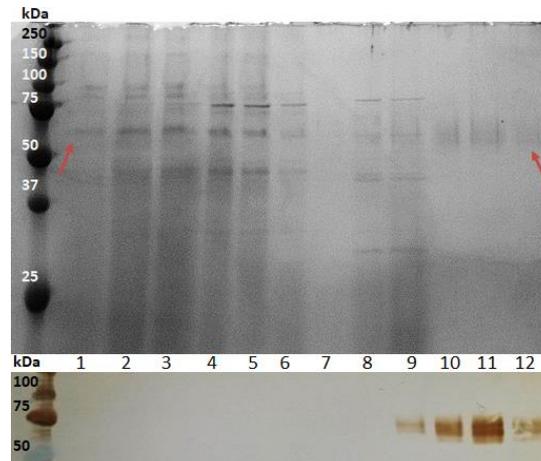


Figure 3-8: 12% SDS gel and immunoblot for the 100 mL scale fermentation and Ni-purification of *rAaeUPO*-H using pre-stained protein standard (Biorad) with: 1 – 24 h, 2 – 48 h, 3 – 72 h, 4 – dialysed sample, 5 – flow-through, 6 – wash unbound I, 7 – wash unbound II, 8 – wash I, 9 – wash II, 10 – fraction 11, 11 – fraction 12, 12 – fraction 13. The arrows are indicating the molecular weight at which the glycosylated protein is observed (55-60 kDa).

Similar results were observed for the construct carrying the N-terminal His-tag. The expression of His-tagged protein over time was monitored using an immunoblot with ECL stain (Figure 3-9, left). The signal for the protein of interest around 55 kDa appeared after 48 h and increased in strength up until the 96 h sample. The

purification of this sample did not show a band with the size of interest in the flow-through, but a signal in the peak fractions (F9 to F11). These findings were supported by ABTS-based activity assays analysing the different stages of this purification (Figure 3-9, right). The flow-through and no-imidazole wash showed only negligible levels of activity, with very little to no change in colour observed. The 30 mM imidazole wash fraction contained some active protein, and the strongest activity was observed in the peak fractions 9 to 12, with residual activity in fraction 13. These results pointed towards the successful expression of His-tagged H-rAaeUPO with very little to no untagged protein in the fermentation supernatant.

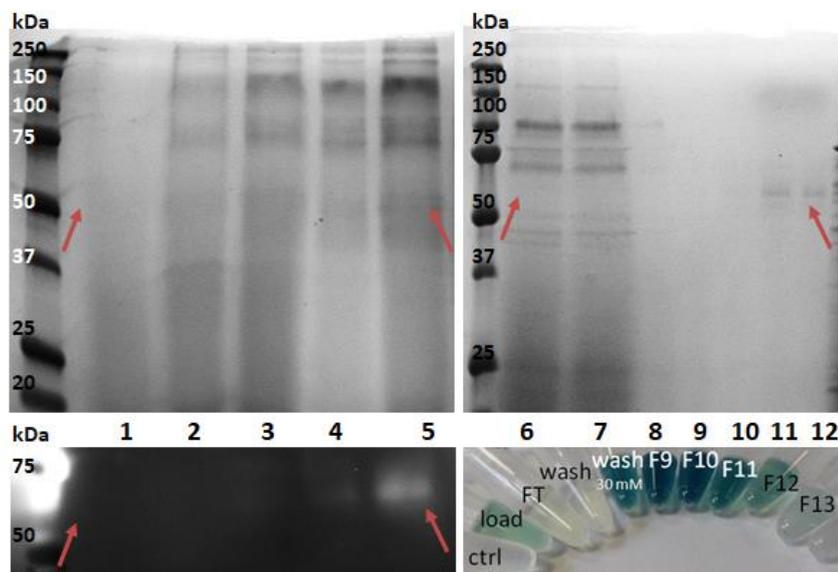


Figure 3-9: 12% SDS gel and immunoblot for the 100 mL scale fermentation of H-rAaeUPO (left) and 12% SDS gel and ABTS-based activity assay for the Ni-purification (right) with: 1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h, 5 – 96 h, 6 – Ni load, 7 – Ni flow-through, 8 – wash, 9 – wash 30 mM imidazole, 10 – fraction 9, 11 – fraction 11, 12 – fraction 13. All gels were run using pre-stained protein standards and the ABTS reaction was run with a no-enzyme control (ctrl) and various samples from the purification.

In comparison, the two rAaeUPO constructs with C- and N-terminal His-tag showed different expression results on a 100 mL-scale shake-flask cultivation. Whilst the C-terminally tagged rAaeUPO-H was prone to cleavage of the His-tag during the process, giving active, but untagged protein, the N-terminally tagged H-rAaeUPO seemed to maintain an intact His-tag, allowing for Ni-purification. Further analysis on the integrity of the expressed rAaeUPO-H will be shown later (see Section 3.3).

3.2.4. Fed-Batch Fermentation

As *Pichia* is well known for growing to high cell densities, which can increase the overall yield of protein, fed-batch fermentations were performed. Using this method,

the fermentation process is much better controlled, by regulating parameters like temperature, dissolved oxygen (DO), pH, agitation, and aeration. During the methanol-fed batch stage especially, the yeasts metabolism is high, which increases the temperature in the reaction vessel and external cooling, to keep a temperature between 25-30 °C, might be necessary. Temperatures above 32 °C are detrimental to the protein expression and can affect cell survival. It is important to control the DO concentration, by agitation and aeration, as oxygen is essential for *Pichia* to metabolise both glycerol and methanol. Concentrations above 20% should be maintained throughout the process. The DO measurement can be used to comment on the health of the culture. In limiting carbon conditions, for example when the feed is paused, the time it takes for the culture to respond to the change, by decreasing its metabolic rate, can be an indication, as to whether the carbon source is limiting (response time <1 min) or accumulating (response time >10 min). Especially during the methanol fed-batch phase it is important to analyse the culture using DO responses to avoid the accumulation of methanol, as high concentrations (>2%) are toxic to the cells.

rAaeUPO-H was fermented in different size fermenters. A comparison for a 0.2 L and a 2 L fermentation will be given. Prior to any fermentation, fresh colonies were grown from the transformant chosen in Section 3.2.2, and the cells were amplified in overnight cultures for inoculation of the fermentation medium (see Figure 3-10).

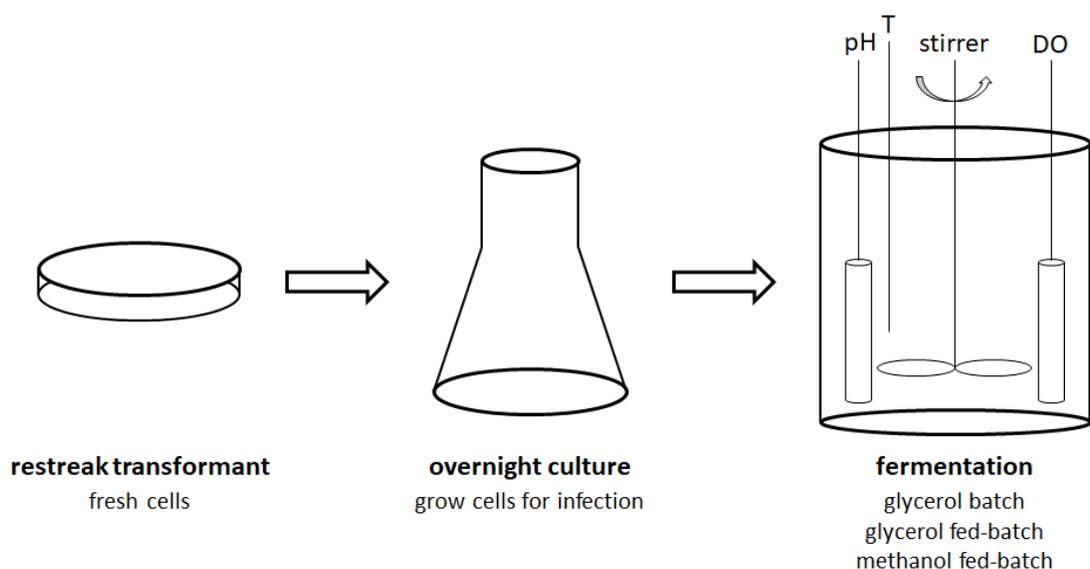


Figure 3-10: Schematic representation of the steps of a *Pichia* fermentation, as performed in this work.

The fermentation process can be summarised in three phases, the glycerol batch phase, the glycerol fed-batch phase, and the methanol fed-batch phase. Once the fermenter was inoculated with the cells, the cells were left to utilise the glycerol provided in the medium. The glycerol batch phase normally took approx. 24 h, and the full utilisation of the carbon source is indicated by a spike in the DO reading (as indicated in Figure 3-11). The cells were grown further, feeding a defined amount of glycerol at a constant rate (glycerol fed-batch phase). The amount of glycerol was calculated from the initial fermentation volume and the duration of the feed was set to 4 h. After the full utilisation of the glycerol provided during the fed-batch phase (approx. 30 h after inoculation), the transition into methanol was started (methanol fed-batch phase) and protein expression was induced (see Figure 3-11). The cultures took about 3 h to fully adapt to utilising methanol, after which point the feed-rate was increased slowly. As mentioned before it was crucial to monitor the DO response to a pause in the feed, to assess the state of the culture and to avoid accumulation of methanol (example of DO spikes in response to feed-pauses indicated in Figure 3-11). In addition to the DO read, changes in temperature and base feed were also indicative of the cultures health. During the methanol fed-batch phase a constant addition of base should be observed to maintain a pH of 5. This is due to formaldehyde, the metabolic by-product of the methanol consumption. Furthermore, ammonium hydroxide was the sole nitrogen source in the process. The wet-cell weight, which was determined twice daily, reported information about the growth of the culture. The methanol fed-batch could last between 72 to 96 h, depending on the protein, and the final wet-cell weight could be as high as 450 g L⁻¹.

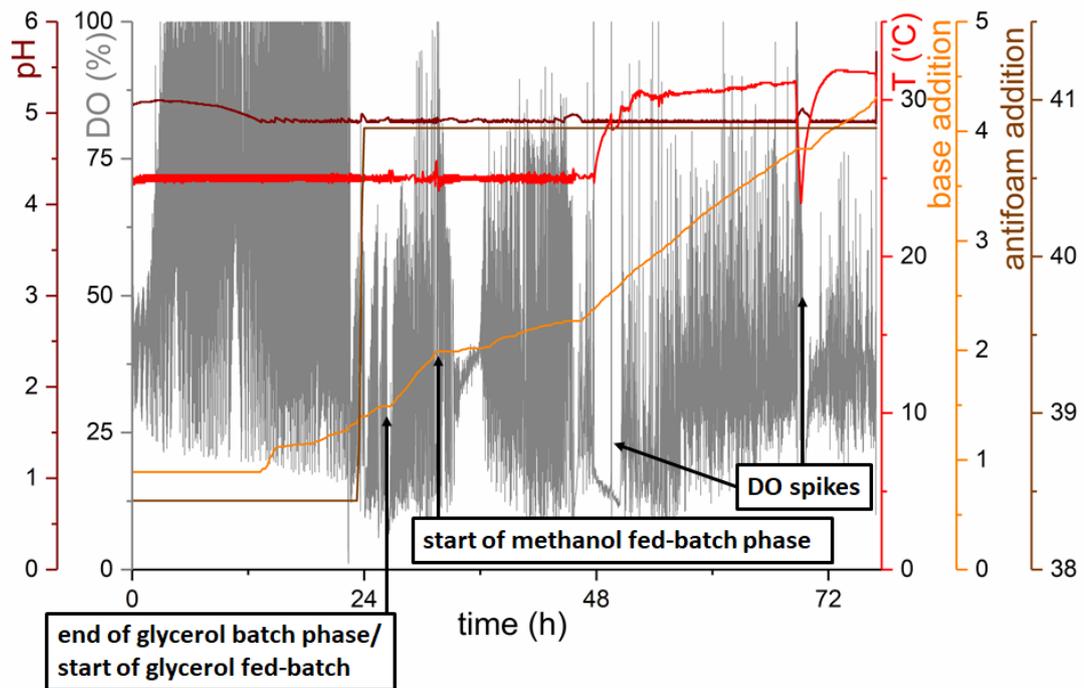


Figure 3-11: Fermentation-log for the *rAaeUPO-H* fermentation run at 0.2 L scale. The pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated are the end of the glycerol batch and beginning of the glycerol-fed batch phase, the beginning of the methanol fed-batch phase, and two examples for DO spike responses to feed-pauses.

The methanol fed-batch phase for the fermentation of *rAaeUPO-H* on 0.2 L scale, as shown in Figure 3-11, lasted for approx. 45 h, before all feeds and logs were stopped, and the cells were separated from the supernatant. At the start of the methanol fed-batch the wet-cell weight was determined to be 230 g L^{-1} , and after 45 h methanol feed, the wet-cell weight was around 290 g L^{-1} . In the first hours of the methanol fed-batch phase, during the adaptation to the new carbon source, the base addition slowed down. Once adapted and with increasing feed-rates, the base addition also increased again. After about 48 h of fermentation, an increase in the feed rate, but also the temperature can be observed. The increased metabolism was the cause for these observations. The increase in temperature was due to insufficient cooling provided by the fermenter and would not be expected in an ideal set-up. After approx. 65 h, where the second DO spike is indicated in Figure 3-11, a sharp decrease in temperature is visible, combined with a stop in base addition. It is a good example to show how the culture reacted to limited carbon conditions, by adjusting its metabolism. The DO concentration measurements during the experiment were changing frequently between approx. 20 and 80%, ideally the concentration would

be kept constant at around 40%. However, the fine-tuning of the algorithm controlling the agitation and aeration was not successful in this case, and led to the results presented in Figure 3-11.

The fermentation was quantified using specific activity measurements, using ABTS and NBD as substrates. Further, gels were run to monitor the expression of protein over time and purification trials using Ni-affinity chromatography were performed. The gel monitoring the expression of *rAaeUPO-H* in the supernatant over time, showed the appearance and accumulation of a protein with a size around 65 kDa (see Figure 3-12, left). The specific activities towards NBD (1 mM) and ABTS (0.2 mM) were determined to be $0.4 \pm 0.1 \text{ U mL}^{-1}$ and $6.8 \pm 0.1 \text{ U mL}^{-1}$, respectively, using 10 μL of the harvested cell-free supernatant in a 1 mL reaction. 100 mL of the harvested cell-free solution was dialysed into a salt-free buffer and then applied onto a Ni-affinity chromatography column. The elution profile from the column is shown below, the elution of multiple peaks overlapping between fractions 8 and 18 was observed (see Figure 3-12, right insert). The absorbance at 280 nm displayed a sharp peak in the fractions 10 and 11 (350 mAU), and a broad peak with a maximum of 200 mAU in the fractions 14 and 15. The 420 nm profile showed a broad peak, slightly shifted compared to the 280 nm absorbance, with an absorption maximum of 40 mAU in fraction 13. All fractions were analysed using SDS-PAGE and the purification of a protein with a molecular weight around 60 kDa could be observed (see Figure 3-12, right). Additionally, another protein with a molecular weight just below the first one, could be observed in the fractions 12 to 17. In addition to the column bound protein, the gels showed a band with the same molecular weight in the column flow-through, and with decreased intensity in the column wash fraction. Subsequent experiments decreasing the amount of sample applied to the column showed, that detection of the protein in the flow-through was not caused by an overloading of the column.

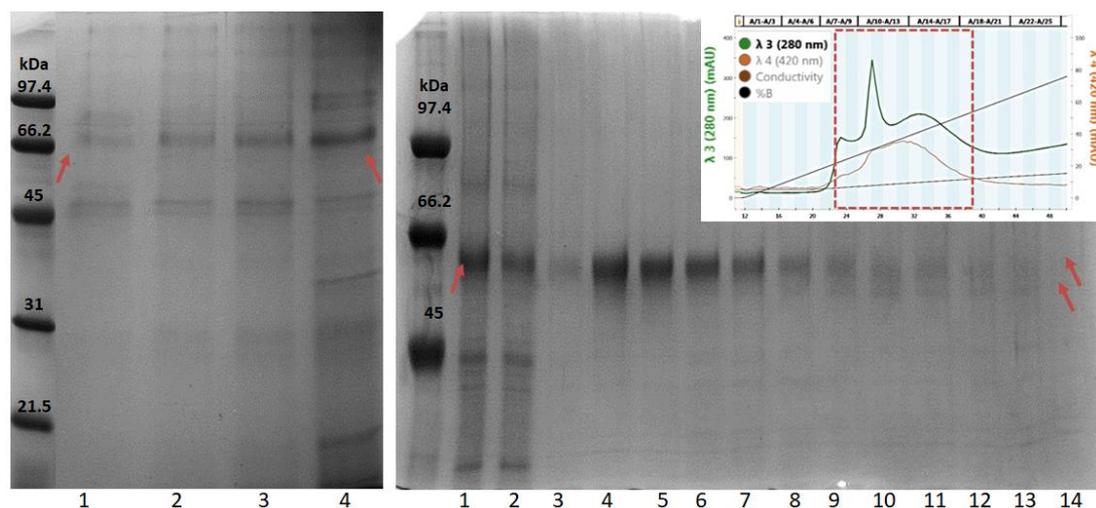


Figure 3-12: 12% SDS gels using the low molecular weight marker, for **left:** the fermentation of *rAaeUPO-H* at 0.2 L scale with supernatant samples taken during the methanol fed-batch phase at : 1 – 0 h, 2 – 13 h, 3 – 21 h, 4 – 37 h; **right:** the Ni-affinity purification of *rAaeUPO-H* harvested after 45 h, with: 1 – sample load, 2 – flow-through, 3 – wash, 4 – fraction 8, 5 – fraction 9, 6 – fraction 10, 7 – fraction 11, 8 – fraction 12, 9 – fraction 13, 10 – fraction 14, 11 – fraction 15, 12 – fraction 16, 13 – fraction 17, 14 – fraction 18. These fractions are highlighted in the chromatogram insert (**top right**) with the absorbance at 280 nm in green and at 420 nm in orange. The expected molecular weight for *rAaeUPO-H* is indicated using the red arrows.

To complement the SDS-PAGE analysis, activity assays were performed with the flow-through, and different fractions. Both the peroxygenase (NBD) and the peroxidase (ABTS) activity were studied and compared to a non-enzyme control sample. Activity towards both substrates was detected in all fractions. For NBD the strongest activities were measured in the flow-through and fraction 14, in addition fraction 11 was also noticeably increased using ABTS. These findings for the expression of *rAaeUPO-H* are in line with results discussed previously, regarding the stability of the C-terminal His-tag. Based in the findings above, it is proposed that His-tag cleavage was observed 45 h after methanol, leaving only a fraction of protein with an intact tag to bind to the affinity chromatography column.

Due to difficulties maintaining the temperature with the 0.5 L bioreactors, large-scale fermentation of *rAaeUPO-H* was attempted in a 7 L fermentation vessel, with 2 L initial fermentation medium. The preparation and fermentation phases were the same as described for the small-scale fermentation. The methanol fed-batch phase was started with 220 g L⁻¹ wet-cell weight, which increased to 240 g L⁻¹ after 38 h methanol addition, and yielded in 290 g L⁻¹ cells at the harvest after 62 h. The specific activities towards NBD (1 mM) and ABTS (0.2 mM) were determined to be 1.4 ± 0.1 U mL⁻¹ and 12.3 ± 0.2 U mL⁻¹, respectively, using 10 µL of the harvested cell-

free supernatant after 62 h in a 1 mL reaction. Comparing these activities to the ones obtained for *rAaeUPO-H* expressed in 0.2 L initial fermentation medium for 45 h, it was visible that the activities increased over time. Maintaining the fermentation for 17 h longer, yielded a doubling in activity reported for ABTS, and a tripling in the activity observed for NBD, indicating the successful expression of more active protein. The SDS-PAGE and immunoblot analysis of the expression of *rAaeUPO-H* in the fermentation supernatant over time is shown below (see Figure 3-13, left). In addition to the time course of the expression, the analysis also included samples from Ni-affinity purifications run after 38 h (20 mL) and 62 h (40 mL) fermentation. The flow-through samples and the peak fractions 5 to 7 were run on the gel and also investigated for their activity towards NDB (shown in Figure 3-13 on the right).

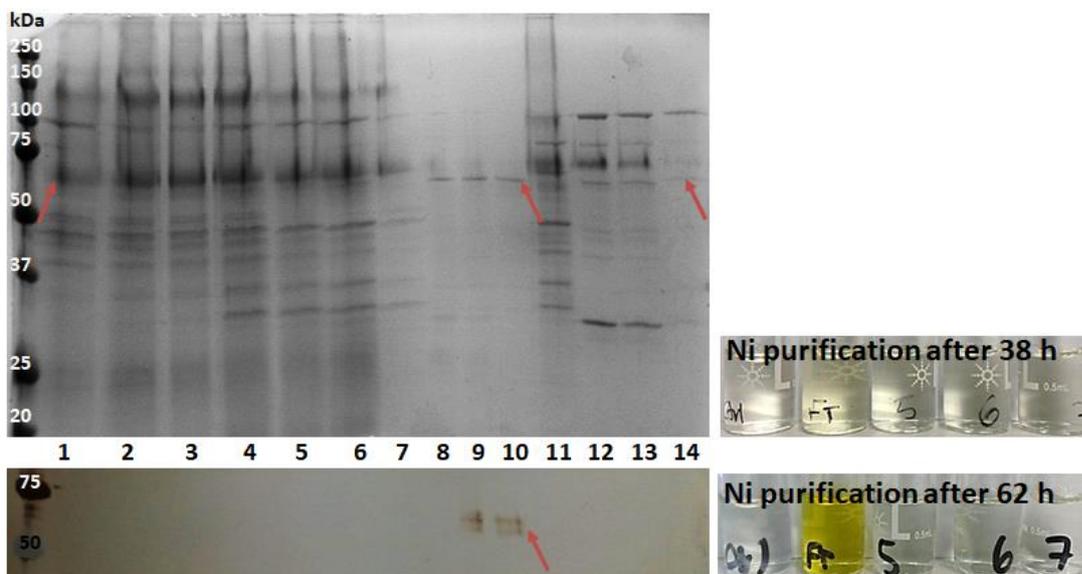


Figure 3-13: left: 12% SDS gel and immunoblot for samples taken during the methanol fed-batch phase of the 2 L fermentation and Ni-affinity purification after 38 h and 62 h fermentation for *rAaeUPO-H* (approx. 60 kDa, highlighted with arrows) with : 1 – 0 h, 2 – 14 h, 3 – 22 h, 4 – 38 h, 5 – 45 h, 6 – 68 h, 7 – flow-through (38 h), 8 – fraction 5 (38 h), 9 – fraction 6 (38h), 10 – fraction 7 (38 h), 11 – flow-through (62 h), 12 – fraction 5 (62 h), 13 – fraction 6 (62 h), 14 – fraction 7 (62 h); right: NBD-based activity tests run on the flow-through and peak fractions of the Ni purifications run after 38 h (top) and 62 h (bottom) fermentation showing a control (ctrl) on the left, followed by the flow-through (FT) and the fractions 5, 6, and 7.

The SDS-PAGE analysis showed the expression of a protein with the molecular weight around 60 kDa, displaying an increasing intensity over time. No signals were recorded for the time-course samples analysed *via* immunoblot. The purification of *rAaeUPO-H* was attempted after 38 and 62 h. For the 38 h sample, the SDS-PAGE analysis showed a strong band at 60 kDa in the flow-through and faint bands at the same molecular weight in the peak fractions 5, 6 and 7. The immunoblot analysis of the

same samples, showed weak responses against the His-tag, in the fractions 5 and 6. No immunoblot signal was detected in the flow-through sample. The activity determined for the Ni-purification samples showed the strongest response in the flow-through, with little activity in the peak fractions (see Figure 3-13, top right). In summary, these findings indicate the presence of a His-tag on a small fraction of expressed and secreted active protein, with a larger fraction not carrying a His-tag after 38 h methanol feed. Purification trials run after 62 h of expression showed very similar results to the ones just discussed above. In the chromatogram only a weak elution response was observed. As the same method was run as before, the fractions which were identified as peak fractions in the 38 h sample (fractions 5 to 7), were chosen to be analysed together with the flow-through. The gel analysis showed a strong band at the expected molecular weight in the flow-through, and only faint bands in the peak fractions. No response to the His-tag was observed in the immunoblot. The activity measurements using NBD showed strong activity in the flow-through, but none in the peak fractions. It can be concluded that while the activity assays would suggest further accumulation of rAaeUPO-H over time, the integrity of the C-terminal His-tag decreases, to the point where no Ni-affinity purification or immunoblot analysis was successful after 68 h.

In addition to the fermentation of rAaeUPO-H, the N-terminally tagged variant, H-rAaeUPO, was also studied in small-scale fed-batch fermentation. The fermentation log is presented below (see Figure 3-14) and was found very similar to the one reported for the C-terminally tagged rAaeUPO. After the initiation of the methanol fed-batch phase, the base addition slowed down during the first hours allowing adaptation. Once fully adapted, the base-feed remained constant and a steady increase in wet-cell weight was observed, from 175 g L^{-1} at the start of the methanol fed-batch phase to 370 g L^{-1} at harvest after 72 h. The temperature control of this process was again not ideal, with a peak in temperature to $37 \text{ }^\circ\text{C}$ after approx. 36 h and subsequent expression maintained around $30 \text{ }^\circ\text{C}$. Limiting the carbon source led to almost immediate temperature decrease and DO spikes, indicating the high metabolic rate the culture was maintaining. The specific activities towards NBD (1 mM) and ABTS (0.2 mM) were determined to be $0.4 \pm 0.1 \text{ U mL}^{-1}$ and $3.1 \pm 0.1 \text{ U mL}^{-1}$, respectively, using $10 \text{ }\mu\text{L}$ of the harvested cell-free supernatant after

72 h in a 1 mL reaction. These results were similar to the results obtained for the rAaeUPO-H after 45 h expression, with respect to the activity determined towards NBD. The ABTS activity for H-rAaeUPO was only half of that determined for the C-terminally tagged protein after 45 h. Due to the poor temperature control within the experiment it was impossible to determine whether the loss of activity was due to reduced expression and secretion of the different construct, or whether the early spike in temperature and maintained higher temperature (48-60 h for H-rAaeUPO versus 24 h for rAaeUPO-H).

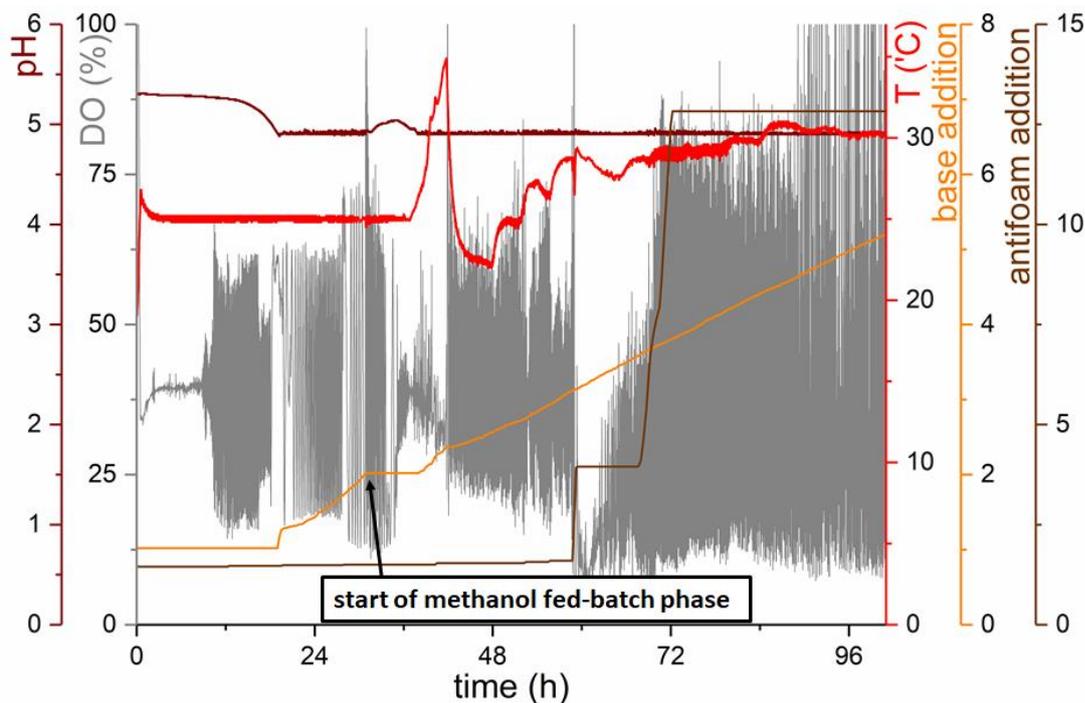


Figure 3-14: Fermentation-log for the H-rAaeUPO fermentation run at 0.2 L scale. The pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated is the beginning of the methanol fed-batch phase.

The expression of H-rAaeUPO was followed over time using SDS-PAGE and immunoblot analysis (see Figure 3-15). In the gel a protein band around 65 kDa intensified over time. In the corresponding immunoblot, targeting the His-tag, a response developed after 24 h, increased in intensity, and then disappeared after the 48 h sample. Ni-affinity purification trials after the 72 h harvest were unsuccessful and no binding to the column was observed. The activity was retained in the column flow-through.

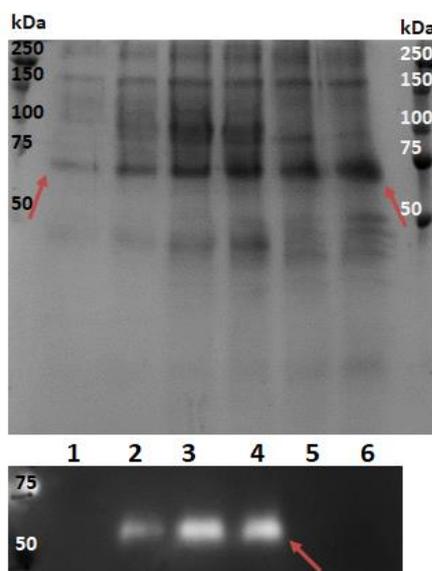


Figure 3-15: 12% SDS gel and immunoblot for the time course of expression for the 0.2 L scale fermentation of H-rAaeUPO with: 1 – 0 h, 2 – 24 h, 3 – 42 h, 4 – 48 h, 5 – 66 h, 6 – 72 h; gels were run using pre-stained protein standards.

In summary, the fed-batch fermentation of both rAaeUPO constructs was successful, however, prolonged fermentation times led to the suspected complete cleavage of the His-tag at both the C- and N-terminus. The amount of rAaeUPO was increased when the methanol fed-batch phase was run for longer, and the importance of fermentation parameter control, like the temperature, was mentioned. The expression levels of rAaeUPO-H were found to be higher, compared to the results for H-rAaeUPO. However, at this point, it was not possible to comment on whether the construct or the fermentation conditions were reason for the decrease in expression/secretion.

3.2.5. Fermentation Optimisation

To study the integrity of the His-tag on the rAaeUPO constructs in more detail, a range of fermentation conditions were tested. The majority of the trials focused on decreasing the native *Pichia* protease activity through changes in pH and the addition of protease substrates. Other tests were performed with increased control over the fermentation process, allowing for example a more stable temperature during the methanol fed-batch phase. The analysis of the different experiments relied on specific activity measurements performed with the fermentation supernatant, SDS-PAGE, and immunoblot analysis, and, where appropriate, purification trials. All

experiments were performed on 0.2 L scale, with the option of running up to three different experiments in parallel.

First, the fermentation guideline by Invitrogen suggested to lower the pH during the methanol fed-batch to reduce the activity of native secreted *Pichia* proteases.¹³⁷ Parallel experiments were performed with decreased pH values at 3 and 4, and compared to a further experiment performed using standard conditions (pH 5). The temperature control during these experiments was better, and for the pH 5 and 3 samples, the temperature remained at 25 °C throughout the whole fermentation, in the case of the pH 4 fermentation, the temperature increased slightly to 27 °C during the methanol fed-batch phase. The cell growth did not seem affected by the decrease in pH. The pH 3 sample increased the wet-cell weight from 230 g L⁻¹ (t = 0 h) to 400 g L⁻¹ after 60 h, and the pH 4 and 5 samples went from 200 g L⁻¹ to 360 g L⁻¹ just before harvest. The colour of the cell-free supernatant was pH dependent; while the pH 5 sample appeared green-brown, the colour changes to yellow at lower pH (see Figure 3-16, top). The addition of the PTM₁ salts, a mixture of trace salts necessary for the growth of the culture, was believed to be responsible for the different colours. The change in pH could affect the metal coordination sphere and ligands, and therefore lead to a change of the solution colour. The colour of the solution was not found to be indicative of the presence or absence of the target protein. Dialysis into salt-free buffer at pH 7.0 showed the diffusion of the metals into the buffer reservoir and the analysed samples only showed minimal amounts of colouration. The activities of the different cultures towards NBD and ABTS were determined in initial trials (see Figure 3-16, bottom). Using NBD, it was found that only the pH 5 sample showed peroxygenase activity, no conversions were observed for the pH 4 and 3 samples. Peroxidase activity towards ABTS was detected in all three samples with decreasing intensity at lower pH. This was possible as the ABTS assay is the more sensitive assay. The specific activities towards both substrates were also quantified using a spectrophotometer. The activities towards ABTS (0.2 mM) were recorded to be 24.0 ± 0.2, 1.9 ± 0.1, and 0.1 ± 0.1 U mL⁻¹ for pH 5, 4, and 3, respectively, and towards NBD (1 mM) values of 3.4 ± 0.1, and 0.3 ± 0.1 U mL⁻¹ were recorded for pH 5 and 4, respectively, no activity was determined for the pH 3 sample. The increased process control at the small-scale allowed for the specific activity towards NBD to

increase from $0.4 \pm 0.1 \text{ U mL}^{-1}$ (45 h) to $3.4 \pm 0.1 \text{ U mL}^{-1}$, and the activity towards ABTS almost quadrupled from $6.8 \pm 0.1 \text{ U mL}^{-1}$ after 45 h fermentation to $24.0 \pm 0.2 \text{ U mL}^{-1}$ after 60 h.

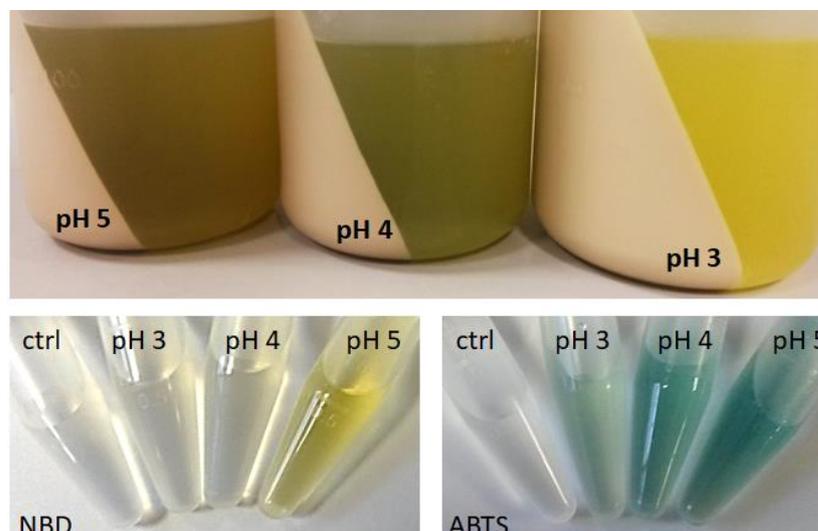


Figure 3-16: top: harvested fermentations of *rAaeUPO-H* performed at varying pH, showing different colours of the fermentation supernatants against the cell pellet in the centrifugation tubes; bottom: activity assays using NBD (left) and ABTS (right) to compare the activity of *rAaeUPO* after 60 h fermentation at different pH.

Purification trials using *rAaeUPO-H* expressed at different pH was unsuccessful and no binding to the Ni-affinity chromatography column was observed. Activity assays confirmed these results, as the activity measured in the column load was equivalent to the activity measured in the flow-through samples.

The addition of casamino acids (CA), protease substrates, was another attempt trying to reduce the native *Pichia* protease activity.¹³⁸ The fermentation process ran smoothly and the temperature was kept at $25 \pm 1 \text{ }^\circ\text{C}$. The addition of 1% CA into the fermentation medium did not affect the cell growth; the wet-cell weight increased from 175 g L^{-1} at methanol induction to 350 g L^{-1} after 72 h expression. The time-course samples taken during the methanol fed-batch phase were analysed using SDS-PAGE (see Figure 3-17), together with the samples obtained for a Ni-affinity purification trial. The harvested cell-free supernatant was dialysed prior to charging the column. The chromatogram showed the presence of a small elution peak ranging from fraction 9 to fraction 13. In addition to the peak fractions, the column flow-through and wash samples without and with 30 mM imidazole were also analysed. In the SDS gel the appearance and intensification of a band around 65 kDa was observed for the fermentation samples. The same band was found in the flow-

through and wash samples from the Ni-affinity chromatography. In the peak fractions a protein with a molecular mass around 60 kDa was identified and an additional protein at approx. 55 kDa appeared in the fractions 11 onwards. The activity of the different purification samples was tested using a ABTS assay in the presence of a non-enzyme control. The strongest activity was identified in the column flow-through and wash samples, and weaker activity was observed in the fractions 9 and 10, with decreasing response in the following fractions. These findings suggested that a small fraction of *rAaeUPO-H* was still His-tagged after 72 h fermentation, however, the majority of the protein, as described before, did not possess an intact His-tag for purification.

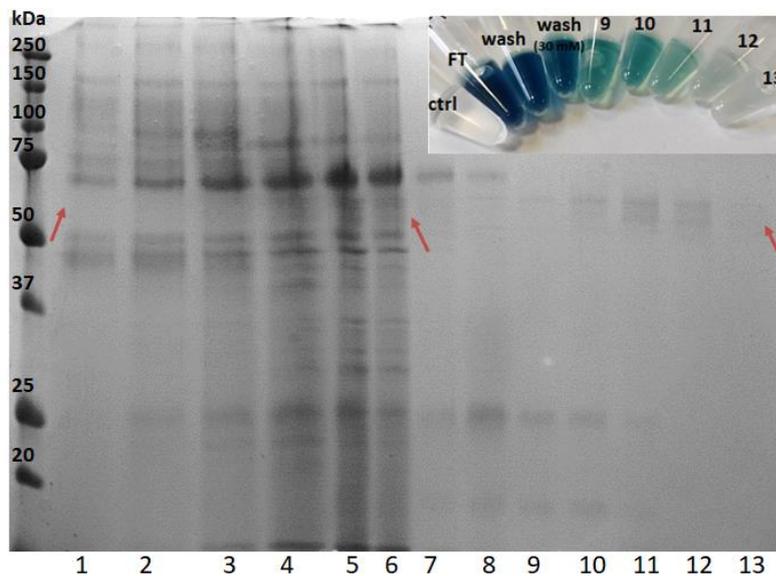


Figure 3-17: 12% SDS-gel with pre-stained ladder for the fermentation (1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h) and Ni-affinity purification (5 – column load, 6 – flow-through, 7 – wash, 8 – wash with 30 mM imidazole, 9 – fraction 9, 10 – fraction 10, 11 – fraction 11, 12 – fraction 12, 13 – fraction 13) of *rAaeUPO-H* grown in the presence of casamino acids; arrows are indicating the expected molecular weight around 65 kDa. **Insert:** ABTS activity assays performed with the Ni-affinity purification samples in the presence of a control (ctrl).

In addition to the analysis of the purification samples, the *rAaeUPO-H* content in the fermentation supernatant was quantified using NBD and ABTS-based specific activity assays. The clarified supernatant showed $3.5 \pm 0.4 \text{ U mL}^{-1}$ for NBD (1 mM) and $22.1 \pm 0.3 \text{ U mL}^{-1}$ for ABTS (0.2 mM). These values were in line with the results reported for successful temperature control before and show that the addition of CA to the fermentation medium did not affect the overall expression and secretion of *rAaeUPO-H*.

The last optimisation experiment discussed here did not aim to increase the stability of the His-tag, but aimed to increase the overall expression and secretion of *rAaeUPO-H*. Therefore, the expression time was increased from a 72 h methanol fed-batch phase to lasting 96 h. The fermentation log for this experiment is shown below (see Figure 3-18), the glycerol batch phase was started upon cell addition at $t = 18$ h, the methanol fed-batch phase was initiated at $t = 48$ h. In the first 24 h of the methanol fed-batch phase the culture adapted to the new carbon source and low metabolic rate allowed for good process control. With increasing methanol feed-rates, the temperature control decreased and an increase in temperature to $30\text{ }^{\circ}\text{C}$ starting day two on methanol ($t = 72$ h) to $35\text{ }^{\circ}\text{C}$ after 48 h ($t = 96$ h) expression was observed. The dissolved oxygen concentration was well controlled after 48 h on methanol ($t = 96$ h) and was kept around 40%. This tight control allowed to observe DO spikes very clearly (as indicated in Figure 3-18). Also coupled to limited carbon conditions was a swift decrease in temperature coinciding with the DO spike, illustrating the quick change in metabolic rate.

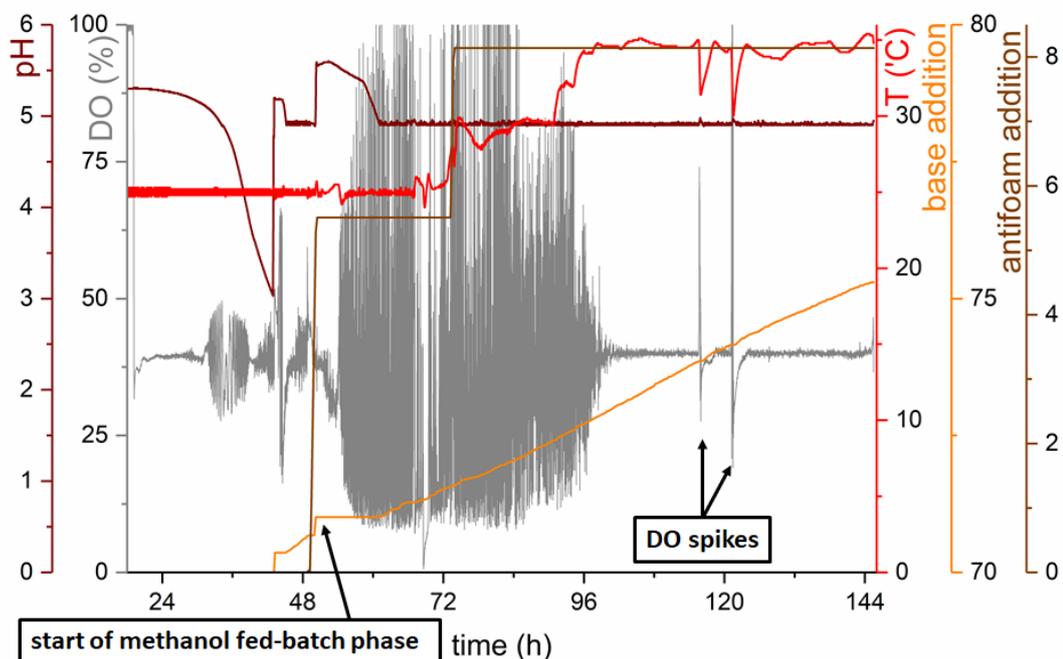


Figure 3-18: Fermentation-log for the *rAaeUPO-H* fermentation run at 0.2 L scale with a 96 h methanol fed-batch phase; the cells were added to the fermenter at $t = 18$ h (start of the glycerol batch phase). The pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated are the beginning of the methanol fed-batch phase, as well as DO spikes.

The growth of the culture was followed over time and the wet-cell weight increased from 175 g L⁻¹ at the start of the methanol fed-batch phase to 225 g L⁻¹ after 24 h (+50 g L⁻¹), to 305 g L⁻¹ after 48 h (+80 g L⁻¹), to 355 g L⁻¹ after 72 h (+50 g L⁻¹) to a final weight of 375 g L⁻¹ (+20 g L⁻¹) at harvest after 96 h methanol addition. Within the last 24 h of methanol feed, the cell growth slowed down. This could be due to limited feed conditions, as no increase in the feed-rate was performed during the last day of expression, to prevent a further rise in temperature. An additional measure to study the expression levels was the determination of the specific activities towards NBD and ABTS over time. The results are summarised below (see Table 3-17).

Table 3-17: Specific activities towards NBD and ABTS during the 96 h methanol fed-batch phase of the fermentation of *rAaeUPO-H*.

MeOH addition time (h)	Specific activity (U mL ⁻¹)	
	NBD (1 mM)	ABTS (0.2 mM)
0	0.0 ± 0.0	0.1 ± 0.1
24	0.2 ± 0.1	1.1 ± 0.3
48	0.7 ± 0.1	4.1 ± 0.2
72	0.8 ± 0.1	4.5 ± 0.7
96	1.0 ± 0.1	7.1 ± 0.3

The specific activities measured in the fermentation supernatant over time are increasing with continued methanol addition. However, coinciding with the rise in temperature to 35 °C after 48 h, the activity almost stagnated between the 48 and 72 h measurements. The increase in the following measurement was interesting, as it suggested that prolonging the methanol fed-batch phase could yield in increased amounts of active, secreted *rAaeUPO*. The specific activities determined after 72 h were not comparable in yields with the results obtained for well controlled fermentations. However, they were similar to other experiments performed, for example the H-*rAaeUPO* fermentation, where the process and temperature control was challenging.

To summarise, the fermentation optimisation experiments were not yielding increased amounts of His-tagged *rAaeUPO-H*; cleavage was still observed in the fermentation supernatant over time. It was possible to show that the methanol fed-batch phase could be run for an additional day (24 h) to increase the overall yield of secreted and active peroxygenase. Furthermore, the number of experiments run

allowed to comment on the importance of temperature control during the fermentation process. While a rise in temperature to 30 °C facilitated cell growth and secretion of *rAaeUPO-H* at decreased levels, further rises in temperature to 35 °C and above led to a halt in expression for approx. 24 h before activity levels rose again. It was possible to show that a well-controlled process yielded supernatant with specific activities towards ABTS of up to 24.0 U mL⁻¹ (pH optimisation experiment run at pH 5), while less controlled fermentations showed decreased activities around 4.5 U mL⁻¹ (prolonged methanol fed-batch experiment, 72 h sample).

3.3. Purification and Biophysical Characterisation of *AaeUPO*

To allow the comparison of the expressed *rAaeUPO* to the literature described protein, it was important to purify and characterise the enzyme. Due to the absence of the His-tag, a different purification route was established. Furthermore, the cleavage of the His-tag made it important to analyse the integrity of the secreted protein, in addition to its kinetic characterisation.

3.3.1. Multistep Purification

As explained in the previous section, the expression of *rAaeUPO-H* with an intact His-tag was unsuccessful and Ni-affinity chromatography was not able to purify the protein of interest. As a result, purification using different chromatography methods was attempted.

Initial studies identified anion-exchange (Q) chromatography and hydrophobic interaction chromatography (HIC) to be effective in the purification of *rAaeUPO-H*. The most effective route was found to be an ammonium sulfate precipitation, followed by HIC, followed by Q-chromatography, with a final size-exclusion chromatography (SEC) step and will be discussed in more detail here. In the precipitation step, the ammonium sulfate concentration was slowly raised to 40% before any precipitants were removed. The supernatant, in which *rAaeUPO-H* remained, was applied onto a HIC column and a gradient from 40% ammonium sulfate to salt-free was run over 15 column volumes (CV). The protein elution was monitored following the absorbance at 280 nm (green in Figure 3-19, top left) and the additional tracking of the absorption at 420 nm (orange) allowed identifying heme-thiolate proteins. While a range of proteins elutes over the course of the

gradient, the 420 nm absorbance placed the heme-thiolate protein in the fractions 14 to 22. These results were confirmed by analysing the activity of the different fractions towards NBD (see Figure 3-19, top right). The strongest activity was found in the fractions 16 to 18, with residual activity in 14 to 22. SDS-PAGE analysis showed similar results. The protein of interest (approx. 60 kDa) showed a band in the fractions 14 to 22. However, proteins of lower molecular weight (approx. 45 and 40 kDa) co-elute with rAaeUPO-H. The fractions 16 to 20 were pooled and dialysed into salt-free buffer prior to the next purification step.

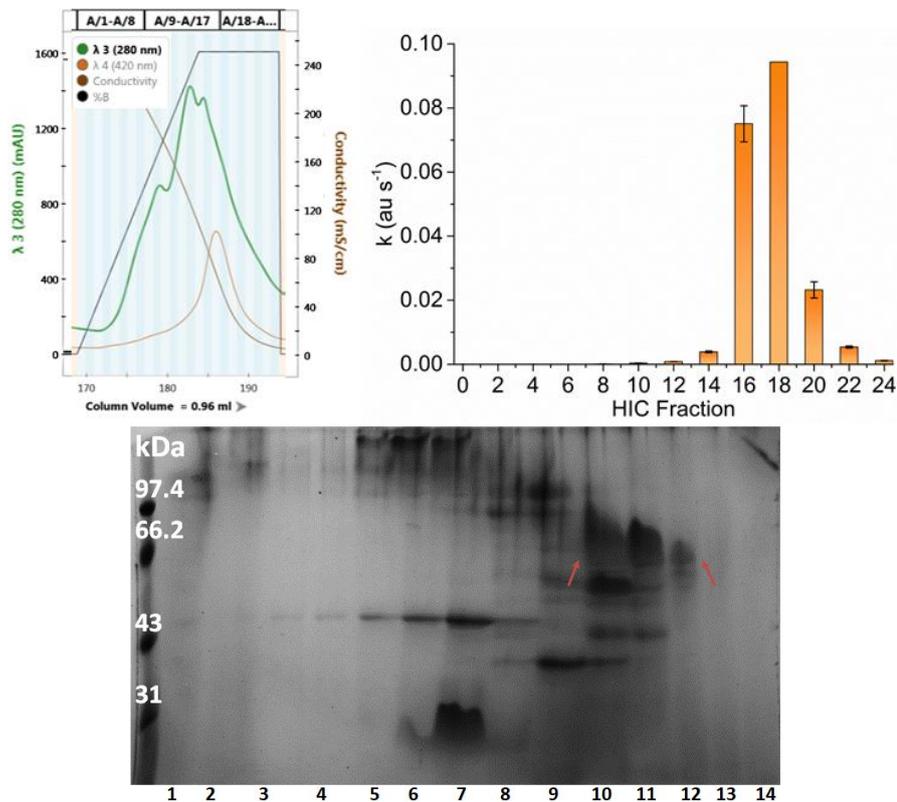


Figure 3-19: top left: Chromatogram for the HIC purification of rAaeUPO-H, with the 280 nm absorbance in green, the 420 nm absorbance in orange, the concentration of buffer B (no salt buffer) in black, and the conductivity in brown; **top right:** reaction rate for the conversion of NBD plotted against the HIC fraction number; **bottom:** 12% SDS gel with low molecular weight marker showing the purification of rAaeUPO using HIC with: 1 – column load, 2 – flow-through, 3 – fraction 2, 4 – fraction 4, 5 – fraction 6, 6 – fraction 8, 7 – fraction 10, 8 – fraction 12, 9 – fraction 14, 10 – fraction 16, 12 – fraction 18, 12 – fraction 20, 13 – fraction 22, 14 – fraction 24; highlighted with arrows the target protein (approx. 60 kDa).

As the pooled fractions from the HIC purification still displayed impurities of lower molecular weight proteins (see Figure 3-20, bottom right, load sample), more steps were needed to isolate the target protein. The salt-free sample was applied to an anion-exchange chromatography column and the elution followed two wavelengths, as described above. A two-step gradient procedure was established, slowly

increasing the salt concentration to 300 mM over 10 CV and then a rapid increase to 2 M sodium chloride over 4 CV. The chromatogram shows the elution of proteins at different salt concentrations; the 420 nm absorbance showed the strongest response in the fractions 7 to 11 (see Figure 3-20, top). Again NBD activity assays helped to identify those fractions containing peroxygenase, which overlapped greatly with the 420 nm absorption peak (see Figure 3-20, bottom left). The SDS-PAGE analysis further confirmed these findings and also showed the effective isolation of rAaeUPO-H from other proteins. Most impurities were found to elute off the Q-column at higher salt concentrations. The fractions 7 to 11 were pooled and concentrated to a final volume of 300 μ L for the subsequent size-exclusion chromatography purification.

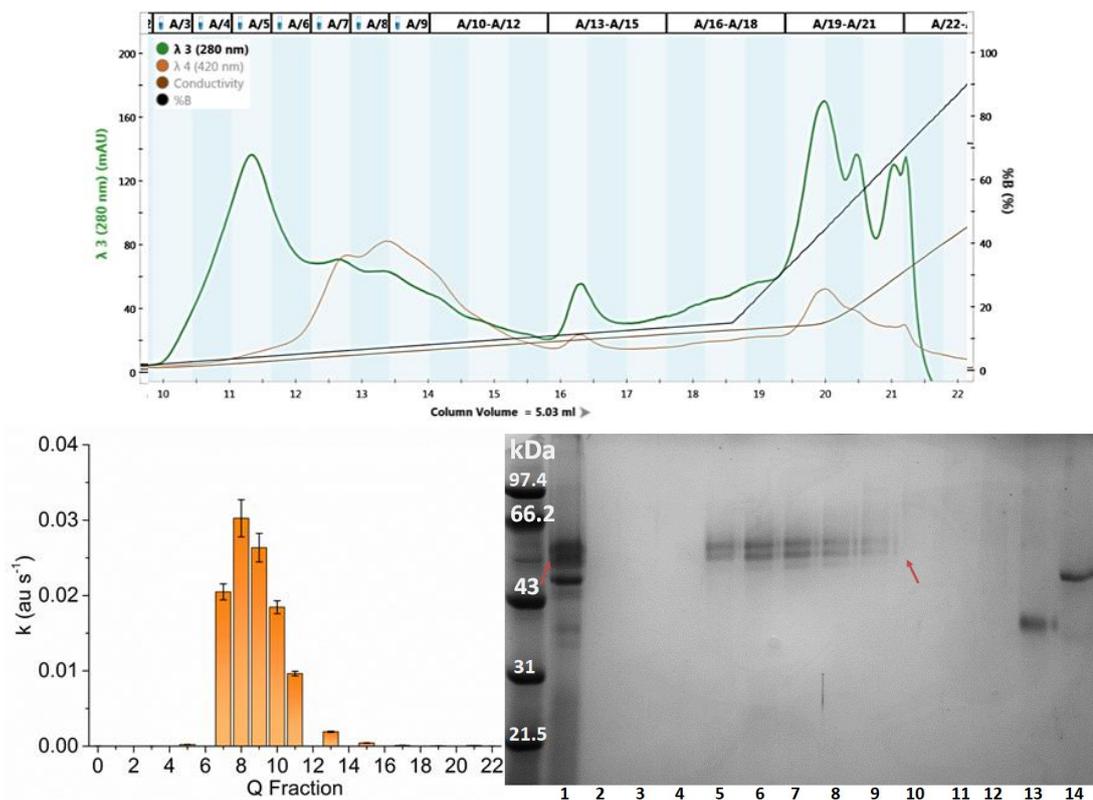


Figure 3-20: **top:** Chromatogram for the Q purification of rAaeUPO-H, with the 280 nm absorbance in green, the 420 nm absorbance in orange, the concentration of buffer B (high salt buffer) in black, and the conductivity in brown; **bottom left:** reaction rate for the conversion of NBD plotted against the Q fraction number; **bottom right:** 12% SDS gel with low molecular weight marker showing the purification of rAaeUPO using Q chromatography with: 1 – column load, 2 – wash, 3 – fraction 3, 4 – fraction 5, 5 – fraction 7, 6 – fraction 8, 7 – fraction 9, 8 – fraction 10, 9 – fraction 11, 10 – fraction 13, 12 – fraction 15, 12 – fraction 17, 13 – fraction 19, 14 – fraction 21; highlighted with arrows the target protein (approx. 55-60 kDa).

The aim of the final purification step using SEC was to fully isolate rAaeUPO from contaminants and gain access to a high purity protein sample. The concentrated

protein was applied to an analytical SEC column using a sample pump and the elution was monitored, as described earlier, using the absorbance at 280 and 420 nm. The chromatogram showed one clear peak eluting in the fractions 8 to 10 (see Figure 3-21, left). The peak absorbance measured at 420 nm was with 0.2 AU double that measured at 280 nm (0.1 AU), indicating the presence of heme-thiolate protein at a good purity (to be discussed in more detail below). The SDS-PAGE analysis further showed *rAaeUPO-H* eluting in high purity, with its two literature reported isoforms around 50-60 kDa. A second much smaller peak was observed in fraction 14; gel analysis did not identify the peak as target protein. The fractions 8 to 10 were pooled and concentrated to a volume around 150 μL . The concentration of the sample was calculated using the absorbance read at 280 and 420 nm, giving a final value around 3 mg mL^{-1} . Judging by the final amount of purified protein, it was estimated that the fermentation of *rAaeUPO-H* on the 2 L scale yielded approx. 3 mg of protein per litre harvested fermentation supernatant.

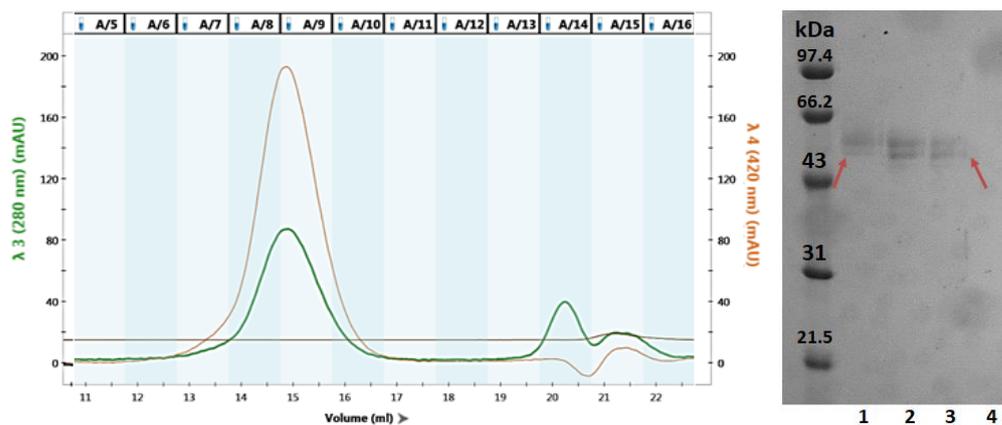


Figure 3-21: left: Chromatogram for the analytical SEC purification of *rAaeUPO-H*, with the 280 nm absorbance in green, the 420 nm absorbance in orange; right: 12% SDS gel with low molecular weight marker showing the purification of *rAaeUPO* using SEC with: 1 – fraction 8, 2 – fraction 9, 3 – fraction 10, 4 – fraction 14; highlighted with arrows the target protein (approx. 55-60 kDa).

Using the purified *rAaeUPO* it was possible to obtain a UV/Vis spectrum (see Figure 3-22), allowing to compare the positioning of the Soret peak identifying the heme-thiolate linkage and the location of the charge-transfer-(CT) or Q-bands to literature known values for the wild-type and *Pichia* expressed *AaeUPO*. In this work the Soret band was found with a maximum at 419 nm, and the CT1 and CT2 bands were located at 568 and 538 nm, respectively. These values are comparable to the literature described findings for *Pichia* expressed *rAaeUPO* with 418 nm for the Soret band and 570 and 537 nm for CT1 and CT2, respectively.¹³³

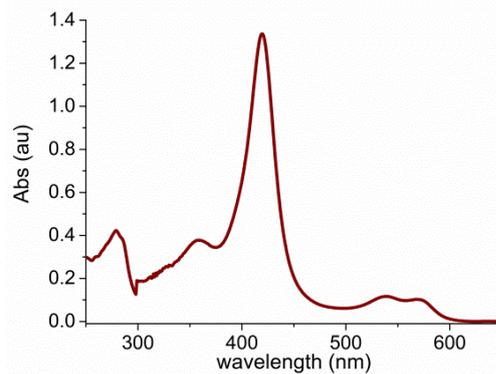


Figure 3-22: UV/Vis spectrum for the purified and concentrated rAaeUPO-H.

From the UV/Vis spectrum it is further possible to determine the Reinheitszahl (Rz). This factor is a measure of purity which is defined as the ratio between the absorbance at 420 nm (heme-thiolate) and the absorbance at 280 nm (protein). In the literature Rz values above 2 were used to describe pure (r)AaeUPO. After the final purification step and concentration of the pooled fractions, a Rz value of 3 was determined from the UV/Vis spectrum shown (Figure 3-22), indicating a highly pure rAaeUPO-H protein sample.

3.3.2. Protein Identification and Glycosylation Analysis

The proteolytic cleavage of the His-tag which was observed during the fermentation of the protein made it essential to study the protein's integrity. Protein identification by mass spectrometry is a simple and quick way to gain insight into the amino acid sequence of a given protein. In this case the purified rAaeUPO-H was isolated on a denaturing gel, and the excised band was sent for protein ID analysis. The sample was subjected to a trypsin digest before fragment analysis using MALDI-TOF/TOF. Analysed fragments covered approx. 70% of the mature rAaeUPO-H sequence. It was possible to detect fragments identifying both the N- and the C-terminal sequence of the protein, confirming that the cleavage of the C-terminal His-tag did not affect the integrity of the mature protein itself.

Further experiments focused on the analysis of the glycosylation of the *Pichia* expressed rAaeUPO-H. These studies were performed to allow a comparison of the expressed protein to the literature described wild-type and yeast expressed variant. The glycosylation pattern of *Pichia* was a further point to be addressed by these experiments, as link between the glycosylation and the folding and activity of AaeUPO was suspected. The glycosylation was analysed in three distinct

experiments. The first experiment used PNGaseF to cleave *N*-linked sugars from the denatured protein. The results were visualised using SDS-PAGE analysis (see Figure 3-23). The molecular weight of *rAaeUPO*-H and the deglycosylated form were calculated using the migrated distances of the ladder to create a calibration curve. The two isoforms of *rAaeUPO*-H appeared with molecular weights of 50 and 55 kDa, the deglycosylated *rAaeUPO*-H displayed a molecular mass of 40 kDa. This is slightly higher than the expected mass, calculated from the amino acid composition (35.9 kDa) plus the prosthetic heme (0.6 kDa). The level of glycosylation was calculated according to the literature described method, by looking at the ratio between deglycosylated and glycosylated protein. From the SDS-PAGE analysis for the deglycosylation of *rAaeUPO*-H, the levels of glycosylation were estimated to be around 25% using the SDS-based molecular mass for the deglycosylated protein, or around 30% using the calculated, theoretical molecular mass.

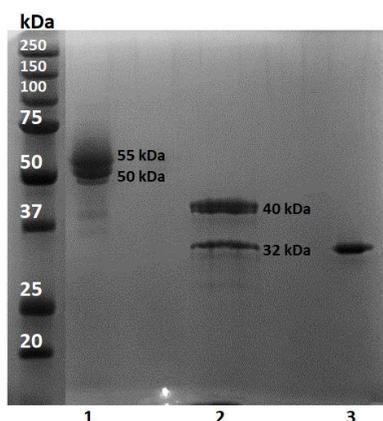


Figure 3-23: 12% SDS gel run with a pre-stained ladder showing the denatured *rAaeUPO* (lane 1), the deglycosylation mixture (lane 2) and PNGaseF (lane 3); the molecular weights for the proteins were calculated using the migration of the ladder to create a calibration curve.

Next, LC-MS and linear MALDI-MS experiments were performed to help characterise the glycoprotein. The LC-MS experiments revealed the presence of a deglycosylated *rAaeUPO*-H, still coordinating the heme (MW = 36.5 kDa). It further identified a peak corresponding to a molecular mass of 42.3 kDa as the dominant glycosylated peak. These findings would suggest a glycosylation degree of approx. 15%, which is significantly lower than the ones reported earlier in this work and in the literature.^{69, 120, 133} However, the nature of the LC-MS experiment did not allow to comment on the presence of further, less dominant glycosylated isoforms and this high-energy method is known to deglycosylate weak sugar linkages early, not allowing the

detection of the native, fully glycosylated rAaeUPO-H. The experiment did however allow to follow the glycan fragmentation, which identified chains of up to ten sugar moieties (see Figure 3-24).

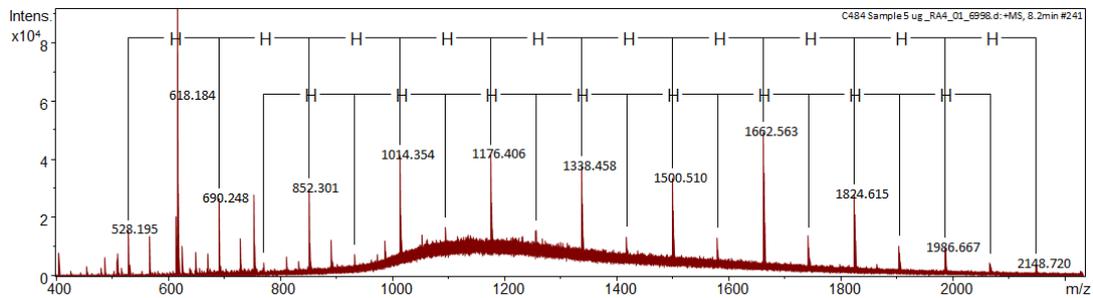


Figure 3-24: Glycan fragmentation for two glycan chains determined in LC-MS experiments, displayed is the C₄ elution under standard source energy; with the m/z^{-1} plotted against the signal intensity, the heme at 618 m/z^{-1} , H for a hexose moiety (approx. 160 Da).

Linear MALDI-MS experiments were performed trying to obtain a mass for the fully glycosylated rAaeUPO-H. Using this method, the peaks observed presented the average mass, rather than the monoisotopic mass which was determined for the LC-MS experiment. It was further possible to observe the multiple charged species. In the spectrum two glycosylated isoforms were observed, though not resolved from each other, with masses of 48.2 kDa and 46.5 kDa (see Figure 3-25). The higher charged species were observed at 23.9 (+2), and 15.6 kDa (+3). The deglycosylated rAaeUPO-H was identified with an average molecular mass of 37.1 kDa, and the multi charged species appeared with 18.4 (+2) and 12.2 kDa (+3). The glycosylation degree which was calculated from the masses obtained using linear MALDI-MS experiments was around 25%. The findings are more in line with the results obtained using PNGaseF, and they are also closer to the findings reported in the literature, with glycosylations degrees of 25% for the wild-type AaeUPO and 30% sugar content for the *Pichia* expressed rAaeUPO.¹³³

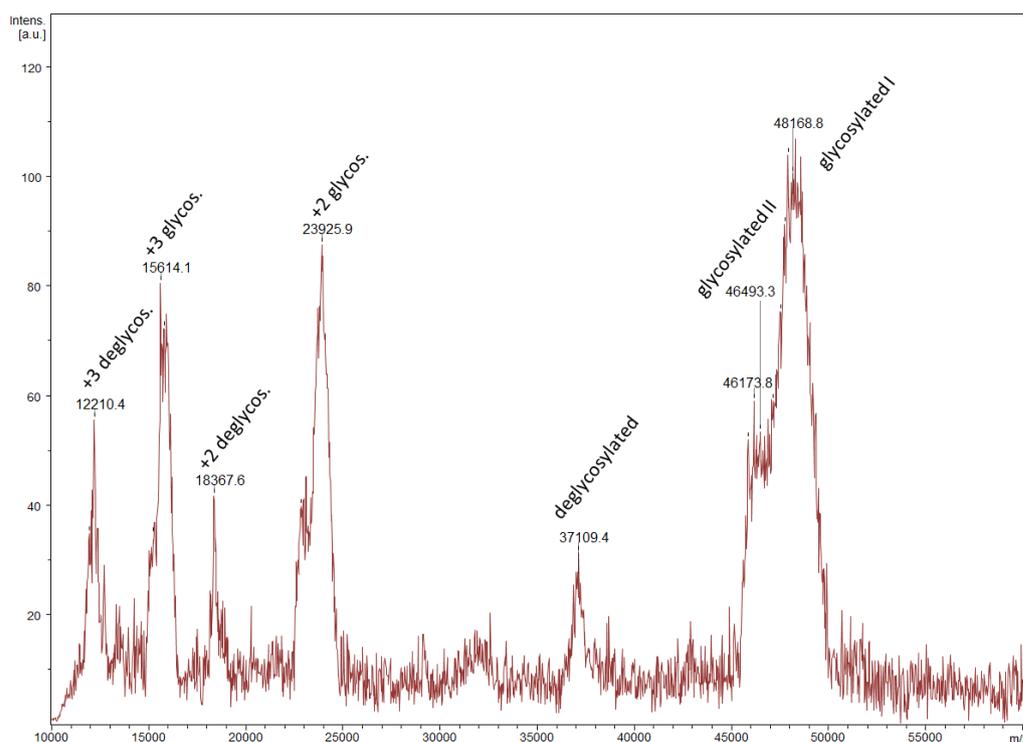


Figure 3-25: Linear MALDI-MS spectrum for rAaeUPO-H, showing peaks for the glycosylated and deglycosylated protein, as well as their higher-charge species.

3.3.3. Kinetic Characterisation

In addition to characterising the glycosylation of rAaeUPO-H, kinetic studies were performed. By determining parameters like the Michaelis constant (K_m) and k_{cat} , it was possible to gain insight into the functionality of the protein and to compare it further to the literature.

The reaction of purified rAaeUPO-H with NBD, ABTS, and veratryl alcohol (VA) at different substrate concentrations were followed using UV/Vis spectrometry. The determined reaction rates were plotted against the substrate concentration to give the Michaelis-Menten plot (see Figure 3-26). From the plot the maximum reaction rate V_{max} and the K_m could be determined (see Equation 3-3). Following Equation 3-4 the turn-over number k_{cat} was determined.

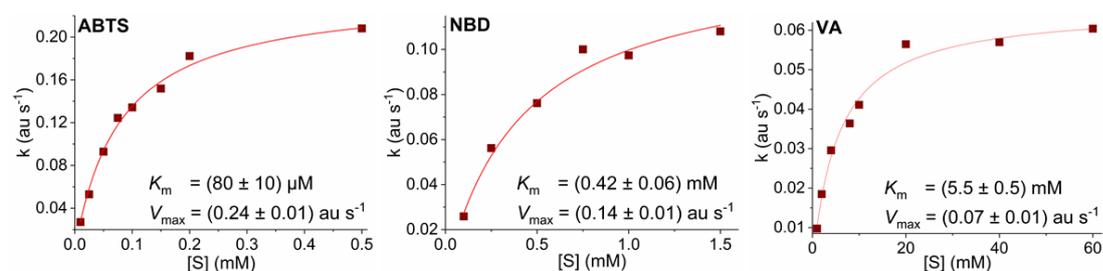


Figure 3-26: Michaelis-Menten plots for the reactions of purified rAaeUPO-H with ABTS, NBD, and VA; with the reaction rate plotted against the substrate concentration. Reactions were run in duplicates and reported with standard deviation.

A summary of the kinetic parameters looking at the peroxygenase (NBD) and peroxidase (ABTS) activity, as well as for veratryl alcohol are presented below (see Table 3-18). In addition, the literature reported parameters for the *Pichia* expressed *rAaeUPO* are also listed for comparison (referred to as literature values).

Table 3-18: Kinetic parameters for the purified *rAaeUPO*-H (this work) and literature values for the *P. pastoris* expressed *rAaeUPO*.¹³³

Substrate	Kinetic constant	<i>rAaeUPO</i> -H	Literature values
ABTS	K_m (μM)	80 ± 10	50 ± 10
	k_{cat} (s^{-1})	600 ± 10	550 ± 20
	k_{cat}/K_m ($\text{s}^{-1} \text{mM}^{-1}$)	$7,700 \pm 700$	$11,000 \pm 2,200$
NBD	K_m (μM)	420 ± 60	850 ± 260
	k_{cat} (s^{-1})	140 ± 10	500 ± 80
	k_{cat}/K_m ($\text{s}^{-1} \text{mM}^{-1}$)	340 ± 70	590 ± 100
VA	K_m (mM)	5.5 ± 0.5	6.0 ± 1.0
	k_{cat} (s^{-1})	70 ± 5	205 ± 15
	k_{cat}/K_m ($\text{s}^{-1} \text{mM}^{-1}$)	15 ± 5	30 ± 5

The conversion of ABTS using *rAaeUPO*-H showed a K_m of $80 \mu\text{M}$, and a k_{cat} of 600 s^{-1} , both values were slightly higher than the literature known values.¹³³ The catalytic efficiency was with $7,700 \text{ s}^{-1} \text{mM}^{-1}$ lower than the one previously reported for *rAaeUPO* expressed in *Pichia* ($11,000 \text{ s}^{-1} \text{mM}^{-1}$). In the case of NBD, the K_m was found to be half that of the literature known value with $420 \mu\text{M}$ versus $850 \mu\text{M}$, respectively. The k_{cat} was also found to be significantly lower, leading to an overall decreased catalytic efficiency of $340 \text{ s}^{-1} \text{mM}^{-1}$ compared to the $590 \text{ s}^{-1} \text{mM}^{-1}$ reported in the literature. For the third substrate, veratryl alcohol, the determined K_m was comparable to the literature, but a three-fold decrease in k_{cat} led to a drop in the catalytic efficiency from $30 \text{ s}^{-1} \text{mM}^{-1}$ to $15 \text{ s}^{-1} \text{mM}^{-1}$.

In addition to determining the kinetic parameters for the *Pichia* expressed *rAaeUPO*-H, pH profiles of the fermentation supernatant (crude) and purified protein were recorded. Both, ABTS and NBD, were considered in this experiment (Figure 3-27). The reaction with ABTS was most active at a pH of 4 for the crude *rAaeUPO*-H and at pH between 4-5 for the purified protein. For NBD, the optimal pH was found to be 7 with relative activities above 60% at pH 6 and 8 for the crude enzyme. All results were comparable to literature reported results for *Pichia* expressed *rAaeUPO*.

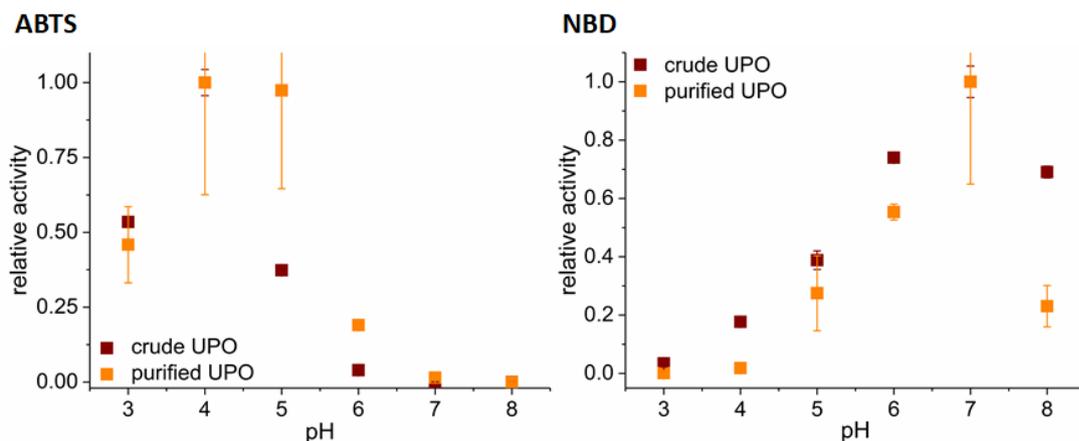


Figure 3-27: pH profiles for the relative activity of the fermentation supernatant (crude, red) and the purified *rAaeUPO-H* (orange), reacting with ABTS (0.2 mM, left) and NBD (1 mM, right). Reactions were run in duplicates and are reported with standard deviations.

3.4. Crystallisation and Structure Analysis of *AaeUPO*

The crystallisation of purified *rAaeUPO-H* was attempted at protein concentrations varying between 2-10 mg mL⁻¹, in the presence and absence of ligands like (*R*)-1-phenylethanol. Initially three 96-well plate screens were set-up to identify conditions suitable for crystallisation. First crystals were observed after six weeks, with more hits appearing over the next six months. An overview highlighting some of the conditions where crystal were obtained is given below (Figure 3-28). *rAaeUPO-H* crystals grew with different morphologies, including columns, broomsticks, branched columns, and needles. Generally, the crystals appeared to have a red colour, due to the presence of heme.

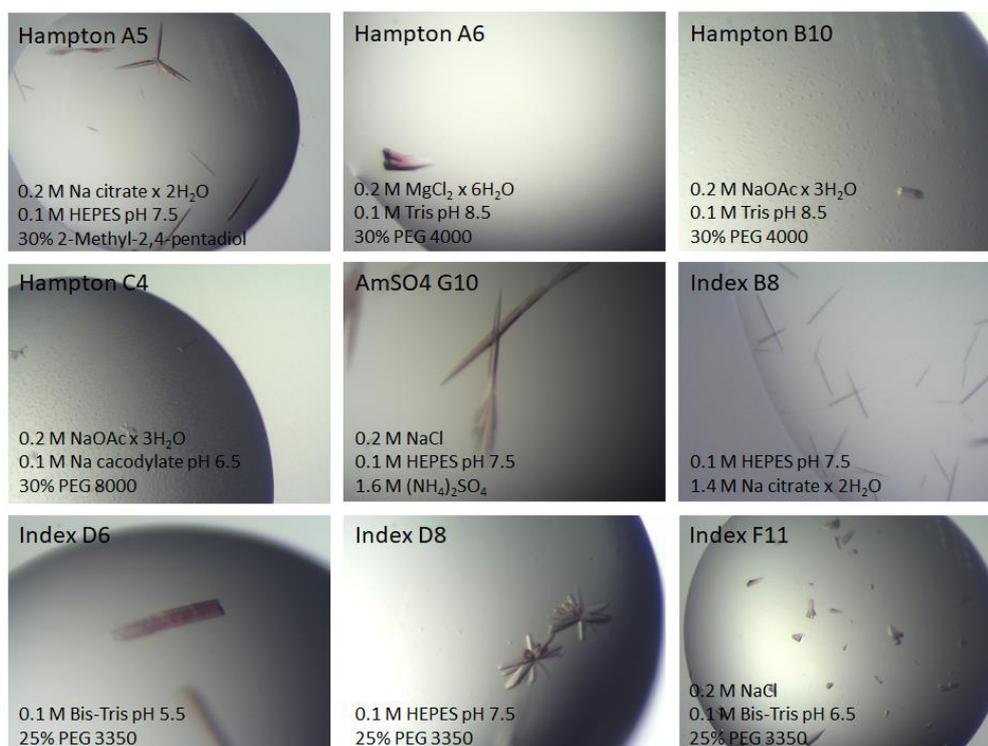


Figure 3-28: *rAaeUPO-H* protein crystals showing different morphologies, obtained in initial 96-well based screens, with the name and composition of the condition given.

The presence of PEG, at different molecular weights, showed a positive effect on the crystal growth (six out of the nine displayed conditions). Crystals were obtained at different pHs, ranging from 5.5 (Index D6) to 8.5 (Hampton A6/B10). Different sodium salts, like chloride, acetate, and citrate, were found in six of the nine conditions, while no salts were added to two conditions (Index D6/D8).

Optimisations for different conditions were performed in both 24-well hanging drop and 48-well sitting drop format. Overall the sitting drop method yielded more crystals. The best results were obtained using 0.1 M Bis-Tris pH 5.5 to 7.5, in the presence or absence of 0.2 M sodium chloride, with PEG-3350 concentrations between 18-32%. The protein concentrations were varied between 2-8 mg mL⁻¹ and co-crystallisation and soaking of obtained crystals was trialled. An overview of the different crystals obtained in the conditions described in given below (Figure 3-29).

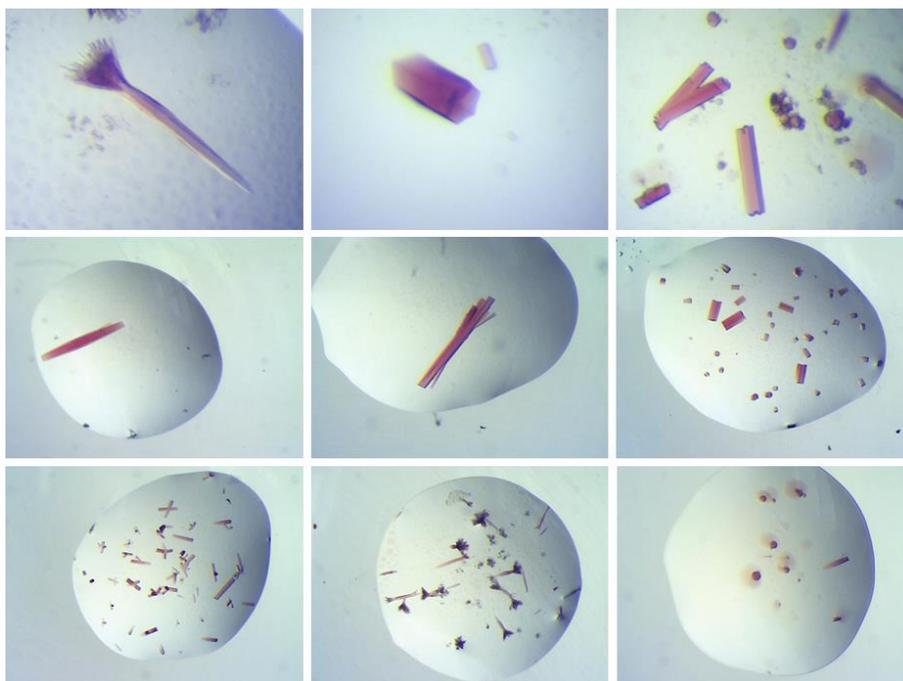


Figure 3-29: *rAaeUPO-H* protein crystals with different morphologies grown in 0.1 M Bis-Tris pH 5.5-7.5, with or without 0.2 M NaCl, PEG-3350 concentrations between 18-32%, using 6.4 mg mL⁻¹ purified protein.

Crystallisation optimisation experiments yielded in *rAaeUPO-H* crystals in slightly different morphologies (Figure 3-29). Longer and shorter hexagonal columns were observed, as well as branching off these columns. Further broomstick like crystals were found in conditions with higher pH and 26-28% PEG-3350. However, a trend between pH, PEG concentration, and salt addition favouring certain morphologies was not observed. Over time the red crystals grew to a length of up to 1 mm, making them visible to the naked eye.

A selection of crystals was flash-frozen and analysed first in-house, and if diffracting well, using the Diamond Light Source. Data was processed and integrated using the Xia2 processing system. The structures were solved using molecular replacement with the structure of *AaeUPO* (PDB code 2YOR) as a model. The structures were refined and after building of the protein backbone and side chains, residual density of the omit map revealed different ligands coordinating to the heme iron and sugar residues in each structure. The structures for *rAaeUPO-H* solved here showed the same overall fold, and *rAaeUPO-H* coordinating to an acetate is presented below (*rAaeUPO-H_ACT*, 5898, Figure 3-30).

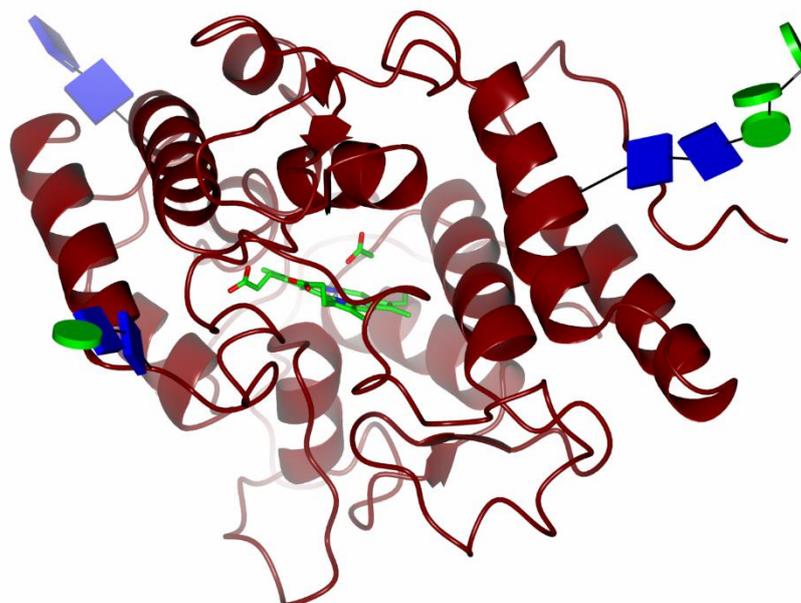


Figure 3-30: Ribbon-structure of the solved *rAaeUPO-H_ACT* (5898); with the protein backbone in dark red, the sugars represented in glycoblocks (*N*-acetylglucosamine (GlcNAc) in blue, mannose (Man) in green), and the heme and acetate ligand coordinating to the iron in cylinder representation (carbons in green, oxygen in red, nitrogen in blue). Created using CCP4mg.

The solved structures for *rAaeUPO-H* were found to be almost identical to the one reported for *AaeUPO*, the wild-type protein. A structural alignment of *rAaeUPO-H* (*rAaeUPO-H_ACT*, 5898) with the wild-type *AaeUPO* (PDB code: 2YOR) using GESAMT analysed the superimposition of 323 amino acids and calculated a root mean square deviation (RMSD) of 0.48 Å, and a Q-score of 0.97.¹³⁹ The mutation sites were found to only have minor influence on the protein fold. An overview highlighting the five mutations on the mature protein, namely V57A, L67F, V75I, I248V, and F311L, and comparing them to the wild-type protein is given below (Figure 3-31). No change in charge or polarity was introduced by the mutations. Due to the non-polar nature of the substitutions, intramolecular forces, like hydrogen-bonding or ionic interactions were not influenced.

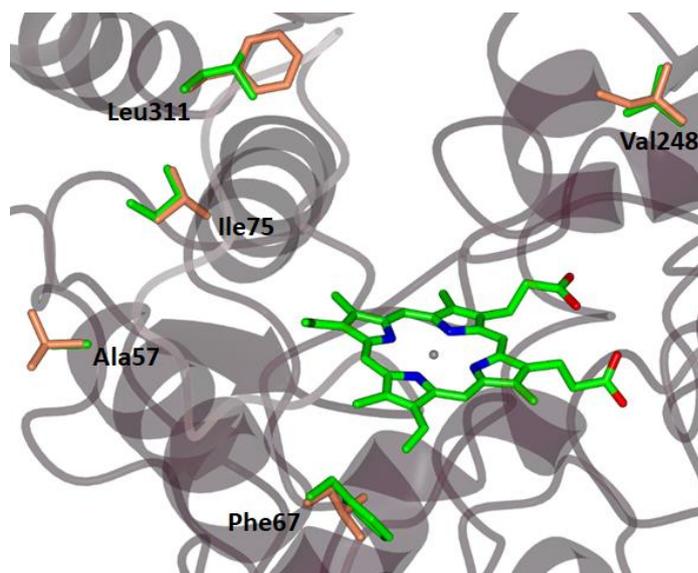


Figure 3-31: Ribbon-structure of rAaeUPO-H_sugars (4181) in dark red, with the heme and mutated side-chains in cylinder representation (C - green, O - red, N - blue), overlaid with the wild-type AaeUPO (2YOR) with crimson cylinders representing the mutated side-chains. Created using CCP4mg.

The previously described *cis*-peptide between Pro108 and Pro109 was found to be conserved in this structure, the halide binding site and the magnesium-ion coordinated by the heme, Glu122, Gly123 and Ser126, were also present in the structure solved here.⁷⁰ While the wild-type AaeUPO crystalized as a dimer revealing five glycosylation sites (Asn11, Asn141, Asn161, Asn182, Asn286), this study showed rAaeUPO to crystalize as a monomer with four glycosylation sites (Asn11, Asn141, Asn161, Asn286). Multiple length sugar chains were observed in different crystals, with the one linked to Asn286 always being the most dominant one, with up to nine residues (see Figure 3-32).

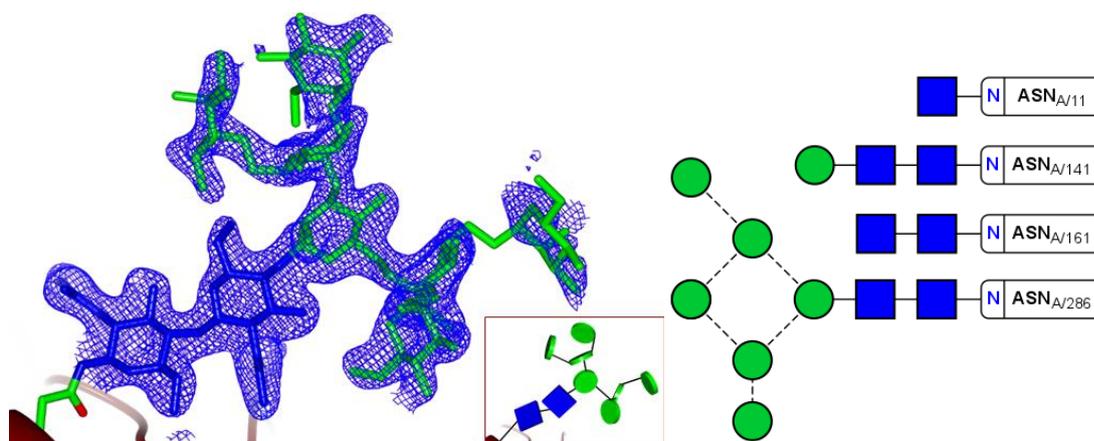


Figure 3-32: left: Ribbon-structure of rAaeUPO-H_sugar (4181) highlighting the sugar-chain off Asn286 (cylinder, C - green, O - red, N - blue); the protein backbone in dark red ribbons, the sugars represented in glycoblocks (*insert*) or cylinders, displaying GlcNAc in blue, and Man in green, and with the electron density corresponding to the $2F_o - F_c$ omit map contoured at 1σ ; right: schematic representation of the different sugar-chains observed in the structure. Created using CCP4mg.

Studies of the wild-type protein showed sugar-chains with up to eight moieties; the longest glycol-chains were found connected to Asn141 and Asn286 (2YOR: Asn141, 2YP1: both). Extensive branching of the sugar-chains was observed for *rAaeUPO-H_sugars*, confirming the expected high-mannose type glycosylation patterns described in the literature.^{86, 120} Surface exposed asparagine residues were found to form an *N*-link to *N*-acetylglucosamine (GlcNAc) residues. Two GlcNAc residues were followed by a mannose (Man) residue, which started to branch out and connect to further Man residues.

Various attempts to co-crystallize or soak formed crystals with enantiopure products or substrates of AaeUPO have been unsuccessful. In most cases acetate was found to coordinate to the heme iron, bridging to Glu196, which is thought to be involved in the acid-base catalysis surrounding the hydrogen peroxide activation (see Figure 3-33).

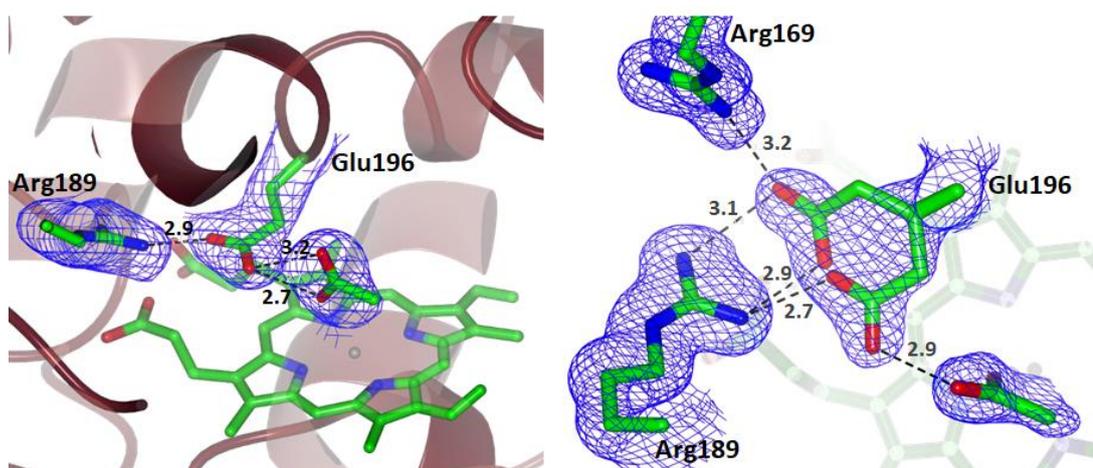


Figure 3-33: Ribbon-structure of *rAaeUPO-H* in dark red, with the heme, acetate, Glu196, Arg169, and Arg189 in cylinder representation (C - green, O - red, N - blue) and the electron density corresponding to the $2F_o-F_c$ omit map contoured at 1σ ; **left:** *rAaeUPO-H_ACT* (5898) structure showing the acetate ligand coordinating to the Glu196; **right:** *rAaeUPO-H_sugar* (4181) with an acetate ligand in the presence of Glu196 in two conformations, interacting with Arg189 and Arg169. Created using CCP4mg.

In *rAaeUPO-H_ACT* an acetate ligand was found to coordinate to the iron at a distance of 3.0 Å. The acetate was further found to interact with the Glu196 with distances of 2.7 Å for the oxygen coordinating the iron, and 3.2 Å for the second oxygen (see Figure 3-33, left). Additionally, an interaction between Glu196 and Arg189 was observed, with 2.9 Å between the acid and base functions. In another acetate-containing structure (*rAaeUPO-H_sugar*) similar interactions between the acetate and the active site residues were observed, with 2.9 Å distance between the

oxygen and the iron, and 3.3 and 2.9 Å distance between the Glu196 and the iron-coordinating oxygen and the second acetate-oxygen, respectively (see Figure 3-33, right). Interestingly in this structure, the Glu196 was observed in two conformations, showing interactions with the literature described base-pair Arg189 at 2.9 and 3.1 Å, but it also identified interactions to a second arginine, Arg169, at 3.1 Å. This movement of the Glu196 could allow for an expansion of the active site, and therefore might hold insight into the substrate scope of *rAaeUPO*, defined by two different acid-base pairs. In the *rAaeUPO-H_NHI* structure an unidentified ligand was found coordinating to the heme (see Figure 3-34, left). In a comparison to the wild-type structures, a similarity to the solved 5-(hydroxymethyl)imidazole ligand was observed. While the *rAaeUPO-H_NHI* structure did accommodate the 5-methylimidazole, no density for the terminal hydroxyl group was observed.

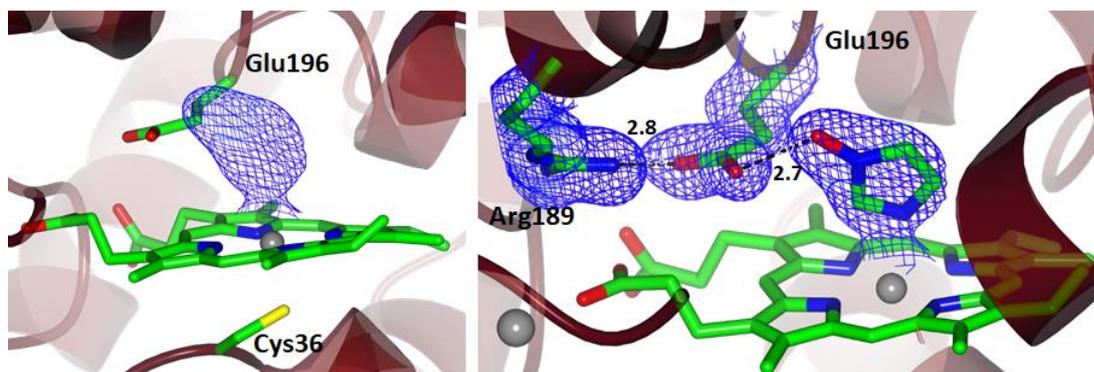


Figure 3-34: Ribbon-structure of *rAaeUPO-H* in dark red, with the heme, ligand, Glu196, and Cys36 in cylinder representation (C - green, O - red, N - blue, S - yellow) and the electron density corresponding to the $2F_o - F_c$ omit map contoured at 1σ ; **left:** *rAaeUPO-H_NHI* (7995) obtained coordinating to an unidentified ligand; **right:** *N*-hydroxyimidazole modelled as possible ligand into *rAaeUPO-H_NHI*. Created using CCP4mg.

Due to the distance to the Glu196 at 2.7 Å, which suggested hydrogen-bonding, it could be suggested that a hydroxylated imidazole was observed within the active site. Literature reported the hydroxylation reactions of other aromatic heterocycles like pyridine in the *N*-position, opened the possibility to model *N*-hydroxyimidazole into the observed density (see Figure 3-34, right).¹⁰¹ The imidazole moiety was further found to coordinate to the iron *via* a nitrogen at a distance of 2.4 Å. Interactions between the Glu196 and the literature reported Arg189 were observed with 2.8 Å between the amine and carboxylic acid residues. The presence of imidazole in the fermentation medium can be explained through the involvement of the molecule in the histidine biosynthesis and breakdown.

3.5. Summary, Discussion and Conclusion

In this chapter the expression of two different *rAaeUPO* constructs in *P. pastoris*, their purification, and characterisation were discussed. The constructs were based on the recently published nine-point *AaeUPO* mutant, which was evolved for yeast expression experiments.¹²⁰ Here, the addition of a C- and N-terminal His-tag to the mature protein were trialled to allow simple and quick protein isolation.

The generation of both constructs was confirmed *via* DNA sequencing and *Pichia* X-33 was transformed with the *rAaeUPO* variants. Expression tests on 2 mL scale helped identifying *Pichia* transformants which incorporated multiple copies of the gene of interest, making them better yielding targets in expression experiments. While the exact number of gene copies was not identified, it was possible, through antibody studies, to identify hit transformants (see Figure 3-7). At this stage, it was also observed that a cleavage of the His-tag occurred, when both fused to the C- and N-terminus of the protein. Over the time the expression was monitored in this experiment, the intensity of the *rAaeUPO* band in the SDS gel was found to increase, while the corresponding signal in the immunoblot decreased after approx. 24 to 48 h. Scale-up experiments in 100 mL shake-flasks showed similar results for *rAaeUPO*-H. The accumulation of protein in the fermentation supernatant was observed over time, however, immunoblot analysis showed a disappearance of the His-tag. These results were strengthened by Ni-affinity chromatography trials. Only a minor amount of protein was found to bind to the column and elute with increasing imidazole concentration, the majority of the *rAaeUPO*-H was found in the column flow-through (see Figure 3-8). For the N-terminally tagged H-*rAaeUPO* the shake-flask experiment showed an improved stability of the His-tag. Immunoblot analysis gave a good signal after 96 h expression and Ni-affinity purification of H-*rAaeUPO* was successful (see Figure 3-9). In addition to the SDS gel, which showed the protein eluting off the column and not in the flow-through, activity assays based on the conversion of ABTS also confirmed the successful purification of H-*rAaeUPO*.

The next stage described the fed-batch fermentation of the protein. The expression was monitored at 0.2 and 2 L scale, and yielded in active *rAaeUPO* variants. A summary of the different fermentation runs and the specific activities towards ABTS and NBD of the obtained for the supernatants is given below (Table 3-19).

Table 3-19: Summary of the specific activities towards ABTS and NBD determined in the supernatants of various fermentations of *rAaeUPO* variants performed in this work.

Fermentation Sample	Specific activity / U mL ⁻¹	
	ABTS (0.2 mM)	NBD (1 mM)
<i>rAaeUPO</i> -H		
0.2 L, 45 h, pH 5 ^T	6.8 ± 0.1	0.4 ± 0.1
2 L, 62 h, pH 5	12.3 ± 0.2	1.4 ± 0.1
0.2 L, 60 h, pH 5	24.0 ± 0.2	3.4 ± 0.1
0.2 L, 60 h, pH 4	1.9 ± 0.1	0.3 ± 0.1
0.2 L, 60 h, pH 3	0.1 ± 0.1	0.0 ± 0.0
0.2 L, 72 h, pH 5, CA	22.1 ± 0.3	3.5 ± 0.4
0.2 L, 96 h, pH 5, CA ^T	+55% (from 72 h)	+25% (from 72 h)
H- <i>rAaeUPO</i>		
0.2 L, 72 h, pH 5, CA ^T	3.1 ± 0.1	0.4 ± 0.1

^T - indicating problems with the temperature control during the fermentation process.

The expression level of *rAaeUPO*-H with a specific activity towards ABTS of 24.0 U mL⁻¹ was found a magnitude lower than the literature reported levels for *rAaeUPO* expression in *Pichia* with 232.0 U mL⁻¹.¹³³ It is possible that the best transformant selected in this work did not incorporate a comparable amount of gene copies, as only a small number of colonies were analysed in initial experiments. The process control, especially with respect to the temperature, could also lead to a discrepancy in specific activities observed. Here, it was possible to show the extent that an increase in temperature had to the production of active *rAaeUPO*. This is especially interesting to follow in the experiment which was run with 96 h protein expression time. The increase in temperature, due to an increased metabolism of the culture and insufficient cooling, coincided with a stagnation in active protein production, which only recovered after approx. one day adjusting to the changed conditions (see Table 3-19). When comparing this data to other experiments studying the integrity of the His-tag in the fed-batch fermentation supernatant over time, it was possible to see, that the increase in temperature after 24-48 h expression, coincided with the loss of the affinity tag (see Figure 3-15). This phenomenon could be explained by the increased cell-lysis occurring in advanced stages of the fermentation, as well as the accumulation of secreted native proteases in the fermentation medium, which are known to facilitate the cleavage of the His-tag.^{138, 140}

Optimisation experiments trying to improve the stability of the His-tag connected to *rAaeUPO-H* included variations in pH and the addition of casamino acids, protease substrates. In no case was an improvement observed. While the addition of casamino acids did not affect the protein expression, lowering the pH from 5 drastically affected the amount of active *rAaeUPO-H* produced in the medium, with a ten-fold decrease at pH 4, and almost no detectable protein at pH 3.

The instability of the His-tag in the fed-batch fermentation process was observed for both the C-terminally tagged *rAaeUPO-H* and the N-terminally tagged *H-rAaeUPO*. The expression level of *H-rAaeUPO* was found to be approx. eight times lower than the level recorded for *rAaeUPO-H*. This difference could be explained by the poor temperature control during the *H-rAaeUPO* fermentation, but it is also possible, that the location of the His-tag, after the modified *rAaeUPO* secretion sequence and before the mature protein, influenced the processing of the protein by the yeast.

Following the successful production of the protein, a purification route, not involving Ni-affinity chromatography, was established. The fermentation supernatant was treated with ammonium sulfate followed by a hydrophobic interaction chromatography step. After dialysis the pooled, protein containing fractions were applied to an anion exchange column, followed by a final size exclusion chromatography step. During the purification the protein containing fractions were identified *via* SDS-PAGE analysis, as well as NBD-based activity assays. It was calculated that the fermentation of *rAaeUPO-H* on the 2 L scale yielded approx. 3 mg purified protein per litre harvested fermentation supernatant.

The purity of the isolated *rAaeUPO-H* was determined using the Reinheitszahl (Rz), relying on the absorbance at 420 nm, corresponding to the heme-thiolate Soret band, and the absorbance at 280 nm, corresponding to the protein back bone. With an Rz of 3.0 the purified *rAaeUPO-H* was comparable to the literature, where a value of 2.4 was described for the *Pichia* expressed *rAaeUPO* (see Table 3-20).¹³³ Furthermore, the UV/Vis spectrum allowed determining the position of the Soret band and the location of the CT- or Q-bands. In this work the Soret band was found with a maximum at 419 nm, and the CT1 and CT2 bands were located at 568 and 538 nm, respectively; all values are comparable to the literature known values (see Table 3-20).

The purified protein was subjected to mass spectrometry based protein identification experiments, which revealed that *rAaeUPO-H* was expressed during the fermentation and that the presence of the proteases and the cleavage of the His-tag left the C-terminus of the peroxygenase intact. An analysis of the glycosylation allowed for a further comparison to the literature. Using MALDI-MS based experiments a glycosylation degree of 25% was determined here, while up to 30% were reported in the literature (see Table 3-20). LC-MS based experiments were not able to capture the mass of the fully glycosylated *rAaeUPO-H*, however, it identified glycosylation branches consisting of up to ten sugar moieties (see Figure 3-24).

Table 3-20: Biochemical and spectroscopic data for the *P. pastoris* expressed *rAaeUPO-H* in comparison to literature known data.¹³³

Feature	<i>rAaeUPO-H</i>	<i>rAaeUPO</i> ¹³³
MW / kDa ^a	48.2	51.1
MW / kDa ^b	35.9	35.9
Glycosylation degree / % ^a	25	30
Rz (A ₄₂₀ /A ₂₈₀)	3.0	2.4
Soret region / nm	419	418
CT1 / nm	568	570
CT2 / nm	538	537

a – determined via MALDI-MS

b – based on the amino acid sequence

The characterisation of *rAaeUPO-H* further included kinetic experiments determining the K_m and k_{cat} values for different substrates. The results obtained in this work for ABTS, NBD, and veratryl alcohol were comparable to the literature described values, with an overview given above in Table 3-18. The overall catalytic efficiencies for *rAaeUPO-H* were slightly lower with $7,700 \text{ s}^{-1} \text{ mM}^{-1}$, $340 \text{ s}^{-1} \text{ mM}^{-1}$, and $15 \text{ s}^{-1} \text{ mM}^{-1}$, for ABTS, NBD, and VA, respectively, compared to the literature reported values for *Pichia* expressed *rAaeUPO* with $11,000 \text{ s}^{-1} \text{ mM}^{-1}$, $590 \text{ s}^{-1} \text{ mM}^{-1}$, and $30 \text{ s}^{-1} \text{ mM}^{-1}$, respectively.

Lastly in this chapter, the crystallisation of *rAaeUPO-H* was described and the solving of three distinct structures, namely *rAaeUPO-H_ACT*, *rAaeUPO-H_sugar*, and *rAaeUPO-H_NHI*. Initial crystallisation screens revealed different conditions yielding in *rAaeUPO-H* crystals in different morphologies. During scale-up diffracting crystals of up to 1 mm in length were obtained in conditions using 0.1 M Bis-Tris pH 5.5 to

7.5, in the presence or absence of 0.2 M sodium chloride, with PEG-3350 concentrations between 18-32%, in sitting drop experiments. The data for the crystals was collected at the Diamond Light Source, processed using xia2, and solved using the wild-type *AaeUPO* structure (2YOR) as a model.^{70, 82} *rAaeUPO-H_ACT* was processed to a resolution of 2.2 Å and revealed an acetate coordinating to the heme iron and the Glu196, described in the acid-base mechanism of the peroxygenase. Interactions of the carboxylic acid residues of Glu196 with the base-partner Arg189, as described in the literature, were also observed (see Figure 3-33).

The *rAaeUPO-H_sugar* structure also contained an acetate ligand in the active site. In addition to the interactions observed in the other structure, Glu196 was observed in two distinct conformations. In the first conformation Glu196 coordinated to the acetate and the Arg189 in a known fashion, in the second conformation Glu196 was rotated away from the active site, not interacting with the acetate anymore, but to both the Arg189 and an additional arginine Arg169. This movement and interaction with a second base could shed insight into the substrate scope of *rAaeUPO* coupled to an extension of the active site. In addition to the ligand binding and acid-base pair interactions, *rAaeUPO-H_sugar* also revealed a well-defined sugar structure of Asn286 (see Figure 3-32). It was possible to model nine glycan moieties, showing the often in the literature described high-mannose type formations.^{70, 86}

The last structure, *rAaeUPO-H_NHI*, revealed a so far unidentified ligand coordinating to the iron (see Figure 3-34). The density observed is similar to the density observed in the wild-type structure of *AaeUPO*, where 5-(hydroxymethyl)imidazole (5-HMI) was modelled as a ligand.⁷⁰ 5-HMI was discussed to be the product of the *AaeUPO* conversion of methylimidazole, which was found in the fermentation medium. The structure reported here would not be able to accommodate the hydroxylated methylimidazole, while methylimidazole could be modelled into the density. The distance between the methyl residue and the Glu196 however, at 2.7 Å suggested a hydrogen-bonding partner, making the methylimidazole coordination less plausible. The coordination of a hydroxylated imidazole would be another possibility. As *AaeUPO* has been reported to hydroxylate pyridine in the *N*-position, *N*-hydroxyimidazole (NHI) was modelled into the density successfully (see Figure 3-34, right).¹⁰¹

Overall the structures obtained for rAaeUPO-H, the mature five-point mutant of AaeUPO, were almost identical with the structures reported for the wild-type, AaeUPO (2YOR). The mutations were not found to influence the overall fold, due to the preservation of the hydrophobic nature of the amino acids mutated in the process. Of the five previously reported glycosylation sites observed in the fungal AaeUPO, only four could be observed in the five-point mutant. Whether *Pichia* expressed rAaeUPO-H with a changed glycosylation pattern remains to be fully investigated. Other structural elements, including a *cis*-peptide, the presence of a chloride, and a magnesium in an octahedral coordination utilising the protein backbone and the heme propionates, were conserved in the rAaeUPO-H structures. In summary, this chapter reported the successful expression of two rAaeUPO variants in *Pichia pastoris*. Fermentation optimisation experiments did not aid the stability of the introduced His-affinity tag, resulting in the untagged, but active rAaeUPO-H five-point mutant. The characterisation of the protein showed high similarities to the literature described, yeast-expressed rAaeUPO. Crystal structure analysis revealed three novel structures of rAaeUPO-H in complex with acetate and a ligand thought to be *N*-hydroxyimidazole. The structures also gave novel insight into a second possible acid-base pair, and described a well-defined nine-glycan long sugar branch.

4. Yeast Expression and Characterisation of the Unspecific Peroxygenase from *Agaricus bisporus*

Since the discovery of *AaeUPO* about a decade ago, different UPO homologs have been reported in the literature. The interest to gain access to these homologs arises from observed changes in selectivity and substrate scope. The first homolog to be mentioned in the literature was *CraUPO*, expressed in its native fungus *Coprinellus radians*, showing conversion products complementary to those observed with *AaeUPO* using polycyclic aromatic hydrocarbon substrates.^{77, 88} The peroxygenases from *Marasmius rotula* and more recently *Chaetomium globosum* have also been successfully expressed in their native fungus and showed improved results in the reactions with human metabolites, terminal oxygenations and the selective functionalisation of testosterone, when compared to *AaeUPO*.^{99, 108, 141} *CciUPO* is up to date the only UPO besides *AaeUPO* that has been described to express in a heterologous host, namely *Aspergillus oryzae* (Novozymes patent WO/2008/119780).⁷⁹ The enzyme has been applied to reactions with aliphatic compounds and different vitamin Ds, and showed different reactivities and selectivities compared to *AaeUPO*.^{79, 100}

The interest in gaining access to an increased number of UPO homologs is dependent on an easily adaptable expression system. While directed evolution experiments gave access to *AaeUPO* in high yields, it is not a feasible process to apply to every identified or putative UPO homolog.¹²⁰ Furthermore, the reactivity of novel homologs is unknown, therefore analytical and screening procedures should not solely rely on (spectrophotometric) assays developed with *AaeUPO* on hand.

In this chapter, the expression of a novel homolog from *Agaricus bisporus* (*AbiUPO*) in *Pichia pastoris* will be described and the development of an expression platform easily adaptable for other UPO homologs.

4.1. Material and Methods

The materials and methods for the cloning, expression and purification of *Abi*UPO followed the procedures described for the *Agrocybe aegerita* UPO homolog (see Section 3.1) unless stated otherwise.

4.1.1. Gene Design and Synthesis for *Agaricus bisporus* UPO (K5XIK0)

The native gene sequence of *Abi*UPO homolog was ordered through GeneArt (Life Technologies Ltd.). Sequence alignment with *Aae*UPO allowed the identification of the signalling and pro-peptide regions, as well as the expected mature protein sequence (described in Section 2.2). The mature protein was amplified using the primers listed below (see Table 4-1).

Table 4-1: Gene sequence for the not optimised *Abi*UPO and list of primers used for amplification of for pPICZ α with N- and C-terminal His-tag and 3C site.

*Abi*UPO (not optimised)

```
ATGTTTAGCCTTCTCAATTCGTCACCTCTCGGACTTGCTTGCACCTGGTCTGTTCTTGCTT
TCCTTCGACATATACTTCATTGGGCGGTTTGCCAGAGAAGAATTGGATAGAATATTGC
CATCTCTCCAGTATCGTTCCTCCGGAGCTCCTCCTGGCCCTCTGAAATTCAATGGCACCA
AGCTCGTCAACGACGATCAACACCCATGGAAGCCTTTGAAACATGGTGACATGCGAGG
GCCATGCCAGGATTGAATACGCTTGCTTCGCATGGATATCTTCCTCGTAATGGTATTGC
GACGCCTGTTCAAATCATCAACGCCGTCCAGGAAGGTTTTAACATGGAGAACAGTGTCG
CGAGATTAGTAACCTATGCTGCCATCTGGTTGATGGTAATTCATCACAGACAACTCA
GCATCGGCGGAAAGAGTCCGTTGACTGGTCCTAGCCCACCGGCCCCAGCTAATGCGGC
AGGTTTGAATACGCATGCTCTGTTTCGAGGGAGATGTAAGCATGACTCGAGCGGATGCC
TTCTTTGGTGACAACCACAGTTTTCAACGAGACACTATTTGACGAGTTTACTGCCTTCAGC
AACCAATTCGGTGCCGAAATAACAACCTTACGGTCGCTGCGGAGTACCGATTTCATCG
GATTCAAGAATCTATCGCGACAAATCCCAACTTTTCCTTCGTGTCCCCTCGGTTTTTCACT
GCTTATGCCGAGTCCGTGTTCCCATCAATTTTTTTCATAGATGGACGTCAGGGAGATGG
CCAATTGGATCTTGACGTTGCTCGTGGCTTCTTCCAGAACATGCGGATGCCGGATGGAT
TCCATAGGGCCAGTATGCCGACGGGGCTTGAGGGCCTCGCCGAAATCGCTAGCGTTCA
CCCCATTTCCCCGGTGCCAATGTCAACGGGGTCAATACTTATACTTTTGATCCCAGCTC
CGTGACTTCACAACCTTCTGTCTGTTGTACGTGAATTTGTAAACCAGACCGTGCGCAG
TCTCTACCCAGAACCAACTGGTAATTTAAAGAAGGCGCTCAAAAAGAATTTGGAATTCTT
ATATGGCCCATTCAGCGATCAATGTTCCAAGTCTTTCCTTATGGAAAGGACAATTGA
```

Creation of In-Fusion overhangs for pPICZ-His (C-terminal)

Forward	ACACTCGAGGCCCGATCCCCGGGAGCTCCTCC
Reverse	ATGATGGTCGACGGCATTGTCCTTCCATAAGGAAAGAC

Creation of In-Fusion overhangs pPICZ-His3C (N-terminal)

Forward CTGTTCCAGGGACAATCCCCGGGAGCTCCTC
Reverse ATGATGGTCGACGGCTCAGTCCTTTCCATAAGGAAAGAC

4.1.2. Purification

The following buffers were prepared for protein isolation and purification (see Table 4-2). All are based on Tris-base at pH 7.0, with variables salts and concentrations depending on the purification method. All were made up with ultra-pure water and filter over 0.2 µm filters prior to use.

Table 4-2: Composition of buffers used in the protein preparation and purification.

Resuspension Buffer	Ni-Purification Buffer
20 mM Tris-Cl	20 mM Tris-Cl
150 mM NaCl	150 mM NaCl
pH 7.0	500 mM Imidazole
	pH 7.0

4.2. Selection and Expression of *AbiUPO*

4.2.1. Construct Design

To gain access to UPO homologs, literature suggesting that signalling sequences evolved for a specific enzyme often showed increased secretion levels for proteins with high sequence identities was taken into consideration.^{142, 143} The constructs created here were therefore carrying the evolved *rAaeUPO* signalling sequence, containing four point mutations. The homolog from *Agaricus bisporus*, identified as described elsewhere through sequence alignment with 65% sequence homology (see Section 2.2), was used for proof-of-concept experiments. Two constructs were prepared to give the N- and C-terminally tagged *AbiUPO*. In both cases, the mature protein was amplified *via* PCR and In-Fusion overhangs were created to allow insertion into the respective vectors. The vectors were prepared from the *rAaeUPO*-H and H-*rAaeUPO* plasmids keeping the N-terminal signalling sequence and C-terminal His-tag or the N-terminal signalling sequence connected to the His-3C, respectively. The annealing reactions of the genes and vectors were performed using In-Fusion cloning and Stellar competent cells were transformed with the resulting plasmids. After overnight growth, colony PCR reactions identified successful transformants. The amplified plasmid was isolated and sequenced. In the case of H-*AbiUPO*, the construct was successfully created carrying the modified *AaeUPO* signalling sequence followed by the His-tag and the 3C cleavage site on the N-terminus of the mature *AbiUPO* gene. In case of *AbiUPO*-H, sequencing revealed that due to primer misconstruction the C-terminal 3C-site and linker between protein and His-tag were deleted. A schematic representation of the final two *AbiUPO* gene constructs is given below (Figure 4-1).

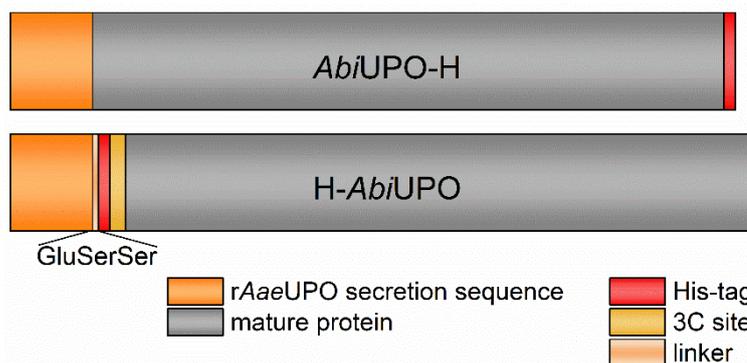


Figure 4-1: Comparison of the yeast constructs *AbiUPO-H* (top) and *H-AbiUPO* (bottom), with the mature protein in grey, the *rAaeUPO* secretion sequence in orange, the His-tag in red, the 3C protease site in yellow, and the linker region in apricot, with description of the three amino acids used as linker.

4.2.2. Test Expression

Similar to the test expression described for *rAaeUPO* variants, small-scale experiments were performed to look at the *AbiUPO-H* and *H-AbiUPO* yeast transformants. The transformation of X-33 with *AbiUPO-H* was not very efficient yielding only a small number of colonies. On a 2 mL scale expression tests were performed, with 72 h incubation on methanol. The dot-blot analysis of these samples using ECL stain was inconclusive (see Figure 4-2, left). It was expected for the functionalised dots to emit light, similar to the results shown for *rAaeUPO* (see Section 3.2.2). However, a signal was only detected on the periphery of the dots, leaving the inside black. This could be due to leakage of the protein solution out of the dot during the functionalising the membrane or due to the uneven utilisation of hydrogen peroxide. UPOs are also known to utilise hydrogen peroxide, another option is that the protein immobilised on the membrane was still functional and competed for the peroxide needed for the luminol oxidation. Due to these ambiguous results, the membrane was further analysed using the SigmaFast tablets to produce the stain. When the substrate/peroxide solution was added to the membrane the evolution of gas around the functionalised dots was observed (see Figure 4-2, middle). This type of reaction was not observed previously and could explain the shortage of hydrogen peroxide seen in the ECL experiments. After incubating the membrane with the stain for 12 h, the typical colorimetric response was detected (see Figure 4-2, right). Strong signals were observed for the 24 h samples, with the intensity decreasing over time. The transformant 1 showed a stronger response and was chosen for further experiments.

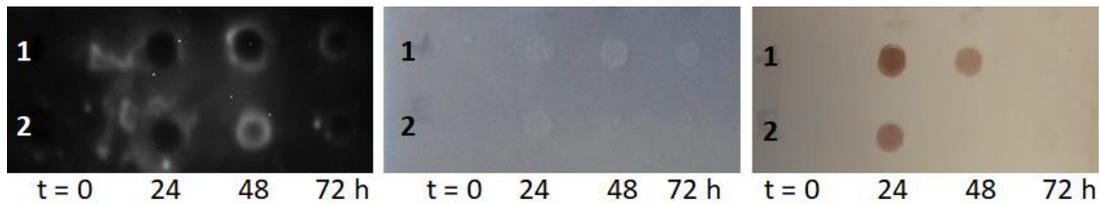


Figure 4-2: Dot-blot analysis of two microfermentations with *AbiUPO-H* over time using: **left** – an ECL-based stain, showing response only on the outside of the functionalised dots; **middle** - SigmaFast tablet-based stain, moments after addition of the stain to the membrane, highlighting gas evolution; **right** - SigmaFast tablet-based stain after 12 h incubation, showing the typical stain response.

The microfermentation of H-*AbiUPO* transformants was run with 96 h on methanol. The results of the dot-blot analysis of H-*AbiUPO* were more comparable to the ones reported for *rAaeUPO*. Using the ECL stain, it was possible to obtain well-defined signals corresponding to His-tagged protein. The functionalisation of the membrane was not ideal, and instead of the protein accumulating in one dot, it leaked and led to blurry signals. However, the strength of response could still be analysed, showing that transformants 3 and 4 had lower expression levels when compared to 1, 2, 5 and 6. Interestingly samples 1 and 2 showed anti-His response after 24 h, which decreased in the 48 and 72 h samples and increased again after 96 h. The other transformants did not show a signal in the 24 h samples, but in the 48 h samples followed by a decreased signal after 72 h and an increase towards the last recorded sample after 96 h. For H-*AbiUPO* the transformants 1 and 6 were used in the following experiments.

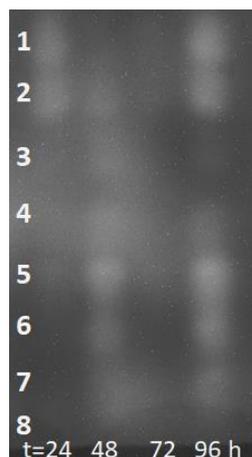


Figure 4-3: Dot-blot analysis of the test expression with H-*AbiUPO P. pastoris* transformants; samples taken after 24, 48, 72, and 96 h protein expression, and visualised using ECL stain.

4.2.3. Shake Flask Expression and Purification Trials

Following the identification of suitable transformants, shake flask experiments on 100 mL scale were performed. The fermentation supernatant of *AbiUPO-H* was

monitored during the expression and analysed using SDS-PAGE (see Figure 4-4). In the samples only one distinct band at approx. 100 kDa could be observed. A minimal increase in intensity of the signals was detected over time. As *AbiUPO* has not been studied previously, no information about possible glycosylation, or degrees of glycosylation were known. According to the amino acid sequence, a molecular weight of 35.8 kDa (not including the His-tag or heme) would be expected. Attempts to identify the protein band of interest using an anti-His immunoblot with the supernatant samples were unsuccessful, as no response was observed on the membrane. In addition, an affinity chromatography purification was attempted with the 100 mL scale *AbiUPO-H* supernatant. The harvested, cell-free solution was dialysed against salt-free buffer and applied to a Ni-column. The chromatogram did not display any peaks, a load sample (labelled dialysis) and the column flow-through were analysed using SDS-PAGE and displayed a similar appearance to the fermentation samples (see Figure 4-4).

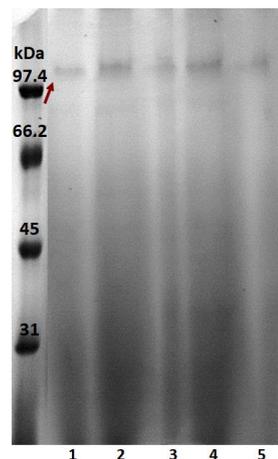


Figure 4-4: 12% SDS gel with low molecular weight marker showing the supernatant samples taken over time from the 100 mL shake flask fermentation of *AbiUPO-H* using *Pichia* (lane 1 – 24 h, 2 – 48 h, 3 – 72 h) and the dialysis (lane 4) and Ni-flow through sample (lane 5).

As during the expression of *rAaeUPO-H* in *Pichia* the introduced His-tag was cleaved over time, additional experiments with *AbiUPO-H* were performed attempting to purify the protein using hydrophobic interaction chromatography. From a second 100 mL shake-flask experiment the cell-free supernatant was applied to a HIC-column and analysed using SDS-PAGE (see Figure 4-5). For the protein fermentation a slightly different expression profile was obtained, compared to the data just discussed (Figure 4-4). In addition to the band around 100 kDa, a band at approx. 65 kDa could be observed, intensifying over time, and an additional band at 30 kDa

appeared after 48 h expression. This improvement in resolution could be due to a lower salt content in the analysed SDS samples, reducing the smearing of the bands. The binding of the proteins to the HIC column was generally good, however, the protein with a molecular weight around 65 kDa was also observed in the column flow-through. The chromatogram showed the elution of proteins with decreasing salt concentration (see Figure 4-5, right). Different to chromatogram recorded for *rAaeUPO-H*, the 420 nm absorbance did not display one sharp peak, and therefore did not allow the identification of the fractions potentially containing *AbiUPO-H*. In the lanes corresponding to the broad elution peak, proteins of different molecular weights eluted at slightly different volumes without identifying one major contributor.

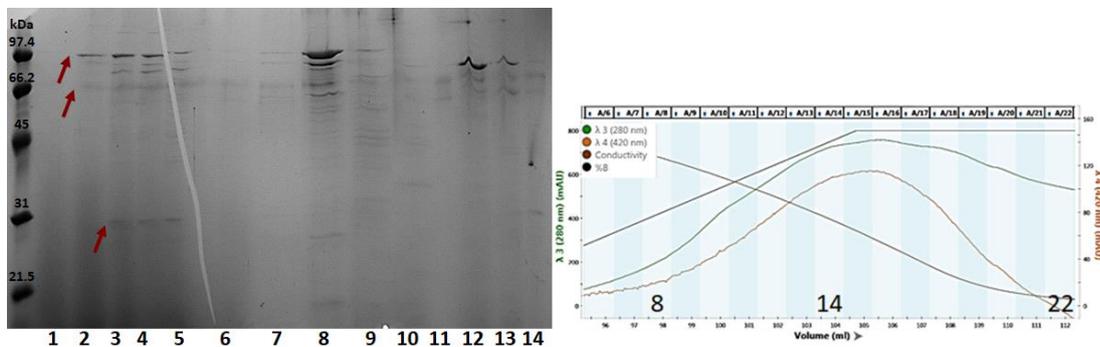


Figure 4-5: left: 12% SDS gel with the low molecular weight marker showing the 100 mL shake flask fermentation of *AbiUPO-H* over time (lane 1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h), the dialysed sample (lane 5) and the analysis of the HIC purification, with: 6 – flow-through, 7 – fraction 8, 8 – fraction 10, 9 – fraction 12, 10 – fraction 14, 11 – fraction 16, 12 – fraction 18, 13 – fraction 20, 14 – fraction 22; right: chromatogram obtained for the purification of *AbiUPO-H* using HIC, with the 280 nm absorbance in green (scale 0-800 mAU) and the 420 nm absorbance in orange (scale 0-160 mAU), with indications of different fractions.

As the HIC purification did not allow the purification of *AbiUPO-H*, the peak fractions 8 to 22 were pooled, dialysed, and applied to an anion exchange column. The results for the purification are shown below (Figure 4-6). While *rAaeUPO-H* was observed to elute within the first gradient, slowly increasing the salt concentration, during this purification a peak was only observed at 2 M sodium chloride. The analysis of the samples on an SDS gel showed that the proteins observed in the load, with approx. 70 to 100 kDa in weight, were also observable in the flow-through and no noticeable binding to the column was detected.

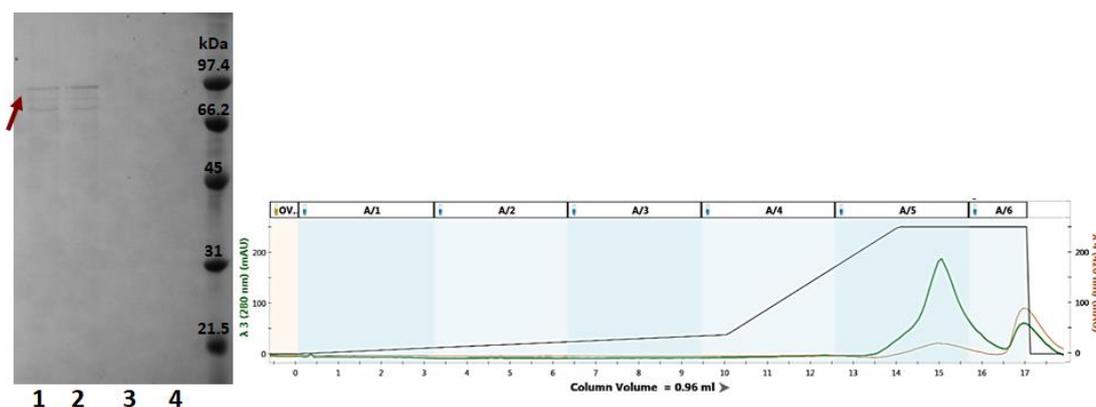


Figure 4-6: left: 12% SDS gel with the low molecular weight marker showing the Q-purification of *Abi*UPO-H, with: 1 – column load, 2 – flow-through, 3 – fraction 5, 4 – fraction 6; right: chromatogram obtained for the Q-purification of *Abi*UPO-H, with the 280 nm absorbance in green (scale 0-250 mAU) and the 420 nm absorbance in orange (scale 0-250 mAU)

The construct with the N-terminal His-tag was also subjected to 100 mL shake flask fermentations and Ni-affinity purification trials. The results obtained using the transformant 6 are shown below (Figure 4-7). Over time several proteins with molecular weights between 75 and 45 kDa showed increasing intensities on the SDS-gel. After dialysis the cell-free supernatant was applied to a Ni-affinity chromatography column and purification of the protein was attempted (see Figure 4-7, right). Two of the strongest bands observed at 75 and 45 kDa did not show binding to the column and were observed in the column flow-through and the wash samples. Proteins with molecular weights around 55, 40 and 30 kDa displayed some binding to the column and eluted in a small peak (fractions 9 to 13). Attempts to positively identify the UPO containing fractions included immunoblots and qualitative activity assays using ABTS and NBD, without success. A quantitative analysis of the fermentation supernatant was performed with increased amounts of protein sample used (20 μ L instead of 10 μ L in a 1 mL reaction). While no activity was recorded towards ABTS, low activities towards NBD were measured at 0.12 ± 0.05 U mL⁻¹. These results encouraged further analysis of the N-terminally tagged H-*Abi*UPO.

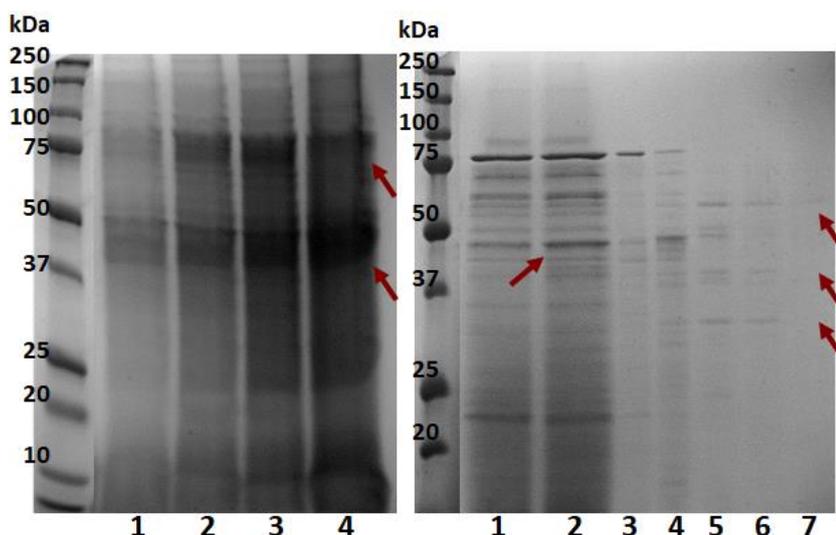


Figure 4-7: 12% SDS-gels with pre-stained markers for **left:** the 100 mL shake flask fermentation of H-AbiUPO (transformant 6) over time, with: 1 – 24 h, 2 – 48 h, 3 – 72 h, 4 – 96 h; **right:** the Ni-affinity purification of H-AbiUPO with the load (lane 1), flow-through (2), wash without (3) and with imidazole (30 mM, lane 4) and the peak fractions 9 (lane 5), 11 (lane 6) and 13 (lane 7).

Following the unsuccessful identification of H-AbiUPO using immunoblot analysis a second transformant was trialled in shake flask experiments (transformant 1). This was due to preliminary studies performed with transformant 1 on the 2 mL test-expression scale which indicated a good response to the immunoblot analysis (data not shown). The experimental grow-up and fermentation was the same as described before and the samples collected over time were analysed using SDS-PAGE and immunoblot analysis (see Figure 4-8). The gel showed proteins with molecular weights between 60 and 100 kDa to increase in intensity over time. For the anti-His antibody analysis the more sensitive ECL stain was chosen, revealing a signal around 60 kDa appearing after 48 h and intensifying over time. These findings, suggesting the presence of H-AbiUPO, were supported by activity assay data. The 72 and 96 h samples were tested for their activity towards ABTS and NBD. Similar to the assays run with the other transformant, no activity towards ABTS was recorded. The specific activity towards NBD were determined to be 0.02 ± 0.01 and 0.05 ± 0.02 U mL⁻¹ for the day-3 and day-4 samples, respectively.

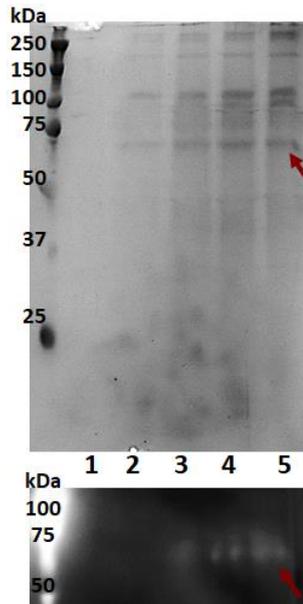


Figure 4-8: 12% SDS gel run with a pre-stained marked and corresponding immunoblot visualised using ECL for the 100 mL shake flask fermentation of H-*AbiUPO* (transformant 1) over time, with: 1 – 0 h, 2 – 24 h, 3 – 48 h, 4 – 72 h, 5 – 96 h.

In summary, the fermentation and purification trials for *AbiUPO*-H were unsuccessful, and no His-tagged protein could be identified. Further activity screens using ABTS and NBD were not showing any activity and therefore not aiding the identification of expressed *AbiUPO*-H. For the N-terminally tagged H-*AbiUPO* two different transformants were tested as initial immunoblot analysis suggested the presence of His-tagged protein in the soluble fraction. Finally, successful expression of the UPO was achieved, aided by activity assays which identified a protein displaying peroxygenase activity towards NBD, while no peroxidase activity towards ABTS was recorded.

4.2.4. Fed-Batch Fermentation

The promising results obtained for H-*AbiUPO* led to the fed-batch fermentation of the protein on 0.2 L scale to increase the amount of protein produced and to allow identifying and characterising the protein. The fed-batch fermentation was run in the presence of 1% casamino acids. The wet cell weight increased from 170 g L⁻¹ at the beginning of the methanol fed-batch phase to 385 g L⁻¹ after 67 h, at harvest. The fermentation log is shown below (see Figure 4-9).

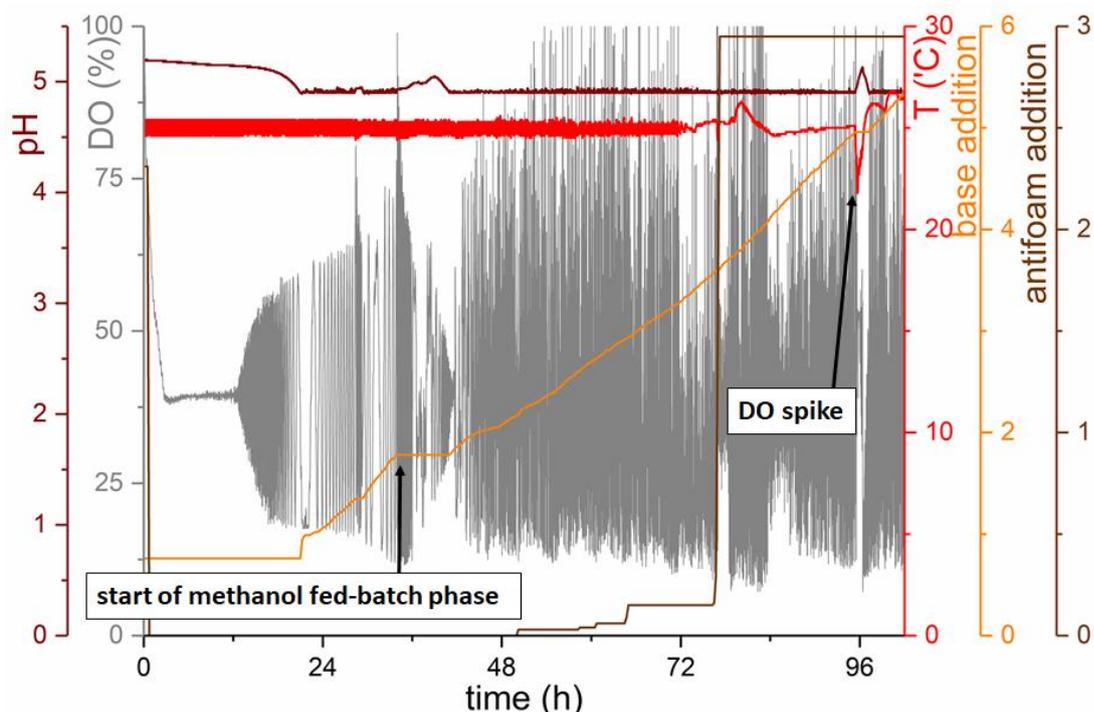


Figure 4-9: Fermentation-log for the H-AbiUPO fermentation run at 0.2 L scale with a 67 h methanol fed-batch phase; with the pH (dark red), DO concentration (grey), temperature (red), base addition (orange) and antifoam addition (brown) plotted over time. The units of the base and antifoam addition are not reported, as the machines internal calibration in mL does not correspond to the actual volume added. Indicated are the beginning of the methanol fed-batch phase, as well as a DO spike.

During the fermentation the temperature was kept at 25 °C, with a small increase to 26.5 °C toward the end of the process. The base addition, after adjusting to the methanol feed, was constant, indicating a well-adapted culture. The feed was stopped to check for a response to limited carbon conditions at around 95 h (indicated in Figure 4-9 with DO spike). A pause in base addition, as well as a drop in temperature correlating to a drop in metabolic rate, were observed. The wet cell weight and the log-data suggested a successful fermentation process. The supernatant samples collected during the methanol fed-batch phase were analysed using SDS-PAGE and immunoblot analysis (see Figure 4-10). The SDS-gel showed different proteins with molecular weights between 50 and 100 kDa with increasing responses over time. The immunoblot coupled to this analysis identified a band corresponding to a molecular weight around 55 kDa with a response appearing after approx. two days of expression which was intensifying over time. The determined specific activity of the fermentation supernatant after 67 h towards NBD was $0.03 \pm 0.01 \text{ U mL}^{-1}$, and no activity towards ABTS was detected.

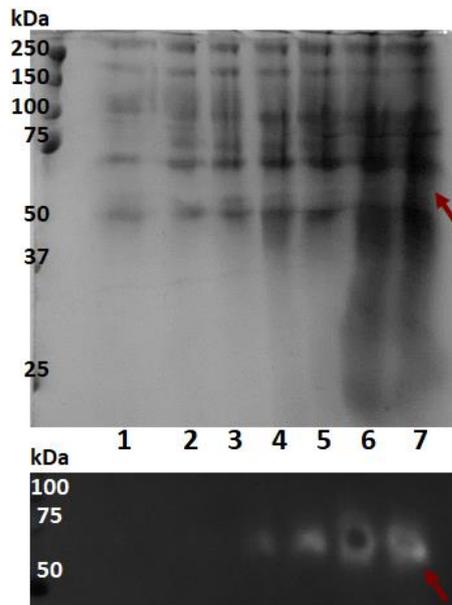


Figure 4-10: 12% SDS gel run with a pre-stained marked and corresponding immunoblot visualised using ECL for the fed-batch fermentation of H-*Abi*UPO (transformant 1) during the methanol fed-batch phase, with: 1 – 0 h, 2 – 13 h, 3 – 22 h, 4 – 37 h, 5 – 45 h, 6 – 61 h, 7 – 67 h.

These results describe the heterologous expression of a novel UPO from *Agaricus bisporus* using *Pichia pastoris*. The modified expression vector, carrying the *rAae*UPO signalling sequence and a C-terminal cleavable His-tag, could allow for quick exchange of the mature UPO sequence and give access to further UPO homologs.

4.3. Purification and Biophysical Characterisation of *Abi*UPO

The dialysed fed-batch fermented protein was subjected to a Ni-affinity purification trial. The SDS-gel of the samples and the chromatogram are presented below (see Figure 4-11). The chromatogram showed a peak in the 280 nm absorbance eluting over fractions 21 to 25, while the absorbance at 420 nm indicated a peak stretching over the fractions 23 to 30. The corresponding gel showed two different proteins with similar molecular weight eluting within the peak fractions (21 to 30). The first peak showed a slightly heavier protein at around 55 kDa, while the second protein eluted with a molecular mass around 50 kDa. Although the 420 nm absorbance peak did not reach the same absorption level as the 280 nm peak, the SDS bands for the 420 nm peak appeared more concentrated.

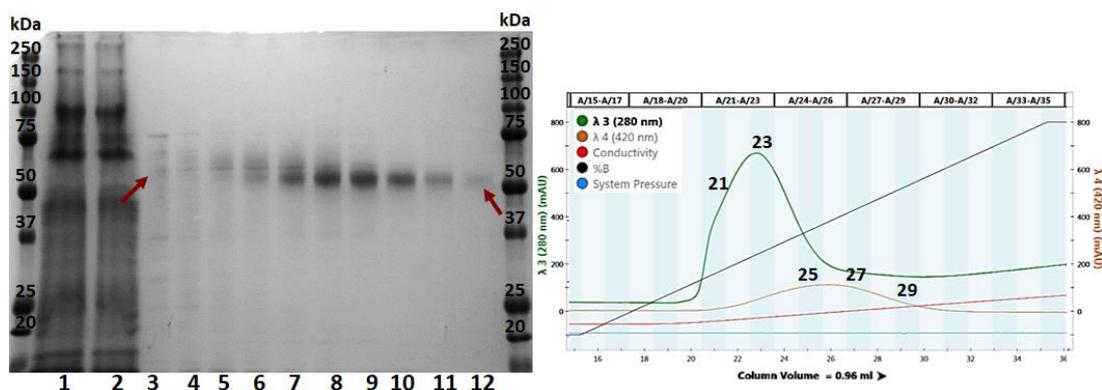


Figure 4-11: left: 12% SDS gel with pre-stained marker showing the purification of H-AbiUPO using Ni-affinity chromatography, with: 1 – column load, 2 – flow-through, 3 – fraction 21, 4 – fraction 22, 5 – fraction 23, 6 – fraction 24, 7 – fraction 25, 8 – fraction 26, 9 – fraction 27, 10 – fraction 28, 11 – fraction 29, 12 – fraction 30; right: chromatogram for the Ni-affinity purification of H-AbiUPO, with the 280 nm absorbance in green, the 420 nm absorbance in orange, and the % imidazole buffer in black.

Activity tests were performed with the different peak fractions to scan for UPO activity. Fractions 21, 23, and 26 showed specific activities towards NBD of 0.03 ± 0.01 , 0.12 ± 0.01 , and $0.10 \pm 0.01 \text{ U mL}^{-1}$, respectively. Towards ABTS no activity was recorded for the sample 21, and for the fractions 23 and 26 the following values were determined: 0.02 ± 0.01 , and $0.03 \pm 0.01 \text{ U mL}^{-1}$, respectively. In addition, an UV/Vis spectrum was recorded for the sample from fraction 26 and is shown below (Figure 4-12).

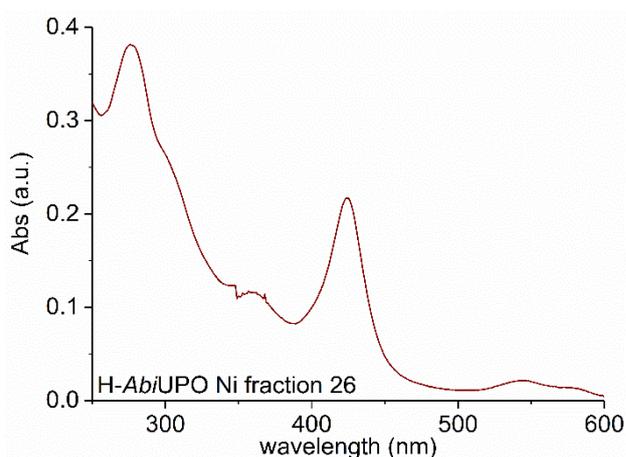


Figure 4-12: UV/Vis spectrum recorded for H-AbiUPO fraction 26 from the Ni-affinity purification.

The spectrum displayed the common UPO features, with a heme-thiolate Soret band at 425 nm, and charge-transfer bands at 577 nm (CT1) and 544 nm (CT2). All bands were red-shifted compared to the data reported for rAaeUPO-H (see Section 3.3) and showed most resemblance to the values reported for CraUPO with 422 nm for the Soret band and 571 and 540 nm for the Q-bands.⁷⁷

To analyse the identity of the protein expressed and purified, mass spectrometry experiments were performed using the excised gel-band of fraction 27. A MASCOT search revealed 34% identity to the sequence of H-*Abi*UPO, without accounting for the secretion sequence. Matching peptides were found between the residues 33 and 237 of the His-tagged mature protein (345 amino acids). No other potential protein hits were identified.

4.3.1. Multistep Purification

Further analysis of H-*Abi*UPO entailed a multi-step purification, where the Ni-affinity chromatography step was followed by size exclusion chromatography (see Figure 4-13). During the Ni-affinity purification a 30 mM imidazole wash step was enforced, changing the chromatogram slightly compared to the results just discussed. As a result, the peaks observed at 280 and 420 nm were overlapping more, and separation from the shoulder, visible in the 280 nm trace, was possible. The fractions 19 to 28 were pooled and prepared for size exclusion chromatography. The sample was brought to a final volume of 350 μ L with a concentration of 2.8 mg mL⁻¹. The chromatogram obtained for the size exclusion step showed one peak with overlapping 280 and 420 nm absorbance. In the 280 nm trace a small shoulder was observed which could correspond to the earlier identified contaminant present in the sample with a slightly increased molecular weight (see Figure 4-11, left).

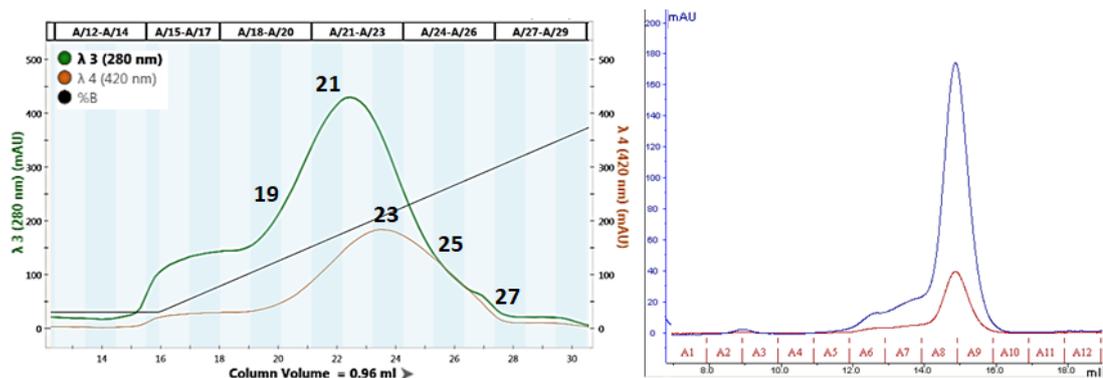


Figure 4-13: Chromatograms recorded for the purification of fed-batch fermented H-*Abi*UPO using **left:** Ni-affinity chromatography, with the 280 and 420 nm in green and orange, respectively, and the imidazole buffer percentage in black; **right:** size exclusion chromatography, with the 280 nm and 420 nm absorbance in blue and red, respectively.

The intensity of the peak observed in the SEC chromatogram with 180 mAU in the 280 nm and 40 mAU in the 420 nm trace was reduced compared to the amount of protein observed in the Ni-affinity chromatogram, with 450 and 200 mAU,

respectively. This observation was confirmed after pooling and concentrating the SEC peak fraction to 200 μL at 0.8 mg mL^{-1} . It seemed like interactions of the H-*Abi*UPO with the SEC column occurred which influenced the amount of protein recovered after the second purification step.

The UV/Vis spectrum recorded for the purified and concentrated H-*Abi*UPO is shown below (Figure 4-14). The location of the Soret and Q-bands were similar to the ones reported for the Ni-affinity purified protein (see Figure 4-12), with 424, 577, and 536 nm, respectively. In contrast to the other spectrum, the 280 nm absorbance was reduced, which gave the H-*Abi*UPO a Reinheitszahl of 1.4, suggesting protein at a good purity level.

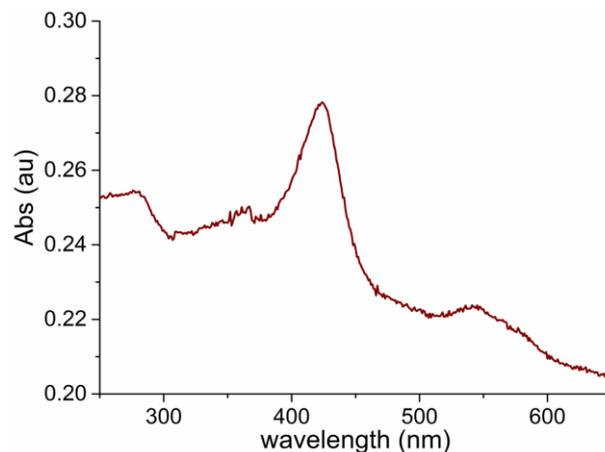


Figure 4-14: UV/Vis spectrum recorded for two-step purified and concentrated H-*Abi*UPO.

Experiments to determine the kinetic constants for H-*Abi*UPO using ABTS, NBD and VA, were performed as they were described for r*Aae*UPO. No determination of the Michaelis-constant or the turnover number was possible, as the experimental set-up did not take the reduced activity towards the mentioned substrates into account. For NBD, the activity decreased from $3.4 \pm 0.1 \text{ U mL}^{-1}$ measured for r*Aae*UPO to $0.02 \pm 0.01 \text{ U mL}^{-1}$ for *Abi*UPO, determined after 72 h expression. Using ABTS as substrate, H-*Abi*UPO did not display any activity. As access to the purified protein was limited and the concentration was low, it was not possible to adjust the reaction set-up to give access to the information in question.

4.4. Summary, Discussion and Conclusion

In summary, this chapter proved that the use of a modified pPICZ vector allowed for the expression of a novel UPO homolog. Previous studies looking at the expression levels of high-redox potential laccases found that the secretion sequence evolved for one laccase also showed increased secretion levels when fused to other laccase genes.^{142, 143} The vector construct here was created around the evolved *rAaeUPO* secretion sequence, carrying four point mutations.¹²⁰ The homolog chosen for expression was from *Agaricus bisporus* which showed the highest sequence similarity to *AaeUPO*, as found *via* BLAST search. Two constructs tagging the protein of interest with a His-tag either at the C- or N-terminus were studied. It was found that only the incorporation of the His-tag between the secretion sequence and the mature protein gave access to soluble protein. It was possible to show that H-*AbiUPO* was functionally expressed by *Pichia*, allowing for the analysis and partial characterisation of a so far unpublished UPO. The identity of the protein was confirmed using mass spectrometry and UV/Vis analysis of the purified H-*AbiUPO* identified the location of the heme-thiolate dependent Soret and Q-bands at 425, 577, and 544 nm. These values are shifted towards longer wavelengths, when compared to *AaeUPO* and other homologs, and are closest related to the values reported for *CraUPO*.⁷⁷ The activity of H-*AbiUPO* towards NBD was reduced when compared to the observations made using *rAaeUPO*-H, with a decrease from 3.4 ± 0.1 to 0.02 ± 0.01 U mL⁻¹ determined in the supernatant of 0.2 L fermentations with expression for 72 h. When using ABTS as substrate, H-*AbiUPO* did not display any activity. Whilst *rAaeUPO*-H shows rapid conversions of ABTS, highlighting the ability to act as a peroxidase, H-*AbiUPO* does not appear to have the same properties. Contrary to the *rAaeUPO* constructs discussed in Chapter 0, H-*AbiUPO* was expressed with an intact His-tag, which allowed for a quick purification of the protein. As the fermentation conditions were maintained, it can only be assumed that the folding of H-*AbiUPO* differs from *rAaeUPO*-H with respect to exposing the introduced 3C protease cleavage site, and therefore protecting the integrity of the tag. Due to the limited access and time constraints of the project, the characterisation of H-*AbiUPO* was not finished. The determination of the kinetic parameters for NBD, VA, and ABTS was attempted using the protocols established for *rAaeUPO*-H. However, due to the decreased activity

observed in specific activity assays, it was not possible to obtain any results and further adjustments of the reactions set-up are needed to allow for the determination of K_m and k_{cat} . Other studies looking at the glycosylation degree and pattern have not been performed on the new H-*Abi*UPO homolog and crystallisation was not attempted so far. Finishing the full characterisation of the homolog and obtaining a crystal structure could give access to more information about the changes in peroxidase/ peroxygenase activity and about the reason behind the integrity of the His-tag. Furthermore, the applicability of the modified UPO expression vector using other homologs opens another field of possible experiments.

5. Biotransformations and Industrial Applications of UPOs

Different UPO homologs are known in the literature to catalyse a wide variety of reactions with a range of selectivities.^{77, 79, 99} Reactions include aromatic, aliphatic, and benzylic hydroxylations, heteroatom oxidations and dealkylations, alcohol oxidations, epoxidations, and one-electron transfer oxidations (for more details see Section 1.2.3). Often reactions proceed with high regio- and stereoselectivity making UPOs industrially interesting targets.

The mutant of *Aae*UPO described in the literature has also been applied to selected biotransformations. However, a comprehensive screen comparing it to the wild-type and establishing its scope have not been reported yet. Also, little information is available on the sulfoxidation activity of *Aae*UPO, or its homologs.

Here, work will be presented obtained during an industrial placement at GSK, a collaborator. This includes a large-scale fermentation and downstream processing of *rAae*UPO-H, and attempts to optimise the handling and applicability of this UPO, through lyophilisation, reaction engineering, and the usage of an *in situ* hydrogen peroxide generation systems. Moreover, a substrate screen including both commercial and GSK confidential substrates helped in identifying novel substrates for both *rAae*UPO-H and the previously unknown H-*Abi*UPO. Furthermore, a design of experiments (DoE) approach helped aid the establishment of a large-scale transformation, permitting a more industrially-applicable process design.

5.1. Materials and Methods

5.1.1. Chemicals and Materials

Chemicals used in this study were purchased from Alfa Aesar (Heysham, UK), Fisher Scientific UK Ltd. (Loughborough, UK), Fluorochem Ltd. (Glossop, UK), Merck Chemicals Ltd. (Nottingham, UK), Scientific Laboratory Supplies Ltd. (Nottingham, UK), Sigma-Aldrich Company Ltd. (Dorset, UK), Tokyo Chemical Industry UK Ltd. (Oxford, UK), and VWR International Ltd. (Lutterworth, UK).

5.1.2. Analytical Instruments

GC experiments (in house) were run on an Agilent 6890N Network GC system fitted with an Agilent 7683B Series Injector using a CFH200 hydrogen generator (Cylinderfree). Achiral analysis was performed using a HP5 column (Agilent Technology, 30 m x 0.32 mm x 0.25 μ m) and chiral analysis was performed with a cyclosilB (C-prefix in method, Agilent Technology, 30 m x 0.25 mm x 0.25 μ m) or BGB175 (B-prefix, BGB Analytik, 30 m x 0.25 mm x 0.25 μ m).

GC experiments (GSK) were run on an Agilent 7890B GC system equipped with a GERSTEL multipurpose sampler and H₂ as carrier gas. Achiral analysis was performed using a HP5 column (Agilent Technology, 30 m x 0.32 mm x 0.25 μ m) and chiral analysis was performed with a cyclosilB (Agilent Technology, 30 m x 0.25 mm x 0.25 μ m) or HYDRODEX β -TBDAC (MACHEREY-NAGEL, 25 m x 0.25 mm x 0.4 mm).

NMR spectroscopy (GSK) was carried out using a Bruker 400 UltraShield B-ACS 60 spectrometer.

5.1.3. Industrial Fermentation and Processing of *rAaeUPO-H*

rAaeUPO was fermented at GSK Worthing by Alison Dann and her team using the procedures described elsewhere (see Section 3.1.6). Four 10 L fermentation vessels were charged with 6 L initial fermentation medium and inoculated with 300 mL YDP overnight culture (5%). The temperature was kept at 25 °C for the duration of the process, and the pH was adjusted to and held at 5 using a 70% NH₄OH feed solution. The dissolved oxygen concentration was kept above 20% by adjusting the agitation (300-950 rpm) or by sparging with air (up to 6 L min⁻¹). The methanol feed rate was

adjusted during the fermentation process to a maximum rate of 52.8 g h⁻¹. The cell-free supernatant was obtained using centrifugation and concentrated by Cyril Boudet, using a 30 kDa hollow fibre membrane (S06-E030-05-N, 0.5 mm, mPES, 2600 cm²). The solution was stored at -20 °C for storage.

5.1.3.1. Protein Lyophilisation

For improved long-term storage and simplified handling of the protein, lyophilisation experiments were performed at GSK Stevenage using a freeze-drying oven. Small amounts of protein were freeze-dried in glass vials overnight using the program “short”, while larger amounts of protein were transferred into shallow basins to increase the surface area and freeze-dried using the “regular” or “weekend” method (see Table 5-1).

Table 5-1: Overview of the different lyophilisation methods, regular, short, and weekend, used in this work, with the temperature, pressure, and duration times of the different steps.

Step	T / °C	Vacuum / mTorr	regular	short	weekend		
			Duration / min	Duration / min	T / °C	Vacuum / mTorr	Duration / min
Thermal Treatment	-29	232	120	60	-40		240
Evacuation	-29				-45	500	
<i>Drying</i>							
Step 1	-24	232	45	25	-40	100	900
Step 2	-20	232	385	200	-20	100	900
Step 3	-16	232	990	500	-20	100	300
Step 4	-13	232	330	170	-15	100	900
Step 5	0	232	180	90	-15	100	180
Step 6	19	232	90	45	4	100	180
Step 7	19	120	30	15	4	500	storage
Step 8	20	120	230	120			

5.1.3.2. Protein Characterisation

The protein obtained during the industrial fermentation and the downstream processing was characterised and compared to previously expressed rAaeUPO-H. SDS-PAGE analysis and specific activity assays towards ABTS and NBD were run as described before (see Section 2.1.7.7 and 3.1.7.6). Bradford assays were performed to determine the protein concentration following the instructions provided with the kit (Coomassie (Bradford) Protein Assay Kit, Thermo Scientific). For the calibration

curve dilutions of a 2 mg mL⁻¹ stock were performed, ranging from 0 to 2 mg mL⁻¹. All protein samples were diluted prior to analysis to obtain absorption within the calibration, routinely 1:2, 1:5, and 1:10 dilutions were prepared. 5 µL sample (protein and standards) was mixed with 250 µL Coomassie stain in a 96-well plate in duplicates and incubated for 15 min at room temperature. Using a plate reader, the absorbance at 595 nm was measured. A linear regression, fitted through the absorbance plotted against the protein concentration of the standards, allowed determining the concentration of unknown protein samples.

Furthermore, the conversion of ethylbenzene was monitored using the different *rAaeUPO-H* samples obtained. Reactions were run in duplicates using 2 mL plastic tubes. Reactions contained 50 mM potassium phosphate (KPi) buffer at pH 7.0 (from 100 mM stock), 10 mM ethylbenzene (0.9 µL), 1 U mL⁻¹ *rAaeUPO-H* sample and 10% acetonitrile (MeCN). The reactions were incubated with shaking at 900 rpm (Thermomixer) at 25 °C. 2 mM hydrogen peroxide (from a 100 mM stock) were added to start the reaction and peroxide was added in 2 mM portions every 15 min for 75 min, followed by two 30 min additions (total addition 1.6 eq). 100 µL samples were taken after 1 and 2.5 h, extracted with ethyl acetate (EtOAc, 200 µL) and analysed using GC (see Table 5-2).

Table 5-2: GC method and condition used for the achiral analysis (HP5) of reactions run at GSK. Numbers refer to structures are displayed in Figure 5-3.

#	Split	Description	Used for
G1	25:1	70 °C to 320 °C at 25 °C min ⁻¹	1-23

5.1.4. High Throughput Substrate and Condition Screening

Both, *rAaeUPO-H* (1-2 U mL⁻¹) and *H-AbiUPO* (20% v/v) were used in screening reactions. The specific activity, relating to the amount of protein used, was determined with respect to ABTS. Screens were performed in either 96-well deep-well plates or in 2 mL plastic tubes, with a reaction volume of 750 µL. All reactions were run in duplicates with a no-enzyme control. In case of the deep-well blocks a Biomek pipetting robot (Beckmann Coulter) was used to prepare the well solutions and for the addition of hydrogen peroxide over time. The block was incubated at 25 °C in an orbital shaker at 220 rpm. For the mixing of components and the

extraction of samples, the block was incubated on a plate mixer at 800 rpm for 5 min before centrifugation (2 min, $2,500 \times g$). For reactions run in 2 mL tubes, the set-up and peroxide addition was performed using multi-dispense pipettes. The tubes were incubated in thermomixers at 900 rpm at 25 °C, unless stated otherwise. For extractions, the samples were vortexed for 15 s and the phases were separated *via* centrifugation (1 min, $16,000 \times g$).

All reactions contained 50 mM buffer (from 100 mM stock, varying pH), substrate in solvent (varying concentrations, from 25-10X stocks), UPO (crude, from different sources) and hydrogen peroxide (1 or 2 mM, from 100 mM stock). Reactions were started through the addition of hydrogen peroxide. Further hydrogen peroxide was added every 30 min until 1.5-2 equivalents were added. Samples were taken by extracting the reaction mixture (usually 100-200 μL) with twice the amount of ethyl acetate. The extraction procedure has been outlined above for the different reaction set-ups. Deep-well block samples were analysed within the block, while 2 mL tube samples were transferred into 1.5 mL glass vials prior to GC analysis (method G1). Unless stated otherwise, substrate, product and, if applicable, by-product markers were analysed on the GC in parallel.

5.1.5. Stirred Small-Scale Reactions

Small-scale stirred biotransformations were performed in 7 mL glass vials in triplicates with no-enzyme control reactions run in parallel. The 2 mL reactions contained rAaeUPO-H (crude, 1.2 U mL^{-1}), 5 mM substrate in 1% ethanol (from 100X/500 mM stock) and were run in 50 mM KPi pH 7.0 (from 100 mM stock). The reactions were started by the addition of 2 mM hydrogen peroxide (from 100 mM stock). Over 2 h H_2O_2 (2 mM, total addition 10 mM, 2 eq.) was added every 30 min. The time course of the reaction was followed by taking 200 μL samples at $t = 0, 30, 60, 120,$ and 180 min and extracting the compounds of interest in equal amounts of ethyl acetate (200 μL). For chiral GC experiments the samples were dried over magnesium sulfate prior to analysis. Details about the different GC methods run for individual samples in-house are given below (see Table 5-3). In case of the ethylbenzene derivatives, (enantiopure) product markers were used for product

identification. For the chiral sulfoxide products, another work was consulted where the retention orders using the BGB175 GC column had been determined.¹⁴⁴

Table 5-3: GC methods and conditions used for the analysis of the ethylbenzene and thioanisole derivatives, with achiral analysis using an HP5 column (numbers only), C-prefix referring to chiral analysis using a cyclosil B column, and B-prefix refers to a chiral BGB-175 column. The numbers refer to the structures displayed in Figure 5-12.

#	Split	Description	Used for
1	1:1	100 °C for 15 min, to 200 °C at 40 °C min ⁻¹	24, 28, 32
2	1:1	90 °C for 10 min, to 200 °C at 40 °C min ⁻¹	1, 26, 30
3	5:1	150 °C for 15 min, to 250 °C at 40 °C min ⁻¹	34, 36, 38
4	1:1	120 °C for 15 min, to 250 °C at 40 °C min ⁻¹	44
5	5:1	150 °C for 10 min, to 250 °C at 40 °C min ⁻¹	40, 42
C1	1:1	100 °C to 145 °C at 2 °C min ⁻¹ , to 200 °C at 40 °C min ⁻¹	24
C2	1:1	90 °C for 35 min, to 200 °C at 40 °C min ⁻¹	1
C3	1:1	90 °C for 50 min, to 200 °C at 40 °C min ⁻¹	26, 30
C4	1:1	110 °C for 30 min, to 150 °C at 1 °C min ⁻¹ , to 200 °C at 40 °C min ⁻¹	28, 32
B1	17.6:1	180 °C for 20 min to 200 °C at 40 °C min ⁻¹	34, 44
B2	17.6:1	180 °C for 30 min, to 200 °C at 40 °C min ⁻¹	42
B3	17.6:1	170 °C for 35 min, to 200 °C at 40 °C min ⁻¹	36, 38, 40

5.1.6. Reaction Optimisation

5.1.6.1. Solvent Screening

The reactions were set up in 2 mL plastic tubes as described above (see Section 5.1.4) and contained 50 mM KP_i buffer pH 7.0 (from 100 mM stock), 10 mM ethylbenzene, rAaeUPO-H (1.3 U mL⁻¹) and 10% solvent. The solvents tested were acetonitrile, dimethyl sulfoxide (DMSO), ethyl acetate, acetone, methyl *t*-butyl ether (MTBE), and ethanol. Hydrogen peroxide (2 mM from 100 mM stock) was added to start the reaction and more additions took place every 15 min for 1 h (1.2 eq total addition). The reactions were incubated shaking at 900 rpm at 25 °C. 100 µL samples were taken after 30 min, and 26 h, extracted with twice the amount of EtOAc (200 µL) and analysed using achiral GC (method G1).

5.1.6.2. Peroxide Addition

Optimisation experiments were run to examine the hydrogen peroxide feed to the reaction mixture. As rAaeUPO-H has different affinities for different substrates, one experiment used ethylbenzene and one used methyl-*p*-tolyl sulfide as substrate.

Using ethylbenzene as substrate, the addition intervals of hydrogen peroxide to the reaction mixture were studied. Reactions were prepared in 50 mL Falcon tubes containing 10 mL reaction volume. Experiments were run in duplicates. The reaction contained 50 mM KPi buffer at pH 7.0, 10% MeCN, 10 mM substrate, 2 U mL⁻¹ rAaeUPO-H and was started through the addition of hydrogen peroxide (2 mM from 100 mM stock). All components were mixed and incubated in a thermomixer (30 °C, 700 rpm). Peroxide was added every 15 or every 30 min and samples were taken every two additions (before the next addition). For sampling 100 μL were extracted with 200 μL EtOAc and analysed using achiral GC (method G1).

Reactions looking at the peroxide loading and addition intervals using methyl-*p*-tolyl sulfide as substrate were run as stirred small-scale reactions (see Section 5.1.5). The reactions of 5 mM substrate with 1.2 U mL⁻¹ rAaeUPO-H were prepared in 50 mM KPi buffer at pH 7.0 in the presence of 1% ethanol. Hydrogen peroxide was added in 1 or 2 mM portions (from a 100 mM stock) every 10 or 30 min. For the 10 min addition experiments, 3.6 eq of H_2O_2 were added (total duration 3 h), and samples (200 μL) were taken every six additions (every hour). In case of the 30 min addition intervals, 2.4 eq of H_2O_2 were added (total duration 6 h), and samples (200 μL) were taken every four additions (every two hours). All samples were extracted with equal amounts of EtOAc (200 μL) and analysed using achiral GC (method 3).

5.1.6.3. In situ Hydrogen Peroxide Generation using PpAOx

As part of the reaction optimisation the *in situ* generation of hydrogen peroxide was considered. The use of an alcohol oxidase from *Pichia pastoris* (PpAOx) has recently been reported in the literature to work with AaeUPO.¹¹¹ A range of experiments was performed here to determine the best UPO:AOx ratio, methanol (MeOH) loading, and reaction co-solvent.

All reactions were set-up in duplicates in 2 mL plastic tubes, sampling and analysis was performed as described Section 5.1.4.

Reactions with varying amounts of PpAOx (0.5, 1, 2, and 4 U mL⁻¹) were prepared containing 50 mM KPi buffer pH 7.0 (from 100 mM stock), 200 mM MeOH, 10 mM ethylbenzene, and 2 U mL⁻¹ rAaeUPO-H. All reactions were made up to 750 μL using ultrapure water. The reactions were incubated shaking at 900 rpm at 30 °C. Samples

(100 μL) were taken after 0, 0.5, 1, 1.5, 2 and 3 h reaction time, extracted with EtOAc (200 μL) and analysed using achiral GC (method G1).

The amounts of *rAaeUPO-H* were varied in another set-up, to 0.5, 1, 2, and 4 U mL^{-1} . The reactions were prepared, run, and sampled as before, using 200 mM MeOH, and 2 U mL^{-1} *PpAOx*.

Experiments with varying amount of methanol (10, 20, 50 mM, and 10%) were prepared with 1 U mL^{-1} *rAaeUPO-H* and 2 U mL^{-1} *PpAOx*, run, and analysed as described above. Samples were taken after 15, 30, 60 and 120 min.

Four solvents (MeCN, acetone, DMSO, and MTBE) were added to the reaction mixture to make up 5 or 10% of the final solution. The reactions contained 1 U mL^{-1} *rAaeUPO-H*, 2 U mL^{-1} *PpAOx*, and 200 mM MeOH. The reactions were incubated shaking at 900 rpm at 30 °C and samples (100 μL) were taken after 1, 2, and 72 h, extracted using EtOAc (200 μL) and analysed using achiral GC (method G1).

Further reactions were set-up with different *rAaeUPO-H* to *PpAOx* ratios and different amounts of protein, as described in Table 5-4.

Table 5-4: *rAaeUPO-H* and *PpAOx* amounts used in experiments run to determine the best protein ratio for the *in situ* generation of hydrogen peroxide.

No.	Amount / U mL^{-1}	
	<i>rAaeUPO-H</i>	<i>PpAOx</i>
1	1	2
2	0.5	1
3	0.37	1.14
4	1	0.05
5	1	0.1
6	1	0.2

All reactions were prepared and run in 2 mL tubes in the presence of 10 mM ethylbenzene and 200 mM MeOH. Reactions were incubated at 900 rpm, 30 °C and samples were taken after 0, 15, 30, 60, 90, and 120 min (no. 1 to 3), or after 15, 30, 60, 120 min (no. 4 to 6) and extracted with EtOAc (200 μL).

5.1.6.4. Substrate Loading

Using the *in situ* peroxide generation system, reactions with different substrate loadings were set-up. All reactions were run in duplicates in 2 mL plastic tubes using

50 mM KP_i buffer pH 7.0, and 200 mM MeOH. Tubes were incubated at 900 rpm, 30 °C.

Ethylbenzene was studied at 20, 30, 50, and 100 mM (from 10X stocks in MeCN) using 2 U mL⁻¹ *PpAOx* and 1 U mL⁻¹ *rAaeUPO-H*. Samples were taken after 1.25, and 3 h (50 µL), extracted with EtOAc (450 µL) and analysed using GC (method G1). After 6 h another portion *rAaeUPO-H* was added (1 U mL⁻¹) and a final sample was taken and analysed after 24 h incubation.

Ethylpyridine was applied to a reaction using 20 mM substrate, 0.02 U mL⁻¹ *PpAOx*, and 1 U mL⁻¹ *rAaeUPO-H*. The reactions were incubated with shaking at 900 rpm, 30 °C and samples were extracted and analysed after 8 h (250 µL sample using 500 µL EtOAc) and 24 h (400 µL sample using 200 µL EtOAc). In addition to the achiral analysis (method G1), chiral analysis was run using a HYDRODEX β-TBDAC column (method G2, see Table 5-3).

Table 5-5: GC method and condition used for the chiral analysis (HYDRODEX β-TBDAC) of 1-(pyridin-2-yl)ethan-1-ol **5** run at GSK.

#	Split	Description	Used for
G2	40:1	60 °C to 210 °C at 4 °C min ⁻¹	5

5.1.6.5. Reaction Scale-up

Scale-up reactions were run in duplicates using 10 mL reaction volume in either 50 mL falcon tubes or 50 mL conical flasks, with a no-enzyme control in parallel. The reactions contained 50 mM KP_i buffer at pH 7.0 (from 100 mM stock), 10 mM substrate in 10% MeCN (from 10X/100 mM stock), *rAaeUPO-H* (2 U mL⁻¹) and were started through the addition of hydrogen peroxide (2 mM from 100 M stock). The falcon tube reactions were incubated shaking at 700 rpm, 25 °C in a thermomixer, while conical flask reactions were incubated in an orbital shaker at 180 rpm, 25 °C. Hydrogen peroxide (2 mM) was added every 30 min and samples were taken after 1, 3, 5, and 96 h (100 µL extracted with 200 µL EtOAc). All samples were analysed using GC (method G1).

5.1.7. Design of Experiments (DoE)

The aim of the design of experiments (DoE) is provide a systematic approach to optimise a process by looking at factor-response relationships. The variables are

defined as factors and the output is referred to as responses. The DoE was set up using Design-Expert software, with the substrate loading, *rAae*UPO-H loading, UPO:AOx ratio, and the temperature as variable factors, and the GC-based conversion to product and by-products as responses. The contribution of the factors to specific response can be calculated using least square estimation, where the different factors and combinations of factors are ranked. Half-normal plots describe the visual output.

As part of the DoE ten 10 mL reactions were prepared and run using the Integrity10 reaction station (Electrothermal, Stone, UK). The exact set-up of the different experiments is given below (see Table 5-6). All reactions were run using 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** as substrate in 50 mM KPi buffer at pH 7.0 (from 100 mM stock) in the presence of 200 mM MeOH and 10% MeCN. The reactions were prepared in glass tubes and fitted with a stir bar. The reaction vessels were placed in the reactor block, brought to the desired temperature, and the MeOH was added to initiate the reaction. The stirring was set to 350 rpm. Samples were taken after 1, 3, 5 and 20 h reaction time. The instrument sampled by mixing 100 μ L of reaction mix with 400 μ L MeCN. For GC analysis (method G1) all samples were diluted to the same molarity, i.e. samples with 50 mM substrate were diluted 1:10 (30 μ L into 300 μ L), 27.5 mM samples were diluted 55 μ L into 300 μ L, and 5 mM samples were run as prepared.

After 20 h the reaction mixture was extracted with ethyl acetate (2 x 10 mL), the organic phases were combined and dried over magnesium sulfate before the solvent was removed *in vacuo*.

Table 5-6: Set-up of the ten DoE runs, with the specifics of the factors 1 to 4, and the space type, with f for factorial and c for centre.

Run	Std.	Space type	Factor 1	Factor 2	Factor 3	Factor 4	
			Substrate / mM	T / °C	rAaeUPO-H / U mL ⁻¹	UPO:AO x	AOx / U mL ⁻¹
1	3	f	5	40	1	20	0.05
2	4	f	50	40	1	5	0.2
3	6	f	50	20	10	5	2
4	10	c	27.5	30	5.5	12.5	0.44
5	1	f	5	20	1	5	0.2
6	5	f	5	20	10	20	0.5
7	2	f	50	20	1	20	0.05
8	8	f	50	40	10	20	0.5
9	7	f	5	40	10	5	2
10	9	c	27.5	30	5.5	12.5	0.44

5.1.7.1. Scale-up Reaction

Following the DoE a 1 g reaction with 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** was set-up in an EasyMax 402 reactor (Mettler-Toledo, Leicester, UK). A 168 mL reaction was set up with 50 mM KP_i buffer pH 7.0 (from 100 mM stock), 10 U mL⁻¹ rAaeUPO-H, 2 U mL⁻¹ PpAOx, 200 mM MeOH and 50 mM **8** (1 g). The reaction was stirred at 350 rpm and kept at 20 °C. Once at temperature, the reaction was started by the addition of MeOH and stopped after 115 h. A sample was taken and analysed using GC (method G1, 100 µL into 300 µL EtOAc). To the reaction mixture 160 mL EtOAc were added and the stirring was continued for 15 min. The phases were separated and the aqueous layer was extracted a second time with EtOAc (160 mL). The combined organic phases were dried over magnesium sulfate and the solvent was removed *in vacuo*.

The crude material (729 mg) was analysed using GC (methods G1 and G3), and ¹H-NMR. The chiral GC analysis of the product, 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol **9**, was performed using a cyclosil B column (method G3, see Table 5-7).

Table 5-7: GC method and condition used for the chiral analysis (cyclosil B) of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-5-ol **9** run at GSK.

#	Split	Description	Used for
G3	48:1	40 °C to 220 °C at 7 °C min ⁻¹ , 2 min at 220 °C	9

The crude mixture was further purified using the Isolera advanced automated flash purification system (Biotage, Uppsala, Sweden) equipped with a 25 g SNAPC18 Ultra

column. Peak fractions for the pure and mixed alcohol products were combined and the solvent was removed *in vacuo*.

The fractions 16 to 28 (pure alcohol) yielded 288 mg and the fractions 12 to 15 (mixed alcohols) yielded in 129 mg. The samples were analysed using GC (methods G1 and G3) and $^1\text{H-NMR}$.

5.2. Characterisation of Lyophilised rAaeUPO-H

The fermentation of rAaeUPO-H at GSK Worthing using four 10 L fermentation vessels was run by Alison Dann and her team. Stocks of *Pichia* transformed with rAaeUPO-H were provided and the established fermentation protocol was followed. Due to on-site safety regulations, the use of oxygen to adjust the DO was not possible. Higher flow-rates of air, with up to 6 L min⁻¹, were able to compensate for this and allowed for the fermentation to proceed without any problems.

The large scale fermentations were analysed using specific activity assays, prior to the processing of the fermentation supernatants. All experiments were performed on-site in Worthing by Cyril Boudet, using the assay described in this work (see Section 3.1.7.6). The four reactors all showed activity towards NBD, ranging from 2.6 to 4.4 U mL⁻¹. As a comparison, the best fermentation run in-house gave a specific activity towards NBD of 3.5 U mL⁻¹ (see Section 3.2.4). A total of 19 L combined cell-free fermentation supernatant was concentrated using tangential flow to approx. 2 L, creating a 10X stock solution with a specific activity towards NBD of 27.1 U mL⁻¹.

The concentrated fermentation supernatant was shipped to GSK Stevenage. Lyophilisation experiments were performed to allow for easier handling of the protein. A small 4.5 mL sample was analysed in initial tests to allow full characterisation and comparison of the lyophilised rAaeUPO-H to the protein in solution. After freeze-drying 0.311 g of protein powder were obtained, yielding in 69 mg mL⁻¹. Further 1.5 L of the concentrated protein solution were lyophilised at a later point, yielding a total of 101 g of protein powder, at 67 g L⁻¹. SDS-PAGE analysis of the concentrated rAaeUPO-H sample and the lyophilised rAaeUPO-H is shown below (see Figure 5-1). The lyophilised protein on the gel was resuspended to a concentration of 10 mg of protein powder per mL using water. Both proteins display a molecular weight of approx. 60 kDa, which is comparable to the results obtained for rAaeUPO-H fermented in-house. rAaeUPO-H accounted for the largest band on the gel in both cases with some impurities at lower molecular weights.

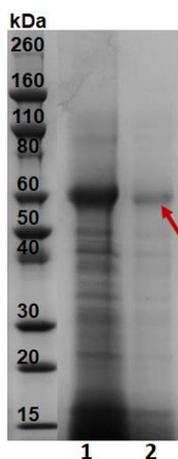


Figure 5-1: 4-12% precast SDS gel with Novex Sharp ladder showing the concentrated fermentation supernatant (10X) of *rAaeUPO-H* (lane 1) and the lyophilised *rAaeUPO-H* sample (lane 2), running at around 60 kDa (arrow).

In addition to the 10 mg mL^{-1} *rAaeUPO-H* sample run on the gel, a sample with a defined volume was lyophilised and readjusted to the original volume. This should allow for a direct comparison of lyophilised protein and protein in solution, without having to account for dilutions in different experiments. The protein concentration of the two samples was determined using Bradford assays, relying on the absorption at 595 nm, after incubating a sample with Coomassie stain. The concentration of the lyophilised samples was calculated to be $5.3 \pm 0.5 \text{ mg mL}^{-1}$, and for the protein in solution a value of $4.9 \pm 0.3 \text{ mg mL}^{-1}$ was determined. The slight deviation could be due to errors introduced during the dilution of the samples for analysis. Further, specific activity assays were run to compare the activity before and after lyophilisation, and whether the freeze-drying procedure causes inactivation of the protein. Towards ABTS specific activities of 180 ± 15 and $155 \pm 5 \text{ U mL}^{-1}$ were determined for the lyophilised and the protein in solution, for NBD these values were 55 ± 1 and $50 \pm 10 \text{ U mL}^{-1}$, respectively. Similar values were obtained in the different experiments run, with slightly increased values recorded for the lyophilised *rAaeUPO-H* sample. In a final experiment the conversion of ethylbenzene **1** (10 mM) was monitored using GC, the chromatograms recorded after 2.5 h reaction time are shown below (see Figure 5-2). The conversions were calculated from the peak areas, giving 74 ± 1 and $75 \pm 1\%$ conversion to 1-phenylethanol **2** from ethylbenzene for the lyophilised *rAaeUPO-H* and the protein in solution, respectively, with $5 \pm 1\%$ overoxidation to acetophenone **3** for both samples. The process of lyophilising

rAaeUPO-H did not seem to affect the activity of the protein and allowed for easier handling and storage of the fermented protein.

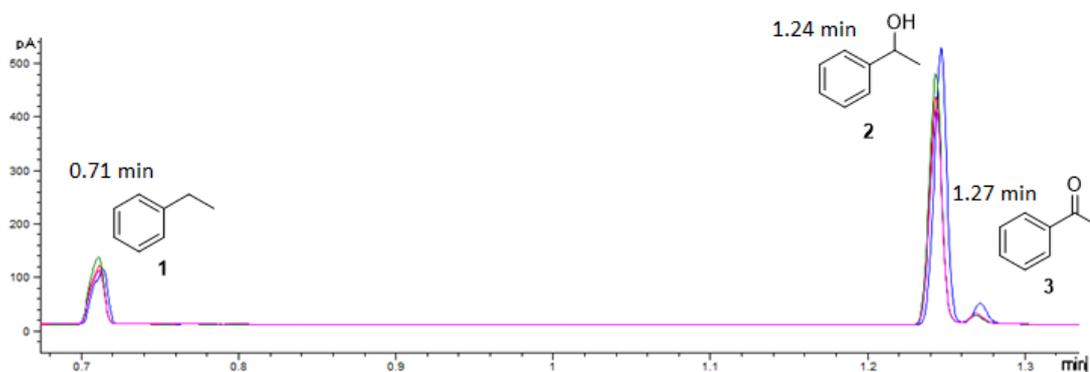


Figure 5-2: GC-chromatogram recorded after 2.5 h reaction of *rAaeUPO*-H with ethylbenzene **1** (0.71 min) yielding in 1-phenylethanol **2** (1.24 min) and acetophenone **3** (1.27 min), with the traces for the lyophilised *rAaeUPO*-H in green and pink and for the protein in solution in blue and red.

5.3. Substrate Screens

The substrate screens were performed as part of the research conducted at GSK, Stevenage and were focussing on a range of commercially available substrates as well as substrates of industrial interest, where the overall structures cannot be shown. Reactions included benzylic, aliphatic, and terminal hydroxylations, alcohol oxidations, and *N*-dealkylation. The different substrates and conversions for reactions with 10 mM loading determined using GC, for both *rAaeUPO*-H and *H-AbiUPO*, are shown below (see Figure 5-3 with conversions in black and dark red for *rAaeUPO*-H and *H-AbiUPO*, respectively).

The conversion of ethylbenzene **1** to 1-phenylethanol **2** was monitored in plastic tubes at pH 7.0 and proceeded with up to 92% conversion in the presence of *rAaeUPO*-H. Overoxidation to acetophenone **3**, as a by-product, was observed at 5%. In the literature *AaeUPO* had been described to catalyse the same reaction at 95% conversion using 1 mM ethylbenzene.⁹⁵ For *H-AbiUPO* no conversion of **1** was observed. The conversion of ethylpyridine **4** to 1-(pyridin-2-yl)ethan-1-ol **5** was monitored at pH 6.4 in plastic tubes and gave 4% of product in the presence of *rAaeUPO*-H determined using GC. The chromatogram did not show the formation of by-products, like the *N*-oxide, which was one major product in the literature described conversion of 3-methylpyridine.¹⁰¹ With the homolog from *Agaricus bisporus* no substrate conversion using **4** was determined. A larger substrate, such as 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8**, was also included in this

work. 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol **9** (16%), its oxidation product **10** (1%), and the alcohol in position C7 **11** (5%), were identified as reaction products when using *rAaeUPO*-H as catalyst at pH 6.4. The application of H-*AbiUPO* at pH 7.0 only showed a conversion of 1% giving **9**, with no peaks in the chromatogram identifying the oxidation product **10**, the alcohol **11**, or the alcohol in the C6-position **12**. This substrate is not believed to have been described as an *AaeUPO* substrate in the literature before.

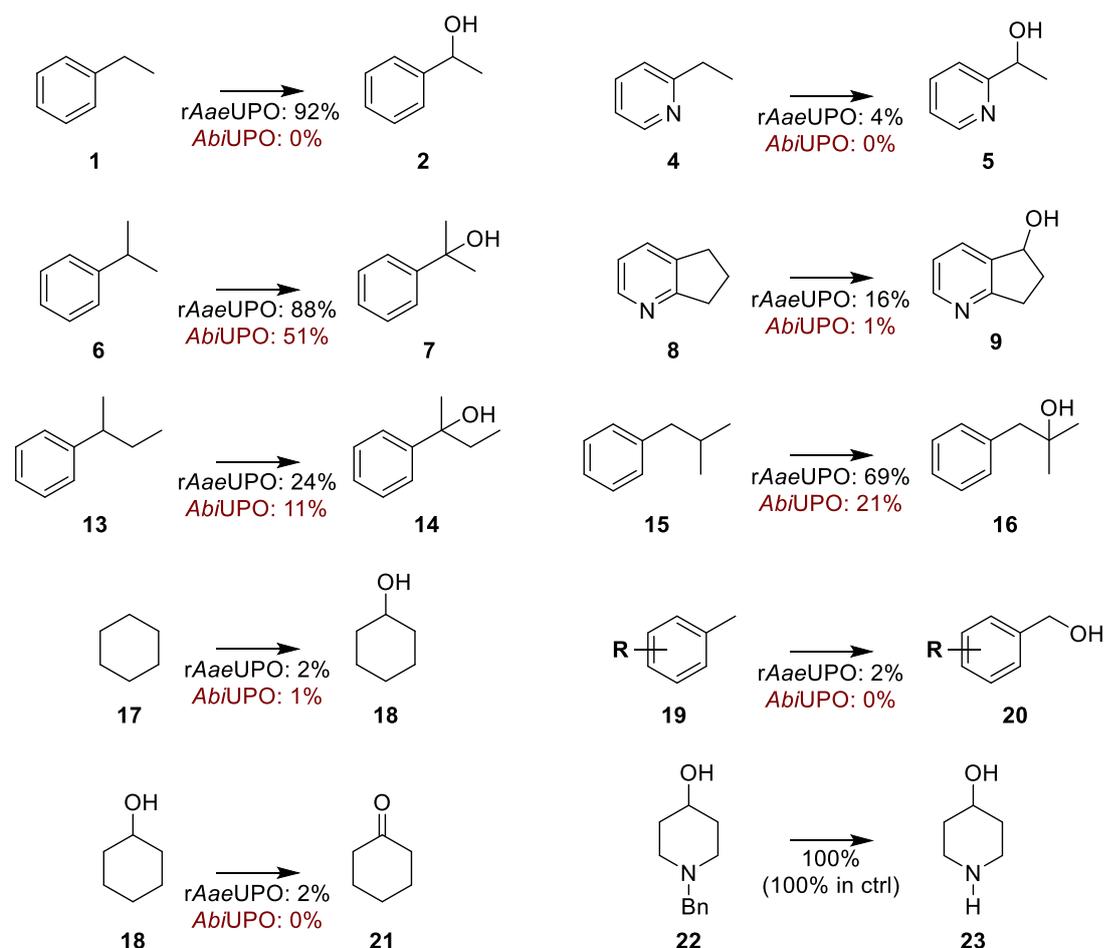


Figure 5-3: Overview over the different screening reactions run with 10 mM substrate in the presence of *rAaeUPO*-H (conversions in black) or H-*AbiUPO* (conversions in dark red); the substituents of the toluene derivatives cannot be further identified. Reactions were incubated overnight (20-24 h), except for the reaction of *rAaeUPO* with **4** (1 h) and **6** (5 h).

Further substrates included cumene **6**, *sec*-butylbenzene **13**, and isobutylbenzene **15** to study the selectivity of UPOs for benzyl and tertiary carbons. Markers for the different possible hydroxylation products were obtained and identified 2-phenylpropan-2-ol **7**, 2-phenylbutan-2-ol **14**, and 2-methyl-1-phenylpropan-2-ol **16**, respectively, as sole reaction products. These results showed a clear preference of

UPOs for the tertiary carbon over the benzylic carbon. For rAaeUPO-H the conversions were calculated to be 88% (pH 6.4, 25 °C), 24% (pH 7.0, 30 °C, PpAOx), and 69% (pH 6.4, 25 °C), respectively, and for H-AbiUPO values of 51% (pH 7.0, 25 °C), 11% (pH 7.0, 30 °C), and 21% (pH 7.0, 25 °C) were obtained. In the literature only *n*-alkylbenzenes have been reported as AaeUPO substrates, and are selectively hydroxylated the benzylic position with decreasing activity when increasing the alkyl-chain length.⁹⁵

The hydroxylation of cyclohexane **17** to cyclohexanol **18** and the oxidation of **18** to cyclohexanone **21** was also studied as part of this work. The reactions only proceeded with low conversions of approx. 2% in the presence of rAaeUPO-H (pH 6.4/7.4), which were decreased when using H-AbiUPO (pH 7.0) to 1% with **17** as substrate, and no conversion was observed when using **18** as substrate. In both cases, using substrates **17** and **18**, background reactions with conversions around 0.8% were observed in the control samples. In the literature, these reactions with AaeUPO have been studied in more detail, suggesting 76% conversion to cyclohexanol and 1% overoxidation to **21**, using **17** as substrate (40 mM, with 4 mM hydrogen peroxide added).⁹⁷ Similar low conversions were observed for a wide range of toluene-derivatives **19**, looking for the terminal hydroxylation products **20**. For H-AbiUPO no conversions were observed, while rAaeUPO-H gives up to 2% conversion in selected cases. The formation of a range of by-products, potentially corresponding to aromatic hydroxylations, were also observed in some cases. Further analysis of these substrates was not pursued. In the literature, experiments with the wildtype AaeUPO reported up to 37% conversion to the benzyl alcohol from toluene; overoxidations products, aromatic hydroxylation products and the methyl-*p*-benzoquinone were observed as reaction by-products.⁷⁶

Looking at the abilities of UPOs to dealkylate heteroatoms, as it has been described in the literature, 1-benzylpiperidin-4-ol **22** was selected as a substrate. Running the reaction at pH 6.4 in the presence of hydrogen peroxide, a full conversion of the substrate giving piperidin-4-ol **23** was observed. In this case however, 100% conversion was also observed in the control samples, suggesting that the *N*-deprotection is catalysed by the peroxide rather than the added rAaeUPO-H. The GC chromatograms obtained for the markers and the different reaction samples are

shown below (see Figure 5-4), supporting the statements just made. The stability of benzylamines in the presence of peracids is known to be poor and handling an aliphatic amine, it is possible that redox potential of hydrogen peroxide is great enough to facilitate the dealkylation shown here.

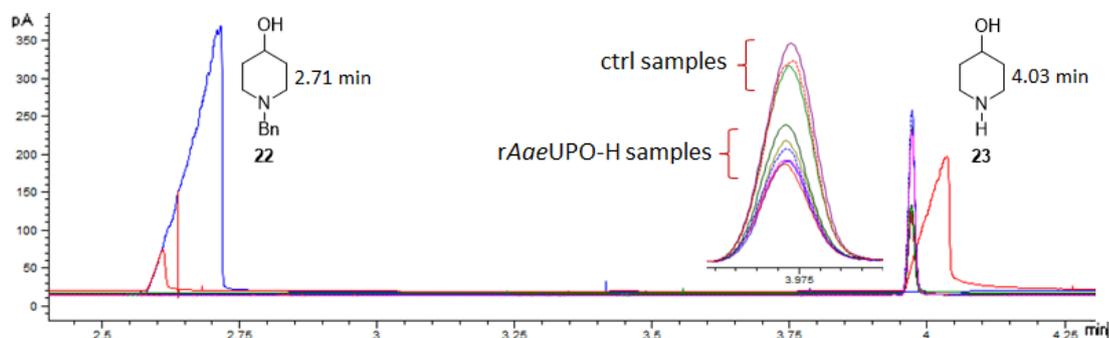


Figure 5-4: GC chromatograms obtained for reactions of 1-benzylpiperidin-4-ol **22** (marker in blue, 2.71 min) in the presence and absence (ctrl) of *rAaeUPO-H* yielding in piperidin-4-ol **23** (marker in red, 4.03 min).

In addition to screening for UPO substrates, different substrate loadings (2, 5, and 10 mM) and pHs (4.4, 5.4, 6.4, and 7.4) were also studied using substrate **1**. These experiments were performed in 96-well deep-well blocks. It was found that conversions yielding in **2** were higher for increased substrate loading, with 48, 52, and 72% using 2, 5, and 10 mM of **1**. The formation of **3** was not observed. By looking at the 10 mM samples at different pH, it was clear that best results were obtained at pH 6.4 (72%), with a 10-15% decrease in conversion at pH 5.4 and 7.4 (61 and 63%, respectively), and more than a 50% drop in conversion at pH 4.4 (34%). As a result of these findings further experiments were mainly performed using 10 mM substrate and a pH between 6.4 and 7.4. The drop in conversion between plastic tubes (92%) and 96-well deep-well blocks (72%) could originate from the different incubation conditions, which could affect the mixing and aeration. Further, as the sampling was performed using a robot, the extraction efficiency could have been influenced, especially in the case of volatile compounds.

5.4. Condition and Reaction Optimisation

Following the substrate and initial condition screens, further optimisation approaches were taken. These experiments were conducted to arrive at a more applicable process, with regards to industrial transformations using *rAaeUPO-H*. Hence, different reaction temperatures were investigated, the tolerance towards co-

solvents was determined and the scalability of the reaction was trialled. Furthermore, experiments looking at hydrogen peroxide addition were conducted followed by studies looking into the *in situ* generation of hydrogen peroxide using an alcohol oxidase from *Pichia pastoris* (PpAOx).

The temperature dependent conversions of cumene **6**, cyclohexanol **18**, and ethylpyridine **4** were tested in the presence of rAaeUPO-H. Reactions were run in duplicates at 20, 25, and 30 °C, with a no enzyme control run at the same conditions. The analysis was based on achiral GC experiments. Cumene showed higher conversions at 25 and 30 °C with 78%, compared to the 20 °C with 70% after 5 h. In addition to the formation of 2-phenylpropan-2-ol **7**, an unidentified by-product was observed chromatogram with conversions of up to 15%. As a marker for 2-phenylpropan-1-ol was available, this compound can be excluded from the list of possible conversion products. Due to the ability of UPOs to catalyse aromatic hydroxylations, it is possible, that the aromatic ring was hydroxylated during the transformation. The reaction samples were left incubated for a total of 72 h, after which all temperatures displayed conversion rates of 82% towards **7**, with 15% conversion to the by-product.

The oxidation of **18** to yield cyclohexanone **21** was also performed at 20, 25, and 30 °C. The determined conversions were only slightly increased from the background reaction visible in the control samples at 0.8%. The strongest result was obtained at 30 °C with 1.4% conversion, whilst the 20 and 25 °C sample showed conversions around 1.1%. A preference for higher temperatures for the conversion of **7** could be suggested, however, the overall low yields and poor selectivity of rAaeUPO-H for the substrate suggested that this was not the best compound to use for further optimisation reactions.

Finally, reactions of rAaeUPO-H with ethylpyridine **4** were subjected to different temperatures. In contrast to the results just discussed, the highest conversions giving 1-(pyridin-2-yl)ethan-1-ol **5** were observed in the 20 °C sample with 6%. The 25 °C sample showed a conversion of 4%, which further decreased for the 30 °C sample to 2%. Using substrate **4** an increase in temperature clearly affects the conversions, where an increase from 20 to 30 °C results in decreasing the conversion by two thirds.

As the results looking at temperature dependence of rAaeUPO-H showed a substrate dependence, further experiments were kept at 25 °C.

As the conversion of **6** yielded good amounts of product, this substrate was chosen for a scale-up reaction. Reactions were set-up in both 50 mL plastic tubes and 50 mL conical flasks using 10 mM substrate. The analysis of the samples over time relied on achiral GC analysis. After the addition of 0.6 eq of peroxide conversions of 58 and 42% were observed in the tube and the flask, respectively. These conversions increased to 84 and 81% after the addition of 1.2 eq, and final conversions of 88 and 86% were observed after 2 eq of hydrogen peroxide were added. The selectivity of 2-phenylpropan-2-ol **7** over the by-product was at 86% in the last measurement. These experiments showed that the conversion of cumene **6** was easily scaled from 750 μ L to a 10 mL reaction volume. The observed conversions for **7** were even improved from 78/82% observed after 5/72 h incubation to 88% observed after a 5 day incubation period.

Next experiments with **1** were performed in the presence of different organic solvents. All reactions were set-up in plastic tubes at pH 7.0 with 10 mM substrate and 10% solvent. The following order was determined from the conversions after 26 h incubation: acetonitrile (91%) = acetone (90%) > DMSO (85%) > ethanol (69%) > ethyl acetate (55%) > MTBE (27%) (see Figure 5-5 for GC chromatograms). rAaeUPO-H showed a good tolerance for organic solvents, at 10% final solvent concentration. The non-miscible solvents, EtOAc and MTBE, seemed to affect the conversion the most, lowering them by 40 and 70% with respect to the conversions observed in MeCN. This drop could be due to interactions of the protein with the solvent, potentially leading to a deactivation of the protein due to changes in folding and stability. Furthermore, the analysis of the samples showed the formation of acetophenone **3**, the overoxidation product of the reaction. The contribution was 5% for the acetonitrile and acetone samples, and about 3% for the DMSO sample. No overoxidation was observed in the other solvents. These findings suggest, that overoxidation only occurs when a critical amount of **2** is formed, which can then be further converted. As a comparison, studies with the *Saccharomyces cerevisiae* expressed rAaeUPO reported lower solvent tolerance levels for acetonitrile, ethanol,

methanol, acetone, and DMSO, where at 10% (v/v) acetone gave the highest observed conversion of 50% when using 0.3 mM ABTS as substrate.¹²⁰

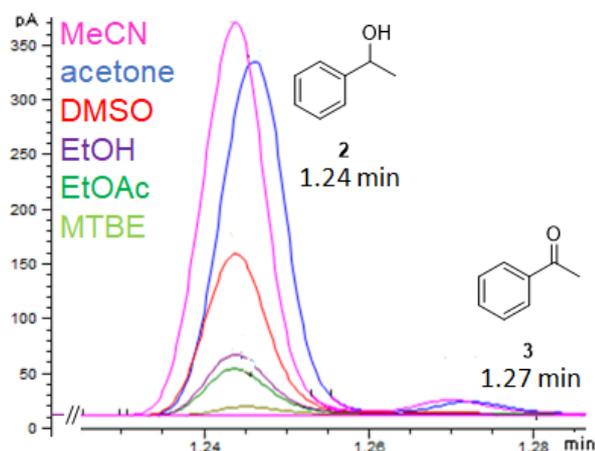


Figure 5-5: GC-chromatograms for the conversion of ethylbenzene **1** to 1-phenylethanol **2** and acetophenone **3** using *rAaeUPO-H* in the presence of different solvents, with acetonitrile in pink, acetone in blue, DMSO in red, ethanol in purple, ethyl acetate in green, and MTBE in light green.

5.4.1. Hydrogen Peroxide Addition

For reactions catalysed by *rAaeUPO-H* to go to completion, equal amounts of hydrogen peroxide need to be supplied to the system. As an excess of hydrogen peroxide can lead to the deactivation of the enzyme, the correct dosing of this co-substrate is very important. Using **1** as a model substrate the addition of 2 mM hydrogen peroxide in 15 and 30 min intervals was studied. Samples were taken after equal amounts of peroxide were added to each reaction to allow for the best comparison. A graphic representation of the product formation (1-phenylethanol **2**) and the observed overoxidation to acetophenone **3** is shown below (see Figure 5-6). After the addition of 0.4 eq of hydrogen peroxide, after 30 and 60 min for the 15 and 30 min addition, respectively, conversions of 83% and 81% were determined from the GC chromatograms, respectively. The level of overoxidation to **3** was at 6% in both cases. The last sample was recorded after the addition of 2 eq hydrogen peroxide (150 and 300 min) and the conversions only slightly increased to 87% for both addition rates with 10% overoxidation to **3**. Using this model system, it was apparent that there was no difference observed for the 15 and 30 min addition rates, allowing to proceed with additions every 15 min, and therefore reducing the overall reaction time by 50%. The data also showed that the production of **2** and the overoxidation to **3** only increased slightly after the first measurement, and settled around 85 and 10%, respectively, with no distinct changes observed over time. This

could be either a result of having driven the reaction to completion or an equilibrium state, where no further conversion was observed, or the stagnation could be due to an inactivation of the enzyme by hydrogen peroxide.

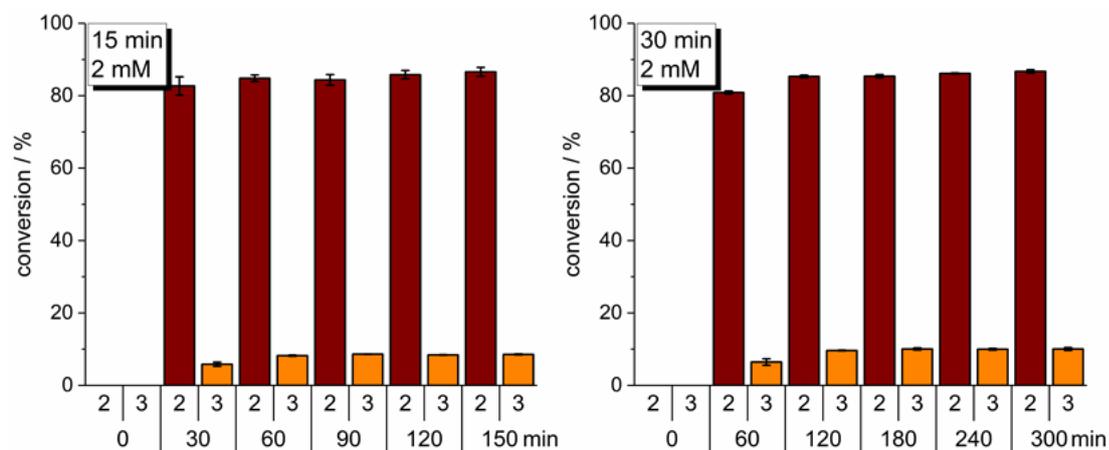


Figure 5-6: Conversions of ethylbenzene **1** to 1-phenylethanol **2** (dark red) and acetophenone **3** (orange) in the presence of *rAaeUPO-H* over time at different hydrogen peroxide addition rates: **left** – every 15 min; **right** – every 30 min. All reactions were run in duplicates and the errors are reported as standard deviation.

In addition to the conversion of **1**, the sulfoxidation reaction of methyl *p*-tolyl sulfide **36** with hydrogen peroxide was also studied in more detail. As the thioanisole derivatives can be oxidised by hydrogen peroxide in the absence of the catalyst, different addition rates, as well as varying amounts of hydrogen peroxide were under investigation. Experiments were performed using 1 or 2 mM hydrogen peroxide added every 10 or 30 min for 3 or 6 h, respectively. The results looking at the formation of methyl *p*-tolyl sulfoxide **37** in the *rAaeUPO-H* and the control sample are shown below (Figure 5-7). For the addition of hydrogen peroxide every 10 min, independent of the amount of peroxide added, a final conversion of 49% was determined in the reactions catalysed by *rAaeUPO-H*. For the 1 mM hydrogen peroxide portions the reaction observed in the control sample after 3 h (3.6 eq of H₂O₂) was at 5% the lowest conversion determined in a control reaction. In the control sample for the reaction, where hydrogen peroxide was added at 2 mM every 10 min, the conversion was at 14% almost three times as high after 3 h, with a total addition of 7.2 eq of H₂O₂. When hydrogen peroxide was added every 30 min at 1 mM a similar 50% conversion was determined after 6 h (2.4 eq of H₂O₂), and when added at 2 mM a conversion of 71% was calculated (4.8 eq of H₂O₂). The highest conversion in the *rAaeUPO-H* catalysed reaction coincided with the highest

conversion in the control sample, with 31% after 6 h. The chiral analysis of these particular samples showed a racemic mixture of **37** in the control reaction, while the enzyme catalysed reactions showed 68% *ee* favouring the (*R*)-enantiomer. Given the high amount of background reaction, the enantiomeric excess was expected to be lower. In comparison to that, the chiral analysis of the samples for the 1 mM peroxide addition every 30 min, also showed a racemic mixture in the control sample, which had a conversion of 13%, while the enzyme catalysed reaction showed 67% *ee* in favour of the (*R*)-enantiomer. As the control reactions were not run in duplicates, it could be possible, that an error in sampling or hydrogen peroxide addition in the 2 mM peroxide every 30 min reaction led to the high conversion determined. From the increase in conversion between the 1 (50%) and 2 mM (71%) hydrogen peroxide added every 30 min, the stable *ee* suggested that the latter amount of hydrogen peroxide was more favourable. This reasoning was also supported by the conversions determined for the 2 mM every 30 min samples after 4 h (3.2 eq of H₂O₂), where the *rAae*UPO-H catalysed reactions showed a mean conversion of 62%, while the conversion in the control was calculated at 10%.

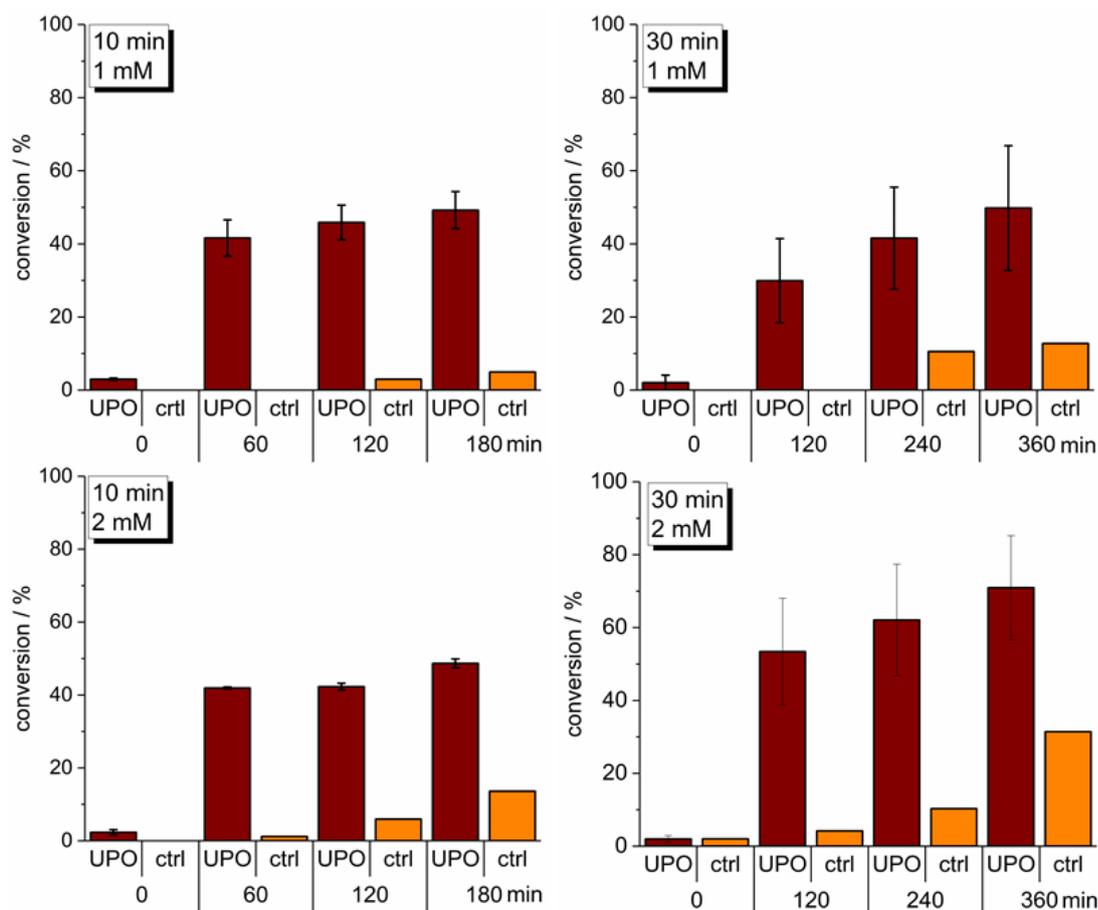


Figure 5-7: Conversions for the formation of methyl *p*-tolyl sulfoxide **37** from the corresponding sulfide **36** in the presence (dark red) and absence (orange) of *rAaeUPO*-H over time at different amounts and addition rates of hydrogen peroxide: **top left** – 1 mM hydrogen peroxide added every 10 min; **top right** – 1 mM hydrogen peroxide added every 30 min; **bottom left** – 2 mM hydrogen peroxide added every 10 min; **bottom right** – 2 mM hydrogen peroxide added every 30 min. All reactions were run in duplicates and the errors are reported as standard deviation.

5.4.2. *In situ* Hydrogen Peroxide Generation

Apart from manually controlling the addition of hydrogen peroxide there is the option of *in situ* generation of the co-substrate. A major advantage of this is that higher substrate loadings can be attempted. In the literature different systems have been described either relying on enzymes like glucose oxidase or alcohol oxidase, or by harnessing photons to activate a cofactor driven production of hydrogen peroxide.^{69, 94, 111} Each system has its advantages and disadvantages. In 2016, a comprehensive study focussed on using the commercially available alcohol oxidase from *Pichia pastoris* (*PpAOx*) in conjunction with other enzymes for the full oxidation of methanol in the presence of *AaeUPO*.¹¹¹ It was reported that reactions in the presence of both *AaeUPO* and *PpAOx* gave high conversion of ethylbenzene in short reaction times.

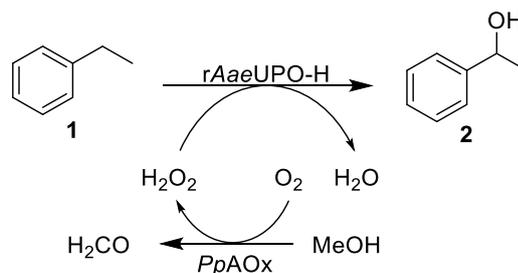


Figure 5-8: Conversion of ethylbenzene **1** to 1-phenylethanol **2** by rAaeUPO-H, using an *in situ* hydrogen peroxide generation system based on the conversion of methanol by PpAOx.

Using this research as a basis, experiments were performed to obtain an *in situ* hydrogen peroxide generation system (see Figure 5-8). For this, a range of scoping reactions were performed to best match the rates of hydrogen peroxide generation from methanol by PpAOx to the peroxide utilisation rates of rAaeUPO-H. The conversion of ethylbenzene **1** to 1-phenylethanol **2** was used as a model reaction, whilst recording the rates of unwanted overoxidation to acetophenone **3**. The results of these initial studies are shown below (Figure 5-9).

In the first instance, the rAaeUPO-H concentration was kept constant at 2 U mL⁻¹, while the PpAOx concentrations were varied (0.5-4 U mL⁻¹, Figure 5-9 top left). It was found that for lower PpAOx concentrations more overoxidation was recorded over time, increasing from approx. 25 to 35% for 0.5 and 1 U mL⁻¹ PpAOx, decreasing the yield of **2**. At higher PpAOx concentrations the overoxidation to **3** showed constant values between 15-20% over time. The highest conversion to **2** was found in the sample using 4 U mL⁻¹ PpAOx at 84%. These findings suggested that the conversion of **1** was complete within the first 30 min of reaction time, as no further increase in conversion yielding **2** was observed. While at high PpAOx concentrations the accumulation of peroxide inactivated the heme, preventing the oxidation of **2** to take place, at lower PpAOx concentration the steady supply of hydrogen peroxide allowed rAaeUPO-H to further react with the obtained 1-phenylethanol **2** giving higher amounts of **3**. Here, the best results in terms of substrate hydroxylation and low alcohol oxidation were obtained using 4 U mL⁻¹ PpAOx (1:2 UPO to AOx ratio).

The next set of experiments was performed at a constant concentration of PpAOx (2 U mL⁻¹) with varying rAaeUPO-H concentrations (Figure 5-9 top right). Here, the opposite behaviour was observed with higher overoxidation to **3** at higher protein concentration. While with 0.5 U mL⁻¹ rAaeUPO-H the conversion rates for **2** increased

over time from 64% (30 min) to 83% (120 min) with only 1% overoxidation after 120 min, at 4 U mL⁻¹ after 120 min only 58% **2** were detected with 40% overoxidation. This set of experiments suggested that when using 2 U mL⁻¹ *PpAOx*, lower amounts of *rAaeUPO-H* were still active, but at reduced rates, while higher amounts of *rAaeUPO-H* were able to stay fully active leading to an oxidation of the alcohol of **2**. The best results here were obtained using 1 U mL⁻¹ *rAaeUPO* (1:2 UPO to AOx ratio).

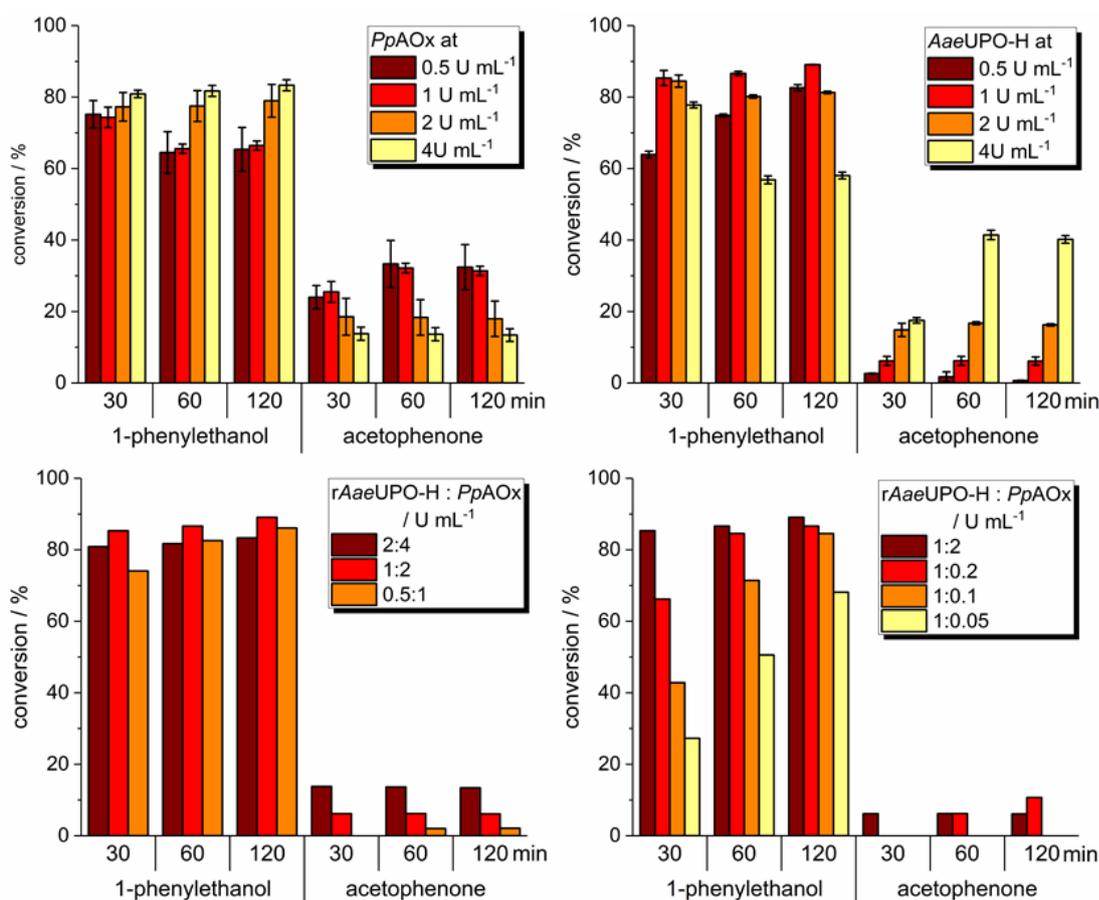


Figure 5-9: Conversions for the formation of 1-phenylethanol **2** and acetophenone **3** plotted over time at: **top left:** varying *PpAOx* concentration at 2 U mL⁻¹ *rAaeUPO-H* and 200 mM MeOH; **top right:** varying *rAaeUPO-H* concentration at 2 U mL⁻¹ *PpAOx* and 200 mM MeOH; **bottom left:** varying enzyme loading at 1:2 *rAaeUPO-H* to *PpAOx* ratio using 200 mM MeOH; **bottom right:** varying *rAaeUPO-H* to *PpAOx* ratios at 200 mM MeOH. Reactions were run in duplicates and the errors are reported as standard deviation.

Having established that the 1:2 ratio between *rAaeUPO-H* and *PpAOx* gave the best results, a series of experiments was performed looking at a constant enzyme ratio with different overall enzyme loading (Figure 5-9 bottom left). The experiments performed at 2:4 and 1:2 U mL⁻¹ UPO to AOx were complemented by a run at 0.5:1 U mL⁻¹ UPO to AOx. The experiments showed that with decreasing amounts of protein the observed overoxidation also decreases, from 13 to 6 to 2%, while the

overall conversion to the desired product **2** stayed within a few percent of each other with 83, 89, and 86%, for the 2:4, 1:2, and 0.5:1 U mL⁻¹ UPO to AOx loadings. This set of experiments further suggested that decreasing the overall amount of protein slowed down the accumulation of peroxide in the solution, however, as overoxidation still occurs, it was likely the best ratios had not been tested yet.

As the overoxidation of 1-phenylethanol to acetophenone was most likely connected to an accumulation of hydrogen peroxide in the system, the last set of experiments was performed with a decreasing amount of *PpAOx* whilst keeping the *rAaeUPO-H* concentration the same (Figure 5-9 bottom right). The results showed that at a 20:1 UPO to AOx ratio (1:0.05 U mL⁻¹) the reaction was slowed down so far, that after 2 h only 68% conversion to **2** was observed, without indications for overoxidation. In the 10:1 UPO to AOx sample (1:0.02 U mL⁻¹) there was also no evidence for overoxidation, while the conversion to **2** has increased over time to 85%. In the next sample with 5:1 UPO to AOx ratio (1:0.02 U mL⁻¹) overoxidation to acetophenone was observed after 60 min incubation, increasing over time (from 6 to 11% after 60 and 120 min, respectively). It became visible from these last experiments that decreasing the amount of *PpAOx* did indeed slow down the generation and accumulation of hydrogen peroxide in the system. To achieve a reasonably short reaction time, without sacrificing the yield, a *rAaeUPO-H* to *PpAOx* ratio between 1:10 and 1:5 seemed advisable.

Some additional experiments with the *in situ* hydrogen peroxide generation system were performed looking at the methanol loadings and the activity in the presence of a small range of solvents. These experiments were all performed using a 1:2 U mL⁻¹ *rAaeUPO-H* to *PpAOx* ratio. With methanol being the substrate converted by *PpAOx* to generate hydrogen peroxide, it was tested whether lowering the amount of methanol in the system could have an influence of the conversion of ethylbenzene **1** or the oxidation of 1-phenylethanol **2** (Figure 5-10, left). It was found that when using 1, 2, or 5 eq of methanol with respect to the amount of substrate used, no changes in the conversion rates were observed. With the option of using methanol as a solvent as well as a co-substrate of the reaction an additional set of experiments were run at 10% (v/v) methanol. A minor decrease in conversion from 89%, measured for

the 50 mM methanol loading, to 85% was observed, keeping the influence of methanol to the overall reaction at a minimum.

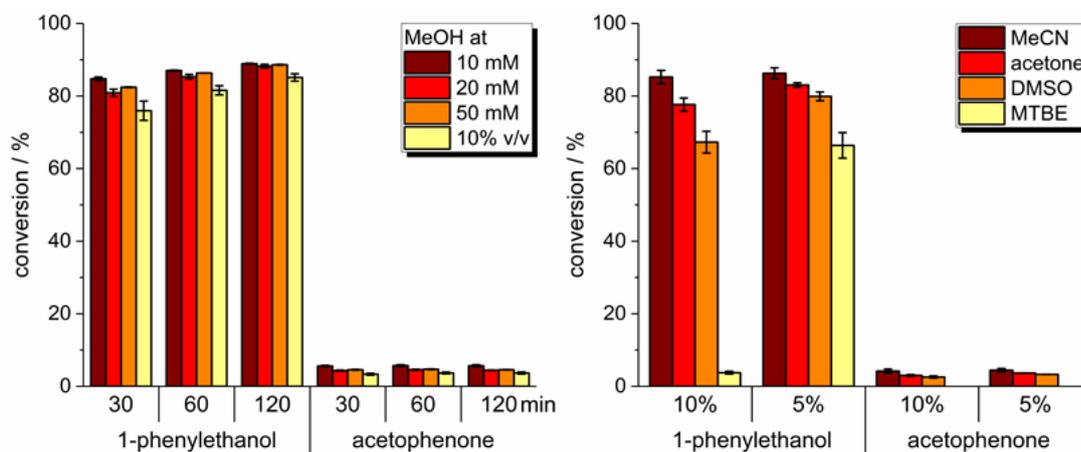


Figure 5-10: Conversions for the formation of 1-phenylethanol **2** and acetophenone **3** plotted over time at: **left:** varying MeOH concentration at 1 U mL^{-1} *rAaeUPO-H* and 2 U mL^{-1} *PpAOx*; **right:** varying solvent and solvent concentration after 2 h incubation with 1 U mL^{-1} *rAaeUPO-H*, 2 U mL^{-1} *PpAOx* and 200 mM MeOH. All reactions were run in duplicates and the errors are reported as standard deviation.

As experiments performed earlier showed that *rAaeUPO-H* was able to catalyse reactions at good rates in the presence of a range of substrates, a small solvent screen was performed here, looking at the influence of solvents on the *in situ* hydrogen generation system. Reactions were run in the presence of 5 and 10% of acetonitrile, acetone, DMSO, and MTBE (Figure 5-10, right). Using acetonitrile, the conversion stayed the same in both 5 and 10% solvent at around 86%. The presence of larger amounts of solvents otherwise led to a decrease in the conversion from 83% (5% solvent) to 78% (10% solvent) in the case of acetone and 80% (5% solvent) to 67% (10% solvent) in the case of DMSO. The largest drop in activity was observed in MTBE, where in the presence of 5% solvent the conversion yielding **2** of 66% was observed, which decreased to 4% in the presence of 10% MTBE. Comparing these results to the previously discussed data where a manual addition of hydrogen peroxide was performed, it was found that *PpAOx* was less active in 10% solvent. With the manual addition acetonitrile, acetone, DMSO, and MTBE yielded in 91, 90, 85, and 27% conversion rates for **2**, where the *in situ* generation system only gave rates of 85, 78, 67, and 4% in the same amount of solvent.

5.4.2.1. Application of the *in situ* H_2O_2 generation system

With an *in situ* hydrogen peroxide system in place running reactions at higher substrate loading became more feasible. As a proof-of concept experiment the 188

conversion of **1** was monitored at 30, 50, and 100 mM substrate loading. The reaction was run using 1 U mL⁻¹ *rAaeUPO-H* and 2 U mL⁻¹ *PpAOx* in 10% acetonitrile. After 3 h, conversions of 72, 55, and 24% were recorded for the 30, 50, and 100 mM loading, respectively. After overnight incubation these conversions increased to 87, 83, and 77%. However, the final 24 h samples need to be regarded with care, as a possible evaporation of the substrate could have influenced the calculated conversions.

Ethylpyridine **4** was used at 20 mM loading with 1:0.02 U mL⁻¹ *rAaeUPO-H* to *PpAOx*. The reaction was run for 8 h and resulted in a conversion rate of 50%. This was a 10-fold increase from the previously observed conversion using the manual peroxide addition. This increase in reaction efficiency allowed the chiral analysis of the product **5**, revealing that the (*R*)-enantiomer was produced with an enantiomeric excess of >99% (see Figure 5-11).

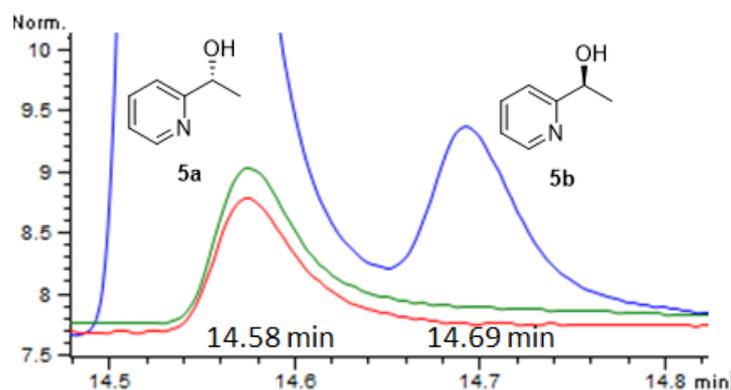


Figure 5-11: Chiral GC-chromatogram for the separation of 1-(pyridine-2-yl)ethan-1-ol **5** into the (*R*)- (**5a**) and (*S*)- (**5b**) enantiomers, with the standard for **5a** in blue, and the reactions catalysed by *rAaeUPO-H* in green and red.

One example allowing the direct comparison between the manual addition and the *in situ* generation system can be described using substrate **8**, 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine. The conversion to the desired alcohol product in position C5 **9** catalysed by *rAaeUPO-H* was around 12% when using the manual peroxide addition and was increased to 24% using the *in situ* generation system (1:0.02 U mL⁻¹ *rAaeUPO-H* to *PpAOx*).

The *in situ* hydrogen generation system was also used with H-*AbiUPO*. A reaction was set up with 20% (v/v) H-*AbiUPO*, 10 mM *sec*-butylbenzene **13** in 10% acetonitrile and 2 U mL⁻¹ *PpAOx*. The determined conversion to **14** was 5% after 20 h incubation. For the same reaction run with the manual addition of hydrogen peroxide over time, total addition of 1 eq hydrogen peroxide with respect to the substrate loading, giving

11% 2-phenylbutan-2-ol **14**. This suggested, that while still functional, a separate optimisation looking at the H-*Abi*UPO to *Pp*AOx ratio would have to be performed to obtain the best results.

5.5. Stirred Small-Scale Reactions

In addition to the broad substrate screen performed in an industrial setting, ethylbenzene derivatives and thioanisole derivatives were studied in-house to gain more insight into side-chain influences and limitations and to obtain a wider insight into the reaction scope of *rAae*UPO-H, the five-point mutant. The full range of substrates and their transformation products is shown below (Figure 5-12).

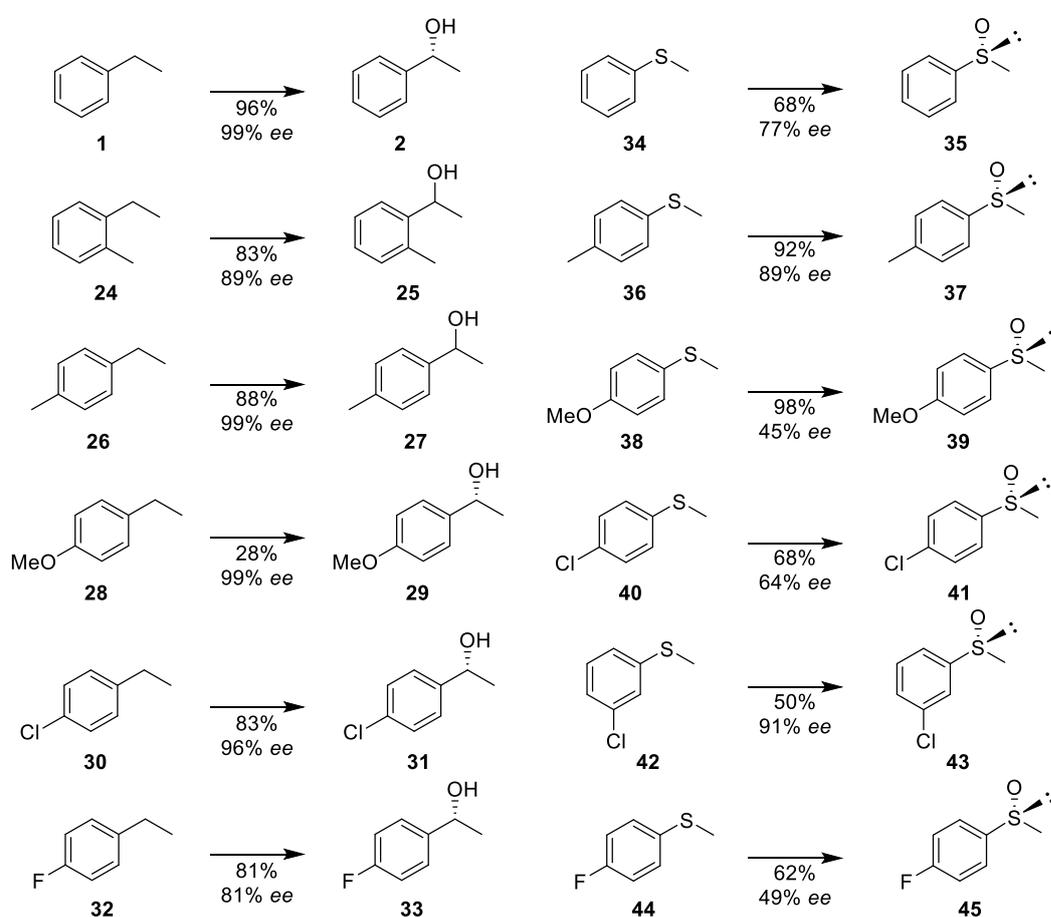


Figure 5-12: Overview over the different ethylbenzene and thioanisole derivatives used in reactions with *rAae*UPO-H and the determined conversions and enantiomeric purity.

All reactions were run with 5 mM substrate in 1% ethanol, for 3 h at room temperature. Samples were analysed over time using chiral and achiral GC.

The hydroxylation reactions using ethylbenzene derivatives were generally giving high conversions of greater than 80%, yielding in the 1-phenylethanol derivatives as products. The only exception was 4-ethylanisole **28**, where only 28% conversion to

1-(4-methoxyphenyl)ethanol **29** was observed after 3 h reaction time. The lower conversion was due to the electron-donating *p*-methoxy group, which created a more negative character at the benzylic carbon centre due to tautomerisation, slowing down the functionalisation at this position. In the literature, studies using wild-type *Aae*UPO showed that **1** was converted with a turnover frequency (TF) of 399 min⁻¹, and **28** and **30** were converted at TFs of 385 and 303 min⁻¹, respectively. The *ees* were reported with 97, 90, and >99%, respectively, favouring the (*R*)-enantiomers.⁹⁴ The sulfoxidation reactions of the thioanisole derivatives displayed conversions ranging from 50% for 3-chlorothioanisole **42** to 98% for 4-methoxythioanisole **38** in the presence of *rAae*UPO-H. The control reactions displayed non-enzymatic sulfoxidation activity, with conversions of up to 5%. The accumulation of hydrogen peroxide in the reaction mixture over time enables the oxidation at the sulfur to take place in the absence of further catalysts. Not much literature can be found looking at the ability of *Aae*UPO to oxidise sulfides; thioanisole has been described to be converted by *Aae*UPO at a TF of 35 min⁻¹, with an *ee* <70%, favouring the (*R*)-enantiomer.⁹⁴

The time-course data for the hydroxylation of **1** and the sulfoxidation of methyl *p*-tolyl sulfide **36** is shown below (Figure 5-13). The GC analysis of the samples showed an accumulation of product over time (yellow) coupled to a decrease in substrate concentration (dark red). No formation of by-products or overoxidation products was observed within the 3 h timescale. The data collected for the other substrates followed similar progressions. For the samples with highest overall conversions, **1** at 96%, **26** at 88%, and **38** at 98%, the conversions only increased by 1-2% within the last hour. A similarly slow increase was also observed for 4-chlorothianisole **40**, with an 1% increase from 67 to 68% conversion to methyl *p*-chlorophenyl sulfoxide **41** after 120 and 180 min reaction time, respectively, and for **42** with a stagnation in conversion at 50% after 120 min. All other substrates showed an increase in product conversion between 5 and 10%. The largest increase of approx. 30% was observed with 1-ethyl-4-fluorobenzene **32** yielding in a conversion to 1-(4-fluorophenyl)ethanol **33** of 81% after 180 min. This data suggested that while some reactions had reached completion after 180 min, the collection of additional time points for some samples could have been beneficial in the determination of the

overall conversion. However, as other experiments with **1** suggest, prolonged incubation times facilitate the formation of by-products, hence, this would also have to be accounted for.

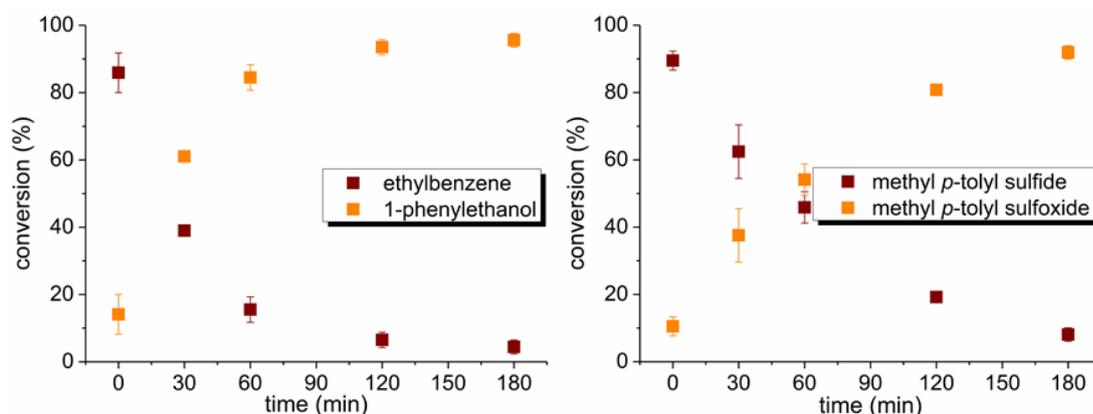


Figure 5-13: GC-conversions of ethylbenzene **1** (left) and methyl *p*-tolyl sulfide **36** (right) in the presence of rAaeUPO-H over time yielding in 1-phenylethanol **2** and methyl *p*-tolyl sulfoxide **37**, respectively. The data represents the mean calculated from independent three reactions. All reactions were run in triplicates and the errors are reported as standard deviation.

The analysis of the total stereochemistry of the product was based on chiral GC experiments. In the case of the ethylbenzene derivatives, product markers were obtained and allowed the determination of the enantiometric excess (*ee*). For the sulfoxidation reactions the analysis relied on previous studies performed in the group, which identified the elution order of the sulfoxide enantiomers using the same chiral column.¹⁴⁴

The hydroxylated products showed excellent *ees* with >99% for (*R*)-1-phenylethanol **2**, 1-(4-methylphenyl)ethanol **27**, and (*R*)-1-(4-methoxyphenyl)ethanol **29**, where only one enantiomer was detected in the chromatogram. An enantiopure marker for **27** could not be obtained for this work, therefore the nature of the enantiomer had not been determined. The elution order for the other 1-phenylethanol derivatives suggested the (*R*)-enantiomer to appear first of the chiral column, suggesting the presence of (*R*)-**27**. For 1-(2-methylphenyl)ethanol **25**, and (*R*)-1-(4-chlorophenyl)ethanol **31** good values of 89 and 96% *ee* were determined. As for **25** the overall stereochemistry was not determined due to no available marker, however, the elution order would suggest the presence of (*R*)-**25**. The reaction of 1-ethyl-4-fluorobenzene **32** yielding in 1-(4-fluorophenyl)ethanol **33** was unique amongst the hydroxylation reactions, as the control reaction displayed 5 and 10% conversion after 120, and 180 min respectively. This was also visible in the chiral GC

chromatogram (see Figure 5-14, bottom left), where a racemic mixture of the product was observed. This unspecific background reaction seemed to effect the enantiomeric excess determined in the reactions with *rAaeUPO*-H (see Figure 5-14, top left), where an average of 88% *ee* towards (*R*)-**33** was measured. For the racemic mixture a baseline-separation of the enantiomers was achieved, whilst the chromatogram for the analysed reactions showed slight overlapping of the peaks, suggesting that the actual *ee* was reduced.

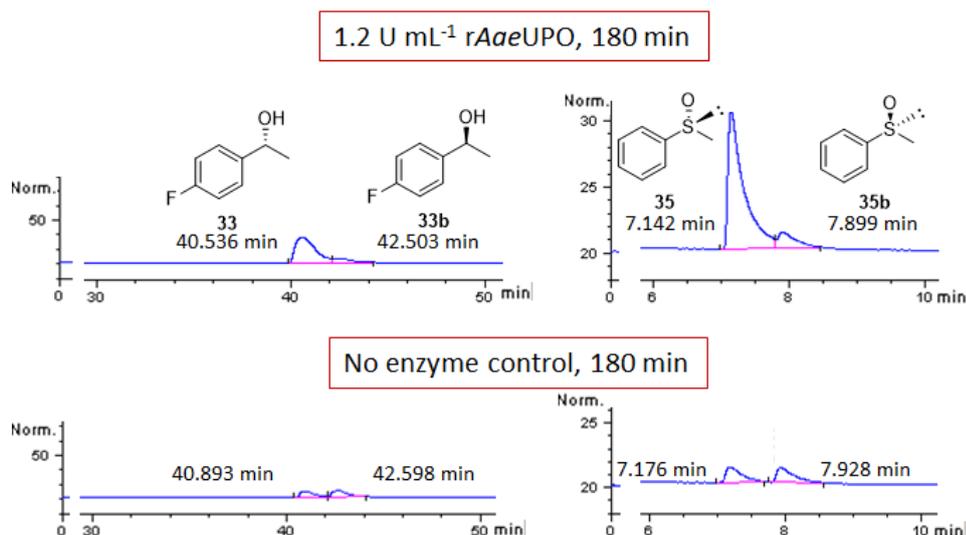


Figure 5-14: Chiral GC-chromatograms for samples containing 1-(4-fluorophenyl)ethanol **33** (left) and methyl phenyl sulfoxide **35** (right), obtained in the presence (top) or absence (bottom) of *rAaeUPO*-H after 180 min reaction time.

The enantiomeric excess determined for the sulfoxidation reactions was generally lower than those observed for the hydroxylated products. The *ees* ranged from 45% for (*R*)-methyl *p*-methoxyphenyl sulfoxide **39** to 91% for (*R*)-methyl *m*-chlorophenyl sulfoxide **43**. The background oxidation of the sulfides in the presence of hydrogen peroxide at around 5% could be influencing the *ee*, however, it seemed that the overall selectivity for these compounds was lower than for the ethylbenzene derivatives. An example for the chiral analysis of methyl phenyl sulfoxide **35** is given in Figure 5-14 (right). A comparison between the samples run in the presence (top) or absence of *rAaeUPO*-H, highlights the influence of the racemic background production of product on the overall enantiopurity of the enzyme catalysed reaction. Similarly, to the observations made for **33**, baseline-separation was observed in the racemic mixture, however, tailing of the (*R*)-enantiomer of **35** led to a merging of the peaks, and therefore the determined *ee* is likely to be calculated too high.

One interesting result to comment on further is the conversion of 3-chlorothioanisole **42**. This substrate was the only substrate tested here carrying a substituent in the *meta* position. Compared to the *p*-substituted **40** with a calculated 68% conversion, the conversion of **42** was at 50% almost 20% lower, and was overall the lowest conversion rate off all sulfide substrates. This could suggest a deactivating effect from substituents in the *meta* position. In comparison the *o*- and *p*-substituted ethylbenzene derivatives, **24** and **26**, only displayed a 5% drop in conversion when changing from the *para* to the *ortho* substituent. Furthermore, the sulfoxide **43**, whilst displaying the lowest conversion, showed the highest enantioselectivity amongst the sulfoxides with 91% *ee* favouring the (*R*)-enantiomer. As no background reaction was observed for this substrate as well as substrate **40**, this could contribute to the high purity, but it could also be that the specificity of the enzyme was increased using substrates with *meta*-substituents.

5.6. Design of Experiments

The design of experiments (DoE) aims to provide a systematic approach to understand and optimise a process by determining factor-response relationships. The variables are the factors and the different outputs are the responses. Here, the conversion of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** was chosen as reaction to be optimised (see Figure 5-15).

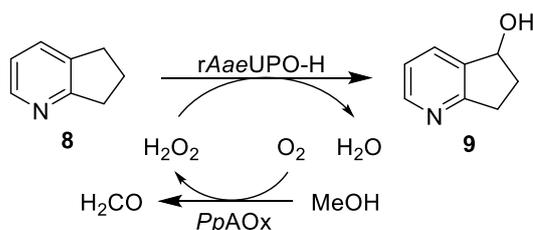


Figure 5-15: rAaeUPO-H catalysed conversion of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** in the presence of the PpAOx and methanol.

As input factors the substrate concentration (5-50 mM), the temperature (20-40 °C), the rAaeUPO-H loading (1-10 U mL⁻¹), and the rAaeUPO-H to PpAOx ratio (20:1-5:1) were chosen. Ten reactions were run using the factors described before (see Table 5-6), containing two centre points, for reproducibility and eight factorial points, with a combination of factors. As responses the GC-based conversion to the desired hydroxylated product 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-5-ol **9** and the GC-based

conversion to by-products (**10-12**) were chosen, to be determined before and after the isolation of the compounds from the reaction mixture. The isolated crude yield was also considered as a response, however, it was deemed to be less reliable than the GC-based conversions (see Table 5-8). The conversion for the by-products presented below describes the sum of conversions registered for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-5-one **10**, the overoxidation product, 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-6-ol **12**, and 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-7-ol **11**, with the biggest contribution from **11** with up to 17%, followed by the **10** with up to 10%, 17% in the case of run 6.

Table 5-8: Responses determined for the ten DoE runs, with the conversions for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-5-ol **9**, and the by-products 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-5-one **10**, 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-6-ol **12**, and 6,7-dihydro-5*H*-cyclopenta[*b*]pyridin-7-ol **11**, before and after extraction, as well as the crude isolated yield.

Run	Response 1	Response 2	Response 3	Response 4	Response 5
	Conversion / % to 9	Conversion / % to 10-12	Isolated crude %	Extracted Conversion / % to 9	Extracted Conversion / % to 10-12
1	24	6	74	78	16
2	1	0	41	2	0
3	35	8	41	42	11
4	26	6	29	45	9
5	73	22	100	72	24
6	66	34	100	73	20
7	7	1	49	10	2
8	2	0	34	2	0
9	73	25	74	70	27
10	29	7	31	55	13

Using the Design-Expert software the different factors contribution to every specific response could be calculated. This was done using least square estimation, where the different factors and combinations of factors were ranked. The visual output, half-normal plots, for the responses 1, 2, 4, and 5, the GC-based conversions, are presented below (Figure 5-16). In all cases the biggest influence on the responses was the substrate concentration, depicted as a negative effect. As high substrate concentrations could lead to an inhibition of the protein, but also could affect the viscosity or integrity of the solution, due to phase separation or precipitation; this result was to be expected. The concentration of *rAaeUPO-H* had a positive effect on

the production of the desired product **9** and an even more pronounced effect on the by-product formation, as shown in the graph at the top right (conversion to by-products after 20 h). This effect can be explained by looking at the ratios of products formed. In initial experiments before the DoE, it was established that **9** was the main reaction product at 24%, with **11** formed at 7% and overoxidation to **10** at 1%. This suggested, that the selectivity for **9** over **11** was around 3:1, or 70%, and overoxidation only occurred with approx. 4% selectivity. From the data collected here, it was found that the selectivity of **9** over **11** was between 75-85%, in all cases. Overoxidation was observed at a selectivity around 3-4% when the conversion to **9** was <60%, usually coinciding with lower *rAaeUPO-H* loading. At higher conversions the amount of overoxidation observed increased to 5-8%, corresponding to a selectivity between 7-11%, and the highest value of 17% conversion to **10** was seen with a selectivity of 25% in the run 6, containing 10 U mL⁻¹ *rAaeUPO-H*. In summary, increasing the protein loading influenced the conversions to both the desired product **9** and the by-products, but while the selectivity for the hydroxylation in the C5 position over the C7 position remained constant, increased overoxidation to **10** was observed. Further, significant factors determined in the half-normal plots were the temperature, the *rAaeUPO-H* to *PpAOx* ratio, and combined effects between the various factors mentioned so far (depicted in the graphs as AC or AB, with the letters corresponding to the factor listed in the respective legend).

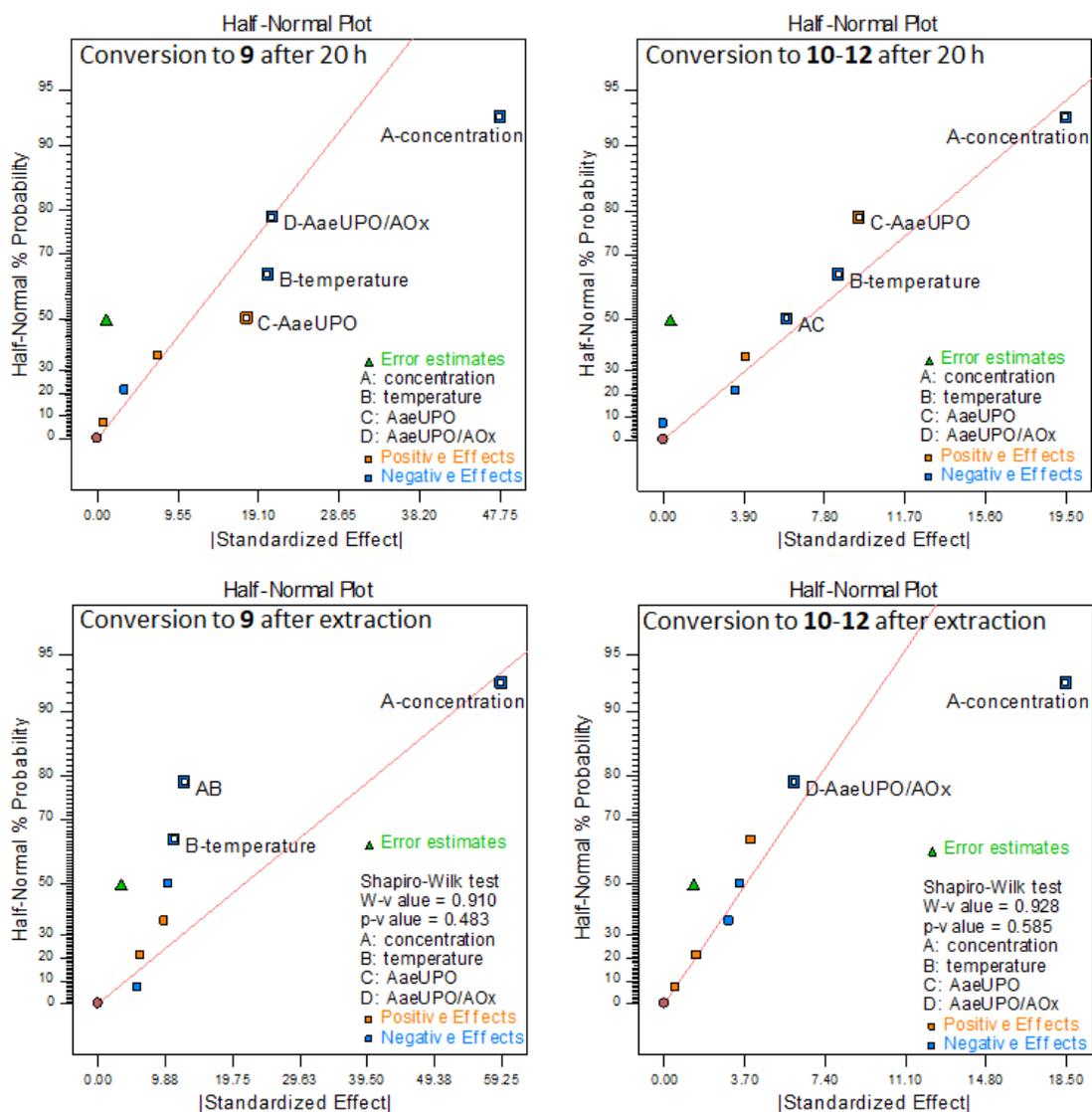


Figure 5-16: Half-Normal plots for the responses 1, 2, 4, and 5 determined using the Design-Expert software, representing the influences of the different factors on the conversion to for 6,7-dihydro-5H-cyclopenta[b]pyridin-5-ol **9** after 20 h (**top left**), and after extraction (**bottom left**), and the sum of conversion rates for the by-products **10-12** after 20 h (**top right**), and after extraction (**bottom right**).

Bringing these factors together, an overlay plot was created where the conversion was to be maximised to >40% before extraction and >50% after extraction, and the by-product formation was to be minimised to <20% before extraction and <15% after extraction (see Figure 5-17). By plotting the *rAaeUPO-H* to *PpAOX* ratio against the substrate concentration and by adjusting the temperature and *rAaeUPO-H* concentration, it was possible to create a window where all the desired responses were met.

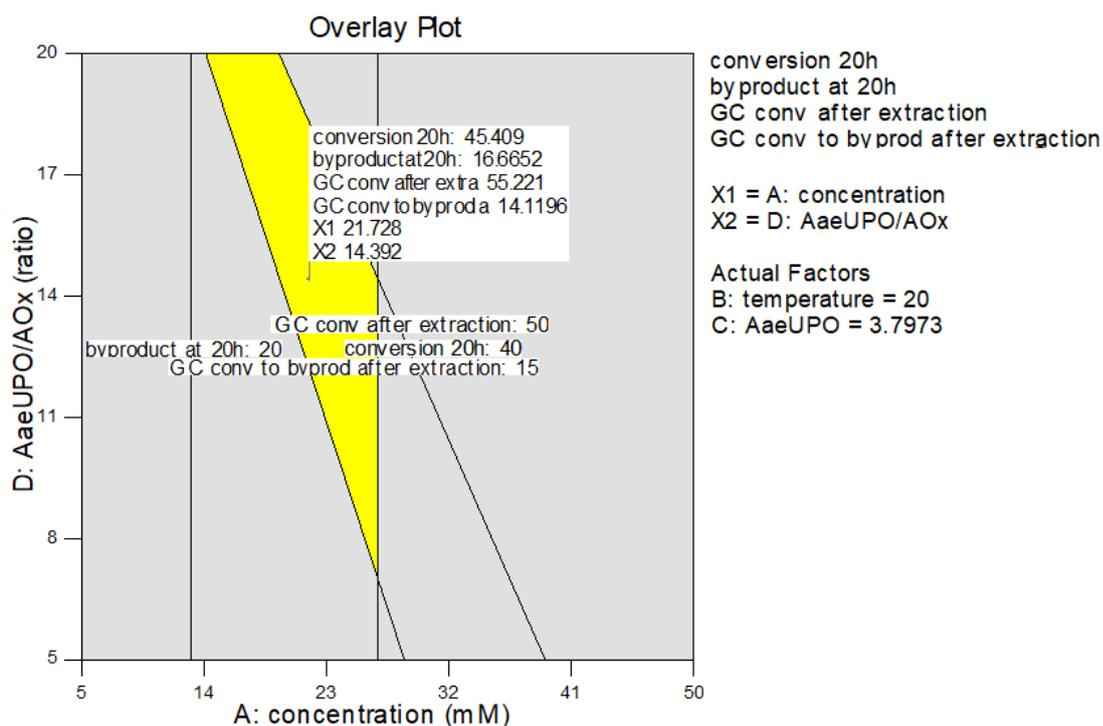


Figure 5-17: Overlay plot showing the *rAaeUPO-H* to *PpAOx* ratio plotted against the substrate concentration aiming to maximise the GC-based conversion towards the desired product **9** (>40% after 20 h, >50% after extraction), and minimising the by-product formation (<20% after 20 h, <15% after extraction) which is represented by the area highlighted in yellow.

Resulting from the overlay plot, conditions meeting the desired outcomes were calculated and listed in order of their desirability. According to the DoE the best results for the reaction of **8** with *rAaeUPO-H* should be obtained at a 40 mM substrate loading at 20 °C, using 10 U mL⁻¹ *rAaeUPO-H* at a 5:1 ratio (2 U mL⁻¹ *PpAOx*). The conversion after 20 h was calculated to be around 50% with approx. 13% by-product formation. The values should decrease to 35% and 10% after extraction.

5.6.1. Scale-up Reaction

Using the output determined in the DoE, a scale-up reaction using 1 g of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** was performed. The reaction was set-up using 50 mM substrate loading, 10 U mL⁻¹ *rAaeUPO-H* at a 5:1 ratio with *PpAOx*, at 20 °C. The conversions were determined using GC analysis before and after the extraction of the compounds from the reaction mixture (see Figure 5-18, left). The desired product **9** was formed at 63% conversion, the overoxidation product **10** was observed with 4%, and the hydroxylations in position C7 (**11**) and 6 (**12**) occurred with 17 and 2% conversion. The determined rates did not change for the extracted sample. The selectivities for the hydroxylation in the different positions C5:C7:C6 (**9:11:12**) were

calculated to 1:0.22:0.03. These results were very similar to the ratios calculated from the $^1\text{H-NMR}$ at 1:0.25:0.1, where all three alcohols could be differentiated from each other through the proton shift corresponding to the hydrogen bound to the hydroxylated carbon (see Figure 5-18, right). The crude yield was determined to be 63%.

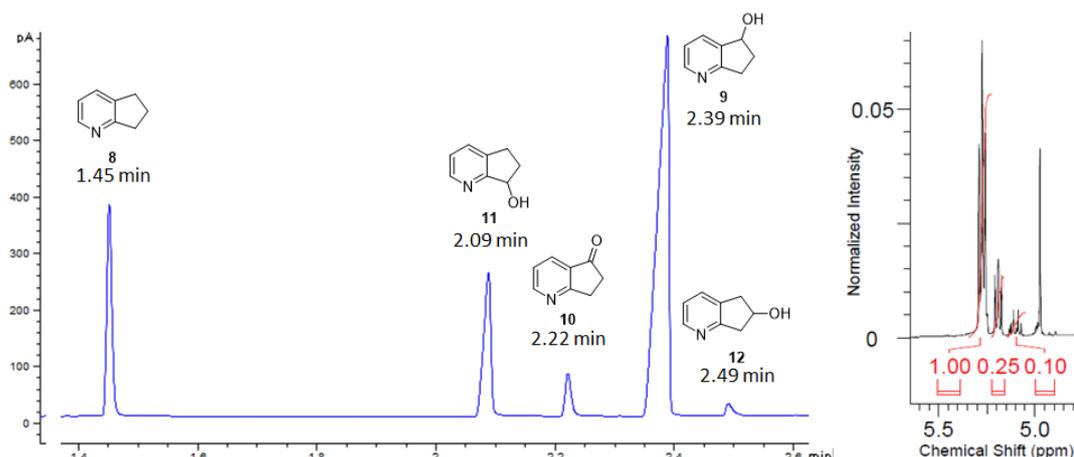


Figure 5-18: 1 g reaction of 6,7-dihydro-5H-cyclopenta[b]pyridine **8** with rAaeUPO-H before extraction: **left** - GC chromatogram; **right** - $^1\text{H-NMR}$ zoomed to the area of interest between 5.0 and 5.5 ppm, showing the peaks for the protons bound to the hydroxylated carbon.

The crude mixture was purified using an automated chromatography system. The product eluted in a broad peak, overlapping with another peak, most likely containing the by-product **11**. As a result, the obtained fractions were combined into two different samples for further analysis, a pure and a mixed product sample. The final GC analysis of these fractions showed the pure fraction with a **9:11** ratio of 1:0.2, and the mixed product fraction with a ratio between **9:11:12** of 1:0.4:0.05. This suggested that the smaller, overlapping peak did indeed mainly contain the product hydroxylated in the position C7, as well as the alcohol in the position C6. The fraction, which was hoped to represent only the desired product **9**, still contained **11**, making up for approx. a fifth of the sample. The isolated yield for the pure fraction was 25%, and for the mixed fraction 11% were calculated. Due to time constraints, further purification trials to yield a better separation and potentially a higher overall extracted yield, were not conducted. The final $^1\text{H-NMR}$ spectrum for the pure fraction is shown below, with all proton assignments (see Figure 5-19).

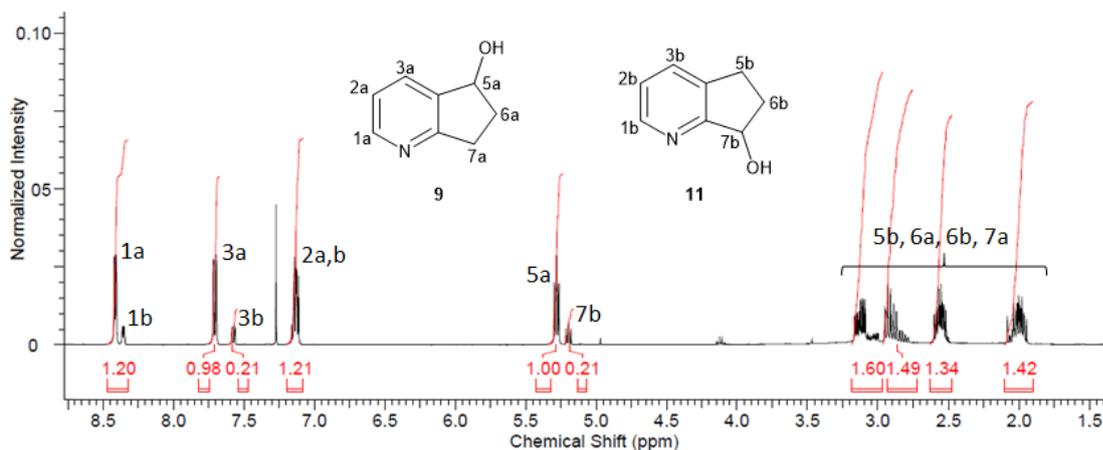


Figure 5-19: $^1\text{H-NMR}$ spectrum obtained for the 'pure'-sample after column chromatography, with the peak assignments for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol **9** and 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-7-ol **11**.

The last part of this experiment included the assignment of the stereochemistry of the alcohol formed in the desired product **9**. Therefore, a chiral GC analysis was performed using a racemic mixture and the enantiopure (*S*)-6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol **9** as standards (see Figure 5-20, left). The (*S*)-enantiomer elutes of the column first, at 16.89 min, and the (*R*)-enantiomer elutes after 16.94 min. As the chromatogram for the racemic mixture showed, the two enantiomers were not fully separated from each other, which will lead to the calculated *ee* to be too high. For all samples run during the purification of the reaction of **8** with rAeUPO-H an enantiomeric excess around 88% was calculated favouring the (*R*)-enantiomer of **9**.

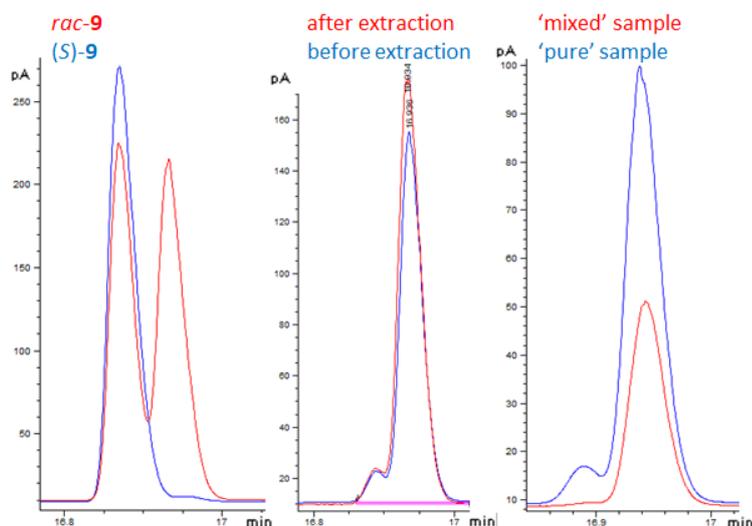


Figure 5-20: Chiral GC-chromatogram for 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol **9**: **left** – racemic mixture (red) and (*S*)-**9** (blue) as standards; **middle** – samples taken before (blue) and after (red) extraction; **right** – final samples for the fractions combined for 'pure' (blue) and 'mixed' (red).

5.7. Summary, Discussion and Conclusion

This chapter gave multiple examples for the application of UPOs, especially *rAaeUPO-H*, in an industrial setting and their flexibility in biotransformations. With help from GSK Worthing it was possible to scale-up the fermentation of *rAaeUPO-H* to a total volume of 24 L, which was concentrated to a 10X solution yielding in 2 L of protein at $155 \pm 5 \text{ U mL}^{-1}$ specific activity towards ABTS. The protein was further subjected to lyophilisation experiments showing that *rAaeUPO-H* did not lose activity in the freeze-drying procedure. No apparent change in activity was observed based on spectrophotometric assays and GC-analysed reactions. The results, obtained in cooperation with an industrial partner highlight the applicability of UPOs in this setting. It further allowed for easier handling and storage of the protein due to the lyophilisation.

Substrate screens identified a wide range of both known and new *rAaeUPO-H* substrates. Best transformation results were obtained at pHs between 6.4 and 7.4 for reactions where benzylic hydroxylations took place. The screening of substituted toluene derivatives showed little to no activity for hydroxylations in the terminal position, while different unidentified by-products could be observed. This is in line with literature studies using toluene as an *AaeUPO* substrate, where besides the benzyl alcohol, five by-products were also formed.⁷⁶ Cyclohexane **17** and cyclohexanol **18** were also found to only yield poor conversion. These results were unexpected, as the wild-type *AaeUPO* had been reported to give up to 76% conversion to **18** using **17** as a substrate, and 1% overoxidation to cyclohexanone **21** was also reported.⁹⁷ In the presence of benzylic and tertiary carbons both *rAaeUPO-H* and *H-AbiUPO* selectively hydroxylated in the tertiary position. For *H-AbiUPO* only substrates carrying a tertiary carbon were identified as substrates. The highest conversion was determined for cumene **6** with 51%. *rAaeUPO-H* was able to catalyse the same reaction with a conversion of 88%.

Optimisation experiments showed that the temperature dependence was coupled to specific substrates, where **6** showed higher conversions at higher temperatures, whilst ethylpyridine **4** showed better rates at lower temperatures. Further experiments showed that reaction containing *rAaeUPO-H* could easily be scaled to larger volumes without decreasing the conversion values. A solvent screen, using a

small library of water miscible and immiscible solvents, showed excellent results in the conversion of ethylbenzene **1** using acetonitrile or acetone (90%), but reduced conversion when using water immiscible solvents like ethyl acetate (55%) or MTBE (27%, see Figure 5-5). In the literature the *rAaeUPO* expressed in *Saccharomyces cerevisiae* showed higher reductions in activity towards ABTS when using 10% solvent.¹²⁰ Only 50% activity was retained for acetone at 10% v/v loading. This suggested, that either the *S. cerevisiae* expressed *rAaeUPO* has lower stability, or that the activity in different solvents is further coupled to the substrate which is used in the experiment.

To further optimise reactions, hydrogen peroxide addition rates were studied. As hydrogen peroxide is known to deactivate the protein in high concentrations, it was crucial to control the addition of this co-substrate.⁹⁷ Using ethylbenzene **1** as model substrate it was found that the addition in 15 and 30 min intervals did not affect the outcome, hence, the reaction could be decreased by picking shorter addition intervals. Using methyl *p*-tolyl sulfide as substrate, the accumulation of peroxide in the reaction mixture did not only affect the stability of the heme but could also lead to background oxidations of the sulfide, which in turn affect the enantiomeric excess of the product. Here, it was shown that longer addition intervals gave higher yields. However, it is important to notice that longer incubation times also increased the substrate conversion observed in the control sample.

An alternative to dosing the manual addition of peroxide to the reaction mixture, an *in situ* hydrogen peroxide generation system could be used. There are different literature known examples for this, relying either on an enzymatic or a photocatalytic component.^{69, 94, 111} Most recently the usage of *PpAOx*, an alcohol oxidase from *Pichia pastoris*, was reported as part of a three-enzyme system fully utilising methanol, to yield three equivalents of hydrogen peroxide.¹¹¹ As only *PpAOx* is commercially available, it was used in experiments here, to generate one equivalent of peroxide per methanol, used as the co-substrate. A wide variety of *rAaeUPO-H* to *PpAOx* ratios was screened to identify the best working condition, with reduced overoxidation, but acceptable reaction time. A 5:1 ratio was found most effective using 1 U mL⁻¹ *rAaeUPO-H* (specific activity towards ABTS). To evaluate the solvent tolerance of this two-enzyme system, different solvents were tested at 5 and 10% (v/v) concentration

(see Figure 5-10). Results showed that compared to the reaction run under comparable conditions with the manual addition of hydrogen peroxide, the conversion rates were decreased, indicating that *PpAOx* did not display as high solvent stability as *rAaeUPO-H*. For acetonitrile the conversion yielding 1-phenylethanol **2** decreased from 91% to 85%, in the case of acetone the drop was from 90% to 78% and the most distinct change was observed using MTBE where a drop from 27% to 4% occurred.

Using other substrates with the *in situ* peroxide generation system and *rAaeUPO-H* it was possible to show improved conversions. For 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** the conversion to the alcohol in position C5 **9** was doubled, from 12% to 24%, and ethylpyridine **4** could be used at 20 mM substrate loading and gave a 50% conversion, improved 10-fold from the experiments run at 25 °C and 10 mM, where results around 4% were determined. The improved outcome of the latter biotransformation further allowed the chiral analysis of the product, revealing that (*R*)-**5** was formed as the sole reaction product (see Figure 5-11). When using *H-AbiUPO* together with *PpAOx* the observed conversions for *sec*-butylbenzene **13** were decreased from 11% to 5%. This suggested that a separate optimisation experiment would be necessary to get the best outcome for reactions run with both *H-AbiUPO* and *PpAOx*.

The establishment of the *in situ* hydrogen peroxide generation system also led to a design of experiments approach to optimise the conversion of 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine **8** using *rAaeUPO-H*. The varying factors in the experiments were the substrate concentration, the temperature, the *rAaeUPO-H* loading, and the *UPO* to *PpAOx* ratio. For the analysis the GC-based conversions to the desired product **9** and the by-products (**10-12**) before and after extraction were considered as response factors. The half-normal plots showed that the temperature and the *rAaeUPO-H* loading had the biggest influence on the reaction outcome (see Figure 5-16). The protein loading in addition to favouring the product formation, also enables higher rates of overoxidation. The ideal experiment, yielding in a conversion to **9** of over 50% and a minimal amount of by-products, was calculated to be run at 40 mM substrate concentration at 20 °C, with 10 U mL⁻¹ *rAaeUPO-H* loading and a 5:1 *UPO* to *PpAOx* ratio. With this in mind, a 1 g reaction of **8** was run, showing 63%

GC-based conversion to **9**, improved from 24% found in the small-scale run experiment (see Figure 5-18 for GC-based conversion). First isolation attempts showed the co-purification of the by-product **11**, which was not successfully removed, as indicated by GC- and ¹H-NMR analysis (see Figure 5-19). Of the 1:0.2 **9:11** ratio sample 288 mg were isolated. Chiral GC analysis of the sample revealed that the (*R*)-enantiomer of **9** was formed with an enantiomeric excess >85%.

Last, a small library of ethylbenzene and thioanisole derivatives was screened using *rAaeUPO-H* as catalyst. Whilst ethylbenzene is a known and established substrate for both the wildtype and the mutant *AaeUPO*, not much research is available looking at the use of UPOs in sulfoxidation reactions.^{94, 104} It was found that the ethylbenzene derivatives were converted with conversions >80%, the only exception being *p*-ethylthioanisole **28**, where a deactivating effect was observed lowering the rate to 28%. Overall good to excellent *ees* were calculated with values >89%, favouring the (*R*)-enantiomer, where markers were available. A drop in *ee* was observed for 1-ethyl-4-fluorobenzene **32**, where background conversion in the control sample took place, effecting the overall determined *ee* for the reaction. The sulfoxidation reactions generally suffered from unspecific background reactions taking place. In the control reactions roughly 5% conversion was observed after 3 h reaction time. The determined conversions in the enzyme catalysed samples were between 50-90%, and the determined *ees* ranged from 45-91%, always in favour of the (*R*)-enantiomer. Generally, the conversions of the ethylbenzene derivatives are similar to literature reported conversion for these types of substrates, with the same or better *ees*.⁸⁰ In the case of sulfoxidations, *rAaeUPO-H* tested here showed good results, however, *CfuCPO* has been described in the literature to catalyse the same or similar reactions with higher enantioselectivity.⁸⁰

In summary, this chapter highlights the versatility and applicability of both *rAaeUPO-H* and *H-AbiUPO* in an industrial environment. It was possible to develop an easier handling system for *rAaeUPO-H* through lyophilisation, and the establishment of a working hydrogen peroxide *in situ* generation system using *PpAOx*. A range of new substrates were identified to be hydroxylated by *rAaeUPO-H*, including **8** which could be scaled to a 1 g reaction. Novel sulfoxidation activity towards a range of thioanisole derivatives was also described. For the so far literature

unknown H-*Abi*UPO good selectivity for substrates with tertiary carbons was determined, whilst other *rAae*UPO-H substrates were not found to be converted by the homolog.

6. Conclusion and Future Work

Over the past years, unspecific peroxygenases received increasing attention due to their wide applicability in biocatalysis. The reaction scope of UPOs covers hydroxylations of aromatic and aliphatic compounds, alcohol and heteroatom oxidations, dealkylations, and epoxidations.⁸⁰ The different literature described UPO homologs display complementary activity towards a range of substrates, allowing access to a wide range of final products.^{77, 97, 99, 104} At the start of this work, these fungal enzymes were produced in the native species with an exception of *Cci*UPO, which has been heterologously expressed in *Aspergillus oryzae* (Novozymes patent WO/2008/119780) before.⁷⁹ This work contributed to the heterologous expression, characterisation and application of unspecific peroxygenases from *Agrocybe aegerita* and *Agaricus bisporus*.

Using *E. coli* as a bacterial expression host, it was possible to generate small amounts of soluble *Aae*UPO, displaying heme-thiolate absorption around 420 nm. However, it was found that the expressed protein did not display any activity towards a small range of known UPO substrates. Furthermore, the inability to purify *Aae*UPO using a range of methods suggested the misfolding of the protein. Bacterial expression hosts have the advantage of being easily manipulated and produce the enzyme of interest in short periods of time. Bacterial hosts are also very simple hosts, not allowing for post-translational modifications, and they have been proven unsuitable for the expression of complex enzymes, like *Cfu*CPO, a protein similar to UPOs.¹¹⁷ The functional expression of *Aae*UPO by *E. coli* was not achieved and the high degree of glycosylation of the protein is thought to be one of the major contributors to this result.

By using a eukaryotic expression system, like the yeast *Pichia pastoris*, post-translational modifications can be achieved. *Pichia pastoris* has the further advantage of growing to high cell densities, allowing for the large-scale production of proteins of interest. Relying on novel findings by another group, published during the first year of this project, it was possible to express and purify *rAae*UPO using yeast hosts.^{120, 133} *rAae*UPO describes a nine-point mutant of the wild-type protein, with four mutations on the signalling peptide and five on the mature protein. The

evolution of the signalling sequence allows for the successful secretion of the protein in the heterologous host, while the mutations on the mature protein are thought to influence the stability of the protein. Here, the successful expression and purification of two His-tagged *rAaeUPOs*, namely H-*rAaeUPO* and *rAaeUPO-H*, was reported. Prolonged fermentation processes led to the cleavage of the purification tag, and various attempts to increase the stability over time, including the repositioning of the tag did not yield the expected results. However, it is important to note that problems with the temperature control of the fermentation vessels have influenced these results, and improvements of the stability of the construct with the N-terminally tagged protein would be expected. The fermentations gave access to the active protein, with k_{cat}/K_m values for NBD, ABTS, and VA of $7,700 \pm 700$, 340 ± 70 and $15 \pm 5 \text{ s}^{-1} \text{ M}^{-1}$, respectively, comparable to literature reported values for *rAaeUPO* expressed in *Pichia*, with $11,000 \pm 2,200$, 590 ± 100 , and $30 \pm 5 \text{ s}^{-1} \text{ M}^{-1}$.^{120, 133} Further characteristics, like the glycosylation and UV/Vis spectrum of the protein, helped paint a comprehensive picture of the expressed protein, which was found to be comparable to previously published results.^{120, 133} In cooperation with an industrial partner, it was further possible to show that the fermentation of the protein was scalable (4 x 6 L), and that downstream processing gave access to highly concentrated *rAaeUPO-H*, which could be lyophilised without the loss of enzyme activity.

Crystallisation studies of the mutant *rAaeUPO-H* were successful and structure determination yielded three distinct structures. *rAaeUPO-H_ACT* crystallised with an acetate coordinating to the heme and the Glu196, one of the acid base-partners, and diffracted to 2.2 Å. The overall mutant structure was found to be almost identical to wild-type, with a 97% Q-score as a result of superimposition using GESAMT.^{70, 139, 145} Interestingly, this structure found the Glu196 in two conformations, coordinating to Arg189 and Arg169. While Arg189 is believed to be the base-partner in the catalysis, this work suggests Arg169 as a second base partner, which would allow for the widening of the substrate-binding pocket. The other two crystals diffracted to 1.6 Å. One of them displayed an extensive nine-glycan long glycosylation extending from Asn286 of the high-mannose type (*rAaeUPO-H_sugars*).⁸⁶ The other high-resolution structure showed an unidentified ligand coordinating to the heme and Glu196

(*rAaeUPO-H_NHI*). During refinement *N*-hydroxy imidazole was modelled into the density with success.

Following on from the successful *Pichia* expression of *rAaeUPO-H*, the expression of the homolog from *Agaricus bisporus* was also attempted. When using the N-terminally tagged variant, *H-AbiUPO*, it was possible to express the homolog with an intact purification tag. The novel homolog was successfully identified using MS experiments. It further showed the UPO characteristic heme-thiolate and Q-bands in the UV/Vis spectrum and displayed decreased activity towards ABTS and NBD, when compared to *rAaeUPO-H*.

The last chapter focused on the application of the two UPOs in biotransformations. Where literature-known *AaeUPO* substrates were used, comparable results were obtained with *rAaeUPO-H*, with an exception of cyclohexane, where activity was largely decreased. Most studied substrates were converted with high regio- and stereoselectivity. Thioanisole derivatives were identified as novel substrates for *rAaeUPO-H*, as well as 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine which after DoE analysis was scaled up to a 1 g reaction, yielding in 63% conversion to 6,7-dihydro-5*H*-cyclopenta[*b*]pyridine-5-ol, improved from 24% before the DoE. The analysis of branched alkylbenzene substrates revealed a preference for tertiary carbons over the benzylic position. Using *H-AbiUPO* in biotransformations, substrate conversion was mainly observed when a tertiary carbon-centre was present. In a direct comparison of the conversions is *rAaeUPO-H* more efficient than *H-AbiUPO*, using cumene, values of 88 and 51% were determined for the formation of 2-phenylpropan-2-ol.

The industrial applicability of *rAaeUPO-H* has been demonstrated and opens many opportunities to produce and apply the protein in even wider substrate screens. The simplified handling of the lyophilised protein further enables to easily scale-up reactions of interest. The immobilisation of the protein could yield reusable and more stable material, for a range of biotransformations. This could be complemented by using an *in situ* hydrogen peroxide generation system, either co-functionalised or as an additive in the reaction medium. The usage of *PpAOx* had been reported in the literature before and was shown to also work on a larger scale together with *rAaeUPO-H*.¹¹¹

The reaction scope looked at in this work was narrowed down to hydroxylations and sulfoxidations. Broadening of this spectrum could identify many different types of substrates and reaction types. CYPs, the closest relative in respect of the reactions catalysed by UPOs, are known to catalyse cyclopropanations, epoxidations, and many more. Completing a full comparison between the two types of enzymes might inform many future experiments. Novel research looking into replacing the CYPs heme-iron with other transition metals allowed for improved conversion rates and novel products and could potentially be applied to UPOs as well.¹⁴⁶

The crystallisation of rAaeUPO-H had been attempted with different ligands, without yielding a substrate or product complex. Future experiments could try to optimise the crystallisation conditions further to give access to such a structure. From there, it would be possible to go down the route of structure-guided evolution, which could allow for substrate-specific tailoring of the protein. The availability of more structures could also shed more light on the relationship between Glu196 and Arg169, and whether there are two separate bases involved in the mechanism. Complementary data could be achieved from site-directed mutants as well.

The expression of H-AbiUPO in a heterologous host was a successful proof-of-concept experiment. Whilst a full characterisation has not been achieved in the time of the project, it will be part of future experiments. Using the construct designed here, carrying a His-tag between the signalling peptide and the mature UPO, might also give access to further UPOs. BLAST searches revealed a large number of putative UPO sequences which could all display varying regio- and stereoselectivities amongst each other. Harnessing this potential in a screen would be valuable for a broad range of application. Apart from biocatalysts, UPOs have been used as sensors, and could even find other applications in the future.¹¹³

7. Abbreviations

%	percent
	average B-factor
<math>1/\sigma(I)>	signal to noise ratio
× g	times gravity
°C	degree Celsius
μ	growth rate, micro (10 ⁻⁶)
¹ O ₂	singlet oxygen
3CP	3-carboxy-PROXYL
³ O ₂	triplet oxygen
Å	Ångström (10 ⁻¹⁰ m)
<i>A. nidulans</i>	<i>Aspergillus nidulans</i>
<i>A. niger</i>	<i>Aspergillus niger</i>
AaeUPO	UPO from <i>Agrocybe aegerita</i>
AaeUPO-I	compound I in AaeUPO
AaeUPO-II	AaeUPO compound II
AbiUPO	UPO from <i>Agaricus bisporus</i>
AbiUPO-H	UPO from <i>Agaricus bisporus</i> with C-terminal His-tag
ABTS	2,2'-Azino-bis(3-ethylbenzthiazoline-6-sulfonic acid)
ACT	acetate
ADH1	alcohol dehydrogenase
AHAS-I	acetoxyacid synthase
ALA	5-aminolevulinic acid
Ala	alanine
ALDH1	artemisinic aldehyde dehydrogenase
amp	ampicillin
AOX1	alcohol oxidase promoter in <i>Pichia pastoris</i>
<i>apo1</i>	gene encoding for AaeUPO
APS	ammonium persulfate
Arg	arginine
AS	ammonium sulfate
Asn	asparagine
Asp	aspartic acid
ATA	transaminase
AU	absorption units
BLAST	Basic Local Alignment Search Tool
BMGY	buffered glycerol complex medium
BMMY	buffered methanol complex medium
bp	base pair

BVMO	Bayer-Villiger monooxygenases
C _{14:1}	Myristoleic acid
C _{18:1}	Oleic acid
CA	casamino acids
cam	chloramphenicol
CC _{1/2}	Pearson correlation coefficient
<i>Cci</i> UPO	UPO from <i>Coprinopsis cinerea</i>
<i>Cfu</i> CPO	CPO from <i>Caldariomyces fumago</i>
<i>Cgl</i> UPO	UPO from <i>Chaetomium globosum</i>
<i>c-myc</i>	epitope encoding EQKLISEEDL for antibody recognition
CO ₂	carbon dioxide
cpd 0	compound 0
cpd I	compound I
cpd II	compound II
CPO	chloroperoxidase
CPR	cognate reductase
<i>Cra</i> UPO	UPO from <i>Coprinellus radians</i>
CT	charge-transfer
CV	column volume
CYB5	cytochrome b ₅
CYP	cytochrome P450 monooxygenase
Cys	cysteine
Da	Dalton
dATP	deoxyadenosine triphosphate
DBT	dibenzothiophene
DdaC	dapdiamide C
DERA	2-deoxy-ribose-5-phosphate aldolase
DMP	2,6-dimethoxyphenol
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DO	dissolved oxygen
DoE	design of experiments
DTT	Dithiothreitol
e ⁻	electron
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	electrochemiluminescence
EDTA	ethylenediaminetetraacetic acid
<i>ee</i>	enantiomeric excess
eq	equivalents
EtOAc	ethyl acetate

EtOH	ethanol
F	Farad, fraction
FAD, FAD ₂	flavin adenine dinucleotide
F _c	structure factor amplitudes calculated from the model
Fe	iron
FeS	iron-sulfur cluster
FMN, FMN ₂	flavin mononucleotide
F _o	structure factors measured from the diffraction patterns
FPLC	fast protein liquid chromatography
FT	flow-through
g	grams
GC	gas chromatography
GlcNAc	<i>N</i> -acetylglucosamine
GloC	<i>trans</i> -4-proline hydroxylase
Glu	glutamic acid
gly	glycine
GSK	GlaxoSmithKline
H ⁺	proton
H ₂ O	water
H ₂ O ₂	hydrogen peroxide
H- <i>Abi</i> UPO	UPO from <i>Agaricus bisporus</i> with N-terminal His-tag
HIC	hydrophobic interaction chromatography
His ₍₆₎ -tag	hexa-histidine tag
HOX	hypohalous acid
HPOPA	2-(4-hydroxyphenoxy) propionic acid
H- <i>rAae</i> UPO	9-point <i>Aae</i> UPO mutant with a N-terminal His-tag
HRP	horseradish peroxidase
Hz	Hertz
Ile	isoleucine
IPTG	isopropyl β-D-1-thiogalactopyranoside
k	kilo (10 ³)
kan	kanamycin
<i>k</i> _{cat}	turn-over number
<i>k</i> _{cat} / <i>K</i> _m	catalytic efficiency
<i>k</i> _H / <i>k</i> _D	intermolecular isotope effect
<i>K</i> _m	Michaelis coefficient
KMO	kynurenine 3-monooxygenase
<i>k</i> _{obs}	reaction rate
KOD	<i>Thermococcus kodakaraensis</i> DNA polymerase
KP _i	potassium phosphate buffer

KRED	ketoreductase
L	litre
LB	Lysogeny broth
LC-MS	liquid chromatography–mass spectrometry
Leu	leucine
LIC	ligation independent cloning
LipL	putative TauD enzyme from the A-90289 gene cluster
m	meter, milli (10^{-3})
M	molar, mega (10^6)
M9	minimal medium
MALDI	matrix-assisted laser desorption/ionization
MALDI-TOF/TOF	matrix-assisted laser desorption/ionization-time of flight
Man	mannose
MCD	monochlorodimedone
<i>m</i> CPBA	<i>m</i> -chloroperoxybenzoic acid
MeCN	acetonitrile
MeOH	methanol
min	minutes
mPES	modified polyethersulfone
<i>Mro</i> UPO	UPO from <i>Marasmius rotula</i>
MTBE	methyl <i>t</i> -butyl ether
NAD(P) ⁺ , NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NBD	5-nitro-1,3-benzodioxole
NHI	<i>N</i> -hydroxyimidazole
n	nano (10^{-9})
NMR	nuclear magnetic resonance
O ₂	oxygen
OD ₆₀₀	optical density at 600 nm
OmpA	outer membrane protein A
<i>P. pastoris</i>	<i>Pichia pastoris</i>
P450	cytochrome P450 monooxygenase
PAH	polycyclic aromatic hydrocarbons
PCR	polymerase chain reaction
PEG	polyethylene glycol
pelB	pectate lyase B precursor
pET22b	expression vector encoding C-terminal His-tag
pETFPP31	expression vector encoding pelB and N-terminal His-tag
pETYSBLIC-3C	here: expression vector encoding cleavable N-terminal His
pH	log ₁₀ of the proton concentration
Phe	phenylalanine
pK _a	logarithmic acid constant

POPA	2-phenoxypropionic acid
<i>PpAOX</i>	alcohol oxidase from <i>Pichia pastoris</i>
pPICZ-3CHis	expression vector with cleavable N-terminal His-tag
pPICZ-His3C	expression vector with cleavable C-terminal His-tag
ppm	parts per million
Pro	proline
psi	pound per square inch
PTM ₁	fermentation trace salts
PVDF	polyvinylidene fluoride
PY	pyridine
Q	anion exchange chromatography
Q5	high-fidelity DNA polymerase
<i>rAaeUPO</i>	9-point <i>AaeUPO</i> mutant
<i>rAaeUPO-H</i>	9-point <i>AaeUPO</i> mutant with a C-terminal His tag
R	structure factor
rmsd	root-mean-square deviation
RNA	ribonucleic acid
rpm	rounds per minute
Rz	Reinheitszahl
s	second
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
Ser	serine
SOC	super optimal broth with catabolite suppression
Sub	substrate
SubO	oxygenated substrate
t	time
T	temperature
T4	DNA ligase
TAE	tris base, acetic acid, EDTA containing buffer
Taq	DNA polymerase from <i>Thermus aquaticus</i>
TBST	tris-buffered saline solution with Tween-20
t _D	doubling time
TEMED	tetramethylethylenediamine
THF	tetrahydrofuran
THF-d ₈	deuterated THF
Tyr	tyrosine
U	Unit (enzyme activity, $\mu\text{mol min}^{-1}$)
U mL ⁻¹	specific activity

UPO	unspecific peroxygenase
UV/Vis	ultraviolet–visible
V	Volt, volume
v/v	volume per volume
VA	veratryl alcohol
Val	valine
V_{\max}	maximum reaction rate
YEP	yeast extract and peptone
YNB	yeast nitrogen base
YPD	yeast extract, peptone and glucose
YPDS	yeast extract, peptone, sorbitol, and glucose
α KGDs	α -ketoglutarate-dependent dioxygenases
α -MF	α -mating factor
ϵ	extinction coefficient
λ_{\max}	wavelength of maximal absorption
Ω	Ohm

8. References

- [1] Anastas, P. T., and Warner, J. C. (1998) *Green Chemistry: Theory and Practice*, Oxford University Press, Oxford, UK.
- [2] Buchholz, K., Kasche, V., and Bornscheuer, U. T. (2012) *Biocatalysis and Enzyme Technology*, 2nd ed., Wiley-VCH, Weinheim, Germany.
- [3] Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C., and Robins, K. (2012) Engineering the third wave of biocatalysis, *Nature* **485**, 185-194.
- [4] Fowler, D. M., Araya, C. L., Fleishman, S. J., Kellogg, E. H., Stephany, J. J., Baker, D., and Fields, S. (2010) High-resolution mapping of protein sequence-function relationships, *Nat. Methods* **7**, 741-746.
- [5] Richmond, K. E., Li, M. H., Rodesch, M. J., Patel, M., Lowe, A. M., Kim, C., Chu, L. L., Venkataramaiah, N., Flickinger, S. F., Kaysen, J., Belshaw, P. J., Sussman, M. R., and Cerrina, F. (2004) Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis, *Nucleic Acids Res.* **32**, 5011-5018.
- [6] Hohne, M., Schatzle, S., Jochens, H., Robins, K., and Bornscheuer, U. T. (2010) Rational assignment of key motifs for function guides in silico enzyme identification, *Nat. Chem. Biol.* **6**, 807-813.
- [7] Turner, N. J. (2009) Directed evolution drives the next generation of biocatalysts, *Nat. Chem. Biol.* **5**, 567-573.
- [8] Choi, J. M., Han, S. S., and Kim, H. S. (2015) Industrial applications of enzyme biocatalysis: Current status and future aspects, *Biotechnol. Adv.* **33**, 1443-1454.
- [9] Reetz, M. T. (2013) Biocatalysis in organic chemistry and biotechnology: past, present, and future, *J. Am. Chem. Soc.* **135**, 12480-12496.
- [10] Rieckenberg, F., Ardao, I., Rujananon, R., and Zeng, A.-P. (2014) Cell-free synthesis of 1,3-propanediol from glycerol with a high yield, *Eng. Life Sci.* **14**, 380-386.
- [11] Wu, J., Liu, C., Jiang, Y., Hu, M., Li, S., and Zhai, Q. (2010) Synthesis of chiral epichlorohydrin by chloroperoxidase-catalyzed epoxidation of 3-chloropropene in the presence of an ionic liquid as co-solvent, *Catal. Commun.* **11**, 727-731.
- [12] Jin, H.-X., Liu, Z.-Q., Hu, Z.-C., and Zheng, Y.-G. (2013) Biosynthesis of (R)-epichlorohydrin at high substrate concentration by kinetic resolution of racemic epichlorohydrin with a recombinant epoxide hydrolase, *Eng. Life Sci.* **13**, 385-392.
- [13] Desai, A. A. (2011) Sitagliptin manufacture: a compelling tale of green chemistry, process intensification, and industrial asymmetric catalysis, *Angew. Chem. Int. Ed.* **50**, 1974-1976.
- [14] Li, T., Liang, J., Ambrogelly, A., Brennan, T., Gloor, G., Huisman, G., Lalonde, J., Lekhal, A., Mijts, B., Muley, S., Newman, L., Tobin, M., Wong, G., Zaks, A., and Zhang, X. (2012) Efficient, chemoenzymatic process for manufacture of the Boceprevir bicyclic [3.1.0]proline intermediate based on amine oxidase-catalyzed desymmetrization, *J. Am. Chem. Soc.* **134**, 6467-6472.

- [15] Huisman, G. W., Liang, J., and Krebber, A. (2010) Practical chiral alcohol manufacture using ketoreductases, *Curr. Opin. Chem. Biol.* **14**, 122-129.
- [16] Jennewein, S., Schurmann, M., Wolberg, M., Hilker, I., Luiten, R., Wubbolts, M., and Mink, D. (2006) Directed evolution of an industrial biocatalyst: 2-deoxy-D-ribose 5-phosphate aldolase, *Biotechnol. J.* **1**, 537-548.
- [17] Simon, R. C., Richter, N., Busto, E., and Kroutil, W. (2013) Recent Developments of Cascade Reactions Involving ω -Transaminases, *ACS Catal.* **4**, 129-143.
- [18] Tauber, K., Fuchs, M., Sattler, J. H., Pitzer, J., Pressnitz, D., Koszelewski, D., Faber, K., Pfeffer, J., Haas, T., and Kroutil, W. (2013) Artificial multi-enzyme networks for the asymmetric amination of sec-alcohols, *Chem. Eur. J.* **19**, 4030-4035.
- [19] Sehl, T., Hailes, H. C., Ward, J. M., Wardenga, R., von Lieres, E., Offermann, H., Westphal, R., Pohl, M., and Rother, D. (2013) Two steps in one pot: enzyme cascade for the synthesis of nor(pseudo)ephedrine from inexpensive starting materials, *Angew. Chem. Int. Ed.* **52**, 6772-6775.
- [20] Akoh, C. C., Chang, S.-W., Lee, G.-C., and Shaw, J.-F. (2008) Biocatalysis for the Production of Industrial Products and Functional Foods from Rice and Other Agricultural Produce, *J. Agric. Food Chem.* **52**, 10445-10451.
- [21] Hori, I., Nihei, K., and Kubo, I. (2004) Structural criteria for depigmenting mechanism of arbutin, *Phytother. Res.* **18**, 475-479.
- [22] Seo, D. H., Jung, J. H., Ha, S. J., Cho, H. K., Jung, D. H., Kim, T. J., Baek, N. I., Yoo, S. H., and Park, C. S. (2012) High-yield enzymatic bioconversion of hydroquinone to alpha-arbutin, a powerful skin lightening agent, by amylosucrase, *Appl. Microbiol. Biotechnol.* **94**, 1189-1197.
- [23] Yachmenev, V. G., Bertoniere, N. R., and Blanchard, E. J. (2002) Intensification of the bio-processing of cotton textiles by combined enzyme/ultrasound treatment, *J. Chem. Technol. Biotechnol.* **77**, 559-567.
- [24] Zhi Fu, G., Chan, A., and Minns, D. (2004) Preliminary Assessment of the Environmental Benefits of Enzyme Bleaching for Pulp and Paper Making (7 pp), *Int. J. Life Cycle Assess.* **10**, 136-142.
- [25] Torres Pazmino, D. E., Winkler, M., Glieder, A., and Fraaije, M. W. (2010) Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications, *J. Biotechnol.* **146**, 9-24.
- [26] Massey, V. (1994) Activation of Molecular Oxygen by Flavins and Flavoprotein, *J. Biol. Chem.* **269**, 22459-22462.
- [27] van Berkel, W. J., Kamerbeek, N. M., and Fraaije, M. W. (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts, *J. Biotechnol.* **124**, 670-689.
- [28] Huijbers, M. M., Montersino, S., Westphal, A. H., Tischler, D., and van Berkel, W. J. (2014) Flavin dependent monooxygenases, *Arch. Biochem. Biophys.* **544**, 2-17.
- [29] Schwarcz, R., Bruno, J. P., Muchowski, P. J., and Wu, H. Q. (2012) Kynurenines in the mammalian brain: when physiology meets pathology, *Nat. Rev. Neurosci.* **13**, 465-477.
- [30] Torres Pazmino, D. E., Dudek, H. M., and Fraaije, M. W. (2010) Baeyer-Villiger monooxygenases: recent advances and future challenges, *Curr. Opin. Chem. Biol.* **14**, 138-144.

- [31] Wallar, B. J., and Lipscomb, J. D. (1996) Dioxygen Activation by Enzymes Containing Binuclear Non-Heme Iron Clusters, *Chem. Rev.* **96**, 2625-2657.
- [32] Kovaleva, E. G., Neibergall, M. B., Chakrabarty, S., and Lipscomb, J. D. (2007) Finding Intermediates in the O₂ Activation Pathways of Non-Heme Iron Oxygenases, *Acc. Chem. Res.* **40**, 475-483.
- [33] Itoh, S., and Fukuzumi, S. (2007) Monooxygenase Activity of Type 3 Copper Proteins, *Acc. Chem. Res.* **40**, 592-600.
- [34] Wang, Y., Li, J., and Liu, A. (2017) Oxygen activation by mononuclear nonheme iron dioxygenases involved in the degradation of aromatics, *J. Biol. Inorg. Chem.* **22**, 395-405.
- [35] Wu, L. F., Meng, S., and Tang, G. L. (2016) Ferrous iron and alpha-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products, *Biochim. Biophys. Acta* **1864**, 453-470.
- [36] Lewis, J. C., Coelho, P. S., and Arnold, F. H. (2011) Enzymatic functionalization of carbon-hydrogen bonds, *Chem. Soc. Rev.* **40**, 2003-2021.
- [37] Holme, E. (1975) A Kinetic Study of Thymine 7-Hydroxylase from *Neurospora crassata*, *Biochemistry* **14**, 4999-5003.
- [38] Houwaart, S., Youssar, L., and Huettel, W. (2014) Pneumocandin Biosynthesis: Involvement of a trans-Selective Proline Hydroxylase, *ChemBioChem* **15**, 2365-2369.
- [39] Yang, Z., Chi, X., Funabashi, M., Baba, S., Nonaka, K., Pahari, P., Unrine, J., Jacobsen, J. M., Elliott, G. I., Rohr, J., and Van Lanen, S. G. (2011) Characterization of LipL as a non-heme, Fe(II)-dependent alpha-ketoglutarate:UMP dioxygenase that generates uridine-5'-aldehyde during A-90289 biosynthesis, *J. Biol. Chem.* **286**, 7885-7892.
- [40] Hollenhorst, M. A., Bumpus, S. B., Matthews, M. L., Bollinger, J. M., Kelleher, N. L., and Walsh, C. T. (2011) The Nonribosomal Peptide Synthetase Enzyme DdaD Tethers N-Fumaramoyl-L-2,3-diaminopropionate for Fe(II)/alpha-Ketoglutarate-Dependent Epoxidation by DdaC during Dapdiamide Antibiotic Biosynthesis, *J. Am. Chem. Soc.* **132**, 15773-15781.
- [41] Green, M. T. (2009) C-H bond activation in heme proteins: the role of thiolate ligation in cytochrome P450, *Curr. Opin. Chem. Biol.* **13**, 84-88.
- [42] van Deurzen, M. P. J., van Rantwijk, F., and Sheldon, R. A. (1997) Selective Oxidations Catalyzed By Peroxidases, *Tetrahedron* **53**, 13183-13220.
- [43] Vaillancourt, F. H., Yeh, E., Vosburg, D. A., Garneau-Tsodikova, S., and Walsh, C. T. (2006) Nature's Inventory of Halogenation Catalysts: Oxidative Strategies Predominate, *Chem. Rev.* **106**, 3364-3378.
- [44] Groves, J. T. (2003) The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies, *Proc. Natl. Acad. Sci.* **100**, 3569-3574.
- [45] Eickhoff, H., Jung, G., and Rieker, A. (2001) Oxidative phenol coupling - tyrosine dimers and libraries containing tyrosyl peptide dimers, *Tetrahedron* **57**, 353-364.
- [46] Hofrichter, M., and Ullrich, R. (2006) Heme-thiolate haloperoxidases: versatile biocatalysts with biotechnological and environmental significance, *Appl. Microbiol. Biotechnol.* **71**, 276-288.

- [47] Guengerich, F. P. (2001) Common and Uncommon Cytochrome P450 Reactions Related to Metabolism and Chemical Toxicity, *Chem. Res. Toxicol.* **14**, 611-650.
- [48] Hofrichter, M., and Ullrich, R. (2014) Oxidations catalyzed by fungal peroxygenases, *Curr. Opin. Chem. Biol.* **19**, 116-125.
- [49] Zamocky, M., Hofbauer, S., Schaffner, I., Gasselhuber, B., Nicolussi, A., Soudi, M., Pirker, K. F., Furtmuller, P. G., and Obinger, C. (2015) Independent evolution of four heme peroxidase superfamilies, *Arch. Biochem. Biophys.* **574**, 108-119.
- [50] Morris, D. R., and Hager, L. P. (1966) Chloroperoxidase. I. Isolation and properties of the crystalline glycoprotein, *J. Biol. Chem.* **241**, 1763-1768.
- [51] Sundaramoorthy, M., Terner, J., and Poulos, T. L. (1995) The crystal structure of chloroperoxidase: a heme peroxidase-cytochrome P450 functional hybrid, *Structure* **3**, 1367-1377.
- [52] Hofrichter, M., Ullrich, R., Pecyna, M. J., Liers, C., and Lundell, T. (2010) New and classic families of secreted fungal heme peroxidases, *Appl. Microbiol. Biotechnol.* **87**, 871-897.
- [53] Wagenknecht, H.-A., and Woggon, W.-D. (1997) Identification of intermediates in the catalytic cycle of chloroperoxidase, *Chem. Biol.* **4**, 367-372.
- [54] Kuhnel, K., Blankenfeldt, W., Terner, J., and Schlichting, I. (2006) Crystal structures of chloroperoxidase with its bound substrates and complexed with formate, acetate, and nitrate, *J. Biol. Chem.* **281**, 23990-23998.
- [55] Pickard, M. A., and Hashimoto, A. (1982) Isoenzymes of chloroperoxidase from *Caldariomyces fumago*, *Can. J. Microbiol.* **28**, 1382-1388.
- [56] Conesa, A., Punt, P. J., and van den Hondel, C. A. (2002) Fungal peroxidases: molecular aspects and applications, *J. Biotechnol.* **93**, 143-158.
- [57] Park, J. B., and Clark, D. S. (2006) Deactivation mechanisms of chloroperoxidase during biotransformations, *Biotechnol. Bioeng.* **93**, 1190-1195.
- [58] Bernhardt, R., and Urlacher, V. B. (2014) Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations, *Appl. Microbiol. Biotechnol.* **98**, 6185-6203.
- [59] O'Reilly, E., Kohler, V., Flitsch, S. L., and Turner, N. J. (2011) Cytochromes P450 as useful biocatalysts: addressing the limitations, *Chem. Commun.* **47**, 2490-2501.
- [60] Fasan, R. (2012) Tuning P450 Enzymes as Oxidation Catalysts, *ACS Catal.* **2**, 647-666.
- [61] Hrycay, E. G., and Bandiera, S. M. (2012) The monooxygenase, peroxidase, and peroxygenase properties of cytochrome P450, *Arch. Biochem. Biophys.* **522**, 71-89.
- [62] Shoji, O., and Watanabe, Y. (2014) Peroxygenase reactions catalyzed by cytochromes P450, *J. Biol. Inorg. Chem.* **19**, 529-539.
- [63] Rabe, K. S., Kiko, K., and Niemeyer, C. M. (2008) Characterization of the peroxidase activity of CYP119, a thermostable P450 from *Sulfolobus acidocaldarius*, *Chembiochem* **9**, 420-425.
- [64] Poulos, T. L., and Johnson, E. R. (2005) Structures of Cytochrome P450 Enzymes, In *Cytochrome P450: Structure, Mechanism, and Biochemistry* (Ortiz de Montellano, P. R., Ed.) 3rd ed., Kluwer Academic / Plenum Publishers, New York, USA.

- [65] Poulos, T. L., Finzel, B. C., Gunsalus, I. C., Wagner, G. C., and Kraut, J. (1985) The 2.6 Å Crystal Structure of *Pseudomonas putida* Cytochrome P450, *J. Biol. Chem.* **260**, 16122-16130.
- [66] Poulos, T. L., Finzel, B. C., and Howard, A. J. (1987) High-resolution Crystal Structure of Cytochrome P450cam, *J. Mol. Biol.* **195**, 687-700.
- [67] Whitehouse, C. J., Bell, S. G., and Wong, L. L. (2012) P450(BM3) (CYP102A1): connecting the dots, *Chem. Soc. Rev.* **41**, 1218-1260.
- [68] Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievens, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., and Newman, J. D. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin, *Nature* **496**, 528-532.
- [69] Ullrich, R., Nuske, J., Scheibner, K., Spantzel, J., and Hofrichter, M. (2004) Novel haloperoxidase from the agaric basidiomycete *Agrocybe aegerita* oxidizes aryl alcohols and aldehydes, *Appl. Environ. Microbiol.* **70**, 4575-4581.
- [70] Piontek, K., Strittmatter, E., Ullrich, R., Grobe, G., Pecyna, M. J., Kluge, M., Scheibner, K., Hofrichter, M., and Plattner, D. A. (2013) Structural basis of substrate conversion in a new aromatic peroxygenase: cytochrome P450 functionality with benefits, *J. Biol. Chem.* **288**, 34767-34776.
- [71] Kluge, M., Ullrich, R., Dolge, C., Scheibner, K., and Hofrichter, M. (2009) Hydroxylation of naphthalene by aromatic peroxygenase from *Agrocybe aegerita* proceeds via oxygen transfer from H₂O₂ and intermediary epoxidation, *Appl. Microbiol. Biotechnol.* **81**, 1071-1076.
- [72] Samanta, S. K., Singh, O. V., and Jain, R. K. (2002) Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation, *Trends Biotechnol.* **20**, 243-248.
- [73] Kastner, M., Streibich, S., Beyrer, M., Richnow, H. H., and Fritsche, D. W. (1999) Formation of Bound Residues during Microbial Degradation of [14C]Anthracene in Soil, *Appl. Environ. Microbiol.* **65**, 1834-1842.
- [74] Pecyna, M. J., Ullrich, R., Bittner, B., Clemens, A., Scheibner, K., Schubert, R., and Hofrichter, M. (2009) Molecular characterization of aromatic peroxygenase from *Agrocybe aegerita*, *Appl. Microbiol. Biotechnol.* **84**, 885-897.
- [75] Omura, T. (2005) Heme-thiolate proteins, *Biochem. Biophys. Res. Commun.* **338**, 404-409.
- [76] Ullrich, R., and Hofrichter, M. (2005) The haloperoxidase of the agaric fungus *Agrocybe aegerita* hydroxylates toluene and naphthalene, *FEBS Lett.* **579**, 6247-6250.
- [77] Anh, D. H., Ullrich, R., Benndorf, D., Svatos, A., Muck, A., and Hofrichter, M. (2007) The coprophilous mushroom *Coprinus radians* secretes a haloperoxidase that catalyzes aromatic peroxygenation, *Appl. Environ. Microbiol.* **73**, 5477-5485.

- [78] Grobe, G., Ullrich, R., Pecyna, M. J., Kapturska, D., Friedrich, S., Hofrichter, M., and Scheibner, K. (2011) High-yield production of aromatic peroxygenase by the agaric fungus *Marasmius rotula*, *AMB Express* 1, 31.
- [79] Babot, E. D., del Rio, J. C., Kalum, L., Martinez, A. T., and Gutierrez, A. (2013) Oxyfunctionalization of aliphatic compounds by a recombinant peroxygenase from *Coprinopsis cinerea*, *Biotechnol. Bioeng.* 110, 2323-2332.
- [80] Bormann, S., Gomez Baraibar, A., Ni, Y., Holtmann, D., and Hollmann, F. (2015) Specific oxyfunctionalisations catalysed by peroxygenases: opportunities, challenges and solutions, *Catal. Sci. Technol.* 5, 2038-2052.
- [81] Wang, X., Peter, S., Kinne, M., Hofrichter, M., and Groves, J. T. (2012) Detection and kinetic characterization of a highly reactive heme-thiolate peroxygenase compound I, *J. Am. Chem. Soc.* 134, 12897-12900.
- [82] Piontek, K., Ullrich, R., Liers, C., Diederichs, K., Plattner, D. A., and 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 Cryst. F* 66, 693-698.
- [83] Wang, X., Ullrich, R., Hofrichter, M., and Groves, J. T. (2015) Heme-thiolate ferryl of aromatic peroxygenase is basic and reactive, *PNAS* 112, 3686-3691.
- [84] Wang, X., Peter, S., Ullrich, R., Hofrichter, M., and Groves, J. T. (2013) Driving force for oxygen-atom transfer by heme-thiolate enzymes, *Angew. Chem. Int. Ed.* 52, 9238-9241.
- [85] Groves, J. T. (2014) Enzymatic C-H bond activation: Using push to get pull, *Nat. Chem.* 6, 89-91.
- [86] Kornfeld, R., and Kornfeld, S. (1985) Assembly of Asparagine-linked Oligosaccharides, *Ann. Rev. Biochem.* 54, 631-664.
- [87] Choinowski, T., Blodig, W., Winterhalter, K. H., and Piontek, K. (1999) The Crystal Structure of Lignin Peroxidase at 1.70 Å Resolution Reveals a Hydroxy Group on the C β of Tryptophan 171: A Novel Radical Site Formed During the Redox Cycle, *J. Mol. Biol.* 286, 809-827.
- [88] Aranda, E., Ullrich, R., and Hofrichter, M. (2010) Conversion of polycyclic aromatic hydrocarbons, methyl naphthalenes and dibenzofuran by two fungal peroxygenases, *Biodegradation* 21, 267-281.
- [89] Kinne, M., Ullrich, R., Hammel, K. E., Scheibner, K., and Hofrichter, M. (2008) Regioselective preparation of (R)-2-(4-hydroxyphenoxy)propionic acid with a fungal peroxygenase, *Tetrahedron Lett.* 49, 5950-5953.
- [90] Kinne, M., Poraj-Kobielska, M., Aranda, E., Ullrich, R., Hammel, K. E., Scheibner, K., and Hofrichter, M. (2009) Regioselective preparation of 5-hydroxypropranolol and 4'-hydroxydiclofenac with a fungal peroxygenase, *Bioorg. Med. Chem. Lett.* 19, 3085-3087.
- [91] Poraj-Kobielska, M., Kinne, M., Ullrich, R., Scheibner, K., Kayser, G., Hammel, K. E., and Hofrichter, M. (2011) Preparation of human drug metabolites using fungal peroxygenases, *Biochem. Pharmacol.* 82, 789-796.
- [92] Otey, C. R., Bandara, G., Lalonde, J., Takahashi, K., and Arnold, F. H. (2006) Preparation of human metabolites of propranolol using laboratory-evolved bacterial cytochromes P450, *Biotechnol. Bioeng.* 93, 494-499.

- [93] Barková, K., Kinne, M., Ullrich, R., Hennig, L., Fuchs, A., and Hofrichter, M. (2011) Regioselective hydroxylation of diverse flavonoids by an aromatic peroxygenase, *Tetrahedron* 67, 4874-4878.
- [94] Churakova, E., Kluge, M., Ullrich, R., Arends, I., Hofrichter, M., and Hollmann, F. (2011) Specific photobiocatalytic oxyfunctionalization reactions, *Angew. Chem. Int. Ed.* 50, 10716-10719.
- [95] Kluge, M., Ullrich, R., Scheibner, K., and Hofrichter, M. (2012) Stereoselective benzylic hydroxylation of alkylbenzenes and epoxidation of styrene derivatives catalyzed by the peroxygenase of *Agrocybe aegerita*, *Green Chem.* 14, 440.
- [96] Gutierrez, A., Babot, E. D., Ullrich, R., Hofrichter, M., Martinez, A. T., and 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.
- [97] Peter, S., Karich, A., Ullrich, R., Gröbe, G., Scheibner, K., and Hofrichter, M. (2014) Enzymatic one-pot conversion of cyclohexane into cyclohexanone: Comparison of four fungal peroxygenases, *J. Mol. Catal. B Enzym.* 103, 47-51.
- [98] Babot, E. D., del Río, J. C., Kalum, L., Martínez, A. T., and Gutiérrez, A. (2015) Regioselective Hydroxylation in the Production of 25-Hydroxyvitamin D by *Coprinopsis cinerea* Peroxygenase, *Chemcatchem* 7, 283-290.
- [99] Kiebitz, J., Schmidtke, K. U., Zimmermann, J., Kellner, H., Jehmlich, N., Ullrich, R., Zander, D., Hofrichter, M., and Scheibner, K. (2017) A Peroxygenase from *Chaetomium globosum* Catalyzes the Selective Oxygenation of Testosterone, *Chembiochem* 18, 563-569.
- [100] Lucas, F., Babot, E. D., Cañellas, M., del Río, J. C., Kalum, L., Ullrich, R., Hofrichter, M., Guallar, V., Martínez, A. T., and Gutiérrez, A. (2016) Molecular determinants for selective C25-hydroxylation of vitamins D2 and D3 by fungal peroxygenases, *Catal. Sci. Technol.* 6, 288-295.
- [101] Ullrich, R., Dolge, C., Kluge, M., and Hofrichter, M. (2008) Pyridine as novel substrate for regioselective oxygenation with aromatic peroxygenase from *Agrocybe aegerita*, *FEBS Lett.* 582, 4100-4106.
- [102] Colby, J., Stirling, D. I., and Dalton, H. (1977) The Soluble Methane Monooxygenase of *Methylococcus capsulatus* (Bath): Its ability to oxygenate n-alkanes, n-alkenes, ethers, and alicyclic, aromatic and heterocyclic compounds, *Biochem. J.* 165, 395-402.
- [103] Hlavica, P., Mietaschk, J., and Baden, I. (1982) Interaction of ligands with cytochrome P-450. On the 442 nm spectral species generated during the oxidative metabolism of pyridine, *Biochem. J.* 204, 425-432.
- [104] Aranda, E., Kinne, M., Kluge, M., Ullrich, R., and Hofrichter, M. (2009) Conversion of dibenzothiophene by the mushrooms *Agrocybe aegerita* and *Coprinellus radians* and their extracellular peroxygenases, *Appl. Microbiol. Biotechnol.* 82, 1057-1066.
- [105] Kluge, M., Ullrich, R., Scheibner, K., and Hofrichter, M. (2014) Formation of naphthalene hydrates in the enzymatic conversion of 1,2-dihydronaphthalene by two fungal peroxygenases and subsequent naphthalene formation, *J. Mol. Catal. B Enzym.* 103, 56-60.

- [106] Peter, S., Kinne, M., Ullrich, R., Kayser, G., and Hofrichter, M. (2013) Epoxidation of linear, branched and cyclic alkenes catalyzed by unspecific peroxygenase, *Enzyme Microb. Technol.* **52**, 370-376.
- [107] Kinne, M., Poraj-Kobielska, M., Ralph, S. A., Ullrich, R., Hofrichter, M., and Hammel, K. E. (2009) Oxidative cleavage of diverse ethers by an extracellular fungal peroxygenase, *J. Biol. Chem.* **284**, 29343-29349.
- [108] Kiebist, J., Holla, W., Heidrich, J., Poraj-Kobielska, M., Sandvoss, M., Simonis, R., Gröbe, G., Atzrodt, J., Hofrichter, M., and Scheibner, K. (2015) One-pot synthesis of human metabolites of SAR548304 by fungal peroxygenases, *Bioorg. Med. Chem.* **23**, 4324-4332.
- [109] Poraj-Kobielska, M., Kinne, M., Ullrich, R., Scheibner, K., and Hofrichter, M. (2012) A spectrophotometric assay for the detection of fungal peroxygenases, *Anal. Biochem.* **421**, 327-329.
- [110] Horst, A. E. W., Bormann, S., Meyer, J., Steinhagen, M., Ludwig, R., Drews, A., Ansorge-Schumacher, M., and Holtmann, D. (2016) Electro-enzymatic hydroxylation of ethylbenzene by the evolved unspecific peroxygenase of *Agroclybe aegerita*, *J. Mol. Catal. B Enzym.*, in press.
- [111] Ni, Y., Fernandez-Fueyo, E., Baraibar, A. G., Ullrich, R., Hofrichter, M., Yanase, H., Alcalde, M., van Berkel, W. J., and Hollmann, F. (2016) Peroxygenase-Catalyzed Oxyfunctionalization Reactions Promoted by the Complete Oxidation of Methanol, *Angew. Chem. Int. Ed.* **55**, 798-801.
- [112] Wu, Y., Wollenberger, U., Hofrichter, M., Ullrich, R., Scheibner, K., and Scheller, F. W. (2011) Direct electron transfer of *Agroclybe aegerita* peroxygenase at electrodes modified with chitosan-capped Au nanoparticles and its bioelectrocatalysis to aniline, *Sens. Actuators B Chem.* **160**, 1419-1426.
- [113] Yarman, A., Grobe, G., Neumann, B., Kinne, M., Gajovic-Eichelmann, N., Wollenberger, U., Hofrichter, M., Ullrich, R., Scheibner, K., and Scheller, F. W. (2012) The aromatic peroxygenase from *Marasmius rutola*--a new enzyme for biosensor applications, *Anal. Bioanal. Chem.* **402**, 405-412.
- [114] Pappa, H. S., and Cass, A. E. G. (1993) A step towards understanding the folding mechanism of horseradish peroxidase, *Eur. J. Biochem.* **212**, 227-235.
- [115] Lindwall, G., Cahu, M.-F., Gardner, S. R., and Kohlstaedt, L. A. (2000) A sparse matrix approach to the solubilization of overexpressed proteins, *Protein Eng.* **13**, 67-71.
- [116] Wang, X. (2013) A Novel Heme-Thiolate Peroxygenase AaeUPO and its Implications for C-H Activation Chemistry, In *Chemistry*, Princeton University, Princeton NJ.
- [117] Zong, Q., Osmulski, P. A., and Hager, L. P. (1995) High-Pressure-Assisted Reconstitution of Recombinant Chloroperoxidase, *Biochemistry* **34**, 12420-12424.
- [118] Conesa, A., van den Hondel, C. A., and Punt, P. J. (2000) Studies on the Production of Fungal Peroxidases in *Aspergillus niger*, *Appl. Environ. Microbiol.* **66**, 3016-3023.
- [119] Conesa, A., van de Velde, F., van Rantwijk, F., Sheldon, R. A., van den Hondel, C. A., and Punt, P. J. (2001) Expression of the *Caldariomyces fumago* chloroperoxidase in *Aspergillus niger* and characterization of the recombinant enzyme, *J. Biol. Chem.* **276**, 17635-17640.

- [120] Molina-Espeja, P., Garcia-Ruiz, E., Gonzalez-Perez, D., Ullrich, R., Hofrichter, M., and Alcalde, M. (2014) Directed Evolution of Unspecific Peroxygenase from *Agrocybe aegerita*, *Appl. Environ. Microbiol.* *80*, 3496-3507.
- [121] Kabsch, W. (2010) XDS, *Acta Cryst. D* *66*, 125-132.
- [122] Evans, P. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics, *Acta Cryst. D* *67*, 282-292.
- [123] Evans, P. (2006) Scaling and assessment of data quality, *Acta Cryst. D* *62*, 72-82.
- [124] Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?, *Acta Cryst. D* *69*, 1204-1214.
- [125] Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction, *J. Appl. Crystallogr.* *43*, 186-190.
- [126] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J. Appl. Crystallogr.* *40*, 658-674.
- [127] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Cryst. D* *66*, 486-501.
- [128] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Cryst. D* *67*, 235-242.
- [129] Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures, *Acta Cryst. D* *67*, 355-367.
- [130] Schuttelkopf, A. W., and van Aalten, D. M. F. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Cryst. D* *60*, 1355-1363.
- [131] Agirre, J., Davies, G., Wilson, K., and Cowtan, K. (2015) Carbohydrate anomalies in the PDB, *Nat. Chem. Biol.* *11*, 303.
- [132] Agirre, J., Iglesias-Fernández, J., Rovira, C., Davies, G. J., Wilson, K. S., and Cowtan, K. D. (2015) Privateer: software for the conformational validation of carbohydrate structures, *Nat. Struct. Mol. Biol.* *22*, 833.
- [133] Molina-Espeja, P., Ma, S., Mate, D. M., Ludwig, R., and Alcalde, M. (2015) Tandem-yeast expression system for engineering and producing unspecific peroxygenase, *Enzyme Microb. Technol.* *73-74*, 29-33.
- [134] Higgins, D. R., Busser, K., Comiskey, J., Whittier, P. S., Purcell, T. J., and Hoeffler, J. P. (1998) Small Vectors for Expression Based on Dominant Drug Resistance with Direct Multicopy Selection, In *Pichia Protocols* (Higgins, D. R., and Cregg, J. M., Eds.), pp 41-53, Humana Press, Totowa, NJ.
- [135] Salamin, K., Sriranganadane, D., Lechenne, B., Jousson, O., and Monod, M. (2010) Secretion of an endogenous subtilisin by *Pichia pastoris* strains GS115 and KM71, *Appl. Environ. Microbiol.* *76*, 4269-4276.
- [136] Sinha, J., Plantz, B. A., Inan, M., and Meagher, M. M. (2005) Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia pastoris*: case study with recombinant ovine interferon-tau, *Biotechnol. Bioeng.* *89*, 102-112.

- [137] Invitrogen. (2002) Pichia Fermentation Process Guidelines.
- [138] Werrten, M. W. T., van den Bosch, T. J., Wind, R. D., Mooibroek, H., and De Wolf, F. A. (1999) High-yield Secretion of Recombinant Gelatins by *Pichia pastoris*, *Yeast* 15, 1087-1096.
- [139] Krissinel, E. (2012) Enhanced fold recognition using efficient short fragment clustering, *J. Mol. Biochem.* 1, 76-85.
- [140] Clare, J. J., Romanos, M. A., Rayment, F. B., Rowedder, J. E., Smith, M. A., Payne, M. M., Sreekrishna, K., and Henwood, C. A. (1991) Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies, *Gene* 105, 205-212.
- [141] Olmedo, A., Aranda, C., del Río, J. C., Kiebish, J., Scheibner, K., Martínez, A. T., and Gutierrez, A. (2016) From Alkanes to Carboxylic Acids: Terminal Oxygenation by a Fungal Peroxygenase, *Angew. Chem. Int. Ed.* 55, 12248-12251.
- [142] Camarero, S., Pardo, I., Canas, A. I., Molina, P., Record, E., Martinez, A. T., Martinez, M. J., and Alcalde, M. (2011) Engineering Platforms for Directed Evolution of Laccase from *Pycnoporus cinnabarinus*, *Appl. Environ. Microbiol.* 78, 1370-1384.
- [143] Mate, D., Garcia-Burgos, C., Garcia-Ruiz, E., Ballesteros, A. O., Camarero, S., and Alcalde, M. (2010) Laboratory evolution of high-redox potential laccases, *Chem. Biol.* 17, 1030-1041.
- [144] Jensen, C. N., Cartwright, J., Ward, J., Hart, S., Turkenburg, J. P., Ali, S. T., Allen, M. J., and Grogan, G. (2012) A flavoprotein monooxygenase that catalyses a Baeyer-Villiger reaction and thioether oxidation using NADH as the nicotinamide cofactor, *Chembiochem* 13, 872-878.
- [145] Piontek, K., Ullrich, R., Liers, C., Diederichs, K., Plattner, D. A., and Hofrichter, M. (2010) Crystallization of a 45 kDa peroxygenase/oxidase from the mushroom *Agrocybe aegerita* and structure determination by SAD utilizing only the haem iron, *Acta Cryst. F* 66, 693-698.
- [146] Key, H. M., Dydio, P., Clark, D. S., and Hartwig, J. F. (2016) Abiological catalysis by artificial haem proteins containing noble metals in place of iron, *Nature* 534, 534-537.